A. Variational Autoencoders

This section reviews VAEs and the tools of Riemannian geometry that support our curvature estimation method. For further background in variational inference and VAEs, we direct the reader to [6, 18].

A *variational autoencoder* (VAE) is a generative deep latent variable model widely used for unsupervised learning [30]. A VAE uses an autoencoder architecture to implement variational inference: it is trained to encode input data in a compact latent representation, and then decode the latent representation to reconstruct the original input data. Consider an N-dimensional dataset of k vectors $x_1, \dots, x_k \in \mathbb{R}^N$. The VAE models each data vector x_i as being sampled from a likelihood distribution $p(x_i|z_i)$ with lower-dimensional unobserved latent variable z_i. The likelihood distribution is usually taken to be Gaussian, so we write the reconstructed input as $x_i^{rec} = f(z_i) + \epsilon_i$ with $\epsilon_i \sim \mathcal{N}(0, \sigma^2 \mathbb{I}_N)$. The function f is here represented by a neural network called the *decoder*.

The VAE simultaneously trains an *encoder* that represents the approximate posterior distribution $q(z|x)$ over the latent variables z. The VAE achieves its objective by minimizing an upper-bound of the negative log-likelihood, which writes as the sum of a reconstruction loss and a Kullback-Leibler (KL) divergence:

$$\begin{aligned} \mathcal{L} &= \mathcal{L}_{rec} + \mathcal{L}_{KL} \\ &= -\mathbb{E}_{q(z)}[\log p(x|z)] + \mathrm{KL}(q(z|x) \parallel p(z)). \end{aligned} \tag{1}$$

We use a similar loss in our experiments, but we adapt the KL term to the topology of the latent space, which we call the template manifold $\mathcal{Z}$.

B. Derivations of the Second Fundamental Form for Surfaces in 3-D

We give additional details on the second fundamental form II, for a surface within $\mathbb{R}^3$, and the associated mean curvature vector.

Definition 7 (Second fundamental form II of a surface in $\mathbb{R}^3$)**.** *Consider a surface S in $\mathbb{R}^3$, as the graph of a twice differentiable function $f(u, v)$. We choose our coordinate system such that $f(u, v) = 0$ defines the tangent plane to S at the point z and (u, v) are the coordinates of a small displacement $x \in T_z S$. By Taylor's theorem, the best approximation of $f(u, v)$ in a small region around z is*

$$f_z(u, v) = \frac{1}{2}(Lu^2 + 2Muv + Nv^2) = \frac{1}{2} x^T \mathrm{II}_z \, x,$$
$$\text{with: } \mathrm{II}_z = \begin{bmatrix} L & M \\ M & N \end{bmatrix}.$$

In this equation, II_z is a matrix of second partial derivatives of f at z, called the second fundamental form of the surface.

The second fundamental form allows to define the mean curvature vector.

Definition 8 (Mean curvature H of a surface in $\mathbb{R}^3$ from its second fundamental form II)**.** *Consider a surface in $\mathbb{R}^3$ represented by its second fundamental form II_z. Then, its mean curvature vector is given by:*

$$H_z = \frac{1}{2} Tr \, \mathrm{II}_z. \tag{2}$$

In the specific case of two-dimensional surfaces immersed in $\mathbb{R}^3$, the mean curvature vector also enjoys an equivalent definition, given below.

Definition 9 (Mean curvature H of a surface in $\mathbb{R}^3$)**.** *Consider a 2-dimensional surface S embedded in $\mathbb{R}^3$. A normal vector at a point $z \in S$ defines a family of normal planes containing this vector, each of which cuts the surface in a direction ψ producing a plane curve. The curvature of such a curve at z is given by $\kappa_z = \frac{1}{R(z)}$, where $R(z)$ is the radius of the osculating circle at that point, i.e. the circle that best approximates this curve locally.*

The mean curvature H_z at the point z is defined as

$$H_z = \frac{1}{2\pi} \int_0^{2\pi} \kappa_z(\psi) \, d\psi, \tag{3}$$

which is the average of the curvatures $\kappa_z(\psi)$ over all directions ψ in the tangent plane at z.

Definition 9 provides the intuition behind the name "mean" curvature, as its defining equation is effectively a mean.

C. Derivations of the Mean Curvature Vectors

C.1. General Formula

We present the general definition of mean curvature, that builds on the definition of second fundamental form. We refer the reader to the next subsections for concrete examples of these definitions in the special case of two-dimensional surfaces in $\mathbb{R}^3$.

Definition 10 (Second fundamental form [3])**.** *Consider the manifold $\mathcal{M}$ represented as the immersion of $\mathcal{Z}$ into $\mathcal{X}$ such that $\mathcal{M} = f(\mathcal{Z})$, $\mathcal{M} \subset \mathcal{X}$. We have:*

$$\begin{aligned} \mathrm{II}(z)^\alpha_{ij} = \nabla^2_{ij} f^\alpha(z) &= \partial^2_{ij} f^\alpha(z) - \Gamma^k_{ij}(z) \partial_k f^\alpha(z) \\ &+ \bar{\Gamma}^\alpha_{\beta\gamma}(f(z)) \partial_i f^\beta(z) \partial_j f^\gamma(z). \end{aligned}$$

Here, Γ^k_{ij} are the Christoffel symbols of $\mathcal{Z}$ for the pullback metric and $\bar{\Gamma}^\alpha_{\beta\gamma}$ the Christoffel symbols of $\mathcal{X}$ for the metric of $\mathcal{X}$. In this formula, i, j are indices for basis elements of $T_{f(z)}\mathcal{X}$, identified with basis elements of $T_z\mathcal{Z}$ since both tangent spaces share the same metric; while α in an index for a basis element of $N_{f(z)}\mathcal{X}$.

