# Non-linear Embeddings in Hilbert Simplex Geometry

Frank Nielsen
Sony Computer Science Laboratories Inc

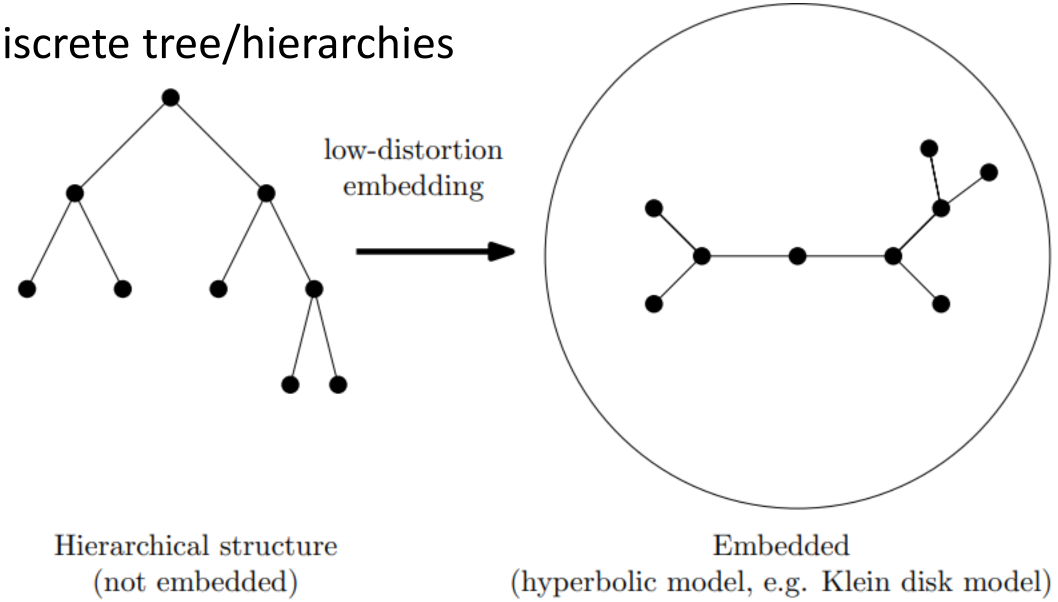
Ke Sun
CSIRO Data 61

2023

arXiv:2203.11434
ICML TAG-ML'23

# Embedding discrete hierarchical structures in hyperbolic spaces: Continuous representations

Discrete tree/hierarchies

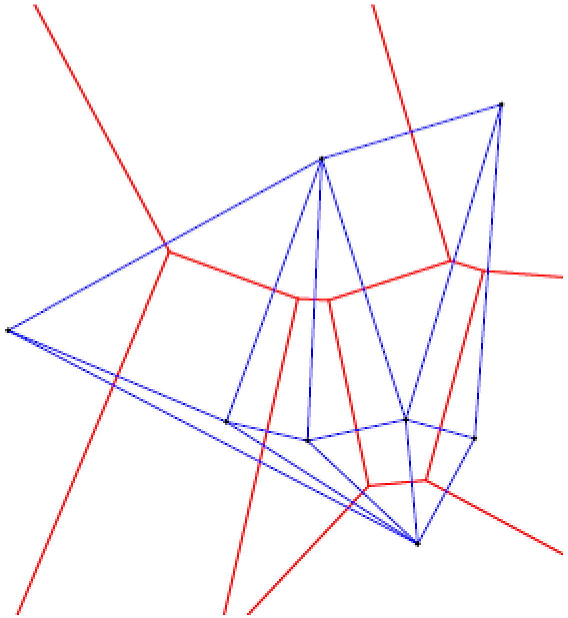



Downstream tasks

