Efficient and Scalable Spatial Regularization of Optimal Transport

LUCAS BRIFAULT, Inria Saclay, France
DAVID COHEN-STEINER, Inria, Université Côte d'Azur, France
MATHIEU DESBRUN, Inria Saclay / Ecole Polytechnique, France

In this paper, we introduce a novel approach to spatial regularization of optimal transport problems. Based on the notion of forward and backward "mean maps" of a transport plan, we introduce a convex formulation of optimal transport problems that incorporates regularization of these mean maps to promote spatial continuity of the resulting optimal plan. Unlike previous regularization approaches that required the optimization of all the transport plan coefficients, our formulation translates into an ADMM-based solver combined with Sinkhorn type algorithms, which drastically reduces the number of variables and scales up to large problems. We demonstrate the usefulness and efficiency of this new computational tool for various applications and for different regularizations.

CCS Concepts: • **Applied computing** → **Transportation**; • **Computing methodologies** → **Optimization algorithms**; **Matching**.

Additional Key Words and Phrases: Optimal transport, convex optimization, diffeomorphic registration.

ACM Reference Format:
Lucas Brifault, David Cohen-Steiner, and Mathieu Desbrun. 2025. Efficient and Scalable Spatial Regularization of Optimal Transport. In *SIGGRAPH Asia 2025 Conference Papers (SA Conference Papers '25), December 15–18, 2025, Hong Kong, Hong Kong*. ACM, New York, NY, USA, 10 pages. https://doi.org/10.1145/3757377.3763976

1 Introduction

The term *optimal transport* (OT) refers to the generic problem of optimizing the cost of transportation and allocation of resources [Villani 2008]. A practical case of optimal transport, mentioned by Monge in the 18th century, consists in determining the most effective way to move a pile of sand into a hole of the same volume, where "most effective" means that the integral of the distances by which the sand is moved (one infinitesimal unit at a time) is minimal. This form of the problem is referred to as Monge's variational formulation and assumes that the sand is moved through a *point-to-point* map called a Monge transport map. This restriction was relaxed by Kantorovich who extended the formulation to essentially allow mass splitting, thus generalizing point-to-point maps into *transport plans* between two probability measures μ and ν, whose set is now convex. In this formulation, a transport plan π is understood as a *probability measure* on a product space whose marginals are μ and ν, and $\pi(x, y)$ specifies the amount of mass transferred from x to y. Specifying a transportation cost $c(x, y)$ between x and y naturally leads to the notion of *optimal* transport, identifying a transport plan with *minimum* total cost which, given weak hypotheses on c, μ and ν, is guaranteed to exist given the convex nature of the problem. With its solid foundations and wide generality, the optimal transport framework has found a large amount of applications in, for instance, machine learning and vision. Graphics has also widely adopted it [de Goes et al. 2011; Lipman and Daubechies 2011; Digne et al. 2014; Solomon et al. 2015; de Goes et al. 2015], partly because optimal transport costs are inherently robust to noise and outliers.

Fig. 1. **Spatially-regularized optimal transport.** We present a convex optimal transport formulation using a mean-map based spatial regularization that can be used for, e.g., surface-to-surface registration. Here, the registration between a dog mesh (left) and a camel mesh (right) results in the deformed dog mesh in the shape of a camel shape (middle) in what amounts to only three calls to Sinkhorn's algorithm.

Computing an OT problem seemingly involves the resolution of a linear program (LP) whose cost is prohibitive for even moderately-sized inputs. Cuturi [2013] proposed an *entropic regularization* which allows the use of an efficient iterated approach (*Sinkhorn's algorithm*) to solve the resulting (now strongly) convex minimization problem often orders of magnitude faster than pure LP solvers, and with good theoretical guarantees of convergence. Additional computational methods were further developed to evaluate optimal transport plans and costs (or approximations and variants thereof) efficiently [Bonneel et al. 2015; Chizat et al. 2018; Bonneel and Coeurjolly 2019; Feydy 2020; Do et al. 2025] — see the recent surveys of Peyré and Cuturi [2019] and Bonneel and Digne [2023].

Entropic regularization of optimal transport allows soft correspondences between discrete measures and is adaptable to a great variety of cost functions, based for instance on geodesic distances in non-flat spaces [Solomon et al. 2015]. Yet, the variational nature of optimal transport alone *does not promote any sort of spatial regularity of the transport plan*. This is a severe limitation in a number of applications: this means that not only a transport plan between two shapes can involve a large amount of mass splitting (i.e., a mass point on the source can *spread* widely over the target), but also that two nearby points on the source measure can be sent to far away regions on the target geometry (i.e., *discontinuity* of the plan). In both cases, deriving some useful form of non-fuzzy maps from such optimal transport plans is challenging, preventing a wider adoption of transport in practice. Surprisingly, only very few approaches

Authors' Contact Information: Lucas Brifault, lucas.brifault@3ds.com, Inria Saclay, Palaiseau, France; David Cohen-Steiner, david.cohen-steiner@inria.fr, Inria, Université Côte d'Azur, Valbonne, France; Mathieu Desbrun, mathieu.desbrun@inria.fr, Inria Saclay / Ecole Polytechnique, Palaiseau, France.

SA Conference Papers '25, Hong Kong, Hong Kong
© 2025 Copyright held by the owner/author(s).
This is the author's version of the work. It is posted here for your personal use. Not for redistribution. The definitive Version of Record was published in *SIGGRAPH Asia 2025 Conference Papers (SA Conference Papers '25), December 15–18, 2025, Hong Kong, Hong Kong*, https://doi.org/10.1145/3757377.3763976.

have been formulated to address spatial regularization, and they require either high computational costs [Solomon et al. 2012, 2013] or non-convex optimizations [Mandad et al. 2017].

Overview. In this work, we introduce a new means of spatial regularization for the optimal transport problem based on any generic cost (based on Euclidean distances, landmarks, or feature descriptors). We show that our spatially-regularized optimal transport formulation is convex, and amenable to efficient and scalable Sinkhorn-like solvers. We demonstrate the usefulness and versatility of this new computational tool by performing volumetric and surface-to-surface shape matching via a Hessian-based spatial regularization.

2 Prior Art

Optimal transport, along with the measure theory it builds upon, have been extensively studied and used in a large array of computational sciences — we provide the necessary background on the foundations of these two topics in App. A of our Supplemental Material. In marked contrast, *spatial regularization* of optimal transport has been given limited attention, be it theoretically or numerically, as we review next to motivate and highlight our contributions.

Convex regularization. A first effort to regularize transport plans in graphics was proposed by Solomon et al. [2012]. They proposed a convex formulation enforcing that nearby points in $\mathcal{X}$ map to nearby distributions in $\mathcal{Y}$ in the sense of the (geodesic) earth movers distance (i.e., 1-Wasserstein distance). As a result, they showed how to compute a "soft map" between a pair of surfaces through a *convex* optimization trading off between continuity, descriptor matching, and spread of the transport plan. A spatial regularization of this first approach through a form of *Dirichlet energy* for measure-based maps (based on 2-Wasserstein distances this time) was later introduced by Solomon et al. [2013]. Their energy, akin to a functional introduced by Brenier [2003], was further studied by Ghoussoub et al. [2020] in an elasticity context where the map is fixed on the boundary of the domain. While this assumption is too restrictive for practical transport problems, Ghoussoub et al. [2020] managed to prove that, in this case, the energy is a *tight* convex relaxation of the Dirichlet energy defined for deterministic maps. The approach used in these two latter works can theoretically be extended to any convex functional of the mapping's derivatives — although the tightness claim seems to be specific to the Dirichlet functional. This may still seem appealing as a generic regularization of transport plans, until one realizes that the numerical implementation of such an approach requires to perform convex optimization on a *quadratic number of variables* w.r.t. the given discrete measures, making it impractical beyond toy models. Despite their convexity, these approaches are thus inappropriate to induce regularization in practice.

Non-convex regularization. In a similar vein, Mandad et al. [2017] proposed to measure the regularity of a transport plan via its local variance, introducing a variance-minimizing energy that favors deterministic maps as opposed to generic transport plans. They proved that the resulting regularized optimal transport minimizes a generalized Dirichlet energy for both the deterministic map and its inverse, allowing the computation of as-conformal-as-possible inter-surface maps satisfying user-prescribed bounds on area distortion. Moreover, the authors implemented their approach as a biconvex problem where iterations of optimal transport are handled via Sinkhorn's algorithm. Alas, their variance-minimizing formulation is highly non-convex, thus requiring complex coarse to fine strategies to avoid getting trapped in local minima and with slow convergence on models with a wide range of feature sizes. Despite the appeal of the expected smooth and bijective maps, it does not offer a practical and robust approach to spatial regularization of transport plans either.