We note that, in the case where $\mathcal{X} = \mathbb{R}^N$ or $\mathcal{X} = \mathbb{R}^N_+$, the Christoffel symbols $\bar{\Gamma}^\alpha_{\beta\gamma}$ are all zeros. Additionally, in the specific case where the manifold $\mathcal{Z}$ is one dimensional, its Christoffel symbols are 0. In other words, for a ring immersed in $\mathbb{R}^N_+$, the Hessian with respect to the pullback metric is the traditional Hessian: $\nabla^2_{ij} f(z) = \frac{\partial^2 f}{\partial x_i \partial x_j}(z)$.

We now give the general definition of the mean curvature vector, for any submanifold of $\mathcal{X}$.

Definition 11 (Mean curvature vector [3])**.** *The mean curvature vector $H(z)$ of $\mathcal{M} = f(\mathcal{Z}) \subset \mathcal{X}$ is defined as:*

$$H^\alpha(z) = \frac{1}{N} Tr\, \mathrm{II}(z)^\alpha = \frac{1}{N} g^{ij} \mathrm{II}(z)^\alpha_{ij},$$

where N is the dimension of $\mathcal{X}$, and the trace Tr is computed with respect to g^{ij}, the inverse of the Riemannian metric matrix of $\mathcal{Z}$.

This leads us to the definition of mean curvature vector of an immersed manifold.

Definition 12 (Mean curvature vector (immersed manifold) [3])**.** *The mean curvature vector $H(z)$ of $\mathcal{M}$ is defined as:*

$$H^\alpha(z) = \frac{1}{N} g^{ij} \Big(\partial^2_{ij} f^\alpha(z) - \Gamma^k_{ij}(z) \partial_k f^\alpha(z) + \bar{\Gamma}^\alpha_{\beta\gamma}(f(z)) \partial_i f^\beta(z) \partial_j f^\gamma(z) \Big),$$

where N is the dimension of $\mathcal{X}$, and g^{ij} is the inverse of the Riemannian metric matrix of $\mathcal{Z}$.

C.2. Mean Curvatures of the Circle

Example 1 (Mean curvatures of the circle immersed in $\mathbb{R}^N$)**.** *We consider a circle C of radius R immersed in $\mathbb{R}^N$. The norms of its mean curvature vector is:*

$$\|H_C(\theta)\| = \frac{1}{R}, \quad \forall \theta \in \mathcal{S}^1. \tag{4}$$

Proof. We compute the mean curvature of a circle immersed in $\mathbb{R}^N$ as:

$$f : \mathcal{S}^1 \mapsto \mathbb{R}^N$$
$$\theta \mapsto f(\theta) = P. \begin{bmatrix} R\cos\theta \\ R\sin\theta \\ 0 \\ \vdots \\ 0 \end{bmatrix} + t,$$

where $P \in SO(N)$ represents a rotation in $\mathbb{R}^N$, and $t \in \mathbb{R}^N$ a translation in $\mathbb{R}^N$ illustrating that the circle can be placed and oriented in any direction in $\mathbb{R}^N$.

We compute the second fundamental form, for $k = 1, 2, 3$:

$$\mathrm{II}_{11}(\theta) = \begin{bmatrix} \frac{d^2 f^1}{d\theta^2}(\theta) \\ \frac{d^2 f^2}{d\theta^2}(\theta) \\ \vdots \\ \frac{d^2 f^N}{d\theta^2}(\theta) \end{bmatrix} = \begin{bmatrix} -R\cos\theta \\ -R\sin\theta \\ 0 \\ \vdots \\ 0 \end{bmatrix}. \tag{5}$$

The mean curvature vector is then:

$$\begin{aligned} H_C(\theta) &= \frac{1}{1} \mathrm{Tr}\mathrm{II}_\theta \\ &= g^{11} \mathrm{II}_{11}(\theta) \\ &= \frac{1}{R^2} \begin{bmatrix} -R\cos\theta \\ -R\sin\theta \\ 0 \\ \vdots \\ 0 \end{bmatrix} \\ &= \frac{1}{R} \begin{bmatrix} -\cos\theta \\ -\sin\theta \\ 0 \\ \vdots \\ 0 \end{bmatrix} \end{aligned}$$

Its norm is: $\|H_C(\theta)\| = \frac{1}{R}$ for all $\theta \in \mathcal{S}^1$. □

C.3. Mean Curvatures of the Sphere

Example 2 (Mean curvatures of the sphere immersed in $\mathbb{R}^N$)**.** *We consider a sphere S of radius R immersed in $\mathbb{R}^N$. The norm of its mean curvature vector is:*

$$\|H_C(\theta, \phi)\| = \frac{1}{R}, \qquad \forall \theta, \phi \in \mathcal{S}^2. \tag{6}$$

Proof. We compute the mean curvature of a sphere of radius R immersed in $\mathbb{R}^N$ as:

$$f : \mathcal{S}^2 \mapsto \mathbb{R}^N$$
$$\theta, \phi \mapsto f(\theta, \phi) = P. \begin{bmatrix} R\sin\theta\cos\phi \\ R\sin\theta\sin\phi \\ R\cos\theta \\ 0 \\ \vdots \\ 0 \end{bmatrix} + t,$$

where $P \in SO(N)$ represents a rotation in $\mathbb{R}^N$, and $t \in \mathbb{R}^N$ a translation in $\mathbb{R}^N$ illustrating that the sphere can be placed and oriented in any direction in $\mathbb{R}^N$.

