Coherence and Manifold Structures in Data:

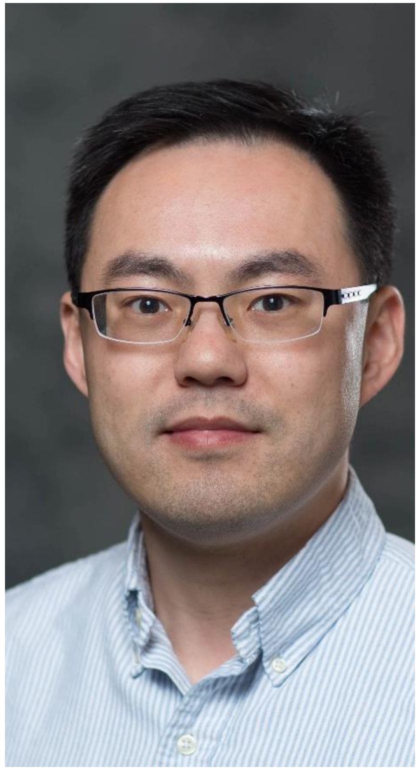
Exploring and Exploiting Geometry in 3D and Beyond

Stefan C. Schonsheck

UC Davis TETRAPODS Institute of Data Science

Krener Assistant Professor of Mathematics

Special Thanks to:



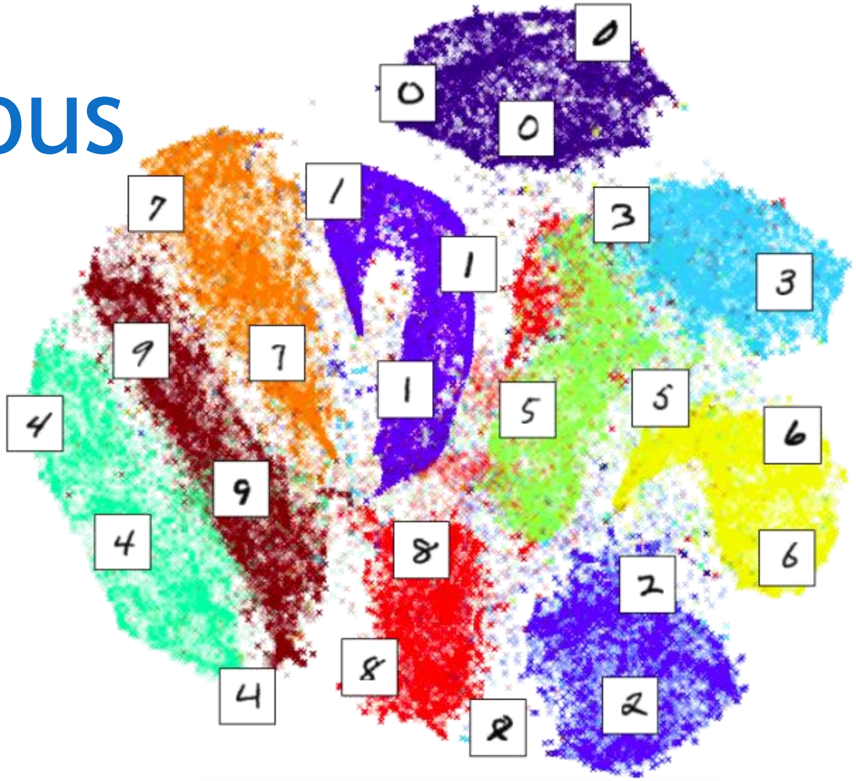
Rongjie Lai
Rensselaer Polytechnic Institute