Our contributions at a glance. In this work, we point out that any transport plan canonically defines forward and backward *mean maps*: from a point on the source, one can associate on the target the *barycenter* of the distribution towards which the transport map carries the source point, thus forming a *forward* mean map (see Fig. 2); and vice-versa for the *backward* mean map between the target and the source. Being actual (point-to-point) maps, these mean maps can be easily *regularized* through any deformation functional. Since mean maps form a convex set, we will show how to devise an efficient *convex formulation* of spatially-regularized optimal transport. Our approach is very generic in that it can use *any convex deformation functional* and *any cost function*. Crucially, our approach is amenable to Sinkhorn-type algorithms, allowing the efficient computation of spatially-regularized transport plans on large models — a property that previous spatial regularization techniques did not enjoy. We demonstrate both the versatility and efficiency of our approach through a variety of examples that our efficient spatial regularization of optimal transport directly induces via regularized forward and backward point-to-point maps.

3 Mean Maps of Transport Plans and their Regularization

We now introduce the notion of "mean maps" of a transport plan π between probability spaces $(\mathcal{X}, \mu)$ and $(\mathcal{Y}, \nu)$ with $\mathcal{X}, \mathcal{Y} \subset \mathbb{R}^d$, and review their key properties. We then discuss how to regularize them to induce continuity of transport plans.

3.1 Mean maps of general transport plans

Given a transport plan $\pi \in \Pi(\mu, \nu)$, we know from the disintegration theorem [Bogachev 2007] that there exists a family of probability measures $(\pi_x)_{x \in \mathcal{X}}$ on $\mathcal{Y}$ (uniquely defined μ-everywhere) such that, for any π-integrable function $f : \mathcal{X} \times \mathcal{Y} \to \mathbb{R}$,

$$\int_{\mathcal{X}\times\mathcal{Y}} f \, d\pi \;=\; \int_{\mathcal{X}} \int_{\mathcal{Y}} f(x,y) \, d\pi_x(y) \, d\mu(x) \tag{1}$$

With these π_x, we can thus define what we will refer to as the *forward mean map* $\overrightarrow{m}_\pi : \mathcal{X} \to \mathbb{R}^d$ (see Fig. 2 for intuition) through:

$$\overrightarrow{m}_\pi(x) \;:=\; \int_{\mathcal{Y}} y \, d\pi_x(y). \tag{2}$$

A forward mean map is well-defined μ-everywhere, and an equivalent definition (less intuitive but useful for computational purposes) is that $\overrightarrow{m}_\pi$ is the unique map in $\mathcal{F}_{\mathcal{X}} = L^2(\mathcal{X}, \mathbb{R}^d; \mu)$ satisfying:

$$\int_{\mathcal{X}} \overrightarrow{m}_\pi \cdot u \, d\mu = \int_{\mathcal{X}\times\mathcal{Y}} u(x) \cdot y \, d\pi(x,y) \qquad \forall u \in \mathcal{F}_{\mathcal{X}}. \tag{3}$$

Reversely, the *backward mean map* can be defined by simply inverting the roles of $(\mathcal{X}, \mu)$ and $(\mathcal{Y}, \nu)$, through $\overleftarrow{m}_\pi(y) := \int_{\mathcal{X}} x \, d\pi_y(x)$.

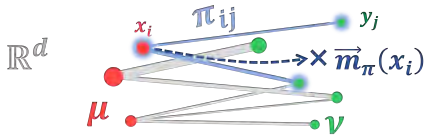


Fig. 2. **Mean maps.** We show the intuition behind a forward mean map with this simple example: from a transport plan π (between red and green masses) bringing the mass at $\mathbf{x}_i$ to two target masses $\mathbf{y}_j$, the image of $\mathbf{x}_i$ for the forward mean map is simply the weighted mean (barycenter) $\vec{m}_\pi(x)$ of these two target masses. The backward mean map is simply the reverse, mapping each $\mathbf{y}_i$ to the barycenter of its associated source masses.

3.2 Properties of (forward and backward) mean maps

A number of properties of these mean maps are worth mentioning as we will leverage them in this work. We go over the properties of *forward* mean maps only as they trivially hold for backward mean maps as well by inverting source and target.

Case of Monge maps: If $\pi \in \Pi(\mu,\nu)$ is induced by a Monge map, i.e., a μ-measurable map $T : \mathcal{X} \to \mathcal{Y}$ such that $T_\sharp \mu = \nu$, then T is the mean map of π. Conversely, if $\vec{m}_\pi$ is a Monge map between μ and ν, then π is the transport plan induced by this map.

Linearity. The map $\pi \mapsto \vec{m}_\pi$ is linear, and the set of forward mean maps $\vec{M}(\mu,\nu) \coloneqq \left\{ \vec{m}_\pi \,\middle|\, \pi \in \Pi(\mu,\nu) \right\}$ of transport plans between μ and ν is a *convex* subset of $\mathcal{F}_\mathcal{X}$, since $\Pi(\mu,\nu)$ is itself convex.

Means: Note that Eq. (3) trivially implies:

$$\mathbb{E}_\mu \left[\vec{m}_\pi \right] \;=\; \bar{\nu}, \tag{4}$$

where the mean $\bar{\nu}$ of the measure ν is simply defined by $\mathbb{E}_\nu \left[id_{\mathbb{R}^d} \right]$.

Variance loss and contraction property. As we prove in App. B of our Supplemental Material, one always has

$$\mathrm{Var}_\mu \left[\vec{m}_\pi \right] \le \mathrm{Var}_\nu. \tag{5}$$

where $\mathrm{Var}_\nu \coloneqq \mathrm{Var}_\nu \left[id_{\mathbb{R}^d} \right]$. Moreover, $\vec{m}_\pi$ is a Monge map between μ and ν iff its variance is maximal, i.e., $\mathrm{Var}_\mu \left[\vec{m}_\pi \right] = \mathrm{Var}_\nu$.

Convex hull property: For any mean map $\phi \in \vec{M}(\mu,\nu)$, $\phi(\mathrm{supp}\,\mu)$ is a subset of the convex hull of $\mathrm{supp}\,\nu$, bounding the map's image.

3.3 Regularizing mean maps

Having derived maps from an optimal transport, we can now define a measure of their regularity via a (preferably convex) deformation functional $\mathcal{J}$, for which a plethora of options is available.

Strain functionals issued from elasticity theory [Landau and Lifshitz 1970] are obvious potential map regularizers as they are isometry invariant, but they are not convex; functionals from linear elasticity, instead, have the advantage of being convex but are not even rotation invariant. As discussed in Sec. 2, the Dirichlet energy $\|\nabla\phi\|_\mu^2$ has often been used, although directly on transport plans instead of on their associated mean maps. However, the identity map is not even a minimizer of this energy, and the Dirichlet energy heavily promotes maps sending every point towards the same destination, which is a significant drawback often translating into pinched maps. A functional of the type $\|\Delta\phi\|_\mu^2$ can remedy this issue, at the cost of more spurious minimizers than just harmonic maps.

A successful alternative to further reduce the set of maps that the regularization does not penalize has involved the Hessian, be it in the context of dimensionality reduction [Donoho and Grimes 2003], boundary conditions [Stein et al. 2018], denoising and interpolation [Rowe et al. 2024], or spectral as-rigid-as-possible shape deformation [Budninskiy et al. 2017] to name a few. Note that only affine maps are in its kernel, and while this does not eliminate the case of constant maps since convexity does not allow their suppression, it effectively removes a number of spurious deformation modes. One of the contributions of our work will be to show that applying a Hessian-based regularization on the mean maps of a transport plan offers a robust approach to shape matching.

4 Optimal Forward/Backward Mean Maps

We are now ready to delve into our convex formulation of spatially regularized optimal transport problems, before discussing how to solve it via an efficient ADMM-based optimization.

4.1 Our generic spatial regularization of optimal transport

In this work, we posit that a spatially-regularized optimal transport between a source shape $(\mathcal{X},\mu)$ and a target shape $(\mathcal{Y},\nu)$ for a given transport cost $c(x,y)$ is the transport plan $\pi \in \Pi(\mu,\nu)$ between source and target shape that has a low total transport cost $\int c(x,y)\, d\pi(x,y)$, and whose mean maps $\vec{m}_\pi$ and $\overleftarrow{m}_\pi$ both have low deformation energy. We thus express our mean-map regularized OT formulation in a way generic enough for a variety of geometric problems where regularized transport plans are called for, using:

$$\boxed{\inf_{\pi \in \Pi(\mu,\nu)} \mathcal{J}_\mathcal{X}(\vec{m}_\pi) + \mathcal{J}_\mathcal{Y}(\overleftarrow{m}_\pi) + \int_{\mathcal{X}\times\mathcal{Y}} c(x,y)\, d\pi(x,y) + \epsilon\, H(\pi),} \tag{6}$$

where $\mathcal{J}_\mathcal{X}$ and $\mathcal{J}_\mathcal{Y}$ are convex lower semi-continuous regularizing functionals on $\mathcal{F}_\mathcal{X}$ and $\mathcal{F}_\mathcal{Y}$ respectively (possibly playing symmetric roles), $c(x,y)$ is a given (but arbitrary) cost function, and where the entropic term (with ϵ small) is added to make the problem well-behaved. This formulation being strongly convex by design, it admits a unique solution. We describe in the remainder of this section a numerical method to efficiently solve it; note that we will only consider the regularization induced by the *forward* mean map $\vec{m}_\pi$ to ease the exposition, but our ADDM-based solver is trivially adaptable to deal with the generic problem in Eq. (6).