We compute the Hessian:

$$\frac{\partial^2 f}{\partial x_i \partial x_j}(\theta,\phi) = \begin{bmatrix} \frac{\partial^2 f^1}{\partial x_i \partial x_j}(\theta,\phi) \\ \frac{\partial^2 f^2}{\partial x_i \partial x_j}(\theta,\phi) \\ \frac{\partial^2 f^3}{\partial x_i \partial x_j}(\theta,\phi) \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \tag{7}$$

where we use the conventions $x_1 = \theta$ and $x_2 = \phi$. In what follows, for conciseness of the derivations, we do not write the components α of f^α for $\alpha = 4, ..., N$, as they only contribute terms equal to 0.

We get:

$$\frac{\partial^2 f}{\partial \theta^2}(\theta,\phi) = \begin{bmatrix} -R\sin\theta\cos\phi \\ -R\sin\theta\sin\phi \\ -R\cos\theta \end{bmatrix},$$

$$\frac{\partial^2 f}{\partial \theta \partial \phi}(\theta,\phi) = \begin{bmatrix} -R\cos\theta\sin\phi \\ R\cos\theta\cos\phi \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix},$$

$$\frac{\partial^2 f}{\partial \phi^2}(\theta,\phi) = \begin{bmatrix} -R\sin\theta\cos\phi \\ -R\sin\theta\sin\phi \\ 0 \end{bmatrix}.$$

We only compute the diagonal terms, avoiding the computation of $\frac{\partial^2 f}{\partial\theta\partial\phi}(\theta,\phi)$ because we only need the diagonal terms in the definition of the trace, given that the inverse of the pullback metric is diagonal.

We compute the Hessian with respect to the pullback metric, again omitting its components for $\alpha > 3$.

$$\mathrm{II}_{ij}(\theta,\phi) = \begin{bmatrix} \frac{\partial^2 f^1}{\partial x_i \partial x_j}(\theta,\phi) - \sum_{k=1}^2 \Gamma_{ij}^k \frac{\partial f^1}{\partial x_k} \\ \frac{\partial^2 f^2}{\partial x_i \partial x_j}(\theta,\phi) - \sum_{k=1}^2 \Gamma_{ij}^k \frac{\partial f^2}{\partial x_k} \\ \frac{\partial^2 f^3}{\partial x_i \partial x_j}(\theta,\phi) - \sum_{k=1}^2 \Gamma_{ij}^k \frac{\partial f^3}{\partial x_k} \end{bmatrix}. \tag{8}$$

For the 2-sphere, the Christoffel symbols are:

$$\begin{aligned} &\Gamma_{11}^1 = \Gamma_{11}^2 = \Gamma_{22}^2 = \Gamma_{12}^1 = \Gamma_{21}^1 = 0, \\ &\Gamma_{22}^1 = -\sin\theta\cos\theta, \\ &\Gamma_{12}^2 = \Gamma_{21}^2 = \frac{\cos\theta}{\sin\theta}, \end{aligned}$$

so that we get $\mathrm{II}_{11}(\theta,\phi)$:

$$\begin{aligned} \mathrm{II}_{11}(\theta,\phi) &= \begin{bmatrix} -R\sin\theta\cos\phi \\ -R\sin\theta\sin\phi \\ -R\cos\theta \end{bmatrix} - \Gamma_{11}^1 \frac{\partial f}{\partial x_1} - \Gamma_{11}^2 \frac{\partial f}{\partial x_2} \\ &= \begin{bmatrix} -R\sin\theta\cos\phi \\ -R\sin\theta\sin\phi \\ -R\cos\theta \end{bmatrix} - 0.\frac{\partial f}{\partial x_1} - 0.\frac{\partial f}{\partial x_2} \\ &= \begin{bmatrix} -R\sin\theta\cos\phi \\ -R\sin\theta\sin\phi \\ -R\cos\theta \end{bmatrix}, \end{aligned}$$

as well as $\mathrm{II}_{22}(\theta,\phi)$:

$$\begin{aligned} \mathrm{II}_{22}(\theta,\phi) &= \begin{bmatrix} -R\sin\theta\cos\phi \\ -R\sin\theta\sin\phi \\ 0 \end{bmatrix} - \Gamma_{22}^1 \frac{\partial f}{\partial x_1} - \Gamma_{22}^2 \frac{\partial f}{\partial x_2} \\ &= \begin{bmatrix} -R\sin\theta\cos\phi \\ -R\sin\theta\sin\phi \\ 0 \end{bmatrix} - (-\sin\theta\cos\theta)\frac{\partial f}{\partial x_1} - 0.\frac{\partial f}{\partial x_2} \\ &= \begin{bmatrix} -R\sin\theta\cos\phi \\ -R\sin\theta\sin\phi \\ 0 \end{bmatrix} + \sin\theta\cos\theta \begin{bmatrix} R\cos\theta\cos\phi \\ R\cos\theta\sin\phi \\ -R\sin\theta \end{bmatrix} \\ &= R\sin\theta \begin{bmatrix} -\cos\phi + \cos^2\theta\cos\phi \\ -\sin\phi + \cos^2\theta\sin\phi \\ -\sin\theta\cos\theta \end{bmatrix} \\ &= R\sin\theta \begin{bmatrix} -\sin^2\theta\cos\phi \\ -\sin^2\theta\sin\phi \\ -\sin\theta\cos\theta \end{bmatrix} \\ &= R\sin^2\theta \begin{bmatrix} -\sin\theta\cos\phi \\ -\sin\theta\sin\phi \\ -\cos\theta \end{bmatrix}. \end{aligned}$$

The inverse of the Riemannian metric matrix is:

$$g_S(\theta,\phi)^{-1} = \begin{bmatrix} \frac{1}{R^2} & 0 \\ 0 & \frac{1}{R^2\sin^2\theta} \end{bmatrix}. \tag{9}$$

