



Riemannian Locally Linear Embedding with Application to Kendall Shape Spaces

Elodie Maignant, Alain Trouvé, Xavier Pennec

► To cite this version:

Elodie Maignant, Alain Trouvé, Xavier Pennec. Riemannian Locally Linear Embedding with Application to Kendall Shape Spaces. GSI 2023: Geometric Science of Information, Aug 2023, Saint-Malo, (France), France. pp.12-20, 10.1007/978-3-031-38271-0_2 . hal-04122754

HAL Id: hal-04122754

<https://inria.hal.science/hal-04122754>

Submitted on 8 Jun 2023



HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution 4.0 International License

Riemannian Locally Linear Embedding with Application to Kendall Shape Spaces

Elodie Maignant^{1,2} , Alain Trounev², and Xavier Pennec¹ 

¹ Université Côte d’Azur and Inria, Epione team, France

`elodie.maignant@inria.fr`

² Centre Borelli, ENS Paris Saclay, France

Abstract. Locally Linear Embedding is a dimensionality reduction method which relies on the conservation of barycentric alignments of neighbour points. It has been designed to learn the intrinsic structure of a set of points of a Euclidean space lying close to some submanifold. In this paper, we propose to generalise the method to manifold-valued data, that is a set of points lying close to some submanifold of a given manifold in which the points are modelled. We demonstrate our algorithm on some examples in Kendall shape spaces.

Keywords: Locally Linear Embedding · Optimisation on Quotient Manifolds · Shape Spaces.

1 Introduction

Dimensionality reduction is a critical issue when it comes to data analysis on complex structures. Especially, the data modelled in the context of shape analysis – for example protein conformations or anatomical shapes – are by nature high-dimensional data. Common tools for dimensionality reduction have been originally designed for data described in a Euclidean space. However, objects like shapes are rather naturally described in a manifold. As an example, Kendall manifolds [3] encode the idea that two configurations of points – eg two protein conformations – should be compared independently of the coordinate system they are written in. We refer to such data as manifold-valued data. A first approach to process manifold-valued data then consists in embedding them in a larger Euclidean space – or equivalently to work extrinsically. This approach has two main drawbacks. First of all, it ignores the structural information contained in the manifold model, which then may not be well recovered in areas of low sampling density. Moreover, there might be a significant gap in dimensionality between the intrinsic and the extrinsic model in some cases. The manifold of unparameterised curves [5, 8] illustrates well this second point as the extrinsic and the intrinsic descriptions differ by the removal of parametrisations – diffeomorphisms – which is an infinite dimensional space. Thus, when the data are modelled in a known manifold, it is relevant to look for a generalisation of existing tools for vector-valued data to manifold-valued data. Locally Linear Embedding (LLE) has been introduced by Roweis and Saul in [7] as a nonlinear

dimensionality reduction tool. Given a set of points of a vector space sampled from some underlying submanifold of lower dimension, the method leverages the locally linear assumption to characterise each data point as a weighted barycentre of the other points nearby and then embed them in a low-dimensional vector space accordingly. Essentially, both the weights and the embedding are written as solutions of a least square problem such that the algorithm is straightforward to write and implement. LLE differs significantly from other dimension reduction methods, firstly as it relies on an intrinsic description of the data which is local – unlike PCA for example – and secondly as it implements a criterion which is affine rather than metric – as opposed to distance-based methods like Multi-Dimensional Scaling (MDS) or Isomap. Therefore, because LLE preserves local affine relationships rather than distances, we expect it to be able to retrieve different information from the data. While there has already been a consequent work about extending PCA [1] and MDS methods to manifold-valued data, LLE has not been yet generalised to our knowledge. In this paper, we propose a new Riemannian formulation of LLE which we refer to as Riemannian Locally Embedding (RLLE) and we detail an algorithm for the weights estimation. We illustrate our method on two examples in Kendall shape spaces and we evaluate RLLE performance in this setting with respect to the LLE one.

2 Riemannian Locally Linear Embedding

In this section, we recall the algorithm implemented by Locally Linear Embedding (LLE). Since it relies on barycentric coordinates, we extend their definition in order to generalise the method to Riemannian manifolds.

