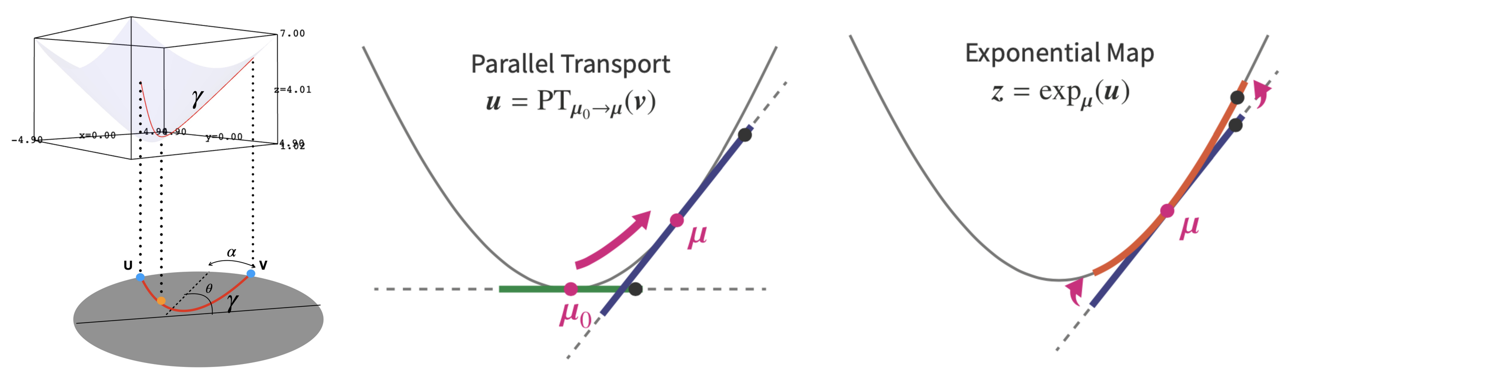
Leveraging Kinematic Space for Deep Representation Learning

Adarsh Jamadandi [1] Uma Mudenagudi[1]

[1]KLE Technological University, Hubli, India.

Wrapped Normal Distribution in Hyperbolic Space

1. Wrapped Normal Distribution in Lorentz Model of hyperbolic geometry [3].
2. Construct Gaussian-like distribution on the tangent space at $\mu_0 = 0$. Use Parallel Transport and Exponential Map to map to a Riemannian manifold.



Application- Hyperbolic Variational Autoencoders (HVAE)

1. Generalization of VAEs with latent space as Lorentz model.
2. Compatible with Riemannian Optimization Tools.

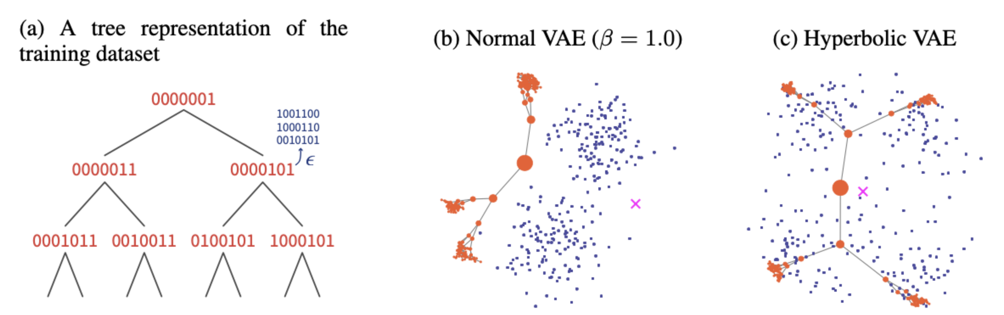


Figure 1. Hyperbolic-VAE automatically discovers hierarchy.

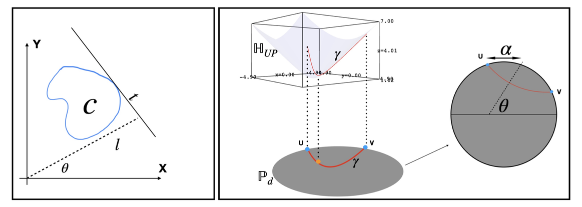
3. Can we go beyond hyperbolic spaces for more powerful representations?