The mean curvature vector is then (omitting its zero com-

ponents):

$$
\begin{aligned}
H_S(\theta,\phi) &= \frac{1}{2}\mathrm{Tr}\mathrm{II}_p \\
&= \frac{1}{2}g^{11}\mathrm{II}_{11}(\theta,\phi) + \frac{1}{2}g^{22}\mathrm{II}_{22}(\theta,\phi) \\
&= \frac{1}{2R^2}\begin{bmatrix} -R\sin\theta\cos\phi \\ -R\sin\theta\sin\phi \\ -R\cos\theta \end{bmatrix} \\
&\qquad + \frac{1}{2R^2\sin^2(\theta)}R\sin^2\theta\begin{bmatrix} -\sin\theta\cos\phi \\ -\sin\theta\sin\phi \\ -\cos\theta \end{bmatrix} \\
&= \frac{1}{2R}\begin{bmatrix} -\sin\theta\cos\phi \\ -\sin\theta\sin\phi \\ -\cos\theta \end{bmatrix} + \frac{1}{2R}\begin{bmatrix} -\sin\theta\cos\phi \\ -\sin\theta\sin\phi \\ -\cos\theta \end{bmatrix} \\
&= \frac{1}{2R}\begin{bmatrix} -2\sin\theta\cos\phi \\ -2\sin\theta\sin\phi \\ -2\cos\theta \end{bmatrix} \\
&= -\frac{1}{R}\begin{bmatrix} \sin\theta\cos\phi \\ \sin\theta\sin\phi \\ \cos\theta \end{bmatrix}.
\end{aligned}
$$

Its norm is: $\|H(\theta,\phi)\| = \frac{1}{R}$, which is the expected formula. □

C.4. Mean Curvatures of the Torus

Example 3 (Mean curvatures of the torus immersed in $\mathbb{R}^N$)**.** *We consider the torus T obtained by rotating a circle of radius b and center $(a,0)$ around the axis z, and immersed in $\mathbb{R}^N$. The norms of its mean curvature vector is:*

$$
\|H_T(\theta,\phi)\| = \frac{R + 2r\cos\phi}{r(R + r\cos(\phi))}, \qquad \forall\theta,\phi \in \mathcal{S}^1 \times \mathcal{S}^1. \tag{10}
$$

Proof. We compute the mean curvature of a torus immersed in $\mathbb{R}^N$:

$$
\begin{aligned}
f : \mathcal{S}^1 \times \mathcal{S}^1 &\mapsto \mathbb{R}^N \\
\theta,\phi &\mapsto f(\theta,\phi) = P.\begin{bmatrix} c(\phi)\cos\theta \\ c(\phi)\sin\theta \\ r\sin\phi \\ 0 \\ \vdots \\ 0 \end{bmatrix} + t,
\end{aligned}
$$

where $c(\phi) = R + r\cos\phi$.

We compute the Hessian:

$$
\frac{\partial^2 f}{\partial x_i \partial x_j}(\theta,\phi) = \begin{bmatrix} \frac{\partial^2 f^1}{\partial x_i \partial x_j}(\theta,\phi) \\ \frac{\partial^2 f^2}{\partial x_i \partial x_j}(\theta,\phi) \\ \frac{\partial^2 f^3}{\partial x_i \partial x_j}(\theta,\phi) \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \tag{11}
$$

where we use the conventions $x_1 = \theta$ and $x_2 = \phi$. In what follows, for conciseness of the derivations, we do not write the components α of f^α for $\alpha = 4, ..., N$, as they only contribute terms equal to 0.

We get:

$$
\begin{aligned}
\frac{\partial^2 f}{\partial\theta^2}(\theta,\phi) &= \begin{bmatrix} -c(\phi)\cos\theta \\ -c(\phi)\sin\theta \\ 0 \end{bmatrix}, \\
\frac{\partial^2 f}{\partial\phi^2}(\theta,\phi) &= \begin{bmatrix} -r\cos\phi\cos\theta \\ -r\cos\phi\sin\theta \\ -r\sin\phi \end{bmatrix}.
\end{aligned}
$$

We only compute the diagonal terms, avoiding the computation of $\frac{\partial^2 f}{\partial\theta\partial\phi}(\theta,\phi)$ because we only need the diagonal terms in the definition of the trace, given that the inverse of the pullback metric is diagonal.

We compute the Hessian with respect to the pullback metric, again omitting its components for $\alpha > 3$.

$$
\mathrm{II}_{ij}(\theta,\phi) = \begin{bmatrix} \frac{\partial^2 f^1}{\partial x_i \partial x_j}(\theta,\phi) - \sum_{k=1}^2 \Gamma_{ij}^k \frac{\partial f^1}{\partial x_k} \\ \frac{\partial^2 f^2}{\partial x_i \partial x_j}(\theta,\phi) - \sum_{k=1}^2 \Gamma_{ij}^k \frac{\partial f^2}{\partial x_k} \\ \frac{\partial^2 f^3}{\partial x_i \partial x_j}(\theta,\phi) - \sum_{k=1}^2 \Gamma_{ij}^k \frac{\partial f^3}{\partial x_k} \end{bmatrix}. \tag{12}
$$