2.1 Outline of Locally Linear Embedding

Consider n points $x_1, \dots, x_n \in \mathbb{R}^m$ sampled from some underlying submanifold of lower dimension d . Then at a sufficiently local scale, the points should lie close to a linear subspace. Under this assumption, each point can be written as a linear combination of its neighbours up to some residual error. Following this observation, LLE implements two main steps. First, compute the approximate barycentric coordinates of x_i with respect to its k nearest neighbours, that is for all i , find w_{ij} which solve

$$\begin{aligned} \min_{w_{i1}, \dots, w_{in} \in \mathbb{R}} \quad & \left\| x_i - \sum_{j=1}^n w_{ij} x_j \right\|^2 \\ \text{subject to} \quad & \sum_j w_{ij} = 1 \end{aligned} \tag{1}$$

such that $w_{ij} = 0$ if x_j is not one of the k nearest neighbours of x_i . This amounts to solving n linear system $\in \mathbb{R}^k$. Then second step consists of finding n new points $y_1, \dots, y_n \in \mathbb{R}^d$ which best retrieve the weights estimated in Problem 1, that is solve

$$\min_{y_1, \dots, y_n \in \mathbb{R}^d} \sum_{i=1}^n \left\| y_i - \sum_{j=1}^n w_{ij} y_j \right\|^2. \tag{2}$$

Problem 2 is equivalent to an eigenvalue decomposition problem.

2.2 Riemannian Barycentric Coordinates

In order to generalise the method to manifold-valued data, we need to rewrite Problem 1 in a general non-Euclidean setting. Precisely, we need to extend the definition of a barycentre and barycentric coordinates. Intuitively, the weighted barycentre of a set of points is the point which minimises the weighted sum of squared distances. Let M denote a Riemannian manifold. We define the following.

Definition 1 (Riemannian barycentric coordinates). *A point $x \in M$ has barycentric coordinates w_1, \dots, w_n with respect to $x_1, \dots, x_n \in M$ if*

$$\sum_{i=1}^n w_i \log_x(x_i) = 0. \quad (3)$$

where \log_x denotes the logarithm map of M at x . A point x which satisfies the previous is called a weighted barycentre of $x_1, \dots, x_n \in M$ affected with the weights w_1, \dots, w_n .

Equation 3 can be interpreted as a first order condition on the minimisation of the weighted sum of squared distances. We can check that both descriptions coincide with the usual definition of barycentric coordinates in a Euclidean setting. Indeed, for $M = \mathbb{R}^m$ and $x, y \in M$, the logarithm map of M at x is simply $\log_x(y) = y - x$. Therefore, according to the previous definition, the barycentre of $x_1, \dots, x_n \in M$ affected with the weights w_1, \dots, w_n is the point x which satisfies $x = \sum_i w_i x_i$. For more details, we refer the reader to [6]. Let us now rewrite Problem 1 for points of a Riemannian manifold.

2.3 Towards Riemannian Locally Linear Embedding

The generalisation of Problem 1 is not straightforward as the previous definition is an implicit definition and does not allow to write a weighted barycentre in closed-form except in the Euclidean case. Instead, one can introduce an auxiliary variable $\hat{x}_i \in M$ satisfying Equation 3 and generalise Problem 1 as a constrained optimisation problem on manifolds. More explicitly, for $x_1, \dots, x_n \in M$, solve

$$\begin{aligned} \min_{\substack{\hat{x}_i \in M, \\ w_{i1}, \dots, w_{in} \in \mathbb{R}}} & d_M(x_i, \hat{x}_i)^2 \\ \text{subject to} & \sum_j w_{ij} \log_{\hat{x}_i}(x_j) = 0 \\ & \sum_j w_{ij} = 1. \end{aligned} \quad (4)$$

It is not trivial however to solve this problem in practice. Especially, the constraint $\sum_j w_{ij} \log_{\hat{x}_i}(x_j) = 0$ lies in the tangent space $T_{\hat{x}_i}M$ which depends itself on the value of \hat{x}_i which we wish to optimise. Rather, we propose to look at the equivalent translated problem

$$\begin{aligned} \min_{\substack{\hat{x}_i \in M, \\ w_{i1}, \dots, w_{in} \in \mathbb{R}}} & d_M(x_i, \hat{x}_i)^2 \\ \text{subject to} & \sum_j w_{ij} P_{\hat{x}_i, x_i}(\log_{\hat{x}_i}(x_j)) = 0 \\ & \sum_j w_{ij} = 1. \end{aligned} \quad (5)$$

where $P_{\hat{x}_i, x_i}$ denotes the parallel transport map of M along the geodesic joining \hat{x}_i and x_i . Since the parallel transport is an isometric map, Problems 4 and 5 are equivalent. In this new formulation however, the constraint lies in the tangent space at x_i , which is independent of the optimisation state.

