Supplementary Materials

1. Background: Manifold of Surface Shapes

A *continuous surface* can be described by a function $q : \mathcal{I} \rightarrow \mathbb{R}^3$, where $\mathcal{I}$ is a two-dimensional space of parameters $(u, v) \in \mathcal{I}$ that parameterize the 3D points $q(u, v) \in \mathbb{R}^3$ on the surface. The space of surfaces is denoted $\mathcal{M} \subset C^\infty(\mathcal{I}, \mathbb{R}^3)$. The space $\mathcal{M}$ is an infinite dimensional manifold immersed in the infinite dimensional vector space $C^\infty(\mathcal{I}, \mathbb{R}^3)$. Recall that one surface shape is represented by one point on the manifold of surface shapes. Thus, the Riemannian metric we choose to equip the manifold $\mathcal{M}$ with defines the distance between its points $q_0, q_1 \in \mathcal{M}$ and also yields the notion of dissimilarity between the two surfaces q_0, q_1. We consider the second-order Sobolev metric [2]:

$$\begin{aligned} G_q(h,k) = \int_M & \big(a_0\langle h,k\rangle + a_1 g_q^{-1}\left(dh_m, dk_m\right) \\ & + b_1 g_q^{-1}\left(dh_+, dk_+\right) + c_1 g_q^{-1}\left(dh_\perp, dk_\perp\right) \\ & + d_1 g_q^{-1}\left(dh_0, dk_0\right) + a_2\left\langle \Delta_q h, \Delta_q k\right\rangle\big) \operatorname{vol}_q, \end{aligned} \tag{1}$$

where h, k are tangent vectors at point $q \in \mathcal{M}$; g_q^{-1} is the pullback metric from $\mathbb{R}^3$ that defines distances on the surface q itself; Δ_q is the Laplacian induced by q; $dh_m, dh_+, dh_\perp, dh_0$ are orthogonal vector-valued one-forms and vol_q is the surface area measure of q. The scalars $a_0, a_1, a_2, b_1, c_1, d_1$ are weighting parameters that define distance between two surfaces based on how they are sheared, scaled, bent, or parameterized with respect to each other. The second-order Sobolev metric cannot distinguish between a surface and its translation, so it inherantly calculates distance without considering the position of the shape in $\mathbb{R}^3$, ensuring that a difference in spatial translation would not affect a distance computation between two shapes. Because of this invariance, mathematically, the space of surfaces is defined as the quotient space: $\mathcal{I}/(\mathbb{R}^3)$.

We can quotient by more actions to give an even more well defined notion of "shape". In the space of surfaces $\mathcal{M}$ described above, two surfaces with the same shape but different orientations would correspond to different points. By contrast, we introduce the space of surface *shapes* $\mathcal{S}$ where two surfaces with the same shape correspond to the same point, regardless of differences in their orientation. Mathematically, the space of surface shapes is defined as the quotient space: $\mathcal{I}/(\operatorname{Rot}(\mathbb{R}^3) \times \mathbb{R}^3)$ —see [2] for details.

In the shape space $\mathcal{S}$, the distance between two surface shapes q_1 and q_2 is given by:

$$d^{\mathcal{S}}(q_1, q_2) = \inf_\phi d(q_1, q_2 \circ \phi) = d(q_1, q_2'), \tag{2}$$

where ϕ represents a choice in orientation. In Eq. (2), the orientation of q_2 is varied until the second-order Sobolev distance d in Eq. (1) between q_1 and q_2 reaches an infimum. This operation matches the orientation of q_2 to the orientation of q_1 so that any remaining discrepancy between them is due to difference in shape.

2. Computation of the Riemannian Gradient

We consider the generative model with Euclidean Gaussian noise, and its associated loss function for geodesic regression. The gradient of the loss function associated with this model can be computed as a Riemannian gradient or as an extrinsic gradient. Here, we give the formula for the Riemannian gradient.

Proposition 1. *The Riemannian gradient is given by:*

$$\nabla_p l = -\sum_{i=1}^n d_p \operatorname{Exp}\left(p, X_i v\right)^\dagger r_i,$$

$$\nabla_v l = -\sum_{i=1}^n x_i d_v \operatorname{Exp}\left(p, X_i v\right)^\dagger r_i,$$

where l is the loss function, $r_i = y_i - \operatorname{Exp}\left(p, X_i v\right)$ are the residuals, d_v and d_p are derivatives, and $\dagger$ is the adjoint. In other words, the Riemannian gradient has the same form as in traditional geodesic regression, only the residuals are given by $r_i = y_i - \operatorname{Exp}\left(p, X_i v\right)$.