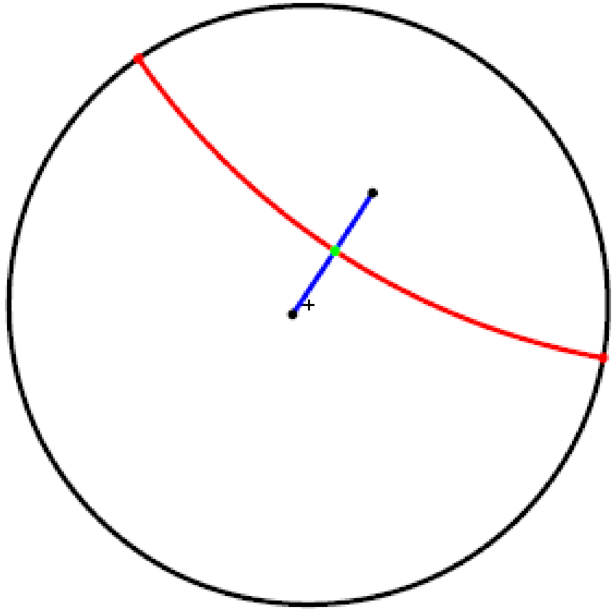
**Theorem: One can embed *any weighted edge tree* as a *Delaunay graph* of embedded tree nodes in *hyperbolic geometry* with *arbitrary small distortion***

**Sarkar, Low distortion delaunay embedding of trees in hyperbolic plane International symposium on graph drawing, 2011**

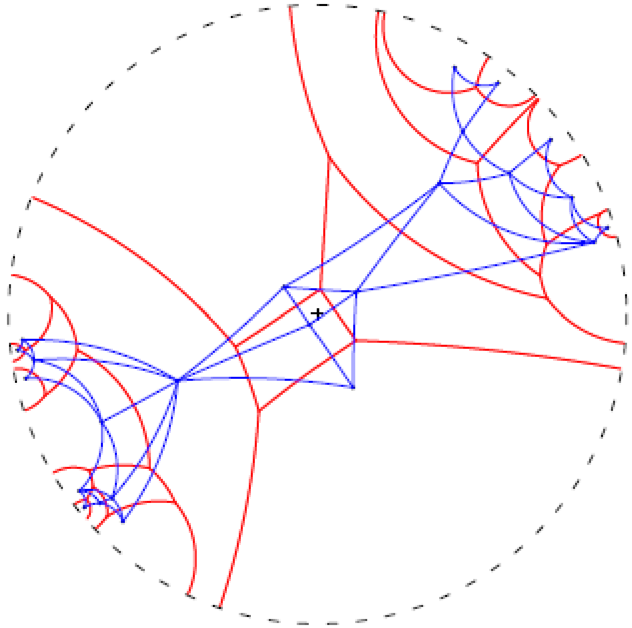
# Hyperbolic Voronoi diagram/Delaunay complex