Riemannian Locally Linear Embedding (RLLE) implements two steps according to the same scheme as LLE. First, the reconstruction step consists in estimating the weights w_{ij} solving Problem 5. Then, the embedding step consists in computing the points y_i solving Problem 2.

3 Algorithm and Implementation

In this section, we provide one possible algorithm to solve the optimisation Problem 2 and describe the implementation of our method. We detail the algorithm for Kendall shape spaces.

3.1 Tangent Space Formulation of the Optimisation Problem

In its current formulation, Problem 5 can be solved using Lagrangian methods for constrained optimisation on manifolds. However, it can be also be formulated alternatively as a vector-valued optimisation problem. Precisely, we keep track of the estimate \hat{x}_i with the tangent vector $v_i \in T_{x_i}M$ such that

$$\exp_{x_i}(v_i) = \hat{x}_i \quad (6)$$

where \exp_{x_i} denotes the exponential map of M at x_i . Additionally, we set

$$u_{ij} = P_{\hat{x}_i, x_i}(\log_{\hat{x}_i}(x_j)). \quad (7)$$

We derive the following optimisation problem

$$\begin{aligned} \min_{\substack{u_{i1}, \dots, u_{in} \in T_{x_i}M, \\ v_i \in T_{x_i}M, \\ w_{i1}, \dots, w_{in} \in \mathbb{R}}} & \|v_i\|^2 \\ \text{subject to} & \sum_j w_{ij} u_{ij} = 0 \\ & \sum_j w_{ij} = 1 \\ & \exp_{[\exp_{x_i}(v_i)]}(P_{x_i, [\exp_{x_i}(v_i)]}(u_{ij})) = x_j \quad (\forall j). \end{aligned} \quad (8)$$

Problem 8 is a priori a vector-valued optimisation problem on the product space $(T_{x_i}M)^n \times T_{x_i}M \times \mathbb{R}^n$. In fact, since the weight w_{ij} is set to be 0 whenever x_j is not a neighbour of x_i , then the correct search space is $(T_{x_i}M)^k \times T_{x_i}M \times \mathbb{R}^k$ where k is the number of neighbours. Now, provided that a basis of $T_{x_i}M^n$ can be explicitly computed, then the search space is the Euclidean space \mathbb{R}^{mk+m+k} , where m is the dimension of M , and the optimisation task is performed using standard Lagrangian methods implemented in most libraries. As a reference, we

use the `SLSQP` solver from `scipy`. Note that in practice, the complexity of the algorithm strongly depends on whether one knows the exponential map and the parallel transport in closed-form as we will discuss later in the paper. In any case, it requires an implementation of both methods which is compatible with automatic differentiation.

Note that another possible way to solve Problem 5 would be to address it directly as a Riemannian constrained optimisation problem using specific optimisation tools like the ones implemented in the `Manopt` library.

3.2 Riemannian Locally Linear Embedding for Quotient Manifolds

In what follows, we propose to detail the algorithm in the concrete case of Kendall shape spaces [3]. We recall that Kendall shape spaces carry a quotient structure. We first describe the algorithm for a general quotient manifold and then give an explicit formulation for Kendall shape spaces. For there does not always exist an explicit description for quotient objects, computations in quotient spaces are generally performed in the top space. It is also often more comfortable. The main motivation of this subsection is to show how to perform the previous optimisation task in the top space.