Proof. Recall that the loss function l is given by:

$$l = \sum_{i=1}^n \|Y_i - \hat{Y}_i\|^2, \tag{3}$$

where $\hat{Y}_i = \operatorname{Exp}\left(p, X_i v\right)$. Following the proof in [3], we compute the gradient of l with two steps: (i) the derivative of the squared Euclidean distance, and (ii) the derivative of the exponential map.

Considering (i), the gradient of the squared Euclidean distance function for a fixed $Y_i \in \mathbb{R}^D$ is:

$$\nabla_{\hat{Y}_i} \|Y_i - \hat{Y}_i\|^2 = -2(Y_i - \hat{Y}_i). \tag{4}$$

We note that the computation of the gradient is simplified compared to the classic geodesic regression case, both in terms of its equation but also in terms of computational complexity.

Considering (ii), the gradient of the exponential map w.r.t. p, v, for any u_1, u_2, is given by:

$$\begin{aligned} d_p \operatorname{Exp}(p, X_i v) \cdot u_1 &= J_1(1), \\ d_v \operatorname{Exp}(p, X_i v) \cdot u_2 &= J_2(1), \end{aligned} \tag{5}$$

where J_1, J_2 are Jacobi fields along the geodesic $\gamma(t) = \operatorname{Exp}(p, tv)$ associated with u_1, u_2 respectively, and defined as follows. Specifically, J_1, J_2 are solutions to the second order equation

$$\frac{D^2}{dt^2}J(t) + R\left(J(t), \quad \gamma'(t)\right)\gamma'(t) = 0,$$

where R is the Riemannian curvature tensor, with initial conditions $J_1(0) = u_1, J_1'(0) = 0$ and $J_2(0) = 0, J_2'(0) = u_2$, respectively [1].

We introduce the adjoint operators defined by, for all w:

$$< d_p\text{Exp}(p, v) \cdot u_1, w > = < u_1, d_p\text{Exp}(p, v)^\dagger w >$$
$$< d_v\text{Exp}(p, v) \cdot u_2, w > = < u_2, d_v\text{Exp}(p, v)^\dagger w > .$$

Using $w = r_i$, we put (i) and (ii) together using the chain rule, and we get:

$$\nabla_p l = -\sum_{i=1}^{n} d_p \,\text{Exp}\,(p, X_i v)^\dagger \, r_i,$$
$$\nabla_v l = -\sum_{i=1}^{n} x_i d_v \,\text{Exp}\,(p, X_i v)^\dagger \, r_i,$$

where now, $r_i = Y_i - \text{Exp}\,(p, X_i v)$.

We find the same formula as in [3], except the geodesic residuals have been replaced by the linear residuals. This concludes the proof. □

3. Probability Density in Projected Gaussian Euclidean Noise

For convenience of notations, we drop the subscripts i.

We compute the formula for the probability density function associated with the model with projected Euclidean Gaussian noise, given by the following proposition.

Proposition 2. *The probability density function associated with the model with projected Euclidean noise is:*

$$p(Y_i \mid X_i;\, p, v) = \frac{1}{\sqrt{2\pi\sigma^2}^m} \exp\left(-\frac{\|P_Y^\perp(\tilde{Y}) - \tilde{Y}\|^2}{2\sigma^2}\right),$$

where we introduce the projection $P_Y^\perp$, different from $\mathcal{P}$, that projects the noiseless data point $\tilde{Y}$ onto the subspace $T_Y\mathcal{M}^\perp$.

Proof. The probability density before projection via $\mathcal{P}$ is:

$$p(\bar{Y} \mid X;\, p, v) = \frac{1}{C(\sigma)} \exp\left(-\frac{\|\bar{Y} - \text{Exp}(p, Xv)\|^2}{2\sigma^2}\right),$$

with $C(\sigma) = \sqrt{(2\pi)^D \sigma^{2D}}$ the normalization constant. We emphasize that the notation $\bar{Y}$ does not denote mean, but rather the data point from the unprojected probability distribution.

The probability density after projection is:

$$p(Y \mid X;\, p, v) = \int_{\bar{Y} \in T_Y\mathcal{M}^\perp} p(\bar{Y}|X) d\bar{Y}, \tag{6}$$

where we integrate over the values of $\bar{Y}$ that will give the same projection Y on $\mathcal{M}$. We refer the reader to Figure 1 for the notations used in this proof.