4.2 Quadratic case and its dual formulation

We first consider a simpler case of the generic problem where the functional $\mathcal{J}$ to regularize $\vec{m}_\pi$ is only quadratic, as we will see that the general case can be solved through repeated instances of this simpler case. Consider a convex quadratic function $Q : \phi \in \mathcal{F}_\mathcal{X} \mapsto \langle \phi, \mathbf{A}\phi \rangle_{\mathcal{F}_\mathcal{X}} + \langle b, \phi \rangle_{\mathcal{F}_\mathcal{X}} \in \mathbb{R}$ where $b \in \mathcal{F}_\mathcal{X}$ and $\mathbf{A}$ is a bounded self-adjoint positive-definite operator on $\mathcal{F}_\mathcal{X}$. When the general regularizer $\mathcal{J}$ is replaced by this quadratic function, we have to solve:

$$\inf_{\pi \in \Pi(\mu,\nu)} Q(\vec{m}_\pi) + \int c\, d\pi + \epsilon\, H(\pi). \tag{7}$$

Introducing a variable ϕ to play the role of $\vec{m}_\pi$, we get the constrained optimization problem:

$$\inf_{\substack{\phi \in \mathcal{F}_\mathcal{X} \\ \pi \in \Pi(\mu,\nu)}} Q(\phi) + \int c\, d\pi + \epsilon\, H(\pi) \quad \text{s.t. } \phi = \vec{m}_\pi \tag{8}$$

Using the method of Lagrange multipliers, we rewrite Eq. (8) as

$$\inf_{\substack{\phi \in \mathcal{F}_\mathcal{X} \\ \pi \in \Pi(\mu,\nu)}} \sup_{\lambda \in \mathcal{F}_\mathcal{X}} \mathcal{L}(\phi, \pi, \lambda), \tag{9}$$

Fig. 3. **Human poses.** Using our Mongified approach, near-isometric maps between a rest pose of a human (left) and nine other poses from FAUST [Bogo et al. 2014] are computed automatically from five landmarks (marked in red on rest pose). We use texture transfer via forward mean maps for visualization.

where the Lagrangian function $\mathcal{L}$ is defined as

$$\begin{aligned}\mathcal{L}(\phi,\pi,\lambda) &:= Q(\phi) + \int c\,d\pi + \epsilon\,H(\pi) + \big\langle \lambda, \phi - \overrightarrow{m}_\pi \big\rangle_{\mathcal{F}_X} \qquad (10)\\ &= Q(\phi) + \langle \lambda, \phi\rangle_{\mathcal{F}_X} + \epsilon\,H(\pi) + \int \big[\underbrace{c(x,y) - \lambda(x)\cdot y}_{=c_\lambda(x,y)}\big]\,d\pi(x,y)\end{aligned}$$

The (Lagrangian) dual problem then reads:

$$\sup_{\lambda\in\mathcal{F}_X} \inf_{\substack{\phi\in\mathcal{F}_X\\ \pi\in\Pi(\mu,\nu)}} \mathcal{L}(\phi,\pi,\lambda). \qquad (11)$$

Remark now that:

$$\inf_{\substack{\phi\in\mathcal{F}_X\\ \pi\in\Pi(\mu,\nu)}} \mathcal{L}(\phi,\pi,\lambda) = \inf_{\phi\in\mathcal{F}_X} \big(Q(\phi) + \langle\lambda,\phi\rangle_{\mathcal{F}_X}\big) + \overbrace{\inf_{\pi\in\Pi(\mu,\nu)} \int c_\lambda\,d\pi + \epsilon\,H(\pi)}^{=E_\epsilon^{\mathrm{OT}}(\lambda)}.$$

It highlights the presence of an optimal transport term, $E_\epsilon^{\mathrm{OT}}(\lambda)$, involving an entropy-regularized transport plan $\pi\in\Pi(\mu,\nu)$ using the cost function c_λ defined in Eq. (10). The other term is an infimum on variable ϕ, for which the optimality condition simply is

$$\frac{\partial}{\partial\phi}\big[Q(\phi) + \langle\lambda,\phi\rangle_{\mathcal{F}_X}\big] = 0 \quad\Leftrightarrow\quad \mathbf{A}\phi + b + \lambda = 0. \qquad (12)$$

We can thus re-express the sup-inf from Eq. (11) as

$$\sup_{(\lambda,\phi)\in\mathcal{F}_X^2} Q(\phi) + \langle\lambda,\phi\rangle_{\mathcal{F}_X} + E_\epsilon^{OT}(\lambda) \quad \text{s.t. } \lambda = -(\mathbf{A}\phi + b) \qquad (13)$$

From there we conclude that Eq. (7) is equivalent to an unconstrained infimum over ϕ of a functional we will denote $\widehat{Q}(\phi)$:

$$-\inf_{\phi\in\mathcal{F}_X} \underbrace{\tfrac{1}{2}\langle\phi,\mathbf{A}\phi\rangle_{\mathcal{F}_X} - E_\epsilon^{OT}(-\mathbf{A}\phi - b)}_{=\widehat{Q}(\phi)}, \qquad (14)$$

Remark that even if we are solving a dual problem, the ϕ that solves (14) is actually the same than the one that solves (8) since strong duality holds and we assumed that $\mathbf{A}$ is non-degenerate.

4.3 A first descent algorithm

The gradient of $\widehat{Q}$ in $\mathcal{F}_X$ happens to have a simple expression:

$$\nabla_X\widehat{Q}(\phi) = \mathbf{A}\big(\phi - \overrightarrow{m}_{\pi[\phi]}\big), \qquad (15)$$

where $\pi[\phi]$ is the optimal transport plan for $OT_\epsilon(\mu,\nu;c_{\lambda(\phi)})$ with $\lambda(\phi) := -\mathbf{A}\phi - b$ (see App. C in Supplemental Material). Since this is an unconstrained convex optimization problem, a simple gradient descent can be implemented. Since $\mathbf{A}$ is invertible, we can even use the scalar product $\langle\cdot,\mathbf{A}\cdot\rangle_{\mathcal{F}_X}$ instead of $\langle\cdot,\cdot\rangle_{\mathcal{F}_X}$, leading to the descent direction $\overrightarrow{m}_{\pi[\phi]} - \phi$ instead of $\mathbf{A}\big(\overrightarrow{m}_{\pi[\phi]} - \phi\big)$. This give us Alg. 1; note that if one prescribes a gradient step with size $0 < t \le 1$, ϕ is always a convex combination of mean maps, and thus, it remains a mean map itself. We estimate convergence is reached when the relative distance between two consecutive functions ϕ is small.