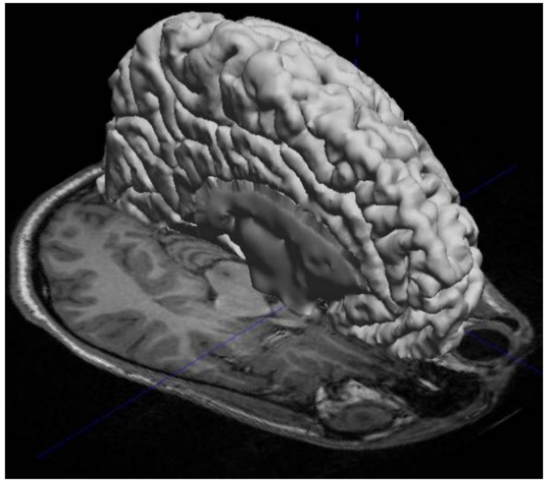
Bin Dong
Peking University

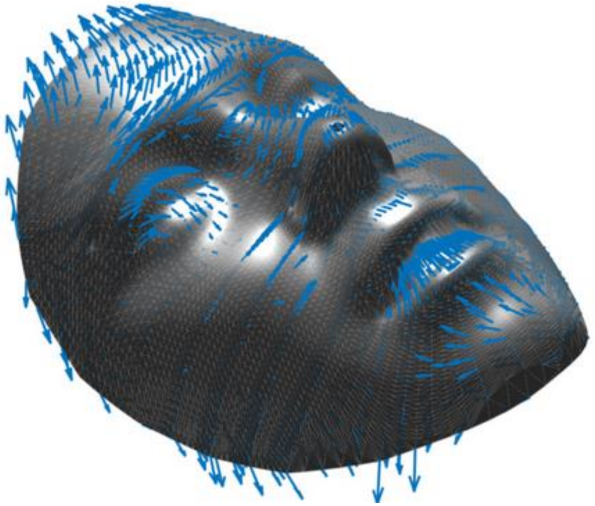
Jie Chen
IBM Watson

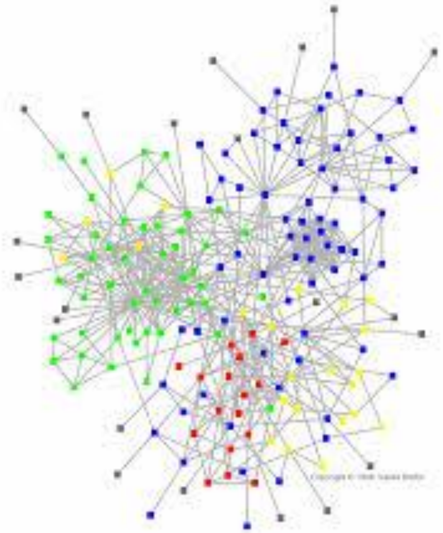
And N. Joey Tatro @ RPI

3+ Dimensional Data is Ubiquitous

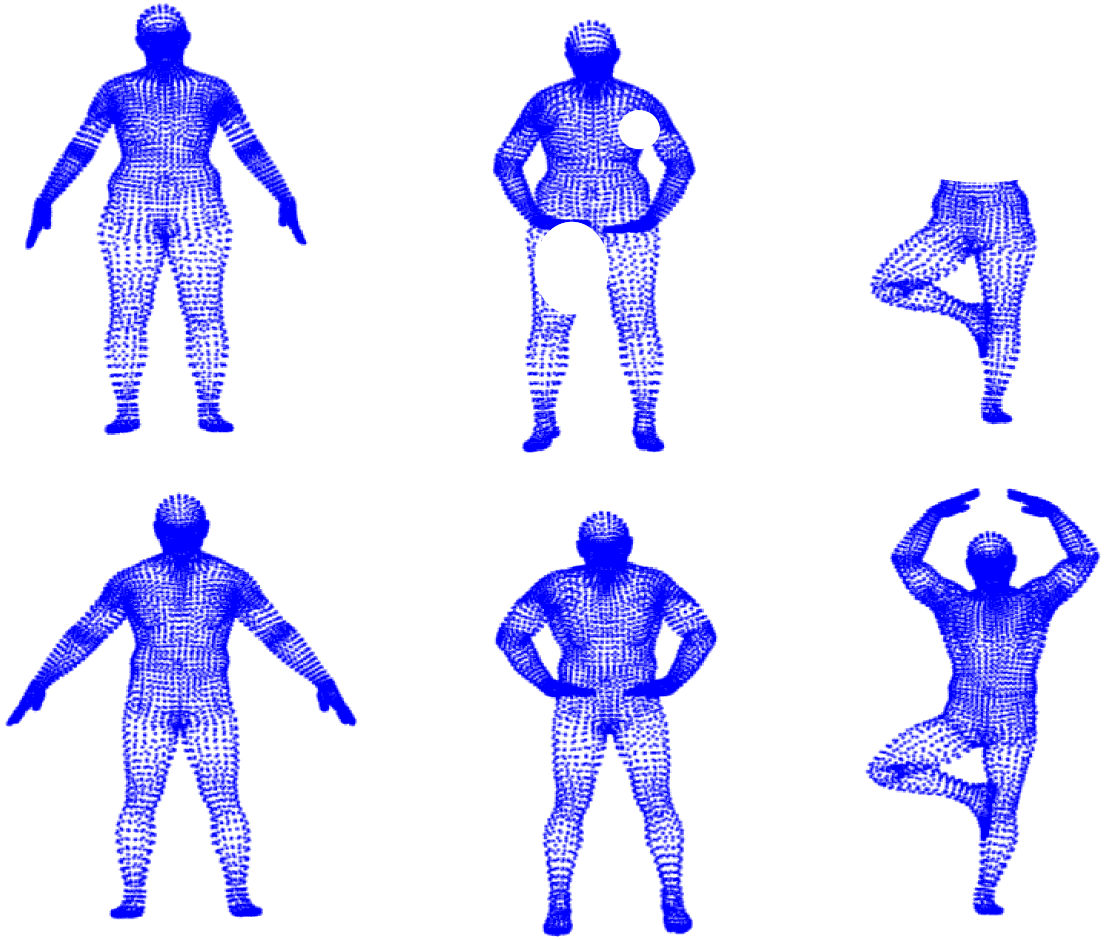


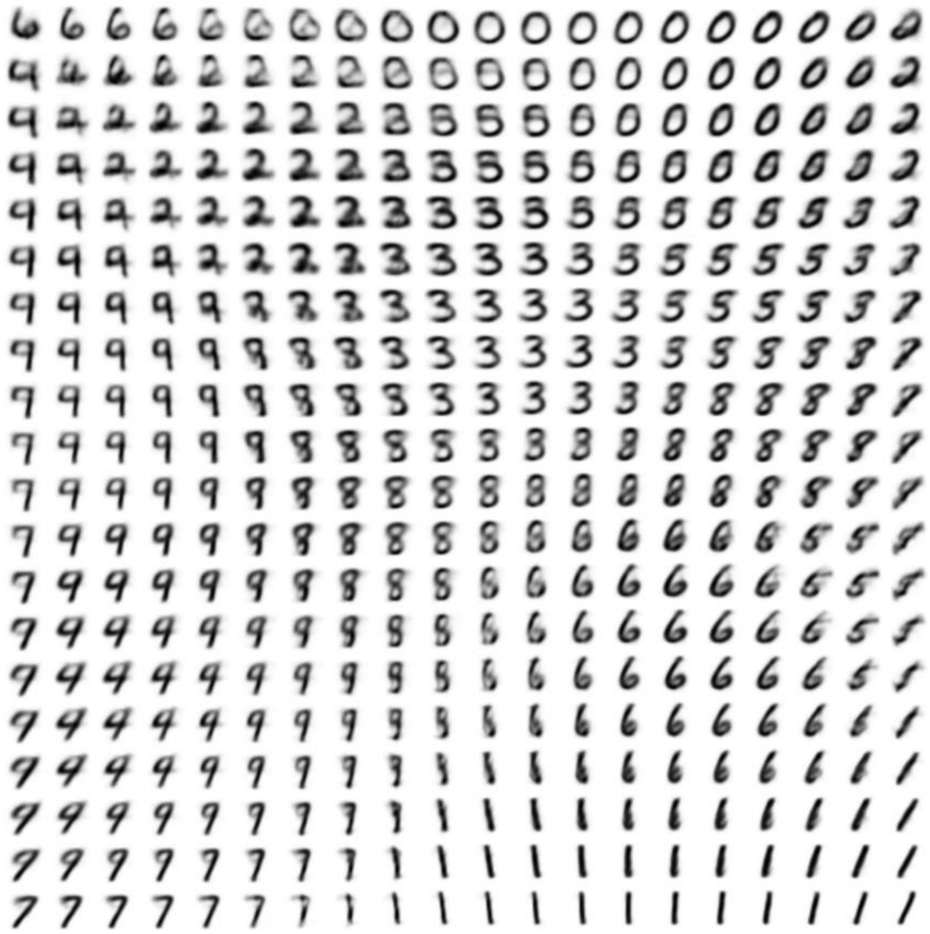






...but is often inherently coherent





This coherence motivates the **Manifold Hypothesis**

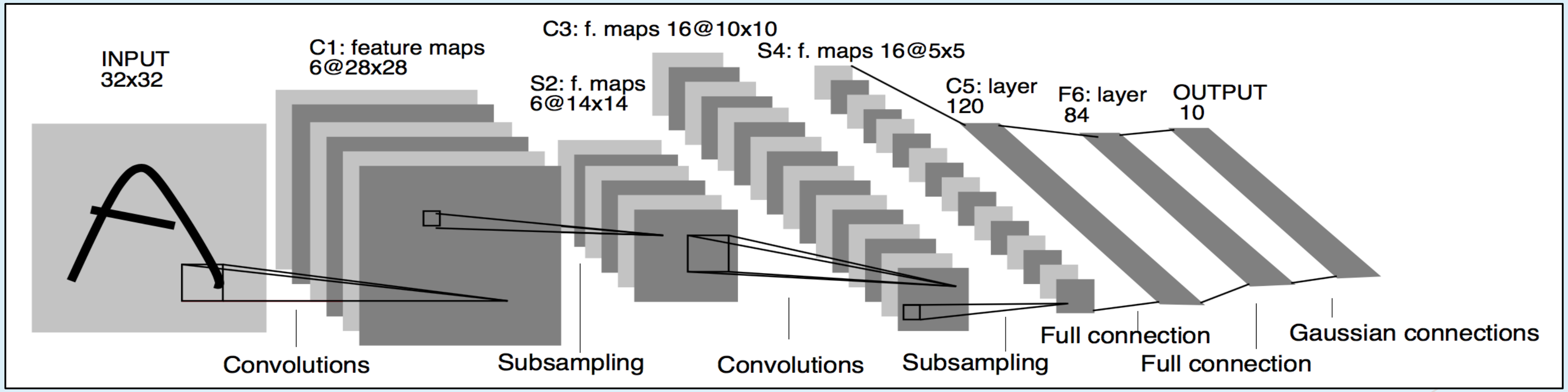
- Parallel Transport Convolution:

SCS., Dong and Lai

Convolution and CNNs

- The use of convolutional layers in neural networks has been one of the most important conceptual developments in deep learning (LeCun89. Why?
  - Shift-invariance
  - Equivariant representations
  - Sparsity of convolution operation

- So can we do convolution on a manifold?

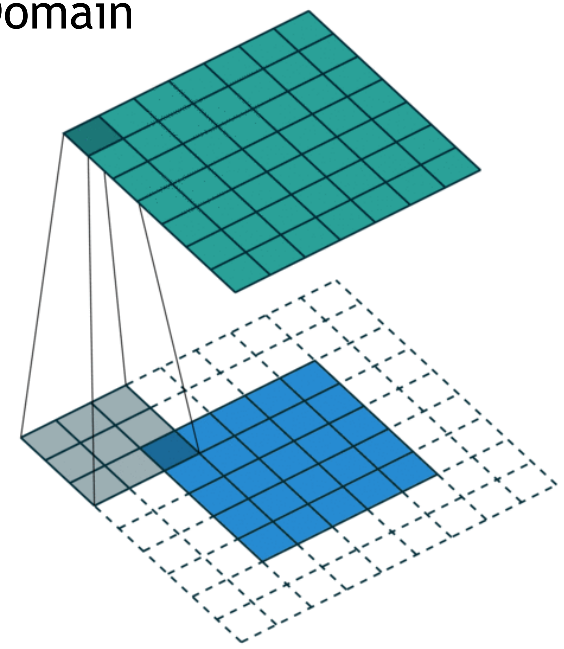


Spatial and Spectral Convolution

$$(f * k)(x) := \int_{\mathbb{R}^n} k(x - y) f(y) dy.$$

$$(f * k) = \mathcal{F}^{-1}\Big(\mathcal{F}(f)\mathcal{F}(k)\Big)$$

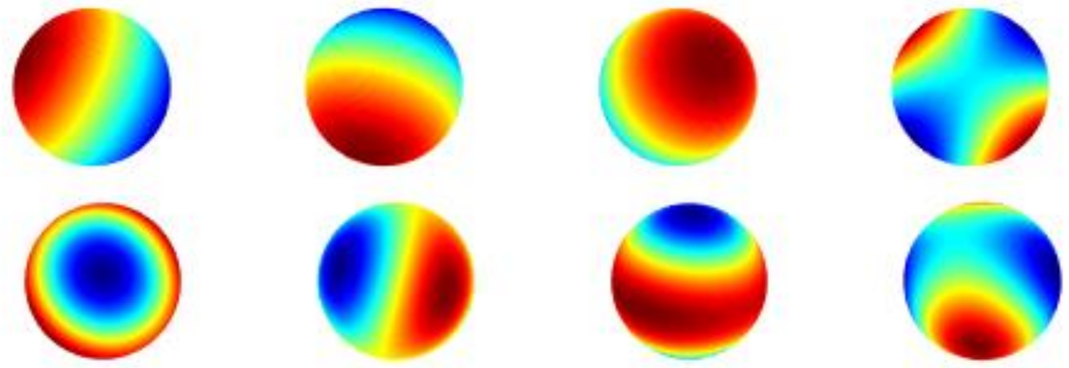
Time Domain



Sparsity = Efficiency

Frequency Domain

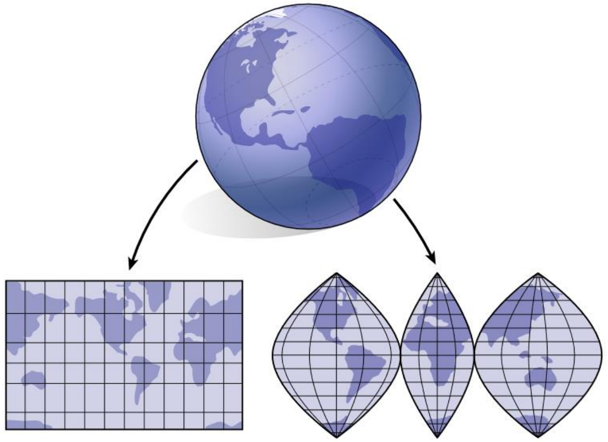
$$\Delta_g \phi = \frac{1}{\sqrt{G}} \sum_{i=1}^{2} \frac{\partial}{\partial x_i} (\sqrt{G} \sum_{j=1}^{2} g^{ij} \frac{\partial \phi}{\partial x_j})$$



Computable on irregular grids

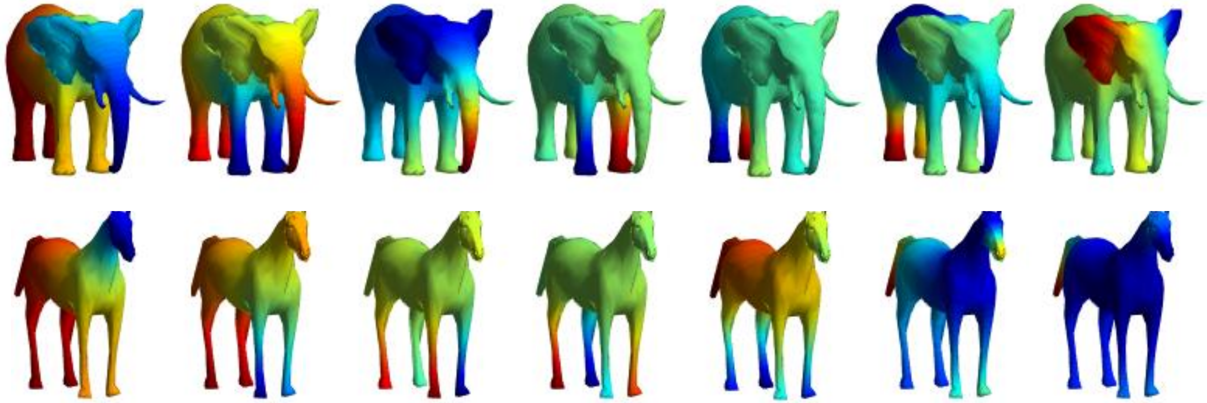
Spatial and Spectral Convolution

$$(f * k)(x) := \int_{\mathbb{R}^n} k(x-y)f(y)dy.$$

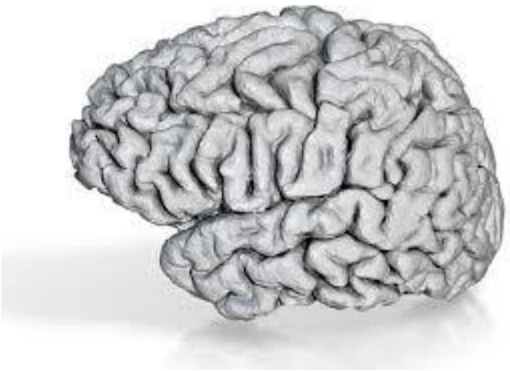


- The minus operation (x-y) encoders a direction…so we need an Atlas

$$(f * k) = \mathcal{F}^{-1}\Big(\mathcal{F}(f)\mathcal{F}(k)\Big)$$



- Sensitive to deformation
- Multiplicity and sign information matters



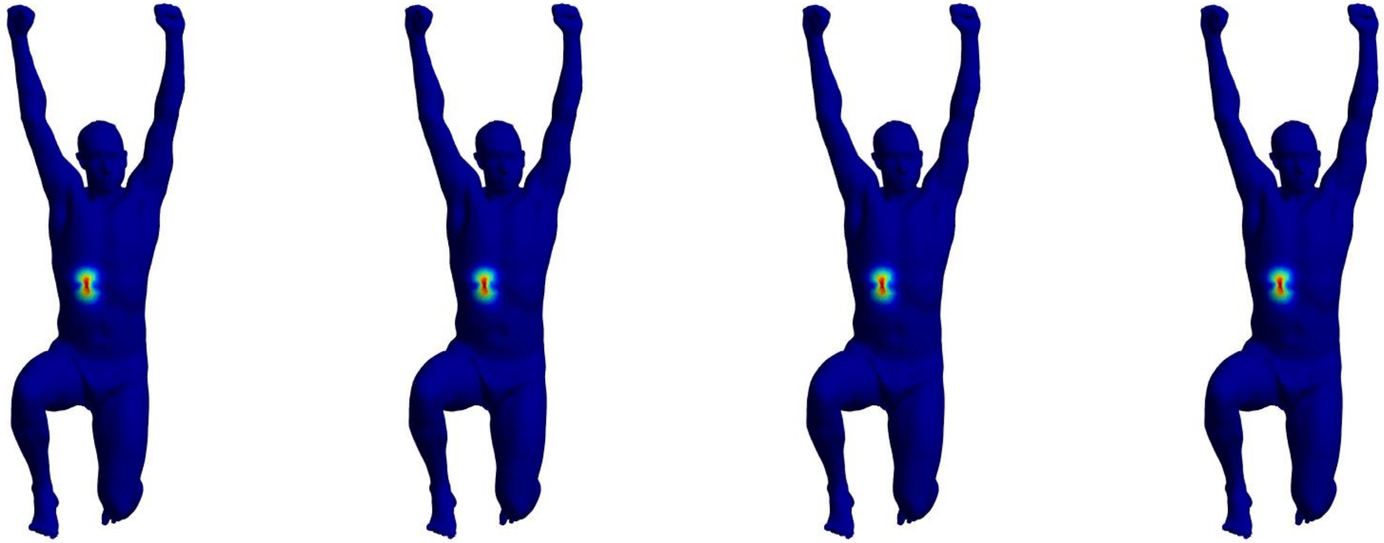
So how to deal with more complicated geometry?

Parallel Transport Convolution (PTC)

Method	Filter Type	Support	Extraction	Directional	Transferable	Deformable
Spectral [5]	Spectral	Global	Eigen	✔	✘	✘
TFG [11]	Spectral	Global	Eigen	✔	✘	✘
WFT [39]	Spectral	Local	Windowed Eigen	✔	✘	✘
GCNN [30]	Patch	Local	Variable	✘	✔	✔
ACNN [3]	Patch	Local	Fixed	✔	✔	✘
PTC	Geodesic	Local	Embedded	✔	✔	✔

TABLE 1

Comparison on different generalizations of convolutional operator on general manifolds.