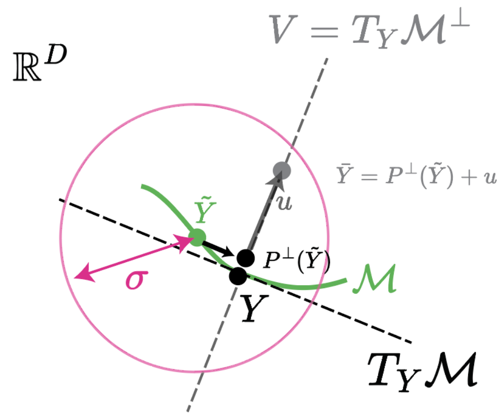

Figure 1. Notations for the Projected Euclidean Noise. $\tilde{Y}$: noiseless data point, $\bar{Y}$: noisy data point before projection, Y: noisy data point after projection via $P_Y^\perp$ on $\mathcal{M}$. The pink circle represents a level set of the isotropic Gaussian noise in ambient space, which has mean $\tilde{Y}$. We have decomposed the ambient space $\mathbb{R}^D$ into $T_Y\mathcal{M}$ and its orthogonal $T_Y\mathcal{M}^\perp$ (dashed lines).

We have:

$$\begin{aligned} & p(Y \mid X;\, p, v) \\ &= \int_{\bar{Y} \in T_Y\mathcal{M}^\perp} p(\bar{Y}|X) d\bar{Y} \\ &= \frac{1}{C(\sigma)} \int_{\bar{Y} \in T_Y\mathcal{M}^\perp} \exp\left(-\frac{\|\bar{Y} - \tilde{Y}\|^2}{2\sigma^2}\right) d\bar{Y}, \end{aligned}$$

where we use the notation $\tilde{Y} = \text{Exp}(p, Xv)$ (see Figure 1).

This equation shows the integral of a multivariate Gaussian distribution in $\mathbb{R}^D$ with mean $\tilde{Y}$ and isotropic variance $\sigma^2 I$ along a subspace V of $\mathbb{R}^D$. Here, V is the subspace perpendicular to $T_Y\mathcal{M}$ at Y: $V = T_Y\mathcal{M}^\perp$. We denote $P_Y^\perp(\tilde{Y})$, the orthogonal projection of the mean $\tilde{Y}$ onto V.

We consider a change of variable $\bar{Y} = P_Y^\perp(\tilde{Y}) + u$, where $u \in \mathbb{R}^{D-m}$ with m the dimension of $\mathcal{M}$ (see Figure 1). Pythagorean theorem in ambient space $\mathbb{R}^D$ gives:

$$
\begin{aligned}
&p(Y \mid X;\ p, v) \\
&= \frac{1}{C(\sigma)} \int_{\bar{Y} \in T_Y \mathcal{M}^\perp} \exp\left(-\frac{\|\bar{Y} - \tilde{Y}\|^2}{2\sigma^2}\right) d\bar{Y} \\
&= \frac{1}{C(\sigma)} \int_{u \in \mathbb{R}^{D-m}} \exp\left(-\frac{\|P_Y^\perp(\tilde{Y}) + u - \tilde{Y}\|^2}{2\sigma^2}\right) du \\
&= \frac{1}{C(\sigma)} \int_{u \in \mathbb{R}^{D-m}} \exp\left(-\frac{\|P_Y^\perp(\tilde{Y}) - \tilde{Y}\|^2 + \|u\|^2}{2\sigma^2}\right) du \\
&= J(\sigma) \int_{u \in \mathbb{R}^{D-m}} \exp\left(-\frac{\|u\|^2}{2\sigma^2}\right) du,
\end{aligned}
$$

where we put the terms that are independent of u outside the integral and where we define:

$$
J(\sigma) = \frac{1}{C(\sigma)} \exp\left(-\frac{\|P_Y^\perp(\tilde{Y}) - \tilde{Y}\|^2}{2\sigma^2},\right) \tag{7}
$$

for convenience of notations. We extract the $D - m$ coordinates of u to get a first analytical expression for the probability density function of the generative model of geodesic regression with projected Euclidean Gaussian noise:

$$
\begin{aligned}
p(Y \mid X;\ p, v) &= J(\sigma) \left(\int_{u' \in \mathbb{R}} \exp\left(-\frac{u'^2}{2\sigma^2}\right) du'\right)^{D-m} \\
&= J(\sigma) \sqrt{2\pi\sigma^2}^{D-m},
\end{aligned}
$$

using the formula for the Gaussian integral: $\int_x \exp(-ax^2)dx = \sqrt{\frac{\pi}{a}}$ with $a = \frac{1}{2\sigma^2}$.