The setting is the following. Let again M be a Riemannian manifold and let G be a group acting on M . Given $x_1, \dots, x_n \in M$, we want to solve Problem 5 for the corresponding data points $\pi(x_1), \dots, \pi(x_n) \in M/G$, where $\pi : M \rightarrow M/G$ is the canonical quotient map. Assume that π is a Riemannian submersion. The vertical space of M at a point x , denoted by $\text{Ver}_x M$, is defined by

$$\text{Ver}_x M = \ker d_x \pi. \quad (9)$$

The tangent space of M at x admits an orthogonal decomposition

$$T_x M = \text{Ver}_x M \oplus \text{Hor}_x M \quad (10)$$

and $\text{Hor}_x M$ is called the horizontal subspace of M at x . A central property is that the tangent space of M/G at a point $\pi(x)$ identifies with the horizontal space of M at x through the tangent map $d\pi$. Moreover, geodesics of M/G correspond exactly to the projection by π of horizontal geodesics of M , that is geodesics spanned by a horizontal vector. Additionally, we define the following

Definition 2 (Horizontal parallel transport). *Let γ be a horizontal curve in M . Then we say that the vector field $t \rightarrow v(t)$ is the horizontal parallel transport of a horizontal vector v along γ if it is horizontal and if its projection to the tangent bundle of M/G is the parallel transport of $d_x \pi(v)$ along $\pi(\gamma)$. We denote the horizontal transport map of M from a point x to a point y by $P_{x,y}^H$.*

Now let us go back to the algorithm. Since the tangent map $d\pi$ allows to identify the tangent spaces of M/G and the horizontal spaces of M , we can lift up

Problem 8 to the top manifold M

$$\begin{aligned}
& \min_{\substack{g_{i1}, \dots, g_{in} \in G \\ u_{i1}, \dots, u_{in} \in \text{Hor}_{x_i} M, \\ v_i \in \text{Hor}_{x_i} M, \\ w_{i1}, \dots, w_{in} \in \mathbb{R}}} \|v_i\|^2 \\
& \text{subject to} \quad \sum_j w_{ij} u_{ij} = 0 \\
& \quad \quad \quad \sum_j w_{ij} = 1 \\
& \quad \quad \quad g_{ij} \cdot \exp_{[\exp_{x_i}(v_i)]} (P_{x_i, [\exp_{x_i}(v_i)]}^H(u_{ij})) = x_j \quad (\forall j).
\end{aligned} \tag{11}$$

The optimisation variables g_{ij} are the elements of the group G lifting up the equality constraint

$$\pi(\exp_{[\exp_{x_i}(v_i)]} (P_{x_i, [\exp_{x_i}(v_i)]}^H(u_{ij}))) = \pi(x_j)$$

to the top space. If G is a matrix Lie group, then it identifies to its Lie algebra \mathfrak{g} through the matrix exponential such that the search space can still be written as a vector space.

3.3 Implementation in Kendall Shape Spaces

The implementation of the algorithm for Kendall shape spaces requires to compute the horizontal spaces, the exponential map and the horizontal parallel transport map. The Kendall shape space Σ_p^q is defined as the quotient

$$\Sigma_p^q = \mathcal{S}_p^q / SO(p) \tag{12}$$

where $\mathcal{S}_p^q = \{x \in M(p, q) \mid \sum x_i = 0 \text{ and } \|x\| = 1\}$ is referred as the pre-shape space [3]. The shape space Σ_p^q describes the possible configurations of a set of points independently of any similarity transformation of the ambient space. The space \mathcal{S}_p^q can be understood as the hypersphere of $\mathbb{R}^{p(q-1)}$ and its exponential map is given by

$$\exp_x(v) = \cos(\|v\|)x + \sin(\|v\|) \frac{v}{\|v\|}. \tag{13}$$

The horizontal subspace at x is described as

$$\text{Hor}_x M = \{v \in M(p, q) \mid \sum x_i = 0 \text{ and } vx^t = xv^t \text{ and } \langle x, v \rangle = 0\}.$$

where $\langle \cdot, \cdot \rangle$ denotes the usual Frobenius scalar product. The horizontal parallel transport can be computed as the solution of a first-order differential equation as described in [4]. Its implementation has already been discussed in the previous work [2] and is available in the library `geomstats`. Finally, the group of rotations $SO(p)$ is a Lie group and its Lie algebra is $\mathfrak{so}(p) = \text{Skew}(p)$.

4 Benchmark Experiment

In this section, we illustrate RLLE on two examples in the shape space Σ_3^3 and we compare its performance to LLE performance. We then discuss more generally the advantages of our method depending on the type of data.