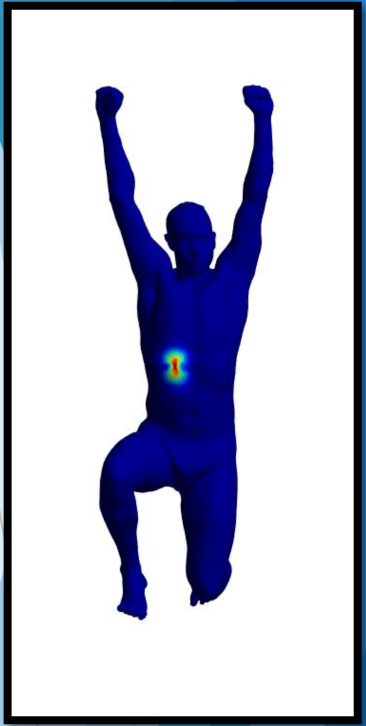
Our Proposal

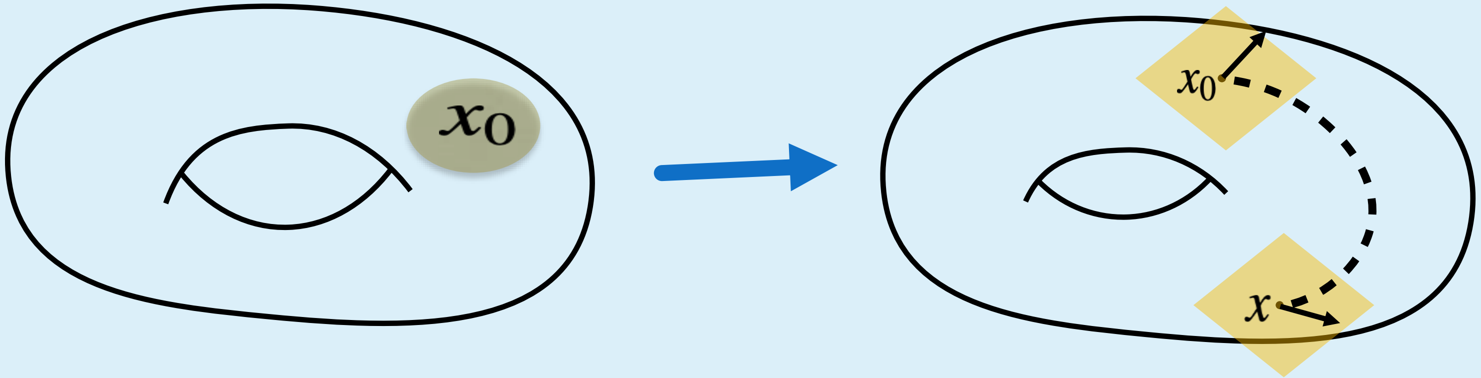
- We would like to: **flatten** → **translate** → **wrap** → **integrate**
  - Flattening and wrapping is easy via the exponential/logarithmic map
- In Euclidean convolution the "minus sign" is indicating a direction and a translation:

$$(f * k)(x) := \int_{\mathbb{R}^n} k(x-y)f(y)dy.$$

  - Since global charts don't exist, we need something to replace this operation
  - Rather than use *translation*, we use *transportation*
- Assuming filter *k* has compact support, we have:

$$(f *_{\mathcal{M}} k)(x) := \int_{\mathcal{M}} f(y)\ k(x,y) d_{\mathcal{M}} y = \int_{\mathcal{M}} f(y)\ k\left(x_0, \exp_{x_0} \circ (\mathcal{P}_{x_0}^{x})^{-1} \circ \exp_x^{-1}(y)\right)\ d_{\mathcal{M}} y$$

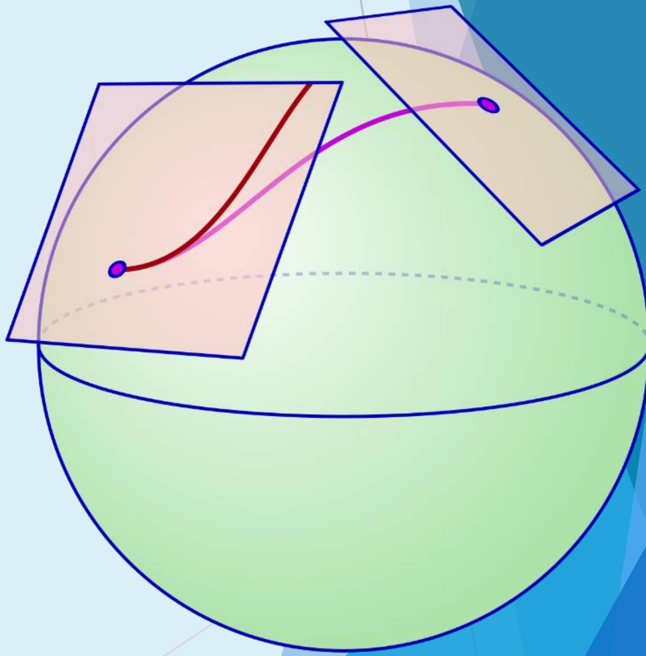




In fact, by given smooth vector fields $\{\vec{u}^1, \vec{u}^2\}$, we can define linear transformation among tangent planes $\mathcal{L}(\gamma)_s^t : \mathcal{T}_{\gamma(s)}\mathcal{M} \to \mathcal{T}_{\gamma(t)}\mathcal{M}$ satisfying:

1. $\mathcal{L}(\gamma)$ is smoothly dependent on γ.
2. $\mathcal{L}(\gamma)_s^s = Id$.
3. $\mathcal{L}(\gamma)_u^t \circ \mathcal{L}(\gamma)_s^u = \mathcal{L}(\gamma)_s^t$.

Parallel transport: $\nabla_{\dot{\gamma}} V = \lim_{h \to 0} \frac{1}{h}(\mathcal{L}(\gamma)_0^h(V_{\gamma(0)}) - V_{\gamma(0)})$



This means we can precompute everything!

Parallel Transportation Convolution

Intuitively, we want a translation-like operation which does not 'rotate' and which uses the most 'stable' path

Stable Path: Geodesic

A curve $\gamma : [0, \ell] \to \mathcal{M}$ on $\mathcal{M}$ is called geodesic if $\nabla_{\dot{\gamma}(t)} \dot{\gamma}(t) = 0$

More Formally:

$$\frac{d^2 x^k(t)}{dt^2} + \sum_{i,j=1}^{2} \Gamma_{ij}^k(t) \frac{dx^i(t)}{dt} \frac{dx^j(t)}{dt} = 0$$

No Rotation: Derivative along connection is 0

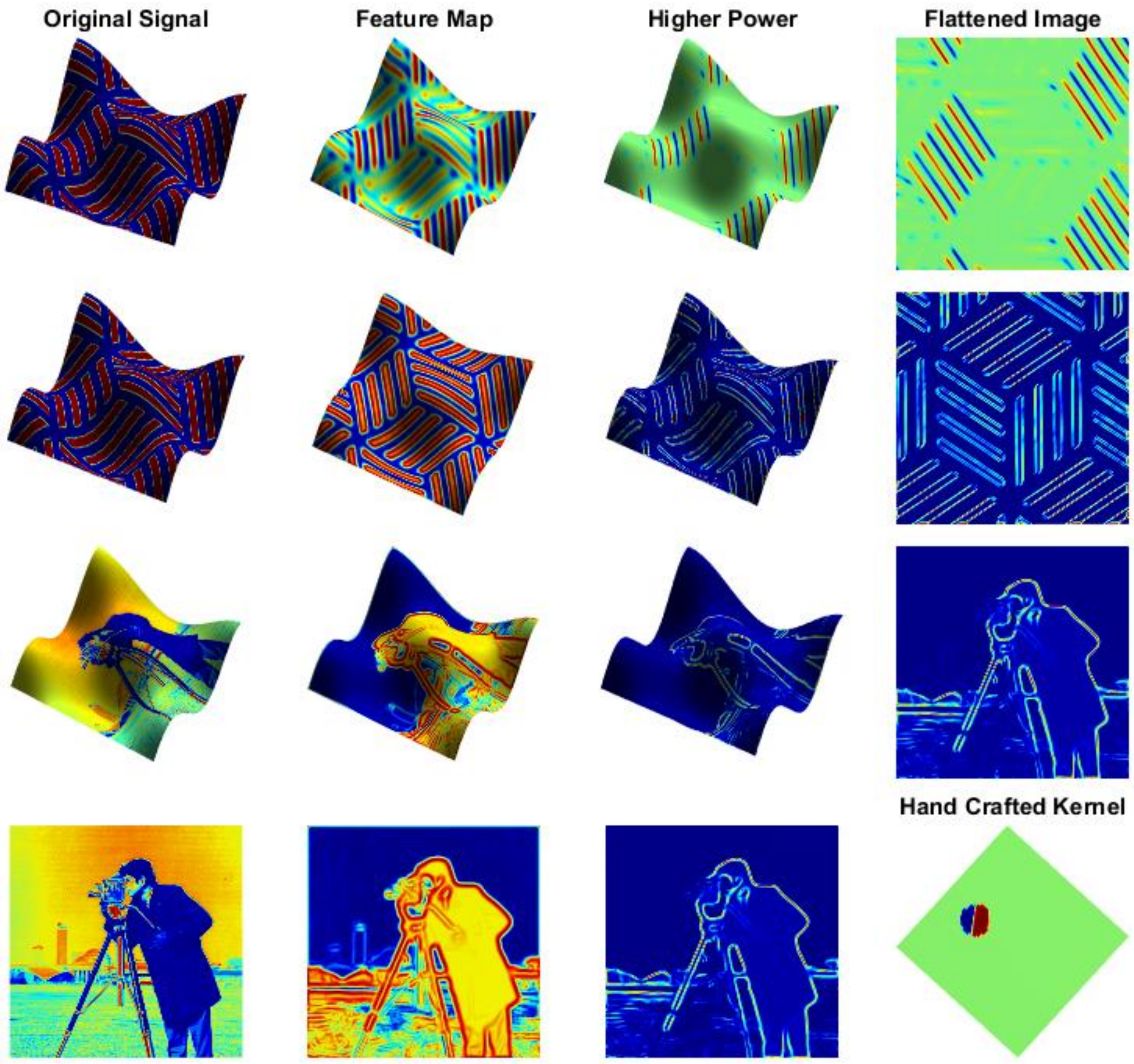
If we write $X(t) = \sum_{i=1}^{2} a^i(t) \partial x^i$, this can be written as the following first order linear system:

$$\begin{cases} \dfrac{da^k(t)}{dt} + \displaystyle\sum_{i,j=1}^{2} \dfrac{d\gamma^i}{dt} a^j(t) \Gamma_{ij}^k = 0, \quad k = 1, 2 \\ \sum_{i=1}^{2} a^i(0) \partial x^i = v \end{cases}$$

Note: In the Euclidean Plane, this all reduces to translation, and PTC becomes traditional convolution

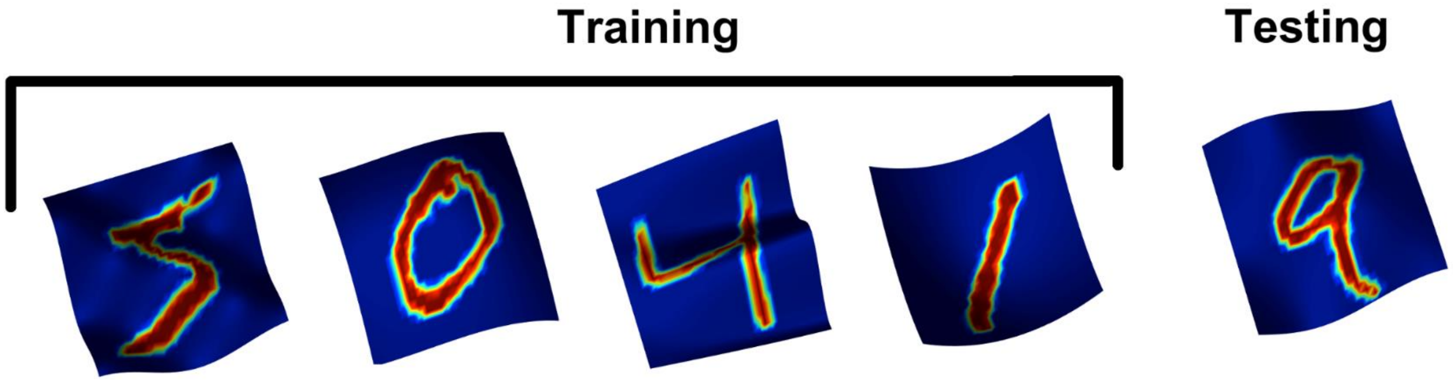
Wavelet-like Operations

- We can use ‘hand crafted’ kennels to preform traditional signal processing tasks (like edge detection) on deformable domains
- However, since most interesting manifolds do no have any homogenous structure, true wavelets remain illusive