Algorithm 1 InteriorDescent($\mathbf{A}, b$)

Require: $0 < t \le 1$
  Compute the solution π of $OT_\epsilon(\mu,\nu;c_{\lambda=0})$
  $\phi \leftarrow \overrightarrow{m}_\pi$
  repeat
    $\lambda \leftarrow -\mathbf{A}\phi - b$
    Compute the solution π of $OT_\epsilon(\mu,\nu;c_\lambda)$
    $\phi \leftarrow (1-t)\phi + t\,\overrightarrow{m}_\pi$
  until Convergence

4.4 Generic case with ADMM

We are now interested in the generic problem of Eq. (6), which we rewrite using a variable ϕ as in the quadratic case above through:

$$\inf_{\substack{\phi\in\mathcal{F}_X\\ \pi\in\Pi(\mu,\nu)}} \mathcal{J}(\phi) + \int c\,d\pi + \epsilon\,H(\pi) \quad \text{s.t. } \phi = \overrightarrow{m}_\pi, \qquad (16)$$

where $\mathcal{J}$ is assumed to be convex and lower semi-continuous. This is trivially equivalent to minimizing a modified functional where the squared norm of the constraint is added, i.e.,

$$\inf_{\substack{\phi\in\mathcal{F}_X\\ \pi\in\Pi(\mu,\nu)}} \mathcal{J}(\phi) + \frac{\rho}{2}\big\|\phi - \overrightarrow{m}_\pi\big\|^2_{\mathcal{F}_X} + \int c\,d\pi + \epsilon\,H(\pi) \quad \text{s.t. } \phi = \overrightarrow{m}_\pi. \qquad (17)$$

We can then form the so-called augmented Lagrangian L_ρ of the problem, yielding the expression:

$$L_\rho(f,\pi,\gamma) := \mathcal{J}(\phi) + \tfrac{\rho}{2}\big\|\phi - \overrightarrow{m}_\pi\big\|^2_{\mathcal{F}_X} + \big\langle\gamma, \phi - \overrightarrow{m}_\pi\big\rangle_{\mathcal{F}_X} + \int c\,d\pi + \epsilon H(\pi). \qquad (18)$$

Applying the Alternating Direction Method of Multipliers (ADMM) method [Boyd et al. 2011] to our generic problem is then simple: it consists in alternatively solving the two following optimizations

$$\inf_{\pi\in\Pi(\mu,\nu)} L_\rho(\phi,\pi,\gamma), \quad \text{and} \quad \inf_{\phi\in\mathcal{F}_X} L_\rho(\phi,\pi,\gamma). \qquad (19)$$

Furthermore, we note that the first one is equivalent to:

$$\inf_{\substack{m\in\mathcal{F}_X\\ \pi\in\Pi(\mu,\nu)}} \frac{\rho}{2}\|m - \phi\|^2_{\mathcal{F}_X} - \langle\gamma, m\rangle_{\mathcal{F}_X} + \int c\,d\pi + \epsilon\,H(\pi) \quad \text{s.t. } m = \overrightarrow{m}_\pi. \qquad (20)$$

This corresponds precisely to the previous problem from Eq. (8) with the quadratic function $Q(m) = \frac{\rho}{2}\|m\|^2_{\mathcal{F}_X} + \langle\gamma - \rho\phi, m\rangle_{\mathcal{F}_X}$, so Alg. 1 is good choice for this first optimization. Let us call $m(\phi,\gamma)$ the resulting (unique) solution for the variable m. The second problem then becomes:

$$\inf_{\phi\in\mathcal{F}_X} \underbrace{\mathcal{J}(\phi) + \langle\gamma,\phi\rangle_{\mathcal{F}_X}}_{\mathcal{J}_\gamma(\phi)} + \frac{\rho}{2}\big\|\phi - \overrightarrow{m}_\pi\big\|^2_{\mathcal{F}_X}. \qquad (21)$$

Note that this problem only depends on the forward mean map $m = \overrightarrow{m}_\pi$ and *does not require the corresponding plan* π. Since $\mathcal{J}_\gamma$ is convex and lower semi-continuous, the unique optimal $\phi(m,\gamma)$ is given by the proximal operator $\mathrm{Prox}[\mathcal{J}_\gamma/\rho](m)$. Remark that if $\mathcal{J}(\phi) = \frac{1}{2}\langle\phi, \mathbf{A}\phi\rangle_{\mathcal{F}_\mathcal{X}} + \langle b,\phi\rangle_{\mathcal{F}_\mathcal{X}}$, then the solution simplifies to

$$\mathrm{Prox}[\mathcal{J}_\gamma/\rho](m) = (\mathbf{A} + \rho\boldsymbol{I})^{-1}(\rho m - \gamma - b), \tag{22}$$

which requires a simple linear solve for many operators $\mathbf{A}$ of interest. Those two solves are performed successively, followed by the usual update $\gamma \leftarrow \gamma + \rho(\phi - m)$. The variable ρ is updated across iterations following an increasing sequence $\rho^1 \le \rho^2 \le \cdots$, although updates based on primal and dual residuals can also be used [Boyd et al. 2010]. An ADMM-based optimization algorithm of our original problem from Eq. (16) follows the pseudocode given in Alg. 2. Note that we estimate that convergence is reached when $\|\phi - m\|$ is small enough.

Algorithm 2 ADMM

$\gamma \leftarrow 0$
$\phi \leftarrow 0$
repeat
  $m \leftarrow m(\phi,\gamma)$ (Eq. (20)) thru InteriorDescent$(\rho I, \gamma - \rho\phi)$
  $\phi \leftarrow \mathrm{Prox}[\mathcal{J}_\gamma/\rho](m)$
  $\gamma \leftarrow \gamma + \rho(\phi - m)$
until Convergence

4.5 Discussion

Given the convexity of our spatially-regularized optimal transport problem, a number of classical optimization tools can be applied to find the unique optimal transport plan, without fear of being trapped in local minima. However, from our own numerical experiments, we found ADMM to be best in terms of convergence rates compared to subgradient algorithms or bundle methods for instance. And while we only described how to introduce spatial regularization of the *forward* mean map $\overrightarrow{m}_\pi$ of a transport plan π from a source shape $(\mathcal{X},\mu)$ to a target shape $(\mathcal{Y},\nu)$ in order to induce spatial continuity, we can of course include regularization of the backward mean map $\overleftarrow{m}_\pi$ as well, even possibly using a different (but still convex) deformation functional. More precisely, instead of the term $\mathcal{J}(\phi)$ in 16, we will have $\mathcal{J}_\mathcal{X}(\phi) + \mathcal{J}_\mathcal{Y}(\phi')$, thus adding a new optimization variable $\phi' \in \mathcal{F}_\mathcal{Y}$, that will be constrained by $\phi' = \overleftarrow{m}_\pi$. This implies modifying the augmented Lagrangian accordingly, introducing another dual variable γ'. In the ADMM algorithm, we will have to compute two proximal operators (of $\mathcal{J}_{\mathcal{X},\gamma}/\rho$ and $\mathcal{J}_{\mathcal{Y},\gamma'}/\rho$) inside each iteration of the main loop, but still only one call to Alg. 1 (with modified parameters). Finally, as discussed early on in this paper, existing spatially-regularized optimal transport approaches [Solomon et al. 2013; Ghoussoub et al. 2020] would involve optimization over the whole set of transport plan coefficients; our formulation thus offers a computational tool for OT that is *significantly more efficient* by bypassing the need for a high-variable-count optimization.

5 Applications

We now demonstrate the versatility and efficiency of optimal mean maps computed through our spatially-regularized optimal transport formulation on a series of examples. After a brief discussion about the implementation of our approach in Sec. 5.1, we focus first on two different regimes of our algorithm which are directly useful to non-rigid registration: first, Sec. 5.2 illustrates how our convex optimization can be used *as is* once a given transport cost function is chosen; then Sec. 5.3 discusses a cost-agnostic alternative, where a single non-convex variance-maximizing term is added to the optimization to help find a Monge map and its inverse that are both spatially regular. Finally, Sec. 5.4 discusses the cases of non-planar parameterizations and dissimilarity measures, which are other straightforward applications of our new approach to spatially-regularized optimal transport. Note that our Supplemental Material provides mesh sizes and computational times for most examples.

5.1 Implementation details

Our ADMM-based solver described in Sec. 4 is basically akin to computing a series of OT solves through the Sinkhorn's algorithm (based on specific, evolving costs), for which existing efficient libraries exists. An important component of our novel approach that we have not described yet is its Hessian-based regularization, which is convex, isometry-invariant, and with a small subspace of minimizers (see Sec. 3.3). We use two different regularization formulation using the Hessian ∇^2 depending on whether we deal with volumetric or surface-to-surface transports, for which we now provide the corresponding proximal operators (see Sec. 4.4).