Our Contributions

1. We introduce Kinematic space, an auxiliary Lorentzian geometry as a powerful transformation tool that can re-interpret geometrical information such as geodesic distance and exponential maps from one geometrical space to another.
2. Leveraging this formulation, we show that learning representations in the upper half-plane model is equivalent to learning in a maximally symmetric pseudo-Riemannian manifold called de Sitter space, where Riemannian optimization methods are applicable.
3. We formulate Wrapped Normal Distribution in Kinematic Space and use it to generalize the notion of Variational Autoencoders, which we call $\mathcal{K}_s-$VAE, with Gaussian-like priors constructed in this newly proposed space. We perform extensive experiments on different standard benchmark datasets and show consistent improvements over existing Euclidean and Hyperbolic versions.

Kinematic Space

1. An auxiliary Lorentzian geometry inspired by Theoretical Physics [1] and Integral Geometry [4].
2. Powerful mathematical formalism that can transform geometrical information such as geodesic distance and exponential map from one space to another.



3. How? Using -

Crofton's Formula

$$\text{Length} = \frac{1}{4}\int_0^{2\pi} d\theta \int_{-\infty}^{+\infty} \eta(\theta, l) dl \tag{1}$$

Length of a curve can be re-interpreted as volume of intersecting lines (geodesics). The space of oriented geodesics - Kinematic space.

de Sitter space ($de\mathbb{S}_2$) as the Kinematic Space

1. We propose to use Poincaré upper half plane model ($\mathbb{H}_{UP}$) of hyperbolic geometry as our latent space for constructing VAEs.
2. Rarely considered in literature - Computationally Intractable.
3. $\mathbb{H}_{UP} \rightleftharpoons_{\mathcal{K}_s}$ de Sitter space.

Learning in $\mathbb{H}_{UP}$ is equivalent to learning in $de\mathbb{S}_2$!

$de\mathbb{S}_2$

The de Sitter space is a maximally symmetric, Lorentzian manifold with constant positive curvature. Let $de\mathbb{S}_2$ be the $(d+1)$ dimensional de Sitter space in the $(d+2)$ dimensional Minkowski space $\mathbb{M}$ visualized as a single sheeted hyperboloid with pseudo-radius λ given by $-z_0^2 + z_1^2 + z_3^2 + ... + z_n^2 = \lambda^2 = \frac{1}{K}$. The induced distance function is given by

$$d_{de\mathbb{S}_2}(\mathbf{x}, \mathbf{y}) = \lambda \operatorname{arcosh}\left(\frac{-\langle \mathbf{x}, \mathbf{y} \rangle_{\mathcal{L}}}{\lambda^2}\right) \tag{2}$$

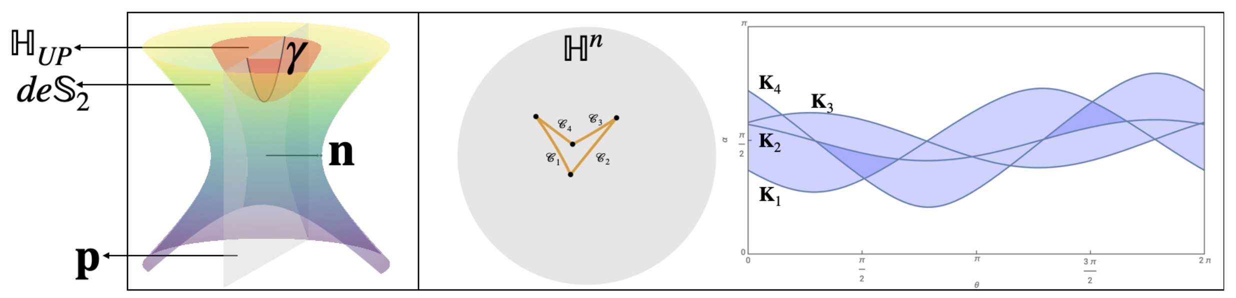


Figure 2. (Left) The de Sitter space can be visualized as a single sheeted hyperboloid in Minkowski space. A geodesic γ drawn on the Upper half plane model can be interpreted as a single point in de Sitter space.(Right) Curves drawn in hyperbolic space $\mathbb{H}^2$ and their corresponding Kinematic space.