4.1 A Swiss Roll example in the shape space Σ_3^3

We run two experiments in the Kendall shape space Σ_3^3 , for which we have an isometric embedding in a hemisphere and therefore a way to visually evaluate the embedding computed by either method. We first illustrate our algorithm on a set of data points generated from a mixture of normal distributions on the hemisphere (Figure 1). We compare the ability of LLE and RLLE to embed such a set in the plane. Our implementation of LLE performs first a Procrustean alignment step before solving Problem 1. Precisely, it aligns each neighbour x_j onto the point x_i by applying the optimal rotation. Note that without this alignment step, the performance of LLE drops significantly. Then we demonstrate numerically the performance of RLLE on an example derived from the "Swiss Roll" data set. Explicitly, given a set of shapes sampled along a logarithmic spiral curve in Σ_3^3 (Figure 2), we evaluate the one-dimensional parametrisation computed by LLE and RLLE with respect to the one given by the arc-length.

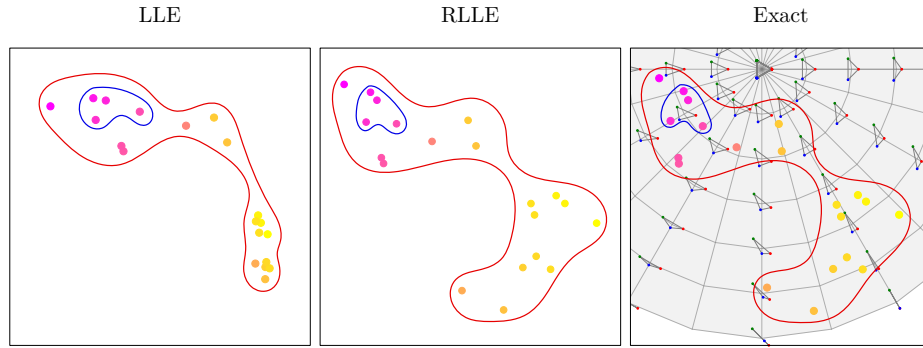


Fig. 1. First experiment. The points are equally sampled from two normal distributions in the hemisphere, flattened into a disk for visualisation. For both methods, the neighbour graph has been generated with the Riemannian distance. The number of neighbours k is chosen to optimise the performance of each method: $k = 9$ for LLE and $k = 10$ for RLLE. Since the methods are by definition invariant by affine transformation, we align each embedding onto the flattened data set. We observe that LLE is able to retrieve very local alignments (blue) but fails at a more global scale (red). This was to be expected as the Euclidean distance approximates the Riemannian distance locally. On the other hand, RLLE is able to retrieve the alignments at every scale.

4.2 Computational complexity

RLLE shares the main drawback of intrinsic manifold learning methods: it is computationally quite expensive. Let us detail this point. We mainly focus on the reconstruction step as the embedding step is common to the LLE method. First, the search space is a space of dimension $km + m + k$, where m is the dimension

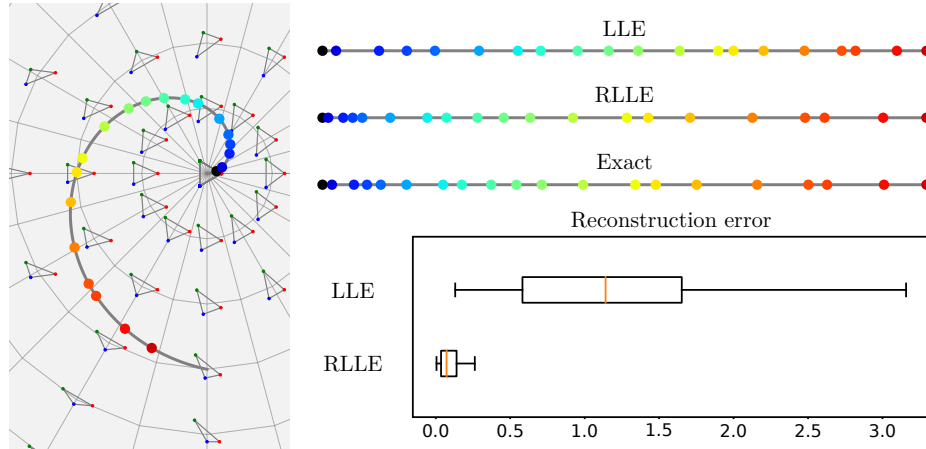


Fig. 2. Second Experiment. The points are sampled along a spiral curve. We fix $k = 3$ for both methods. We illustrate the experiment for one sample in the left and upper-right sub-figures. Both methods are ordering the points correctly. Relative distances however are better preserved by RLLE. We estimate the absolute error between the exact embedding and the one computed by RLLE (resp. LLE). The experiment is run 100 times. We summarise the performance of each method in a box plot. We observe that RLLE performs significantly better and moreover is more stable.