For the torus, the Christoffel symbols are:

$$
\begin{aligned}
\Gamma_{11}^1 &= \Gamma_{22}^1 = \Gamma_{22}^2 = 0, \\
\Gamma_{12}^1 &= \Gamma_{21}^1 = -\frac{r\sin\phi}{c(\phi)}, \\
\Gamma_{11}^2 &= \frac{1}{r}\sin\phi c(\phi),
\end{aligned}
$$

so that we get:

$$\begin{aligned}
\mathrm{II}_{11}(\theta,\phi) &= \begin{bmatrix} -c(\phi)\cos\theta \\ -c(\phi)\sin\theta \\ 0 \end{bmatrix} - \Gamma^1_{11}\frac{\partial f}{\partial x_1} - \Gamma^2_{11}\frac{\partial f}{\partial x_2} \\
&= \begin{bmatrix} -c(\phi)\cos\theta \\ -c(\phi)\sin\theta \\ 0 \end{bmatrix} - 0\frac{\partial f}{\partial x_1} \\
&\qquad - \frac{1}{r}\sin\phi c(\phi)\begin{bmatrix} -r\sin\phi\cos\theta \\ -r\sin\phi\sin\theta \\ r\cos\phi \end{bmatrix} \\
&= c(\phi)\begin{bmatrix} -\cos\theta + \sin^2\phi\cos\theta \\ -\sin\theta + \sin^2\phi\sin\theta \\ -\sin\phi\cos\phi \end{bmatrix} \\
&= c(\phi)\begin{bmatrix} -\cos^2\phi\cos\theta \\ -\cos^2\phi\sin\theta \\ -\sin\phi\cos\phi \end{bmatrix} \\
&= c(\phi)\cos\phi\begin{bmatrix} -\cos\phi\cos\theta \\ -\cos\phi\sin\theta \\ -\sin\phi \end{bmatrix},
\end{aligned}$$

and

$$\begin{aligned}
\mathrm{II}_{22}(\theta,\phi) &= \begin{bmatrix} -r\cos\phi\cos\theta \\ -r\cos\phi\sin\theta \\ -r\sin\phi \end{bmatrix} - \Gamma^1_{22}\frac{\partial f}{\partial x_1} - \Gamma^2_{22}\frac{\partial f}{\partial x_2} \\
&= \begin{bmatrix} -r\cos\phi\cos\theta \\ -r\cos\phi\sin\theta \\ -r\sin\phi \end{bmatrix} - 0\frac{\partial f}{\partial x_1} - 0\frac{\partial f}{\partial x_2} \\
&= -r\begin{bmatrix} \cos\phi\cos\theta \\ \cos\phi\sin\theta \\ \sin\phi \end{bmatrix}.
\end{aligned}$$

The inverse of the Riemannian metric matrix is:

$$g_S(\theta,\phi)^{-1} = \begin{bmatrix} \frac{1}{(R+r\cos\phi)^2} & 0 \\ 0 & \frac{1}{r^2} \end{bmatrix}. \tag{13}$$

The mean curvature vector is then:

$$\begin{aligned}
H_S(\theta,\phi) &= \frac{1}{2}\mathrm{Tr}\mathrm{II}_p \\
&= \frac{1}{2}g^{11}\mathrm{II}_{11}(\theta,\phi) + \frac{1}{2}g^{22}\mathrm{II}_{22}(\theta,\phi) \\
&= \frac{1}{2c^2(\phi)}c(\phi)\cos\phi\begin{bmatrix} -\cos\phi\cos\theta \\ -\cos\phi\sin\theta \\ -\sin\phi \end{bmatrix} \\
&\qquad + \frac{1}{2r^2}(-r)\begin{bmatrix} \cos\phi\cos\theta \\ \cos\phi\sin\theta \\ \sin\phi \end{bmatrix} \\
&= \frac{\cos\phi}{2c(\phi)}\begin{bmatrix} -\cos\phi\cos\theta \\ -\cos\phi\sin\theta \\ -\sin\phi \end{bmatrix} - \frac{1}{2r}\begin{bmatrix} \cos\phi\cos\theta \\ \cos\phi\sin\theta \\ \sin\phi \end{bmatrix} \\
&= -\left(\frac{\cos\phi}{2c(\phi)} + \frac{1}{2r}\right)\begin{bmatrix} 2\cos\phi\cos\theta \\ 2\cos\phi\sin\theta \\ 2\sin\phi \end{bmatrix} \\
&= -\left(\frac{\cos\phi}{c(\phi)} + \frac{1}{r}\right)\begin{bmatrix} \cos\phi\cos\theta \\ \cos\phi\sin\theta \\ \sin\phi \end{bmatrix} \\
&= -\frac{r\cos\phi + R + r\cos\phi}{rc(\phi)}\begin{bmatrix} \cos\phi\cos\theta \\ \cos\phi\sin\theta \\ \sin\phi \end{bmatrix} \\
&= -\frac{R + 2r\cos\phi}{r(R + r\cos(\phi))}\begin{bmatrix} \cos\phi\cos\theta \\ \cos\phi\sin\theta \\ \sin\phi \end{bmatrix}.
\end{aligned}$$

Its norm is: $\|H(\theta,\phi)\| = \frac{R+2r\cos\phi}{r(R+r\cos(\phi))}$, which is the expected formula. □

D. Invariance under Reparameterizations

We give the proof for Lemma 1.