ConvNets

Using a simple 2 Layer PTConv network, we classify handwritten digits from the MNIST database



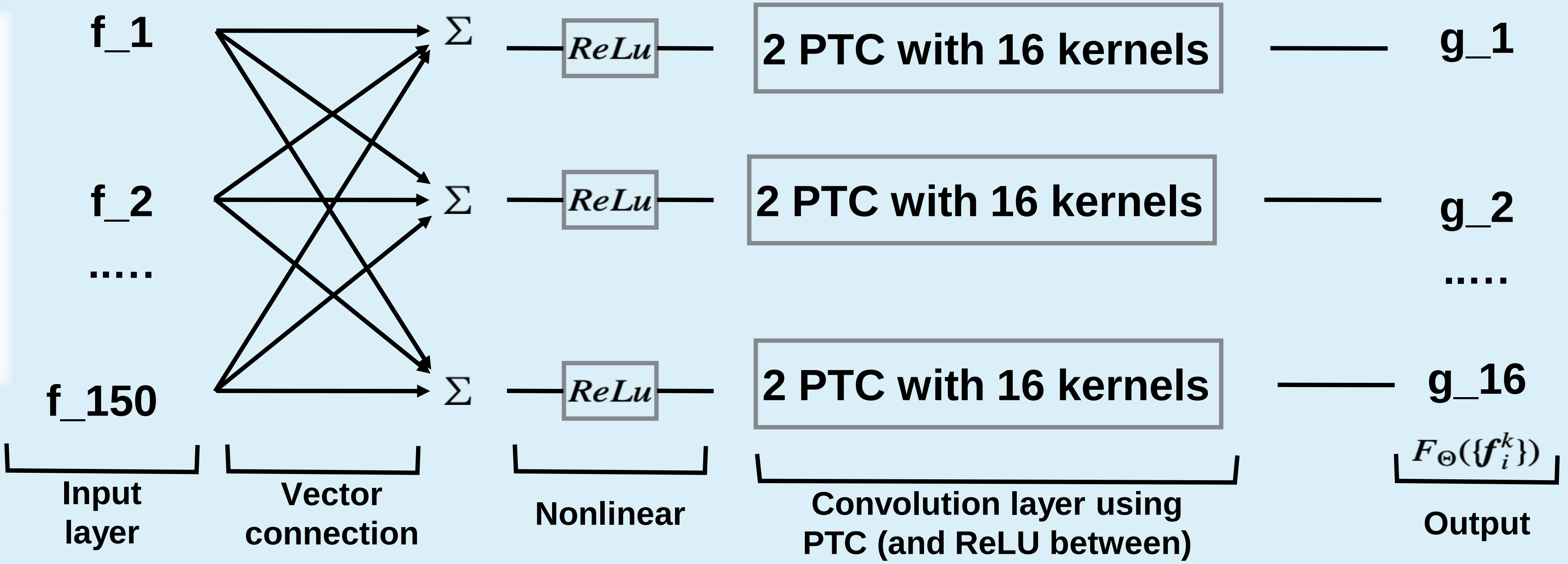
Network	Domain	Accuracy
Traditional	Euclidean	**98.85**
Flat PTCNet	Euclidean	98.10
Spectral	Manifold	95.35
PTCNet	Manifold	97.96

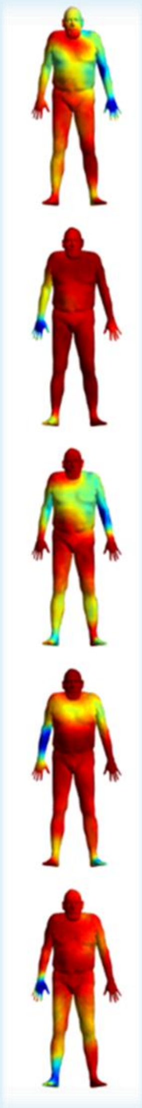
Training	Accuracy
Spectral	88.50
Single Manifold	95.65
Multiple Manifolds	**97.32**

Correspondences via shapeNet architecture

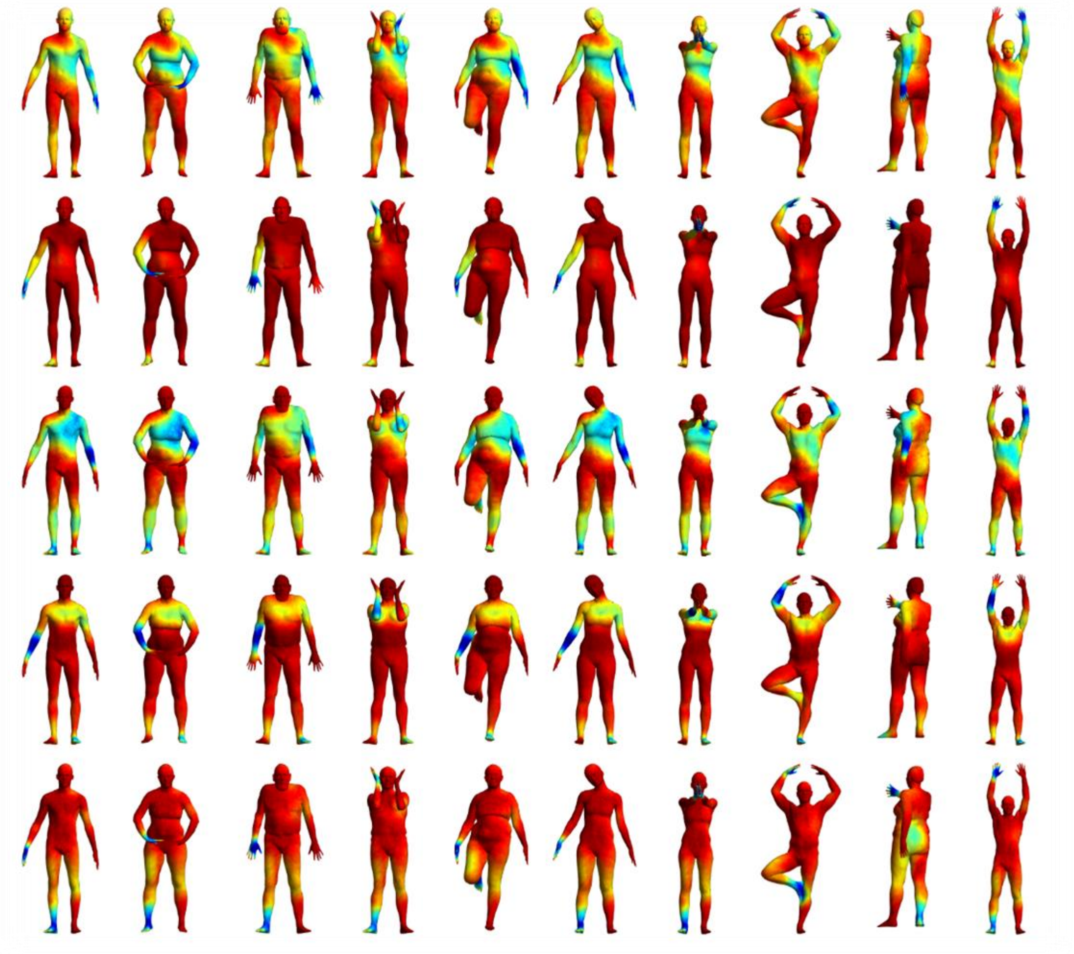
"Triplet" Loss:

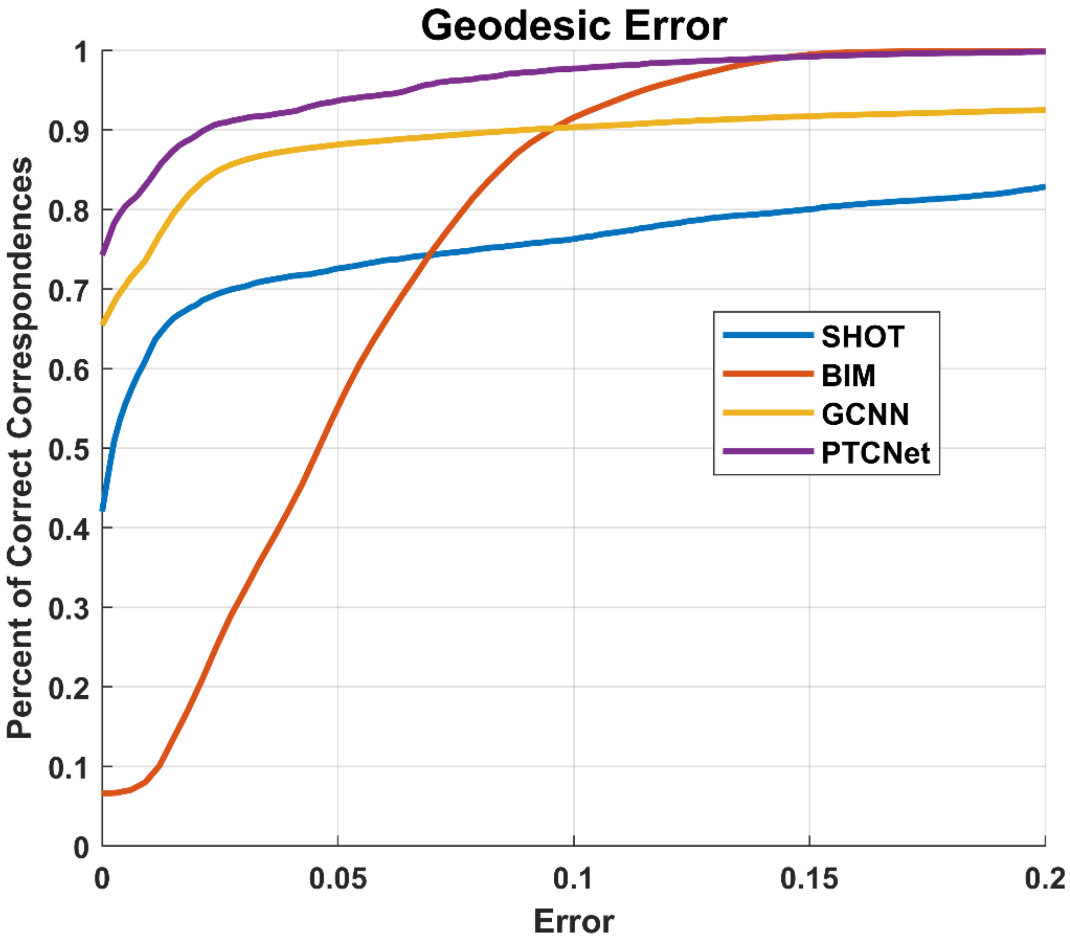
$$L(S;\Theta) = \sum_{x_1,x_2\in S\times S} ||F(x_1;\Theta)-F(x_2;\Theta)||^2+\lambda\sum_{P\in\Pi}(\mu_1-||F(x_1;\Theta)-F(x_3;\Theta))||)^2$$