Wrapped Normal Distribution in $de\mathbb{S}_2$

We construct Gaussian-like distribution on de Sitter space. The steps involved -

1. Sampling a vector $\mathbf{v}$ from the Gaussian distribution $\mathcal{N}(0, \Sigma)$ defined over $\mathbb{R}^n$.
2. Parallel transporting $\mathbf{v}$ from the tangent space $\mathbf{o}$ to the tangent space of new point $\mathbf{u}$ to obtain $\mathbf{j}$ by using the formula,

$$PT_{\mathbf{o}\to\mathbf{u}}(\mathbf{v}) = \mathbf{v} + \frac{K\langle y, u \rangle_{\mathcal{L}}}{1 + K\langle o, u \rangle_{\mathcal{L}}}(o + u) \tag{3}$$

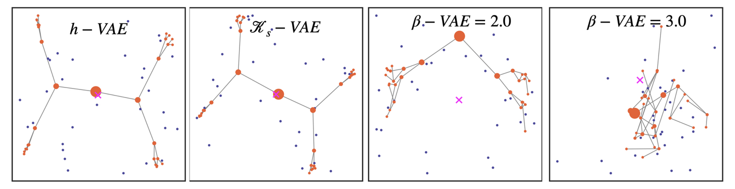
3. Map the point $\mathbf{j}$ to the manifold using the exponential map at $\mathbf{u}$ given by the Equation ??.

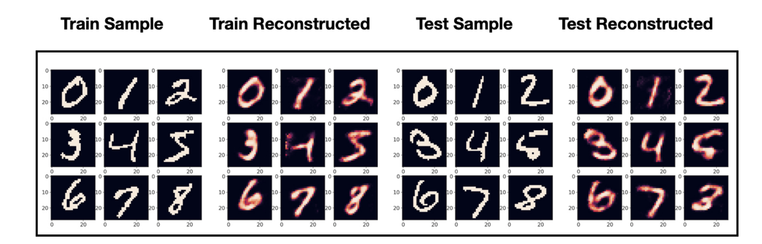
To calculate the probability density of $\mathcal{G}_{de\mathbb{S}_2}(\mu, \Sigma)$,

$$\log g(\mathbf{z}) = \log g(v) - (n-1)\log\left(\frac{\sinh ||\mathbf{j}||}{||\mathbf{j}||}\right) \tag{4}$$

where, $\log g(\mathbf{z})$ is the wrapped normal distribution and $\log g(v)$ is the normal distribution in tangent space of $\mathbf{o}$.

Application - Kinematic Space VAE ($\mathcal{K}_s$-VAE)





	Model	Correlation	Correlation with Noise
Vanilla [2]	β=1.0	0.86 ± 0.06	0.70 ± 0.00
	β=2.0	0.74 ± 0.00	0.59 ± 0.04
	β=3.0	0.35 ± 0.09	0.06 ± 0.00
	h-VAE [3]	0.87 ± 0.00	0.64 ± 0.00
	Ours	**0.89± 0.00**	**0.72 ± 0.05**

Dimension	β-VAE[2] (LL)	h-VAE[3](LL)	**Ours**(LL)
2	-143.16 ± 2.72	-145.61 ±0.45	**-140.88 ± 4.04**
5	-105.78±40.51	-109.22±0.61	**-106.09 ± 0.72**
10	-86.25±0.52	86.40±0.28	**-79.44 ± 0.62**
20	-77.89±0.36	-79.23±0.20	**-61.99±0.18**

References

[1] Bartlomiej Czech, Lampros Lamprou, Samuel McCandlish, and James Sully.
Integral Geometry and Holography.
JHEP, 2015.

[2] Irina Higgins, Loïc Matthey, Arka Pal, Christopher Burgess, Xavier Glorot, Matthew M Botvinick, Shakir Mohamed, and Alexander Lerchner.
beta-vae: Learning basic visual concepts with a constrained variational framework.
In *ICLR*, 2017.

[3] Yoshihiro Nagano, Shoichiro Yamaguchi, Yasuhiro Fujita, and Masanori Koyama.
A wrapped normal distribution on hyperbolic space for gradient-based learning.
In *ICML*, 2019.

[4] Luis A. Santaló and Mark Kac.
Integral Geometry and Geometric Probability.
Cambridge University Press, 2004.