Volumetric case. If we compute a transport plan between two volumes (or surfaces in 2D), we simply use the regularization energy:

$$\mathcal{J} : \phi \mapsto \begin{cases} \frac{1}{2}\left\|\nabla^2\phi\right\|^2_{\mathcal{F}_\mathcal{X}} & \text{if } \phi \in H^2(\mathcal{X},\mathbb{R}^d;\mu) \\ +\infty & \text{else} \end{cases} \tag{23}$$

The corresponding proximal operator $\mathrm{prox}[\mathcal{J}_\gamma/\rho](m)$ is then given by the solution to, for $A = (\nabla^2)^*\,\nabla^2$ being the squared Hessian,

$$\inf_{\phi\in H^2} \ \tfrac{1}{2}\langle\phi, \mathbf{A}\phi\rangle_{\mathcal{F}_\mathcal{X}} + \langle\gamma,\phi\rangle_{\mathcal{F}_\mathcal{X}} + \tfrac{\rho}{2}\|\phi - m\|^2_{\mathcal{F}_\mathcal{X}}\,. \tag{24}$$

So we simply have to solve the linear system:

$$(\mathbf{A} + \rho\mathbf{I})\phi \ = \rho m - \gamma \tag{25}$$

Surface-to-surface case. If, instead, a transport plan between two surfaces $\mathcal{X}$ and $\mathcal{Y}$ in $\mathbb{R}^3$ is computed, we provide in App. D of the Supplemental Material a functional behaving in a similar manner to Eq. (23), with the same desirable properties. Its final expression is

$$\mathcal{J}(\phi) = \frac{1}{2}\sum_{i=1}^{3} \inf_{h^i \in L^2_\mu} e(\phi^i, h^i) \tag{26}$$

where the $e(\phi^i, h^i)$ term is defined as:

$$e(\phi^i, h^i) = \left\|\nabla_\mathcal{X}(\nabla_\mathcal{X}\phi^i + h^i\boldsymbol{n})\right\|^2_{\mathcal{F}_\mathcal{X}} + \left\|(\nabla_\mathcal{X}\phi^i + h^i\boldsymbol{n})\cdot\boldsymbol{n}\right\|^2_{\mathcal{F}_\mathcal{X}} \tag{27}$$

where $\boldsymbol{n}\colon \mathcal{X} \to \mathbb{R}^3$ is the Gauss map of the surface and $\nabla_\mathcal{X}$ is the tangential gradient. Since this energy can be written in the form $\phi \mapsto \inf_h (\phi, h)\mathbf{A}(\phi, h)^T$, the computation of its proximal operator is again a simple linear solve. Details of our energy discretization on surface meshes are available in the Supplemental Material.

5.2 OT-based convex shape matching

Shape matching, also known as non-rigid registration, is a cornerstone of geometry processing aiming to establish correspondences between two shapes or domains — typically called *source* and *template* — that may differ rather significantly in structure or form. While rigid registration relies solely on translation and rotation to align two shapes [Chen and Medioni 1991; Besl and McKay 1992],

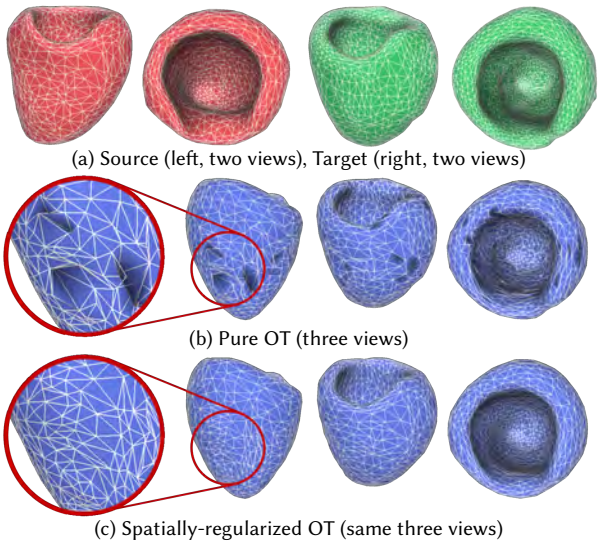


Fig. 4. **Left ventricle registration.** For a left ventricle from the heart of a source patient (a) to a target patient (b), a pure OT registration with no landmarks captures roughly the shape but exhibits numerous and obvious artifacts. Our convex approach to spatial regularization fixes the issues in just three iterations of the ADMM algorithm (hence, in 3 Sinkhorn solves).

non-rigid registration must consider possibly far more complex deformation to align them best. The literature on non-rigid registration is voluminous, varying in the type of mapping between domains that one considers (point-to-point vs. soft), its assumed properties (bijectivity, local injectivity, ...), its priors (isometry, conformality, etc), as well as in the numerical techniques employed in practice. Often, the correspondence map between two geometric domains is represented using point-to-point mappings, spectral-based functional maps, or smooth, invertible transformations encoded through the flow of a vector field, from which multimodal information of the two input shapes — such as color, geometric descriptors, or user-specified landmarks — allows to find a meaningful correspondence map. However, most existing methods for non-rigid registration are inherently non-convex, necessitating computationally intensive multiscale optimization strategies that often lack guarantees of finding an optimal registration. Our approach directly provides a convex approach to shape matching. Assuming that the OT problem for a specially-designed cost function is able to find a few key correspondences between two shapes, our regularization will automatically create both the forward (source-template) and backward (template-source) maps, in often just a few Sinkhorn's algorithm calls (3 to 6 in our tests) as we now describe for various cost strategies.

Euclidean distance cost. An OT approach based on just the squared Euclidean distance for $c(\cdot,\cdot)$ (leading to the 2-Wasserstein metric) can easily handle the registration of a source (red) and target (green) heart ventricle in Fig. 4: the pure OT solve without *any* landmarks already captures most of the registration, yet suffers from numerous irregularities, typical of transport-based solutions. With our spatial regularization where only the forward map is regularized, the optimal forward mean map already leads to a visually perfect registration between these irregular meshes, as shown in Fig. 11.

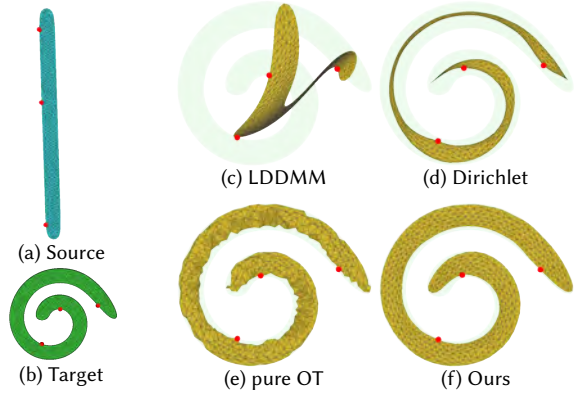


Fig. 5. **Bar to Spiral.** While LDDMM (c) fails to match a bar (a) to a spiral (b) shape given three landmarks, pure OT optimization for costs based on intrinsic distances to landmarks captures the overall bending but with large irregularities and flips. Dirichlet-based regularization provides a cleaner map (d), but suffers large shrinking. Instead, *a Hessian-based regularization of the forward mean map* captures this example very well (f).

Landmark-based cost. Knowledge of k landmarks $(x_1, ..., x_k) \in \mathcal{X}^k$ and their correspondences $(y_1, ..., y_K) \in \mathcal{Y}^k$ can also be exploited to help establish a global registration through a cost of the form

$$c(x,y) = \sum_{k=1}^{K} \left|\theta(d_{\mathcal{X}}(x_k,x)) - \theta(d_{\mathcal{Y}}(y_k,y))\right|^2, \tag{28}$$

to guide correspondences, where $\theta : \mathbb{R}_{>0} \to \mathbb{R}$ is any continuous monotonous function while $d_{\mathcal{X}}$ and $d_{\mathcal{Y}}$ are the geodesic distances of $\mathcal{X}$ and $\mathcal{Y}$ respectively. This is what is used for the 2D illustrative test in Fig. 5, where we show that a bar can be mapped to a spiral using only three landmarks (marked as red disks). Here, the OT solve solely based on the transport costs based on the intrinsic distance to the landmarks succeeds at capturing the deformation (e), but with significant discontinuities and flips. Regularizing just the forward mean map fixes all these issues, with only a small shrinkage effect due to the variance loss of the mean map (App. B of Supplemental Material). Instead, an approach based on *large deformation diffeomorphic metric mapping* (LDDMM [Beg et al. 2005; Pennec et al. 2006] – see details about the regularization and data attachment we used in the Supplemental Material) fails to capture the proper registration, while an OT problem with a Dirichlet-based regularization introduces large shrinking artifacts.

Hybrid cost and unbalanced transport. Finally, we show that an unbalanced optimal transport [Chizat et al. 2018] can also capture a rather challenging deformation between a dog and a camel (see Figs. 1 and 13), where this time 7 landmarks are provided (one on each ear, one on the top of the head, one on the snout, one at the end of the tail, one on the front of the neck and one on the back of the neck), and the transport cost we use is the sum of the landmark-derived cost and squared Euclidean distance. Regularizing the forward mean map establishes a natural correspondence map between the two, and the dog mesh is correctly mapped onto the camel by the optimal map in just three iterations of our ADMM-based solver, amounting to three calls to Sinkhorn's algorithm.

Parameters. All the results shown in the section were first normalized in a unit bounding box, then computed using a step $t = 0.1$

in Alg. 1, an entropy coefficient $\epsilon = 10^{-4}$, a constant $\rho = 1$, without any further tuning. The unbalanced transport was computed with [Chizat et al. 2018] using a multiplicative factor of the KL divergence equal to 10. Note finally that we used a single iteration in Alg. 1 in practice as it proved sufficient.