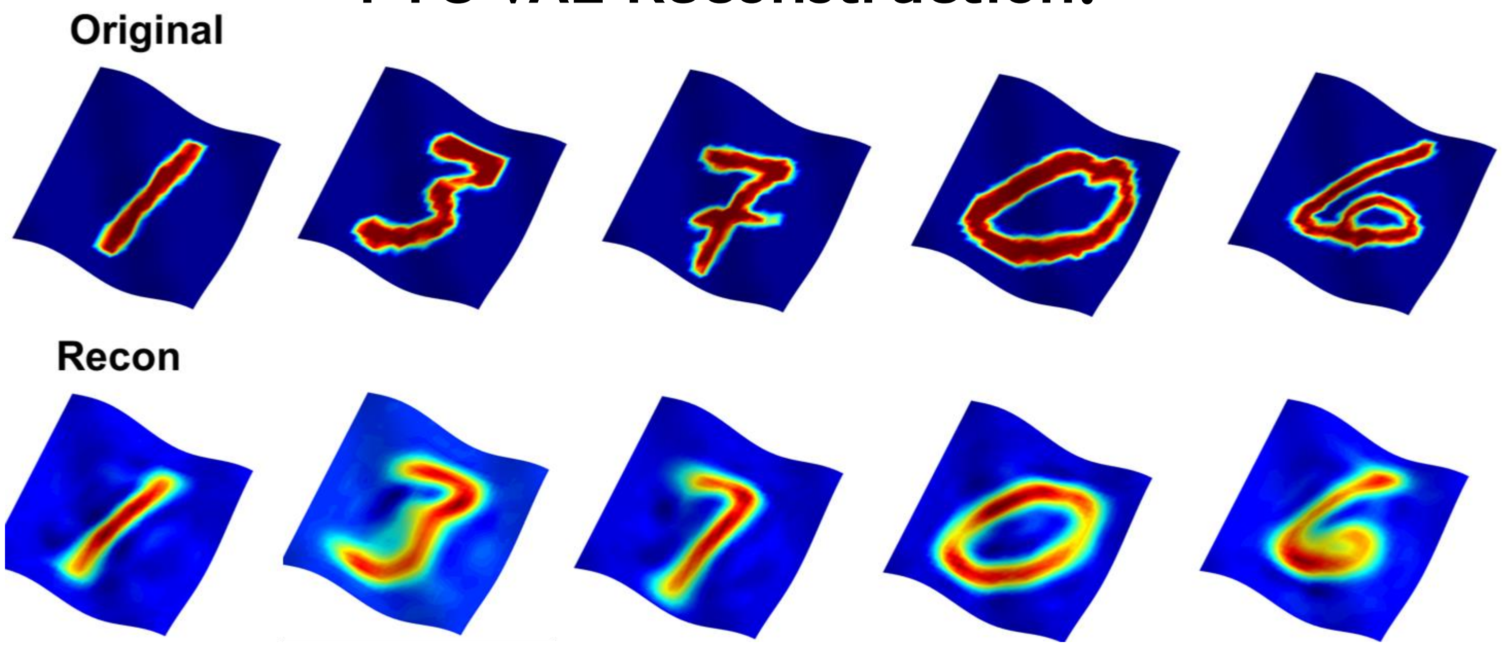
Correspondences via shapenet architecture

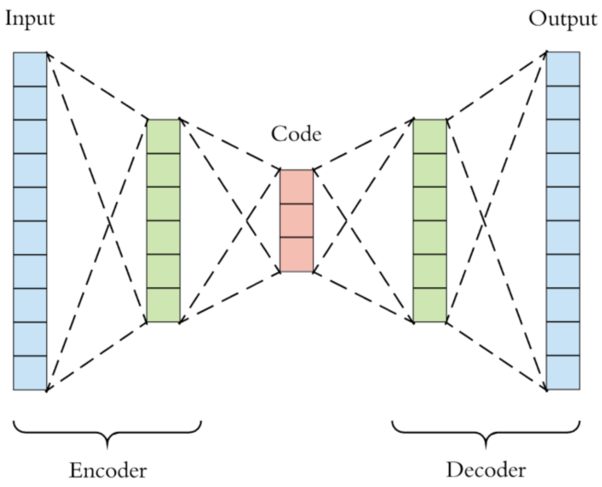




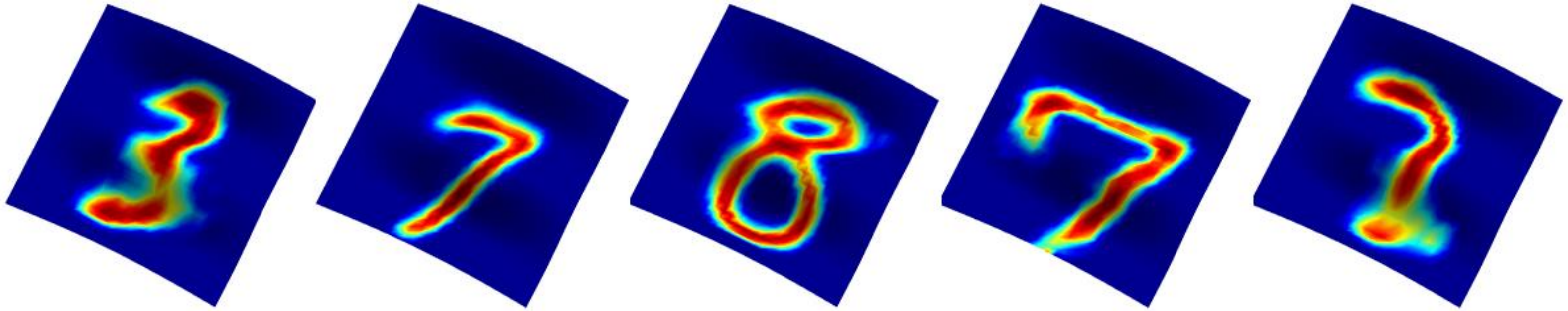
Auto-Encoding on new surfaces

PTC VAE Reconstruction:





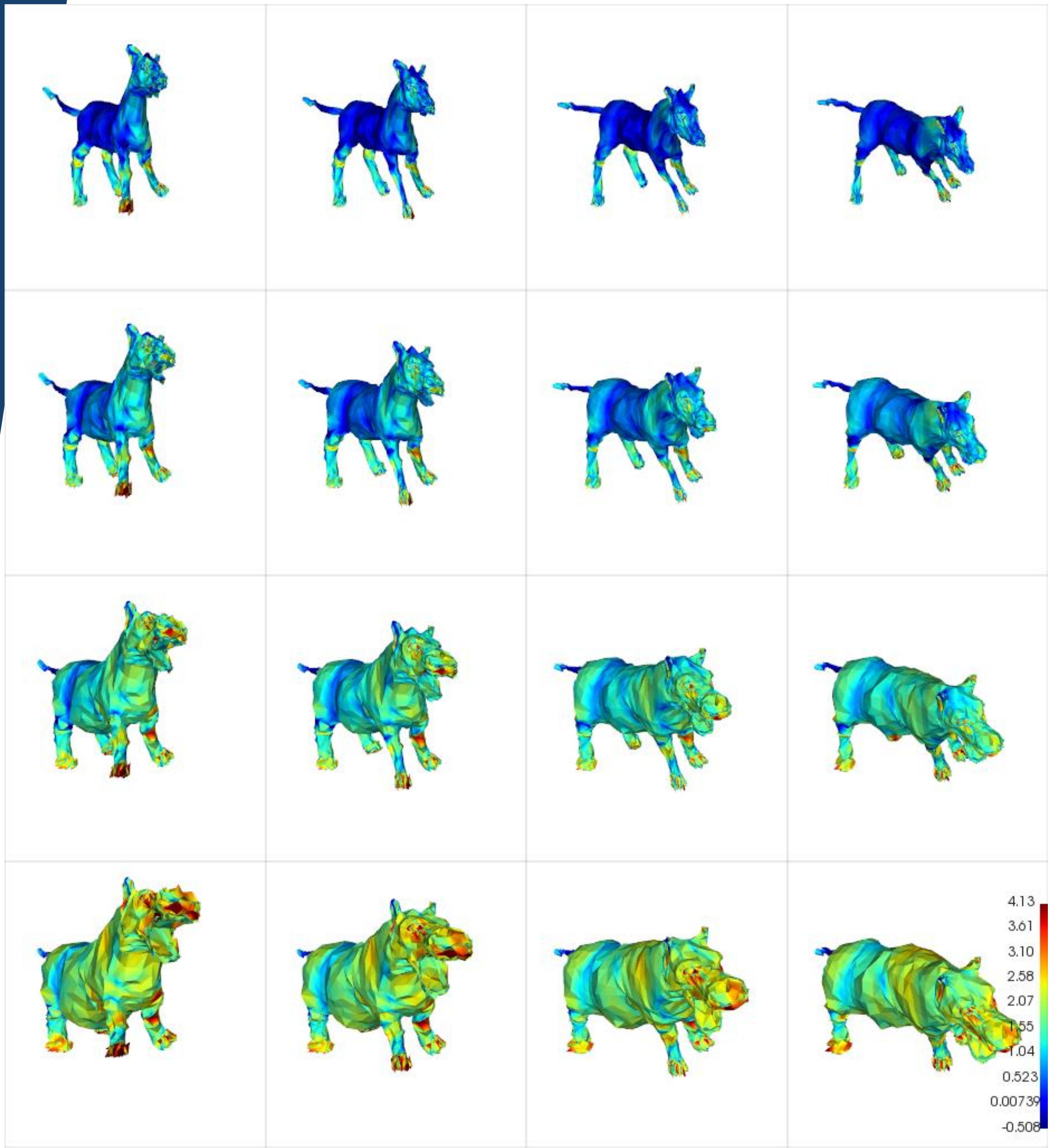
New Surface--New Images:



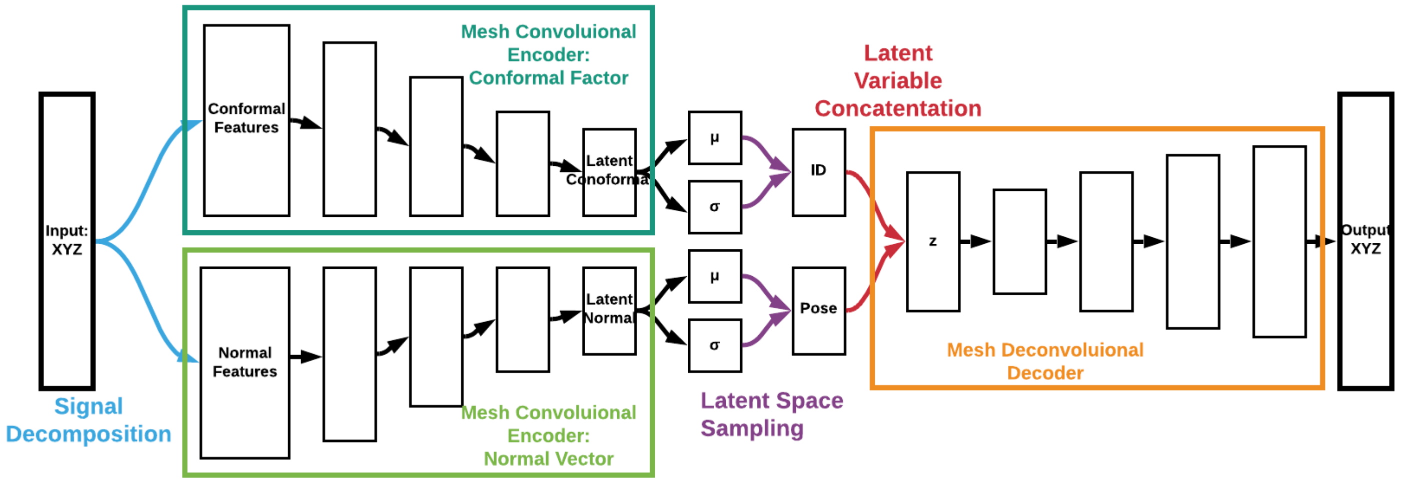
Geometric Disentanglement via PTC-VAE

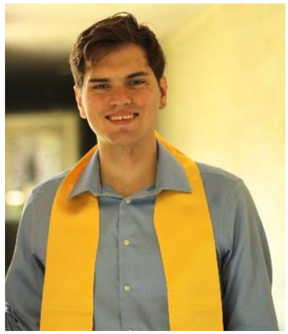
Tatro, **SCS.**, and Lai 2020

- Since PTC is intrinsically defined is automatically disentangles information when put into a standard VAE
- Each row should show the same individual in different ‘poses’ and each column should be the same ‘pose’ of a different animal