of the manifold M and k is the number of neighbours. For the Kendall shape space of parameters p and q , we have $m = p(q - \frac{1}{2}(p - 1)) - 1$. Then the number of optimisation problems to solve is n . Finally, we need to take into account the computational cost of the exponential and the parallel transport methods. In the case of Kendall shape spaces, while the first one is free, the latter performs in $\mathcal{O}(2sp^3 + pq)$ as detailed in [2], where s is the number of integration steps. Each evaluation of the constraint – and so each step of the optimisation task – costs the same. Finally, our methods implements a SQP method to solve each optimisation problem. As a comparison, the solution of Problem 1 is equivalent to a matrix inversion of dimension k . Therefore, LLE performs in roughly $\mathcal{O}(nk^3)$, such that its complexity does not depend on the dimension m of the data.

4.3 Further discussion

Given the computational cost of RLLE, it is important to understand for which type of data it is particularly suited. Typically, the locally linear assumption made by LLE may be valid for large and well-concentrated data sets. In these cases, both methods should perform the same. Moreover, non-local methods like PCA or its manifold generalisations provide a correct estimation whenever the data are sufficiently concentrated. Finally, LLE and RLLE seem of particular use in cases where a distance-based method does not perform well. These remarks suggest that RLLE is more specifically designed for small sample size data sets

with large dispersion, and provides an embedding which might allow to observe more informative patterns than the ones characterised by the distance only.

5 Conclusion

As for now, RLLE has been implemented for Kendall shape spaces only. We are contemplating a general implementation of the method into the library `geomstats` for various manifolds and quotient manifolds – for example the space of unlabelled graphs. Applications to real data sets will also be developed in future works. Especially, we wish to further investigate the analysis of protein conformations using Kendall’s framework following one of our previous works.

Acknowledgements: We thank G-stats team for the fruitful discussions. This work was supported by the ERC grant #786854 G-Statistics from the European Research Council under the European Union’s Horizon 2020 research and innovation program and by the French government through the 3IA Côte d’Azur Investments ANR-19-P3IA-0002 managed by the National Research Agency.

References

1. Fletcher, P.T., Lu, C., Pizer, S.M., Joshi, S.: Principal geodesic analysis for the study of nonlinear statistics of shape. *IEEE transactions on medical imaging* **23**(8), 995–1005 (2004)
2. Guigui, N., Maignant, E., Trouvé, A., Pennec, X.: Parallel transport on kendall shape spaces. In: *Geometric Science of Information: 5th International Conference, GSI 2021, Paris, France, July 21–23, 2021, Proceedings 5*. pp. 103–110. Springer (2021)
3. Kendall, D.G.: Shape manifolds, procrustean metrics, and complex projective spaces. *Bulletin of the London mathematical society* **16**(2), 81–121 (1984)
4. Kim, K.R., Dryden, I.L., Le, H., Severn, K.E.: Smoothing splines on riemannian manifolds, with applications to 3D shape space. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* **83**(1), 108–132 (2021)
5. Lahiri, S., Robinson, D., Klassen, E.: Precise matching of PL curves in \mathbb{R}^N in the square root velocity framework. *arXiv preprint arXiv:1501.00577* (2015)
6. Pennec, X.: Barycentric subspace analysis on manifolds. *Annals of Statistics* **46**(6A), 2711–2746 (2018)
7. Roweis, S.T., Saul, L.K.: Nonlinear dimensionality reduction by locally linear embedding. *science* **290**(5500), 2323–2326 (2000)
8. Tumpach, A.B., Preston, S.C.: Quotient elastic metrics on the manifold of arc-length parameterized plane loops. *arXiv preprint arXiv:1601.06139* (2016)