Lemma 3 (Invariance with respect to reparameterizations)**.** *The curvature profile is invariant under reparameterizations $f \to f \circ \varphi^{-1}$ of the neural manifold $\mathcal{M}$.*

Proof. The distance between two points on the latent manifold is given by:

$$\mathrm{dist}(z_0, z) = \int_0^1 d\tau \sqrt{\frac{d\gamma^c}{d\tau}\frac{d\gamma^d}{d\tau}g_{cd}(\gamma(\tau))}$$

Where γ is a geodesic and g is the pullback metric induced by f. Consider a reparameterization of the latent space $\tilde{z} = \varphi(z)$. The distance between $\tilde{z_0}$ and $\tilde{z}$ is then given by

$$\mathrm{dist}(\tilde{z_0}, \tilde{z}) = \int_0^1 d\tau \sqrt{\frac{d\tilde{\gamma}^a}{d\tau}\frac{d\tilde{\gamma}^b}{d\tau}\tilde{g}_{ab}(\tilde{\gamma}(\tau))}$$

Which we can write as

$$\mathrm{dist}(\tilde{z_0}, \tilde{z}) = \int_0^1 d\tau \sqrt{\frac{d\tilde{\gamma}^a}{d\gamma^c}\frac{d\gamma^c}{d\tau}\frac{d\tilde{\gamma}^b}{d\gamma^d}\frac{d\gamma^d}{d\tau}\tilde{g}_{ab}(\tilde{\gamma}(\tau))}$$

Via the metric tensor transformation law,

$$g_{cd}(\gamma) = \frac{d\tilde{\gamma}^a}{d\gamma^c}\frac{d\tilde{\gamma}^b}{d\gamma^d}\tilde{g}_{ab}(\tilde{\gamma}(\tau))$$

we conclude that $\text{dist}(\tilde{z_0}, \tilde{z}) = \text{dist}(z_0, z)$ when the latent manifold is endowed with the pullback metric. Thus, if we consider the mean curvature vector $H(z) = H(\text{dist}(z_0, z))$ using a reference point $z_0 = \tilde{z}_0$, we obtain a reparameterization-invariant curvature profile on the latent manifold. □

E. Examples of Latent Losses

We illustrate the latent loss term presented in the main text with its explicit formulae for manifolds with topology S^1 or S^2.

Example 4 (Latent loss terms for $\mathcal{Z} = \mathcal{S}^1, \mathcal{S}^2$)**.** *The latent loss terms for neural manifolds parameterized by template manifolds $\mathcal{S}^1$, $\mathcal{S}^2$ are given by:*

$$\begin{aligned}\mathcal{L}^{\mathcal{S}^1}_{latent} &= (1 - \cos(\theta_{gt} - \hat{\theta}))^2,\\ \mathcal{L}^{\mathcal{S}^2}_{latent} &= (1 - \cos(\theta_{gt} - \hat{\theta})\\ &\quad + \sin(\theta_{gt})\sin(\hat{\theta})(1 - \cos(\phi_{gt} - \hat{\phi})))^2.\end{aligned}$$

We implement these loss terms when we seek to enforce a canonical parameterization of the latent space, informed by outside world's task variables.

F. Invariance under Permutations

We give the proof for Lemma 2.

Lemma 4 (Invariance of topology and geometry)**.** *Consider a neural manifold $\mathcal{M}$ embedded in neural state space $\mathbb{R}^N_+$ corresponding to the recording of N neurons. Permuting the order of the N neurons: (i) leaves the topology of $\mathcal{M}$ invariant, (ii) leaves the geometry of $\mathcal{M}$ invariant.*

Proof. Consider a neural manifold $\mathcal{M}$ embedded in neural state space $\mathbb{R}^N_+$ corresponding to the recording of N neurons. Consider S_N the group of permutations of the set $\{1, \ldots, N\}$ labelling the N neurons. The group S_N acts on $\mathbb{R}^N_+$ by permutating of the order in which neurons are recorded, *i.e.*, by permuting the axes of the space $\mathbb{R}^N_+$ as:

$$\begin{aligned}S_N \times \mathbb{R}^N_+ &\mapsto \mathbb{R}^N_+\\ \sigma\ ,\quad (x_1, \ldots, x_N) &\mapsto (x_{\sigma(1)}, \ldots, x_{\sigma(N)})\end{aligned}$$

Consider one permutation $\sigma \in S_N$. We show that the topology and geometry of $\mathcal{M}$ is invariant with respect to σ, by showing that σ is a linear, thus continuous, isometry of $\mathbb{R}^N_+$.

By properties of permutation, σ can be written as a product of transpositions τ_t's:

$$\sigma = \Pi_{t=1}^{T}\tau_t \tag{14}$$

where the product is taken in the sense of the composition. A given transposition τ_t exchanges only two neurons. If each transposition leaves the topology of $\mathcal{M}$ invariant, then so does σ. Thus, we show that any transposition τ_t leaves the topology of $\mathcal{M}$ invariant.

Without loss of generality, we can prove it for a transposition τ_t exchanging neuron 1 and neuron 2, which will simplify the notations. The transposition $\tau_t = \tau_{12}$ exchanges neuron 1 and 2, which corresponds to exchanging the first two axes, axis x_1 and axis x_2 in $\mathbb{R}^N_+$, while keeping all other axes invariant. The action of this transposition corresponds to the symmetry of hyperplane $x_1 = x_2$ within $\mathbb{R}^N_+$, as:

$$\begin{aligned}\mathbb{R}^N_+ &\mapsto \mathbb{R}^N_+\\ x = (x_1, x_2, \ldots, x_N) &\mapsto (x_2, x_1, \ldots, x_N) = T_{12}.x\end{aligned}$$

where T_{12} is the permutation matrix:

$$T_{12} = \begin{pmatrix} 0 & 1 & 0 & \cdots & 0\\ 1 & 0 & 0 & \cdots & 0\\ 0 & 0 & 1 & \cdots & 0\\ \vdots & \vdots & \vdots & \ddots & \vdots\\ 0 & 0 & 0 & \cdots & 1 \end{pmatrix}.$$

We see that the transposition τ_{12} of two axes is a linear map expressed by matrix T_{12}. As a linear map, τ_{12} is continuous. Any continuous map acting on a manifold preserves the topology of this manifold. Consequently, any transposition, and thus any permutation, preserves the topology of the neural manifold $\mathcal{M}$.