5.3 Cost-agnostic shape matching through Mongification

For shape matching tasks involving widely different sources and targets, there may not exist an obvious transport cost which reliably provides solid correspondences — a few failure cases of the purely convex formulation are provided in the Supplemental Material. Remarkably, our work can handle these cases as well, albeit with the addition of a single non-convex term to our optimization strategy.

Remember from Sec. 3.2 that *maximizing a mean map's variance* renders the transport plan as close to a *Monge map* as possible. This means that if we add a maximization of the variance of each mean map, we force them to become Monge maps, thus indirectly favoring a *diffeomorphism* between the two shapes. And in this case, a very well-behaved (regularized) correspondence will be found *even if we select a transport cost c of zero!* However, this cost-agnostic approach, which we dub *Mongification*, is using variance which is a non-convex term, ruining the convexity property we had established. (Note here that maximizing the variance of the mean maps is equivalent to the variance minimization of each π_x, but without needing the actual transport plan directly; it is therefore far more efficient.) Still, it turns out to be a quite interesting alternative: if we use a factor $\alpha\rho$ to this variance term, where $\alpha \in [0, 1]$ and ρ being the constant used in the augmented Lagrangian of Eq. (18), then the case $\alpha = 1$ actually guarantees that Alg. 1 only needs one iteration: the quadratic term having canceled, we are left with an OT problem with a linear functional, which amounts to an OT problem with modified costs. However, we recommend picking $\alpha = 1/2$ instead, even if we still use a *single* iteration in Alg. 1 (just like in the previous purely-convex case) as it has proven sufficient in all our tests. With this added non-convex term, our approach becomes able to deal efficiently and robustly with cases where the pure OT problem is totally unhelpful, partially because the resulting ADMM algorithm *still* consists in purely convex problems to solve at each iteration. We prove robustness of our approach on several examples where both forward and backward mean maps are regularized. First, Fig. 6 shows a registration using this modified spatially-regularized OT problem for two meshes of a horse in different poses; see texture transfer in Fig. 7. Only three landmarks were used on this example. Fig. 9 shows a registration between two poses of a human, with two landmarks (right knee, right shoulder); texture transfer is shown in Fig. 10. Finally, we show registrations between ten human poses in Fig. 3, and a male and a female in Fig. 12.

Parameters. Mongification uses the same parameters as in Sec. 5.2, but starting with $\rho = 0.01$ and doubling ρ after each ADMM iteration. All examples converged in less than 30 ADMM iterations.

5.4 Other examples of usage

Another direct application of our approach is spherical parameterization — or in fact, any *non-planar parameterization*. By using a fine mesh of a sphere shape as target shape with the same area as the source, our Mongified OT-based optimal maps provide regularized area preserving parameterizations. For a Dirichlet energy regularization, we obtain an as-conformal-as-possible area-preserving map for the mean maps, see Fig. 14. Note that compared to state-of-the-art methods for this specific parameterization such as [Nadeem et al. 2017], we do not require any mesh cutting or multiple optimization problems to solve. An area-preserving map of a genus-one mesh onto a torus for a Hessian regularization is also shown in Fig. 8.

Our spatially-regularized OT method can also be used as a *dissimilarity measure* between geometric objects. While optimal transport costs are often used to estimate (dis)similarities between a variety of data [Alvarez-Melis and Fusi 2020; Lipp and Vermeesch 2023] (e.g., for t-SNE or Isomap), transport costs alone are not always representative of differences in the data. In Fig. 15 we compute the standard Euclidean OT costs between two sets of MNIST digits, and the left array (visualizing these costs in grey levels after thresholding at 110% of the maximum cost on the diagonal) indicates very poor separation between digit classes; instead, the right array uses our Mongified forward-backward Hessian-regularized Euclidean cost (with the same visualization), leading to *only one* classification error if a simple single-linkage clustering is performed.

6 Conclusions

In this paper, we provide a scalable convex optimization to spatially regularize optimal transport problems. Our approach is very general as one can regularize the forward, the backward, or both mean maps using arbitrary convex deformation functionals. Our work calls for a number of follow up investigations. First, exploring in greater details its advantages for various registration tasks (maybe by considering not only positions, but normals as well) would be obviously valuable. Second, we believe that our optimal mean maps could be useful in other contexts; e.g., it would be interesting to come up with a notion of *shape average* that exhibit more natural deformations than current OT barycenters, and which may show connections to the "barycentric projection" of Cazelles et al. [2021]. Third, it is quite possible that using other regularization functionals (*e.g.* based on the ℓ_1 norm of derivatives) may either facilitate registration or allow for new applications, such as the computation of fluid flows from known correspondences. Note that we go through the dual of our formulation to solve our registration problem; this means that even if we use non-convex regularization terms (typically, those issued from elasticity for instance), we will end up solving a *concave* dual problem. Whether the non-convexity of the primal problem will create convergence issues needs to be investigated, but it could be an interesting approach to "convexify" other functionals as well. Finally, using a coarse-to-fine solver would likely improve timings significantly for fine meshes.

Acknowledgments

All meshes of human characters (Figs. 3,9,10,12) are from the FAUST dataset [Bogo et al. 2014]; the Horse mesh from Figs. 6 and 7 is from [Sumner and Popović 2004]; meshes from Figs. 4 and 11 are from the ImageCAS dataset of Zeng et al. [2023]; Dog v2 and Camel v201 meshes from Fig. 1 and 13 are from `printable_models` at Free3D; the Bob mesh from Fig. 8 and the Spot mesh from Fig. 14

are from Keenan Crane; and the images from Fig. 15 are from the MNIST dataset [LeCun et al. 2002]. All mesh visualizations used Polyscope [Sharp et al. 2019]. MD benefited from the generous support of the MediTwin consortium (funded by the French government as part of France 2030), from an Inria chair, and from Ansys.

References

David Alvarez-Melis and Nicolo Fusi. 2020. Geometric Dataset Distances via Optimal Transport. In *NeurIPS*, Vol. 33. 21428–21439.

M Faisal Beg, Michael I Miller, Alain Trouvé, and Laurent Younes. 2005. Computing large deformation metric mappings via geodesic flows of diffeomorphisms. *Int. J. Comput. Vision* 61 (2005), 139–157.

P.J. Besl and Neil D. McKay. 1992. A method for registration of 3-D shapes. *IEEE Trans. Pattern Analysis and Machine Intelligence* 14, 2 (1992), 239–256.

Vladimir I. Bogachev. 2007. *Measure theory*. Vol. 2. Springer.

Federica Bogo, Javier Romero, Matthew Loper, and Michael J. Black. 2014. FAUST: Dataset and evaluation for 3D mesh registration. In *Proceedings IEEE Conf. on Computer Vision and Pattern Recognition*.

Nicolas Bonneel and David Coeurjolly. 2019. SPOT: Sliced Partial Optimal Transport. *ACM Transactions on Graphics* 38, 4, Article 89 (2019).

Nicolas Bonneel and Julie Digne. 2023. A survey of Optimal Transport for Computer Graphics and Computer Vision. *Comp. Graph. Forum* 42, 2 (2023), 439–460.

Nicolas Bonneel, Julien Rabin, Gabriel Peyré, and Hanspeter Pfister. 2015. Sliced and Radon Wasserstein Barycenters of Measures. *Journal of Mathematical Imaging and Vision* 51, 1 (2015), 22–45.

Stephen Boyd, Neal Parikh, Eric Chu, Borja Peleato, and Jonathan Eckstein. 2010. Distributed optimization and statistical learning via the alternating direction method of multipliers. *Foundations and Trends in Machine Learning* 3, 1 (2010), 1–122.

Steven Boyd, Neal Parikh, Eric Chu, Borja Peleato, and Jonathan Eckstein. 2011. Distributed Optimization and Statistical Learning via the Alternating Direction Method of Multipliers. *Foundations and Trends in Machine Learning* 3, 1 (2011), 1–122.

Yann Brenier. 2003. *Extended Monge-Kantorovich Theory*. Springer Berlin Heidelberg, 91–121.

Max Budninskiy, Beibei Liu, Yiying Tong, and Mathieu Desbrun. 2017. Spectral Affine-Kernel Embeddings. *Comp. Graph. Forum* 36, 5 (2017), 117–129.

Elsa Cazelles, Felipe Tobar, and Joaquin Fontbona. 2021. A novel notion of barycenter for probability distributions based on optimal weak mass transport. In *NeurIPS Proceedings*. Article 1040.

Yulu Chen and Gérard Medioni. 1991. Object modeling by registration of multiple range images. In *Proceedings of IEEE Int. Conf. on Robotics and Automation*, Vol. 3. 2724–2729.

Lenaic Chizat, Gabriel Peyré, Bernhard Schmitzer, and François-Xavier Vialard. 2018. Scaling algorithms for unbalanced optimal transport problems. *Math. Comp.* 87, 314 (2018), 2563–2609.