Ordinary Voronoi/Delaunay



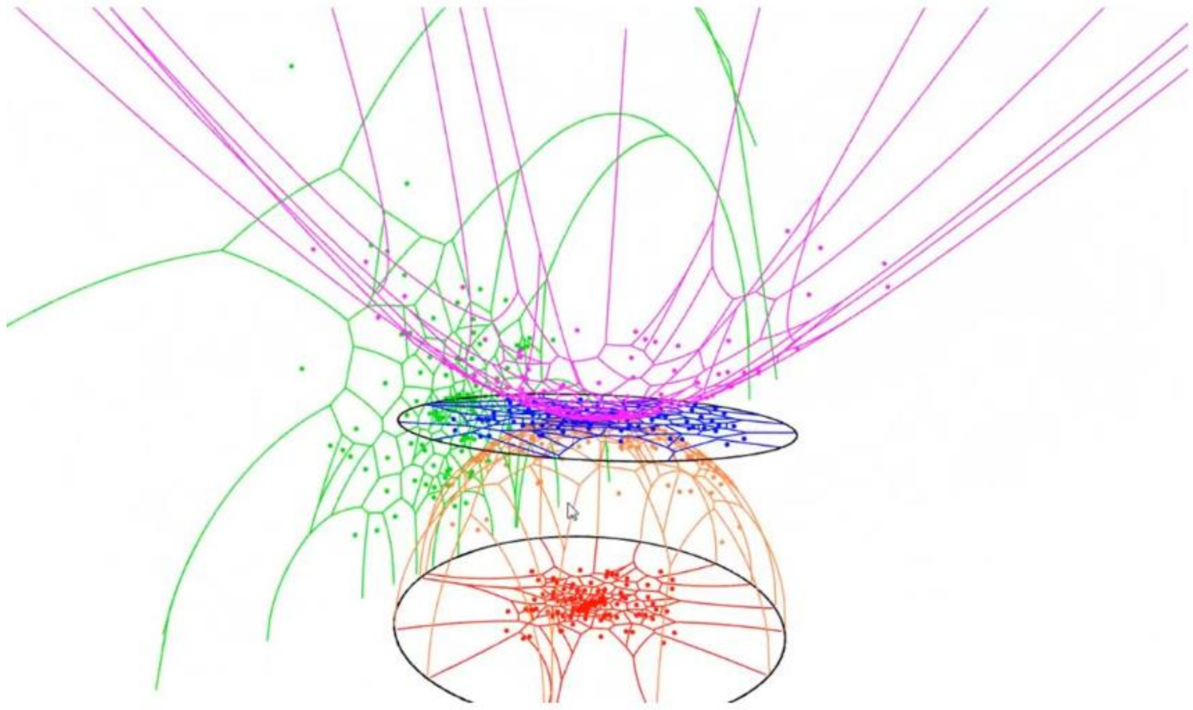
Poincaré bisector/geodesic

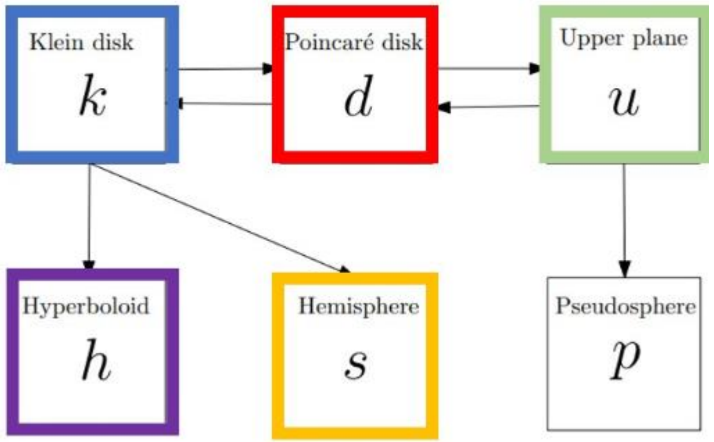


Poincaré Voronoi/Delaunay

**On Voronoi diagrams on the information-geometric Cauchy manifolds**
**Entropy 22.7 (2020)**

# Various models of hyperbolic geometry






$$
\begin{aligned}
k \to h &: \frac{1}{\sqrt{1-\|k\|^2}}(k_1, k_2, 1) \\
k \to d &: \frac{1-\sqrt{1-\|k\|^2}}{\|k\|^2}(k_1, k_2) \\
d \to k &: \frac{2}{1+\|d\|^2}(d_1, d_2) \\
d \to u &: \frac{i(1+d)}{1-d} \\
u \to d &: \frac{u-i}{u+i} \\
k \to s &: \frac{1}{\sqrt{1-\|k\|^2}}(k_1, k_2) \\
u \to p &: \left(\frac{1}{u_2}\cos\frac{u_1}{M}, \frac{1}{u_2}\sin\frac{u_2}{M}, \operatorname{arccosh}(u_2) - \sqrt{1-\frac{1}{u_2^2}}\right)
\end{aligned}
$$

Minkowski/Lorentz model is well-suited for optimization since the domain is uncontrained