Additionally, the matrix T_{12} is orthogonal as we can show that $T_{12}^T.T_{12} = I_N$ where I_N is the identity matrix of shape $N \times N$. Consequently, T_{12} is an isometry of $\mathbb{R}^N_+$ that preserves the geometry of $\mathcal{M}$ embedded in $\mathbb{R}^N$. Consequently, any transposition, and thus any permutation, preserves the geometry of the neural manifold $\mathcal{M}$ in the sense that $d(p, q) = d(\sigma*, \sigma q)$: the distances along the manifold $\mathcal{M}$ are invariant. □

G. Synthetic Datasets

We detail generation of the synthetic datasets in our experiments of distorted circles, spheres and tori.

Distorted Circle Datasets The distorted circle datasets are created from the immersion $f^{\mathcal{S}^1}_{\text{synth}}$:

$$\begin{aligned}&\mathcal{S}^1 \to \mathbb{R}^N\\ &\theta_i \mapsto \mathcal{R} * [A(\theta_i)(\cos\theta_i, \sin\theta, 0, ..., 0)] + \eta_i,\end{aligned} \tag{15}$$

with θ_i uniformly distributed on $\mathcal{S}^1$ for $i = 1, \ldots, n$. In Eq. 15, $A(\theta) = 1 + \alpha[\exp(-5(\theta - \pi/2)^2) + \exp(-5(\theta - 3\pi/2)^2)]$ where the parameter α modulates the amplitude around the ring, creating extrinsic curvature in the vicinity of $\pi/2$ and $3\pi/2$.

Distorted Sphere Datasets The distorted 2-sphere datasets are created via the immersion $f^{\mathcal{S}^2}_{\text{synth}}$:

$$\begin{aligned}&\mathcal{S}^2 \to \mathbb{R}^N\\&(\theta_i, \phi_i) \mapsto \mathcal{R} * \big[A(\theta_i, \phi_i)\\&\qquad (\sin\theta_i \cos\phi_i, \sin\theta_i \sin\phi_i, \cos\theta_i, ..., 0)\big] + \eta_i,\end{aligned} \tag{16}$$

with θ_i, ϕ_i uniformly distributed on $\mathcal{S}^2$ for $i = 1, \ldots, n$. In Eq. 16, $A(\theta, \phi) = 1 + \alpha \exp(-5(\theta)^2) + \alpha \exp(-5(\theta - \pi/2)^2)$ with the parameter α introducing curvature in the vicinity of the north and south poles of $\mathcal{S}^2$.

Distorted Torus Datasets The distorted 2-torus datasets are created via the immersion $f^{\mathcal{T}^2}_{\text{synth}}$:

$$\begin{aligned}&\mathcal{T}^2 \to \mathbb{R}^N\\&(\theta_i, \phi_i) \mapsto \mathcal{R} * \big[A(\theta_i, \phi_i)\\&\quad ((R - r\cos\theta_i)\cos\phi_i, (R - r\cos\theta_i)\sin\phi_i, r\sin\theta_i, ..., 0)\big]\\&\quad + \eta_i,\end{aligned} \tag{17}$$

with θ_i, ϕ_i uniformly distributed on $\mathcal{T}^2$ for $i = 1, \ldots, n$. Here, R and r are the major and minor radii of the torus; these are assumed to carry no relevant information, and are both set to unity. The amplitude function $A(\theta, \phi)$ in Eq. 17 is given by

$$\begin{aligned}A(\theta_i, \phi_i) = 1 + \alpha \exp(-2(\theta - \pi)^2)[\exp(-2(\phi - \pi/2)^2)\\ + \exp(-2(\phi - 3\pi/2)^2)]\end{aligned}$$

with the parameter α introducing extrinsic curvature by stretching the torus on opposite sides at $(\theta, \phi) = (\pi, \pi/2)$ and $(\theta, \phi) = (\pi, 3\pi/2)$

Validation of Learned Topology We show here the results of TDA applied to the synthetic dataset of the distorted T^2. This validates that the first step of our pipeline can effectively capture the topology of the manifold, as we observe the two holes know to characterize the torus topology in Fig. 8.

Effect of Noise on Curvature Estimation Error for Distorted Spheres We quantify the curvature estimation error as we vary the noise levels for distorted spheres, to complement the similar experiments presented in the main text

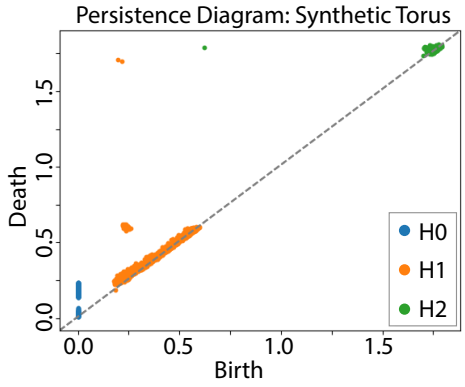

Figure 8. Persistence diagram for synthetic dataset on the torus, illustrating that TDA is an appropriate tool to compute the topology of a neural manifold and to constrain the latent space to be a given template $\mathcal{M}^*$.

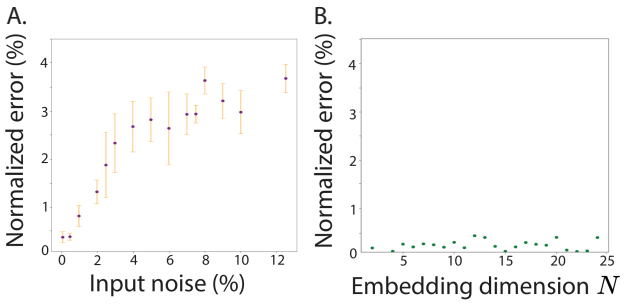

Figure 9. Curvature estimation error on distorted circles. A. While the error increases with the noise level, it does not go over 4% for a range of noise levels corresponding to realistic values observed in neuroscience. Each experiment is repeated 5 times. The vertical orange bars show the -/+ 1 standard deviations of the errors. B. The error shows minimal variations with respect to the number of recorded neurons N. The vertical axis is shared across both plots for ease of comparison.