Marco Cuturi. 2013. Sinkhorn distances: Lightspeed computation of optimal transport. *NeurIPS* 26 (2013), 2292–2300.

Fernando de Goes, David Cohen-Steiner, Pierre Alliez, and Mathieu Desbrun. 2011. An Optimal Transport Approach to Robust Reconstruction and Simplification of 2D Shapes. *Comp. Graph. Forum* 30, 5 (2011), 1593–1602.

Fernando de Goes, Corentin Wallez, Jin Huang, Dmitry Pavlov, and Mathieu Desbrun. 2015. Power particles: an incompressible fluid solver based on power diagrams. *ACM Trans. Graph.* 34, 4, Article 50 (2015).

Julie Digne, David Cohen-Steiner, Pierre Alliez, Fernando de Goes, and Mathieu Desbrun. 2014. Feature-Preserving Surface Reconstruction and Simplification from Defect-Laden Point Sets. *Journal of Mathematical Imaging and Vision* 48, 2 (2014), 369–382.

Khoa Do, David Coeurjolly, Pooran Memari, and Nicolas Bonneel. 2025. Linear-Time Transport with Rectified Flows. *ACM Trans. Graph.* 44, 4 (2025).

David L. Donoho and Carrie Grimes. 2003. Hessian eigenmaps: Locally linear embedding techniques for high-dimensional data. *Proceedings of the National Academy of Sciences* 100, 10 (2003), 5591–5596.

Jean Feydy. 2020. *Geometric data analysis, beyond convolutions*. Ph. D. Dissertation. Université Paris-Saclay.

Nassif Ghoussoub, Young-Heon Kim, Hugo Lavenant, and Aaron Zeff Palmer. 2020. Hidden convexity in a problem of nonlinear elasticity. arXiv:2004.10287 [math.AP]

Lev D. Landau and Evgeniĭ M. Lifshitz. 1970. *Theory of elasticity*. Pergamon.

Yann LeCun, Corinna Cortes, and Chris Burges. 2002. MNIST handwritten digit database. https://web.archive.org/web/20200430193701/http://yann.lecun.com/exdb/mnist/.

Yaron Lipman and Ingrid Daubechies. 2011. Conformal Wasserstein distances: Comparing surfaces in polynomial time. *Advances in Mathematics* 227, 3 (2011), 1047–1077.

A. Lipp and P. Vermeesch. 2023. The Wasserstein distance as a dissimilarity metric for comparing detrital age spectra and other geological distributions. *Geochronology* 5, 1 (2023), 263–270.

Manish Mandad, David Cohen-Steiner, Leif Kobbelt, Pierre Alliez, and Mathieu Desbrun. 2017. Variance-minimizing transport plans for inter-surface mapping. *ACM Trans. Graph.* 36, 4, Article 39 (2017).

Saad Nadeem, Zhengyu Su, Wei Zeng, Arie Kaufman, and Xianfeng Gu. 2017. Spherical Parameterization Balancing Angle and Area Distortions. *IEEE Trans. Vis. Comp. Graph.* 23, 6 (June 2017), 1663–1676.

Xavier Pennec, Pierre Fillard, and Nicholas Ayache. 2006. A Riemannian Framework for Tensor Computing. *Int. J. Comput. Vision* 66 (2006), 41–66.

Gabriel Peyré and Marco Cuturi. 2019. Computational Optimal Transport: With Applications to Data Science. *Foundations & Trends in Machine Learning* 11, 5-6 (2019), 355–607.

Dylan Rowe, Alec Jacobson, and Oded Stein. 2024. Sharpening and Sparsifying with Surface Hessians. In *SIGGRAPH Asia 2024 Conference Papers*. Article 92.

Nicholas Sharp et al. 2019. Polyscope. www.polyscope.run.

Justin Solomon, Fernando De Goes, Gabriel Peyré, Marco Cuturi, Adrian Butscher, Andy Nguyen, Tao Du, and Leonidas Guibas. 2015. Convolutional Wasserstein distances: Efficient optimal transportation on geometric domains. *ACM Trans. Graph.* 34, 4 (2015), 1–11.

Justin Solomon, Leonidas Guibas, and Adrian Butscher. 2013. Dirichlet Energy for Analysis and Synthesis of Soft Maps. *Comp. Graph. Forum* 32, 5 (2013), 197–206.

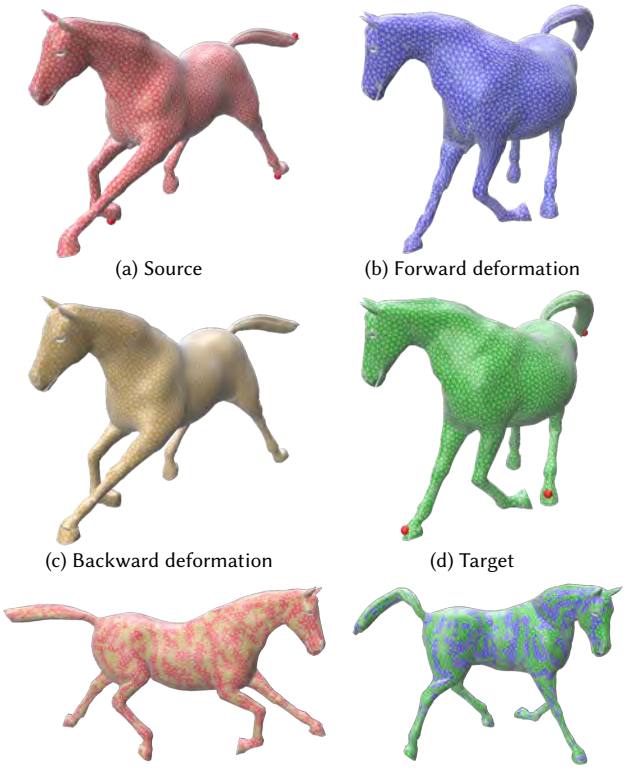
Justin Solomon, Andy Nguyen, Adrian Butscher, Mirela Ben-Chen, and Leonidas Guibas. 2012. Soft Maps Between Surfaces. *Comp. Graph. Forum* 31, 5 (2012), 1617–1626.

Oded Stein, Eitan Grinspun, Max Wardetzky, and Alec Jacobson. 2018. Natural Boundary Conditions for Smoothing in Geometry Processing. *ACM Trans. Graph.* 37, 2, Article 23 (2018).

Robert W. Sumner and Jovan Popović. 2004. Deformation transfer for triangle meshes. *ACM Trans. Graph.* 23, 3 (2004), 399–405.

C. Villani. 2008. *Optimal Transport: Old and New*. Springer Berlin Heidelberg.

An Zeng, Chunbiao Wu, Guisen Lin, Wen Xie, Jin Hong, Meiping Huang, Jian Zhuang, Shanshan Bi, Dan Pan, Najeeb Ullah, et al. 2023. ImageCAS: A large-scale dataset and benchmark for coronary artery segmentation based on computed tomography angiography images. *Computerized Medical Imaging and Graphics* 109 (2023), 102287.

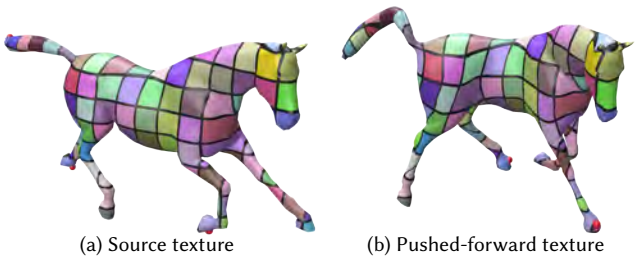


(a) Source (b) Forward deformation

(c) Backward deformation (d) Target

(e) Superposition source/backward (f) Superposition target/forward

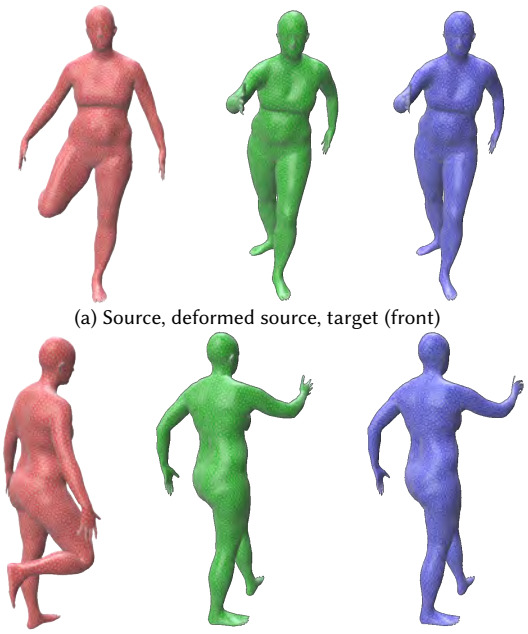
Fig. 6. **Horse poses.** This example of two near-isometric meshes of a horse in very different poses (red and green) is handled with our registration method where a term of variance maximization is added; the blue mesh represents the original red mesh pushed forward by the optimal forward map. Three landmarks (red balls) are used: one on an anterior hoof, one on a posterior hoof, and one at the end of the tail.