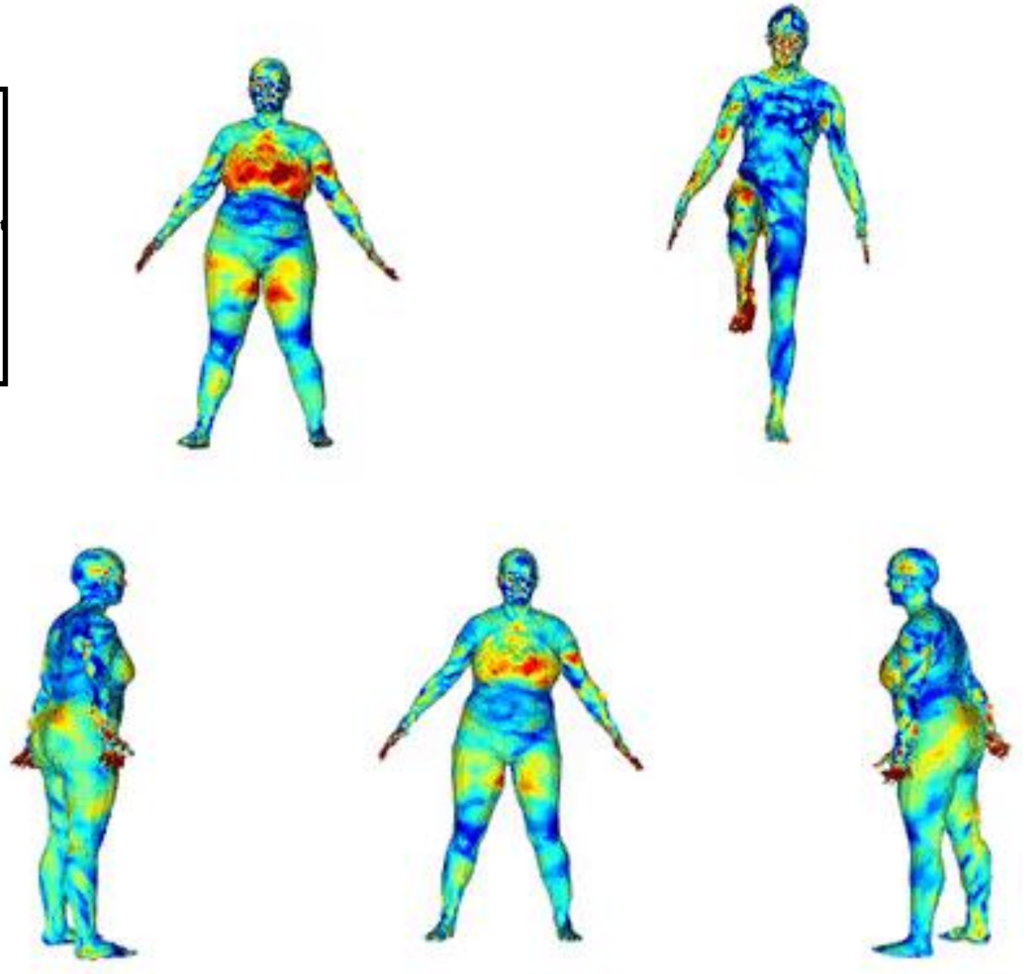


By interpolating in the latent spaces, we can easily change Identity or Pose independently





Thanks to Joey for the movie!



Limitations and Other Work

- **Singularities in vector fields may cause inconsistent frames:**
  - Our solution: Use multiple vector fields
- **Equivariant Convolution** (Ovsjanikov ‘19): Split network into patches and then do PTC on each patch
  - Then need some type of canonical segmentation
- **Gauge Networks** (Cohen ‘19): Exploit symmetry of manifolds to define gauge transition between different patches
  - Most interesting manifolds don’t have this level of symmetry
- **Narrowband PTC** (Jin ‘19): Extend PTC into reach of manifold to better deal with point cloud discretization

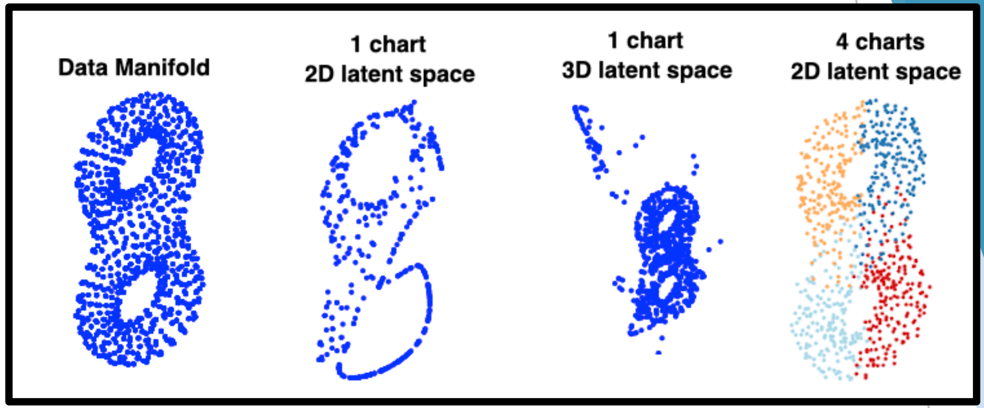
Chart Based Auto-Encoders for Manifold Structured Data

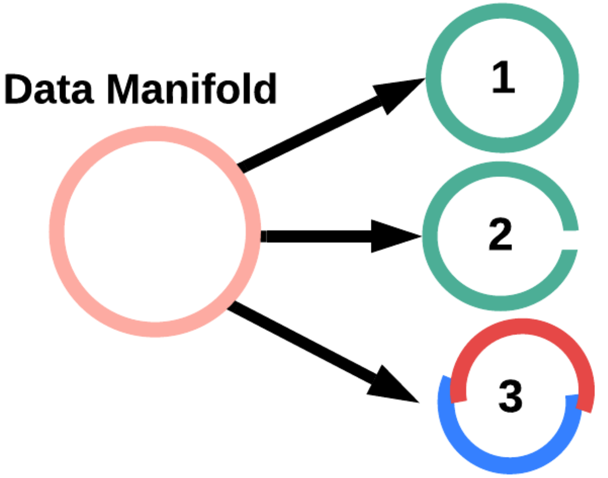
SCS., Chen and Lai

Chart Based Parameterization of Data

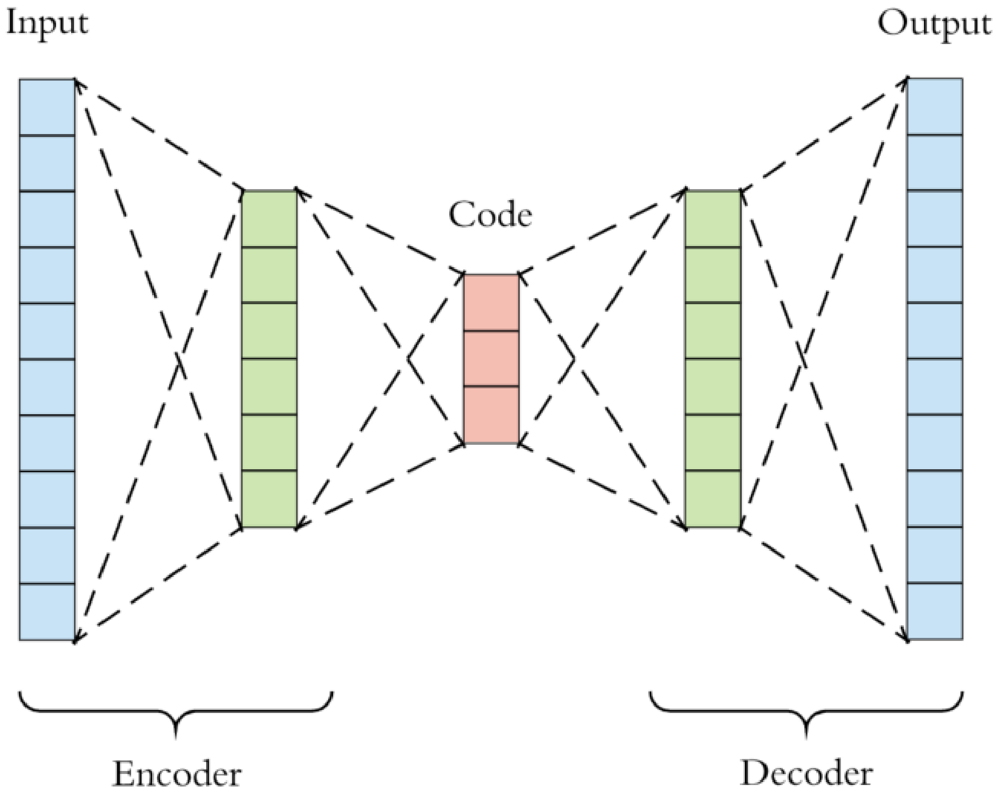
What is a good representation?

- Is it easy to learn?
- Is it easy to use?
- Does it capture the manifold *structure* of the data?





Parametrization	Latent Space	Encoder	Finite VC	Continuous Rep.	Finite Rep.
1	$(-\infty, \infty)$	(cox(z), sin(z))	No	Yes	No
2	$[0,2\pi)$	$(\cos(z), \sin(z))_{[0,2\pi)}$	Yes	No	Yes
3	$(-\delta, \pi+\delta) \cup (-\delta, \pi+\delta)$	$(\cos(z), \sin(z))_{(-\delta,\pi+\delta)}$, $(-\cos(z), -\sin(z))_{(-\delta,\pi+\delta)}$	**Yes**	**Yes**	**Yes**



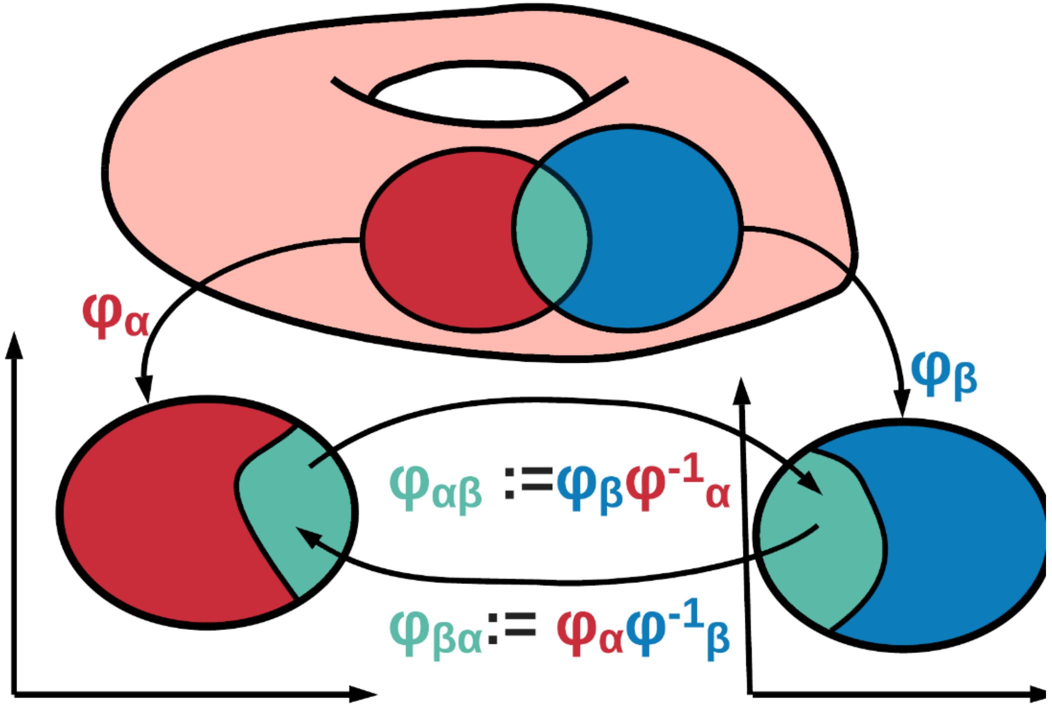


Chart Parameterization or Auto-Encoding?

Faithful Representation:

Definition 1 (Faithful Representation)**.** An auto-encoder $(\mathcal{Z}; \mathbf{E}, \mathbf{D})$ is called a faithful representation of $\mathcal{M}$ if $x = \mathbf{D} \circ \mathbf{E}(x), \forall x \in \mathcal{M}$. An auto-encoder is called an ϵ-faithful representation of $\mathcal{M}$ if $\sup_{x \in \mathcal{M}} \|x - \mathbf{D} \circ \mathbf{E}(x)\| \leq \epsilon$. If $dist(\mathcal{M}, \mathbf{D}(\mathcal{Z})) = \sup_{x \in \mathcal{M}, z \in \mathcal{Z}} \|x - \mathbf{D}(z)\| \leq \epsilon$, then the decoder $(\mathcal{Z}, \mathbf{D})$ is called an ϵ-faithful representation of $\mathcal{M}$.

Definition 2 (Reach of a manifold)**.** Given a d-dimensional compact data manifold $\mathcal{M} \subset \mathbb{R}^m$, let $\mathcal{G} = \left\{ y \in \mathbb{R}^m \mid \exists p \neq q \in \mathcal{M} \text{ satisfying } \|y - p\| = \|y - q\| = \inf_{x \in \mathcal{M}} \|x - y\| \right\}$. The reach of $\mathcal{M}$ is defined as $\tau(\mathcal{M}) = \inf_{x \in \mathcal{M}, y \in \mathcal{G}} \|x - y\|$.