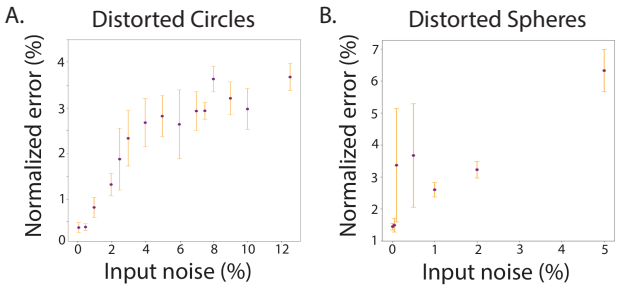

Figure 10. Curvature estimation error on distorted circles (A) and spheres (B). The number of neurons is fixed at $N = 2$ for the distorted circles, and $N = 3$ for the distorted spheres. Each experiment is repeated 5 times. The vertical orange bars show the -/+ 1 standard deviations of the curvature estimation errors.

for distorted circles. Fig. 10 compares the curvature error for the circles (A) and the spheres (B).

In these experiments, the number of neurons N is fixed at $N = 2$ for the circles and $N = 3$ for the spheres. For each value of σ, 5 synthetic manifolds are generated and estimated. The vertical orange bars represent +/- standard deviation. We observe that the error is approximately twice as important in the case of the spheres than in the case of the circles. While the results in the main text seem to indicate that the estimation error does not depend on the number of neurons, we could conjecture that it depends linearly in the dimension of the manifold.

H. Experimental Place Cells (12 neurons)

We used data from 12 place cells within one session, whose neural spikes are binned with time-steps of 1 second, to yield 828 time points of neural activity in $\mathbb{R}_+^{12}$. Our results show that the reconstructed activations match the recorded (ground-truth), see Fig. 7 (A): even if we cannot observe the neural manifold in $\mathbb{R}_+^{12}$, we expect it to be correctly reconstructed. The canonical parameterization is correctly learned in the VAE latent space, as shown in Fig. 7 (B). The curvature profile is shown in Fig. 7 (C) where the angle is the physical lab angle. As for the simulated place cells, we observe several peaks which correspond to the place fields of the neurons: e.g. neuron 4 shown in Fig. 7 (C) which codes for one of the largest peaks, which is expected as it has the strongest activation from Fig. 7 (A). We reproduce this experiment on another dataset with 40 place cells recorded from another animal and find similar results in the supplementary materials. We emphasize that the goal of this experiment is not to reveal new neuroscience insights, but rather to show how the results provided by the curvature profiles are consistent with what one would expect and with what neuroscientists already know.

I. Experimental Place Cells (40 neurons)

We perform the same experiment on real place cell data as in subsection 5.2.2, this time recording from 40 neurons. In this experiment, after the temporal binning, we have 8327 time points. Similarly to the previous experiment, the reconstruction of the neural activity together with the canonical parameterization of the latent space are correctly learned by the model. As expected, we observe a neural manifold whose geometry shows more "petals" which intuitively correspond to the higher number of neurons recorded by this experiment. We locate a place cell whose place field provides one of the highest peaks, neuron 22, and color the curvature profile based on the activity of this neuron.

This demonstrates that our method can be applied to real neural datasets, providing geometric results that match the intuition. In the case of hippocampal place cells, the geometry of the neural manifold depends on the number of neurons, whether they tile the physical space where the animal is moving. It also depends on the spiking profile of their place fields, including its amplitude and width.

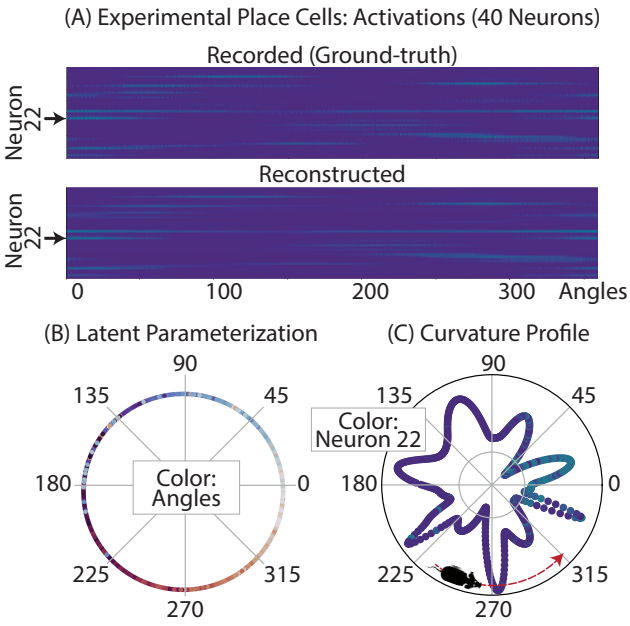

Figure 11. Neural geometry of 40 experimental place cells as an animal moves along a circle. (A) Recorded versus reconstructed neural activity of 12 place cells with respect to the positional angles of the animal in lab space. (B) Latent space's parameterization: the angular latent variables are colored by the corresponding positional angles of the animal in lab space. (C) Curvature profile of the neural manifold in log scale: the angles represent the physical lab space angles, colored by the reconstructed activation of neuron 22.