(a) Source texture (b) Pushed-forward texture

Fig. 7. **Visualization of horse registration.** Texture transfer via the forward mean map helps visualize the correspondences from Fig. 6.

Fig. 8. **Toroidal parameterization**. Our Mongified approach can also directly compute a forward-backward Hessian-regularized area-preserving map from the Bob mesh (left) to a torus mesh (right), here visualized through texture transfer using the optimal forward mean map.



(a) Source, deformed source, target (front)

(b) Source, deformed source, target (back)

Fig. 9. **Pose transfer of human mesh.** Our Mongified spatially-regularized optimal transport approach finds a surface-to-surface map between two human poses (red and green). The blue mesh represents the original red mesh pushed forward by the optimal forward map. Three landmarks were used here: one on each shoulder and one on the right knee.

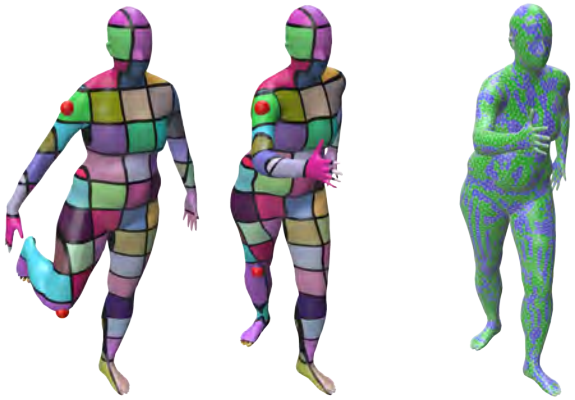


Fig. 10. **Visualization of human correspondences.** Texture transfer (source texture (left) pushed-forward by the forward mean map (middle)) helps visualize the correspondences from Fig. 9, and superposing the target mesh with the forward-deformed source mesh confirms accuracy.

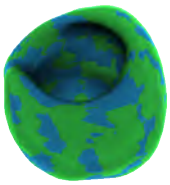


Fig. 11. **Evaluating ventricle registration.** For the registration from Fig. 4, we superpose the target and the forward-deformed source.

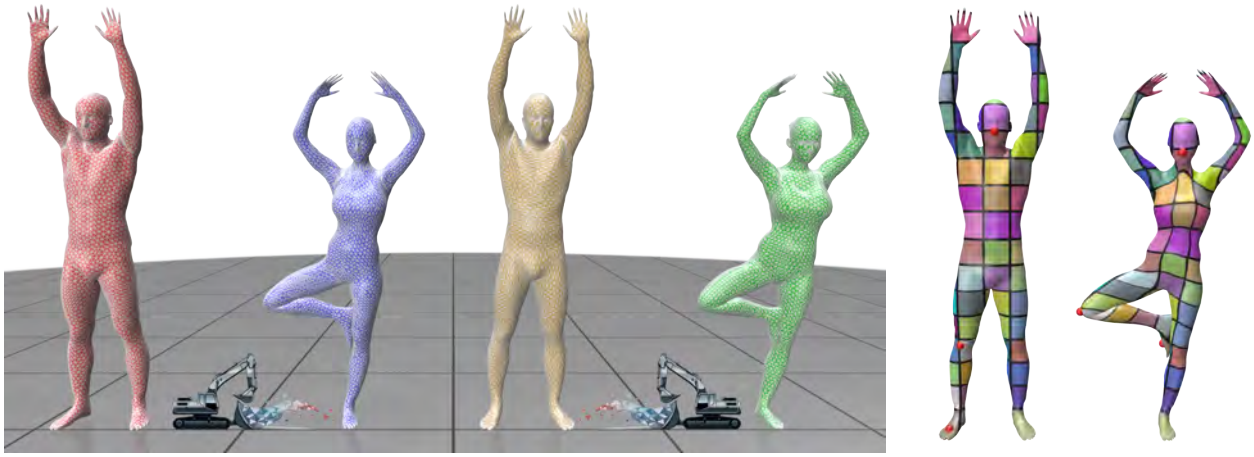
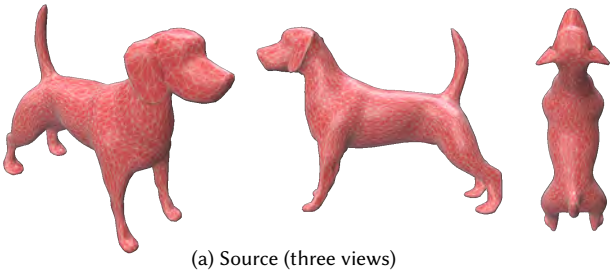
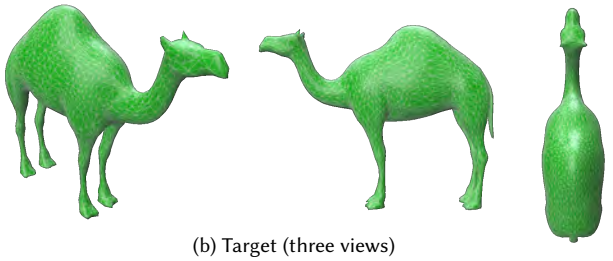


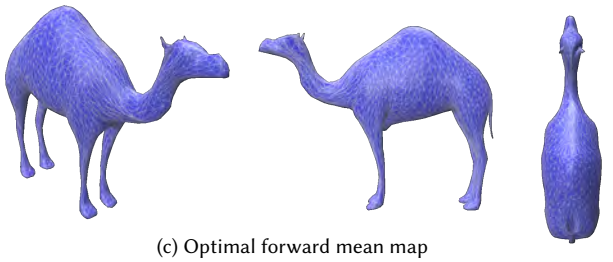
Fig. 12. **Near-isometric forward-backwards maps**. A Mongified forward-and-backward spatially-regularized OT-based registration finds a surface-to-surface map between the meshes of a man (red) and a woman (green) from the three landmarks provided (in red, nose, right knee and right foot). The optimal forward map (blue) sends the original man mesh onto the woman shape, while the optimal backward map (yellow) sends the original woman mesh onto the man shape. Texture transfer via the forward mean map is used to visualize the correspondences of the registration (right).



(a) Source (three views)



(b) Target (three views)



(c) Optimal forward mean map

Fig. 13. **Dog to camel.** In this example from Fig. 1, an unbalanced optimal transport for a spatially-regularized *forward* mean map leads to a registration of a dog (a) onto a camel (b), with only four landmarks on the face, two on the neck and one on the tip of the tail. The dog mesh displaced into the camel shape by the resulting forward mean map is shown in (c). This result converged in 3 ADMM iterations, hence after 3 Sinkhorn's algorithm calls.

Fig. 14. **Spherical parameterization**. Using our Mongified spatially-regularized optimal transport formulation with the Dirichlet energy provides a direct way to construct as-conformal-as-possible area-preserving maps from a mesh to another *and* the other way around as both mean maps are regularized — here, the Spot mesh is mapped to a sphere (no need to fix a single vertex), the parameterization being visualized (front and back) through texture transfer using the optimal forward mean map .

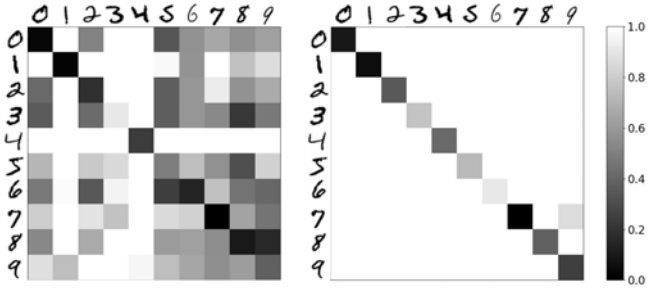


Fig. 15. **Dissimilarity measure.** From two sets of MNIST digits, pure OT costs do not cluster digits very well as visualized on a dissimilarity matrix (left) where we threshold off-diagonal pairwise-comparisons based on 110% of the maximum of the diagonal pairwise-comparisons and rescaled all values between 0 and 1. Our Mongified Hessian-regularized Euclidean OT costs are much more discriminatory than standard Euclidean OT costs given the same matrix visualization (right), leading to improved clustering: only a (7,9) pair is guessed similar, as it is close in value to the dissimilarity between the two 6s, which happen to be indeed topologically different, hence slightly more dissimilar than other same-digit pairs.