Putting these computations together, we get:

$$
\begin{aligned}
&p(Y \mid X;\ p, v) \\
&= \frac{1}{C(\sigma)} \exp\left(-\frac{\|P_Y^\perp(\tilde{Y}) - \tilde{Y}\|^2}{2\sigma^2}\right) \sqrt{2\pi\sigma^2}^{D-m} \\
&= \frac{\sqrt{2\pi\sigma^2}^{D-m}}{\sqrt{2\pi\sigma^2}^{D}} \exp\left(-\frac{\|P_Y^\perp(\tilde{Y}) - \tilde{Y}\|^2}{2\sigma^2}\right) \\
&= \frac{1}{\sqrt{2\pi\sigma^2}^{m}} \exp\left(-\frac{\|P_Y^\perp(\tilde{Y}) - \tilde{Y}\|^2}{2\sigma^2}\right),
\end{aligned}
$$

where we note that the dependency in Y is within the projection $P^\perp$ onto the subspace $V = T_Y\mathcal{M}$, and the dependency in X is within $\tilde{Y} = \text{Exp}(p, Xv)$. □

4. Probability Density in Metric Deformed Gaussian Euclidean Noise

We drop the subscripts i for convenience of notations.

We provide the probability density function associated with the model with deformed Euclidean Gaussian noise in the proposition below.

Proposition 3. *The probability distribution associated with the generative model with deformed Euclidean Gaussian noise is:*

$$
p(Y_i \mid X_i;\ p, v) = \frac{1}{C(\sigma)\sqrt{\det G(Y_i)}} \exp\left(-\frac{\|Y_i - \tilde{Y}_i\|^2}{2\sigma^2}\right),
$$

where $C(\sigma) = \sqrt{(2\pi)^D \sigma^{2D}}$ is the normalization constant, and G represents the matrix of the Riemannian metric of $\mathcal{M}$ at Y_i.

Note that this probability distribution differs from the classical linear regression, through the term $\sqrt{\det G(Y_i)}$. We provide a proof below.

Proof. The probability density before deformation is:

$$
p(\bar{Y} \mid X;\ p, v) = \frac{1}{C(\sigma)} \exp\left(-\frac{\|\bar{Y} - \text{Exp}(p, Xv)\|^2}{2\sigma^2}\right),
$$

with $C(\sigma) = \sqrt{(2\pi)^D \sigma^{2D}}$ the normalization constant.

The "deformation" amounts to considering a different metric on the manifold $\mathcal{M}$ in which Y takes values. This implies that we consider a different measure on the manifold $\mathcal{M}$. The measure associated with Euclidean metric, used to write the Gaussian distribution above is dY. By definition, the measure associated with a Riemannian metric is then: $\sqrt{\det(G(Y))}$ where G is the matrix of the inner-product of the Riemannian metric at point Y.

With respect to this new measure, the probability density function is therefore:

$$
p(Y_i \mid X_i;\ p, v) = \frac{1}{C(\sigma)\sqrt{\det G(Y_i)}} \exp\left(-\frac{\|Y_i - \tilde{Y}_i\|^2}{2\sigma^2}\right).
$$

□

5. Why don't we test geodesic noise, Euclidean noise, and projected Euclidean noise generative models on mesh data?

We do not test the geodesic noise model on mesh data because as described in the introduction, geodesic noise is never added to mesh data.

We cannot test the Euclidean noise model on mesh data because the curvature of the manifold of surface shapes is introduced by the second-order Sobolev metric [2], which quotients by translation. A distance between two meshes on the manifold is measured by disregarding the positions of their barycenters. Therefore, any computation done on this manifold using the second-order Sobolev metric would implicitly submerse linear noise to the manifold, and we would actually be testing Submersed Euclidean Noise.

We cannot test projected Euclidean noise on mesh data because as mentioned above, meshes are brought to the manifold by disregarding the positions of their barycenters, so bringing any mesh to the manifold of surface shapes actually entails a manifold submersion, which is not the same as a "projection", which for example brings a point in embedding space to the surface of a hypersphere.

References

[1] Manfredo Perdigao Do Carmo and J Flaherty Francis. *Riemannian geometry*, volume 6. Springer, 1992. 2

[2] Emmanuel Hartman, Yashil Sukurdeep, Eric Klassen, Nicolas Charon, and Martin Bauer. Elastic shape analysis of surfaces with second-order sobolev metrics: a comprehensive numerical framework. *International Journal of Computer Vision*, 131(5):1183–1209, 2023. 1, 3

[3] P Thomas Fletcher. Geodesic regression and the theory of least squares on riemannian manifolds. *International journal of computer vision*, 105:171–185, 2013. 1, 2