Theorem 0 (**Homeomorphic latent space**)**.** Let $\mathcal{M}$ be a d-dimensional compact manifold. If an auto-encoder $(\mathcal{Z}; \mathbf{E}, \mathbf{D})$ of $\mathcal{M}$ is an ϵ-faithful representation with $\epsilon < \tau(\mathcal{M})$, then $\mathcal{Z}$ and $\mathbf{D}(\mathcal{Z})$ must be homeomorphic to $\mathcal{M}$. Particularly, a d-dimensional compact manifold with nontrivial topology can not be ϵ-faithfully represented by a vanilla auto-encoder with a latent space $\mathcal{Z}$ being a d-dimensional simply connected domain in $\mathbb{R}^d$.

Theoretical Properties of Patch Encoding

Theorem 1 (Universal Manifold Approximation Theorem). Consider a d-dimensional compact data manifold $\mathcal{M} \subset \mathbb{R}^m$ with the reach τ and $C = \dfrac{vol(\mathcal{M})\Gamma(1 + d/2)}{\pi^{d/2}}$. Let $X = \{x\}_{i=1}^n$ be a training data set drawn uniformly randomly on $\mathcal{M}$. For any $0 < \epsilon < \tau$, if the number of training set X satisfying

$$n > \beta_1\Big(\log(\beta_2) + \log(1/\nu)\Big) \tag{1}$$

where $\beta_1 = C\left(\frac{\delta}{4}\right)^{-d}\Big(1 - (\frac{\delta}{8\tau})^2\Big)^{-d/2} = O(\delta^{-d})$ and $\beta_2 = C\left(\frac{\delta}{8}\right)^{-d}\Big(1 - (\frac{\delta}{16\tau})^2\Big)^{-d/2} = O(\delta^{-d})$, then based on the training data set X, there exists a CAE, there exists a CAE $(\mathcal{Z}, \mathbf{E}, \mathbf{D}) = \{(\mathcal{Z}_\ell, \mathbf{E}_\ell, \mathbf{D}_\ell)\}_{\ell=1}^L$ which is an ϵ-faithful representation of $\mathcal{M}$ with probability $1 - \nu$. Moreover, the encoder $\mathbf{E}$ and the decoder $\mathbf{D}$ has at most $O(Lmdn^{1+d/2})$ parameters and $O(\frac{d}{2}\log_2(n))$ layers.

Construction of Simplicial Faithful Rep.

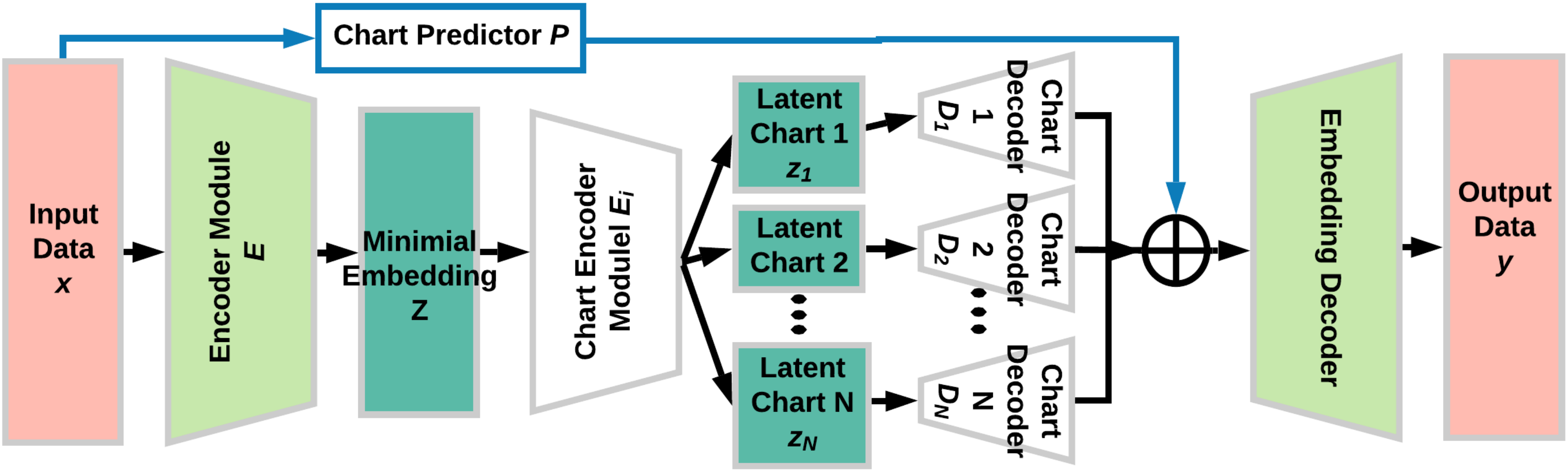
Theorem 2 (Local chart approximation). Consider a geodesic neighborhood $\mathcal{M}_r(p) = \{x \in \mathcal{M} \mid d(p, x) < r\}$ around $p \in \mathcal{M}$. For any $0 < \epsilon < \tau(\mathcal{M})$, if $X = \{x_i\}_{i=1}^n$ is a ϵ-dense sample drawn uniformly randomly on $\mathcal{M}_r(p)$, then there exists an auto-encoder $(\mathcal{Z}, \mathbf{E}, \mathbf{D})$ which is ϵ-faithful representation of $\mathcal{M}_r(p)$. In other words, we have

$$\sup_{x \in \mathcal{M}_r(p)} \|x - \mathbf{D} \circ \mathbf{E}(x)\| \leq \epsilon$$

Moreover, the encoder $\mathbf{E}$ and the decoder $\mathbf{D}$ has at most $O(mdn^{1+d/2})$ parameters and $O(\frac{d}{2}\log_2(n))$ layers.

Theorem 3 (Simplicial representation). Given a d-dimensional simplicial complex $\mathcal{S} = \bigcup_\alpha \mathcal{S}_\alpha$ with n vertices $\{v_\ell\}_{\ell=1}^n$ where each S_α is a d-dimensional simplex. Then, for any given piecewise linear function $f : \mathcal{S} \to \mathbb{R}$ satisfying f linear on each simplex, there is a ReLU network representing f. Moreover, this neural network has $n(K(d+1) + 4(2K-1)) + n$ paremeraters and $\log 2(K) + 2$ layers, where $K = \max_i |\mathcal{N}(v_i)|$ which is bound above by the number of total d-simplices in $\mathcal{S}$.

Architecture and Loss



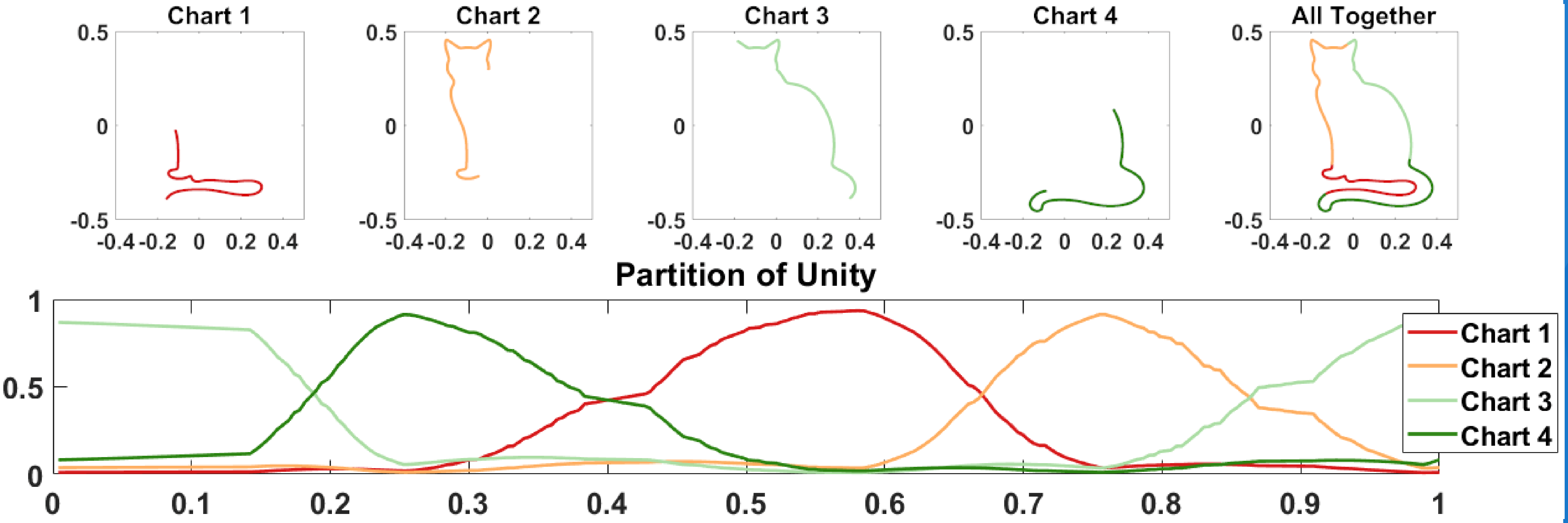
$$\mathcal{L}_{CP}(x,\Theta) := \left(\min_{\alpha} e_{\alpha}\right) - \sum_{\beta=1}^{N} \ell_{\beta} \log(p_{\beta}),$$

Reconstruction Error

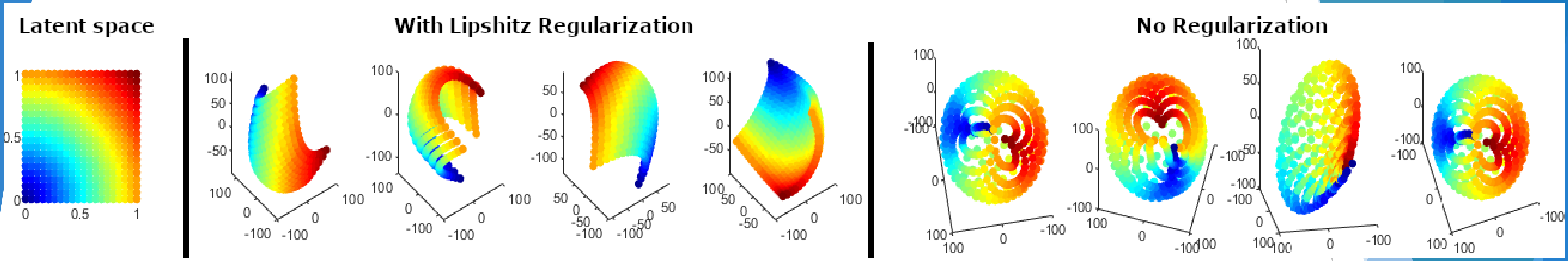
Log-Likelihood of Chart Prediction
—Weighted by 1/Recon. Error