**Visualizing hyperbolic Voronoi diagrams, Symp. Computational Geometry 2014**

**Representation tradeoffs for hyperbolic embeddings, ICMLR 2018**

# Outline and contributions

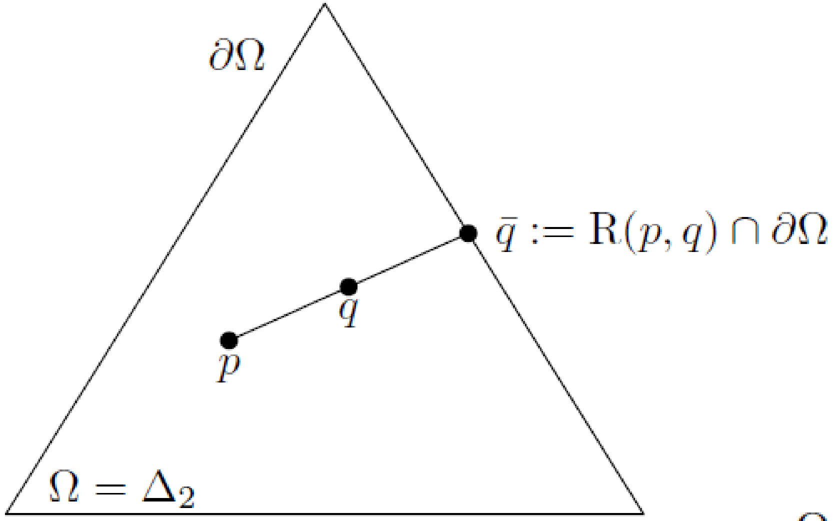
- Present **Hilbert simplex geometry** (HSG) for embeddings of graphs

Contributions:

- Simple proof of monotonicity of Hilbert distance
- Connection of Hilbert distance with Aitchison distance
- Differentiable approximation of Hilbert distance for machine learning
- Application to non-linear HSG embedding: experimentally fast, robust, and competitive

# Funk and Hilbert distance

- Consider an open bounded convex set Ω of a space (eg open standard simplex)
- Funk asymmetric distance (weak distance satisfies the triangular inequality):




$$\rho_{\mathrm{FD}}^{\Omega}(p,q) := \log \frac{\|p-\bar{q}\|}{\|q-\bar{q}\|} = \log \max_i \frac{p_i}{q_i}$$

$$\rho_{\mathrm{FD}}^{\Omega}(p,q) \neq \rho_{\mathrm{FD}}^{\Omega}(q,p)$$

Independent of the chosen norm since

$$\rho_{\mathrm{HG}}^{\Omega}(p,q) = \rho_{\mathrm{HG}}^{\Omega_{pq}}(p,q)$$

$$\Omega_{pq} := \Omega \cap (pq)$$

- Hilbert metric distance (1895) is the symmetrization of Funk distance:

$$\begin{aligned}\rho_{\mathrm{HG}}^{\Omega}(p,q) &:= \rho_{\mathrm{FD}}^{\Omega}(p,q) + \rho_{\mathrm{FD}}^{\Omega}(q,p), \\ &= \begin{cases} \log \frac{\|p-\bar{q}\|\,\|q-\bar{p}\|}{\|p-\bar{p}\|\,\|q-\bar{q}\|}, & p \neq q, \\ 0 & p = q. \end{cases} = \log \frac{\max_{i\in\{1,\ldots,d\}} \frac{p_i}{q_i}}{\min_{i\in\{1,\ldots,d\}} \frac{p_i}{q_i}}\end{aligned}$$