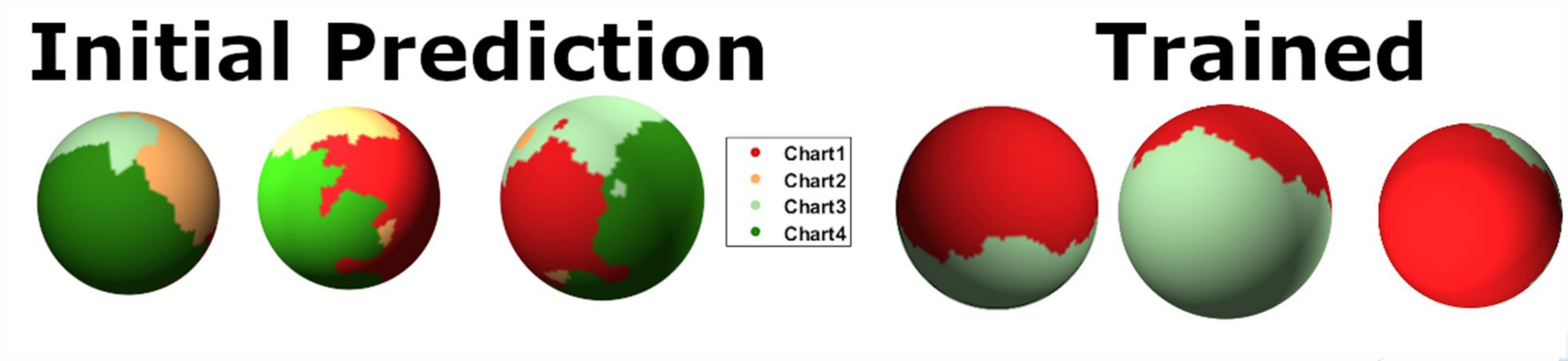
Transitions

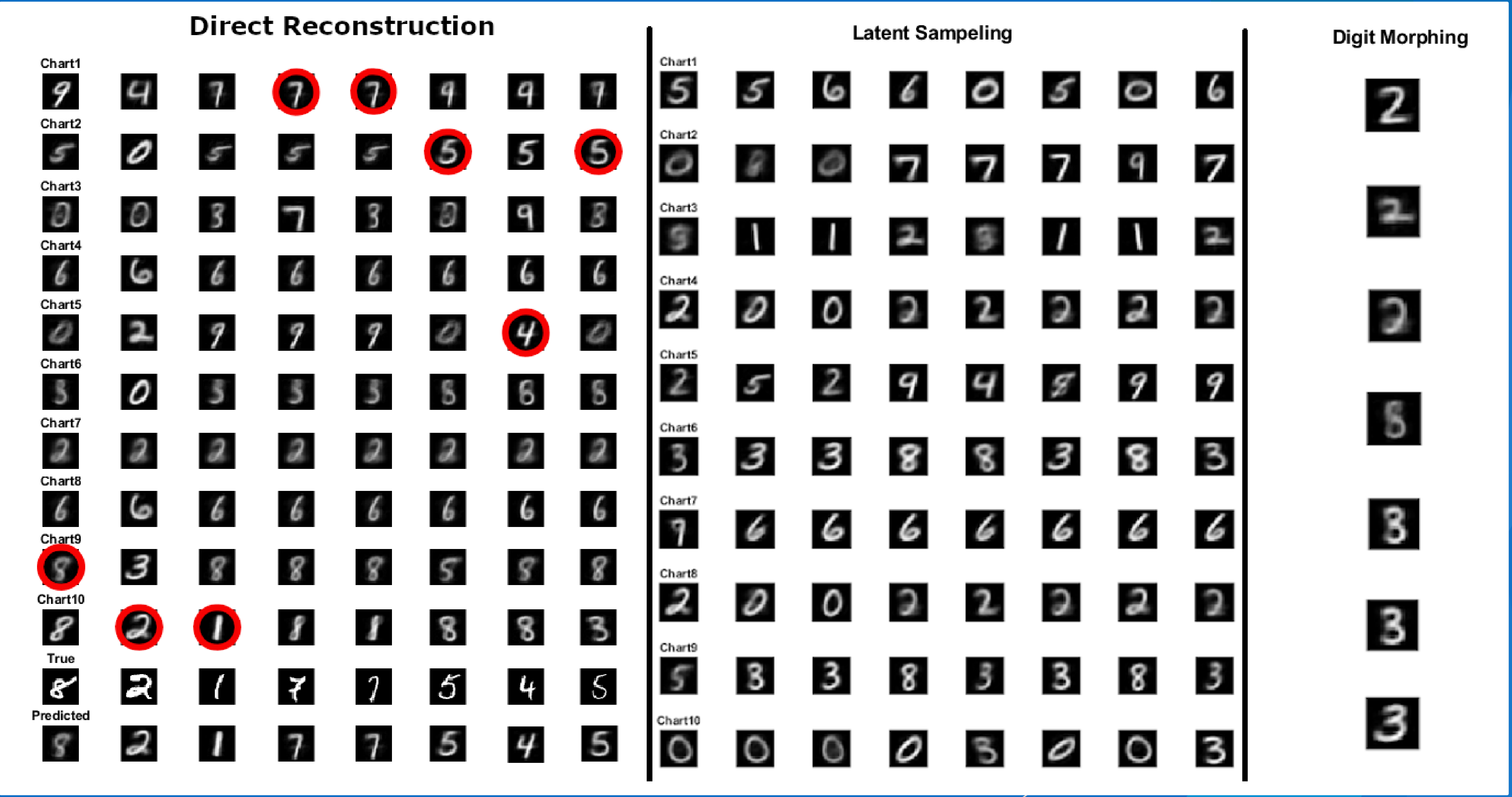
Visualizing the training of a 1d manifold can help us to understand the behavior of these networks

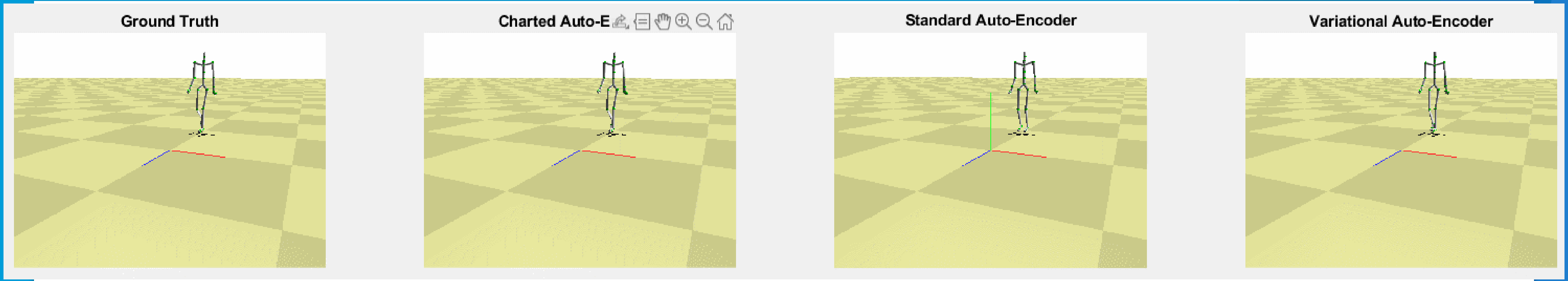


Learning a Sphere

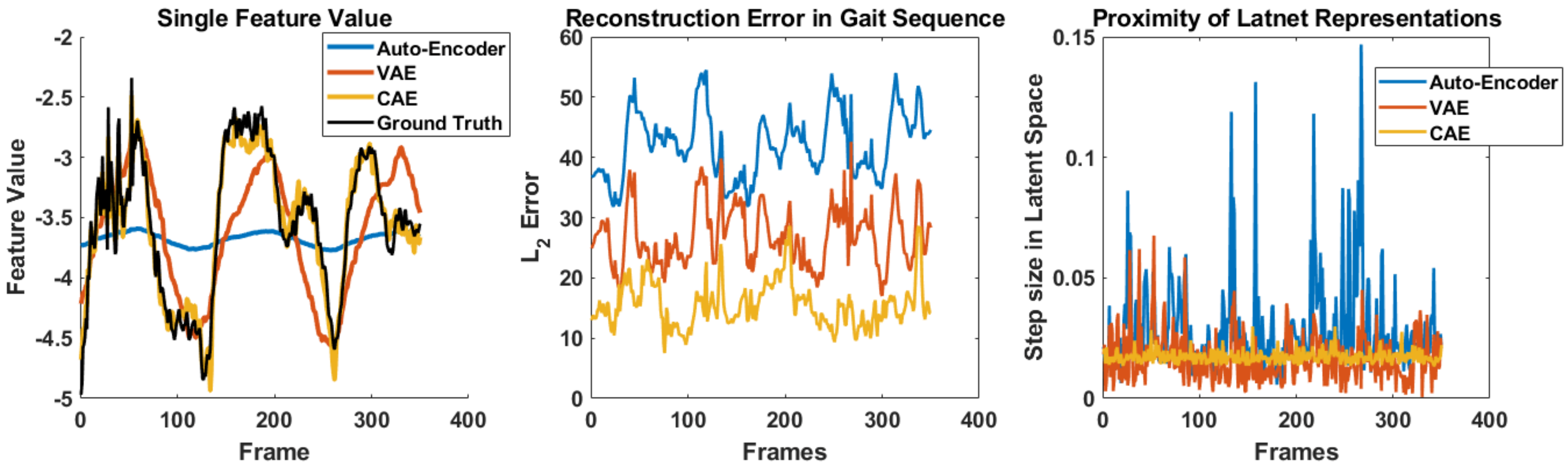





Direct Reconstruction
Chart1
Chart2
Chart3
Chart4
Chart5
Chart6
Chart7
Chart8
Chart9
Chart10
True
Predicted
Latent Sampeling
Chart1
Chart2
Chart3
Chart4
Chart5
Chart6
Chart7
Chart8
Chart9
Chart10
Digit Morphing



Interpolation between frames doesn't work with standard autoencoder because the linear path between representations doesn't respect the cyclic nature of human gait



Ongoing and Future Work

Conormal LB Eigen Systems for Graphs:

- Conformal deformations ≈ reweighted graph
- ML on deformed graphs ≈ Attention models?

Applications of PTC

- Shape analysis—PTC is intrinsic can be used to analyze geometric info
- Generative models for computer graphics/VR/AR
- Stability Analysis for deforming manifolds
- ‘Scale-invariant’ convolutions

Application in Computational Chemistry

- CAE gives an idea of ‘part’ based parameterization
- PTC gives tool to analyze geometric structure of molecules/proteins

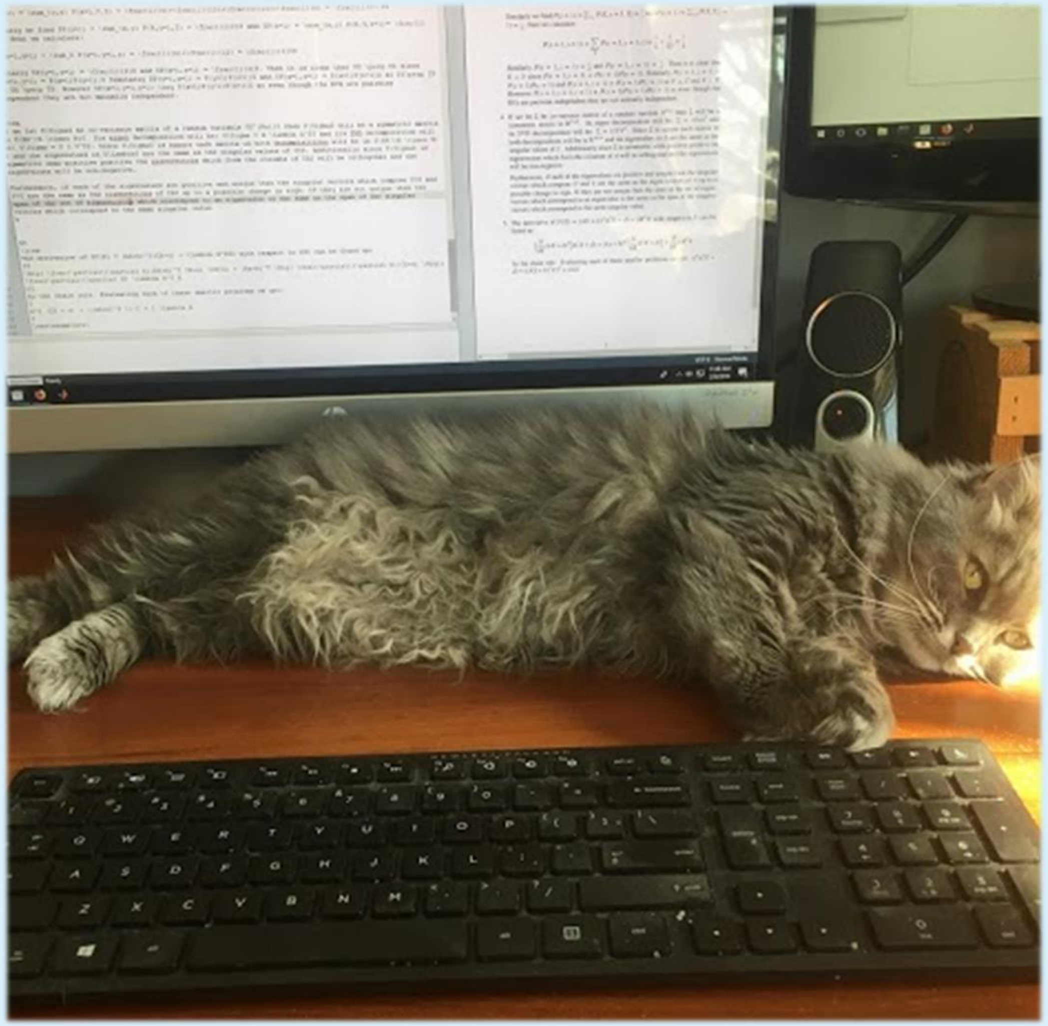
Theoretical Analysis of Deep Learning

- Similar manifold analysis for GAN/wGAN models
- Using non-parametric statistics for better bounds

Thank You

Questions?

- Comments?
- Concerns?
- Ideas for future work?



Slides available at: https://sites.google.com/view/stefancschonsheck