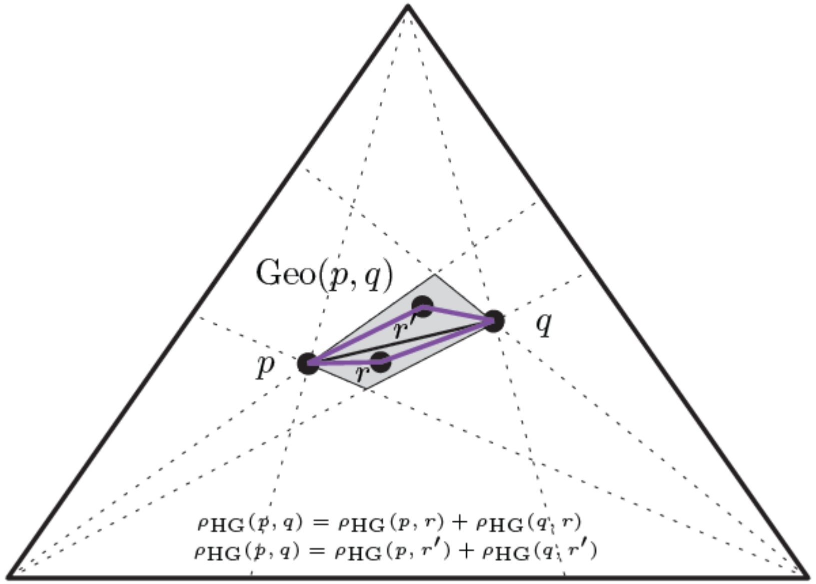
# Hilbert distance: The log cross-ratio metric

$$\rho_{\mathrm{HG}}^{\Omega}(p,q) = \begin{cases} \log \mathrm{CR}(\bar{p}, p; q, \bar{q}), & p \neq q \\ 0 & p = q \end{cases}$$

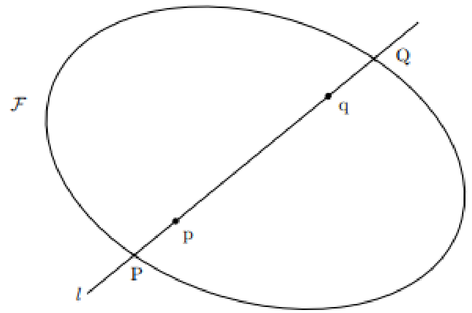
A **metric distance** which satisfies the triangle inequality:

$$\forall r \in [pq], \quad \rho_{\mathrm{HG}}^{\Omega}(p,q) = \rho_{\mathrm{HG}}^{\Omega}(p,r) + \rho_{\mathrm{HG}}^{\Omega}(r,q)$$

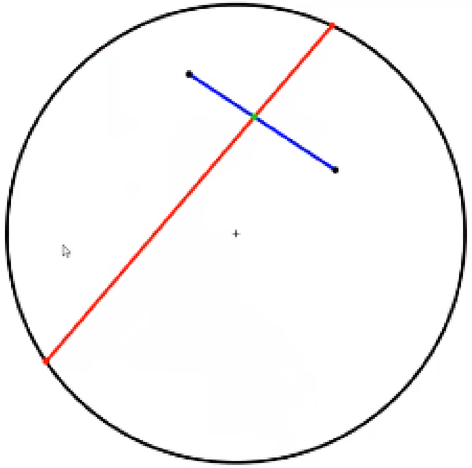
**Straight lines are geodesics** but geodesics are not unique:

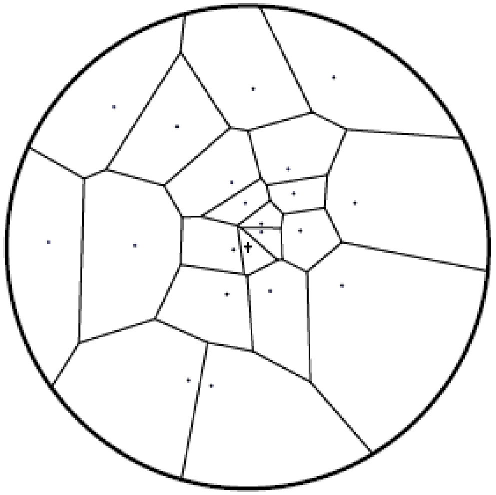


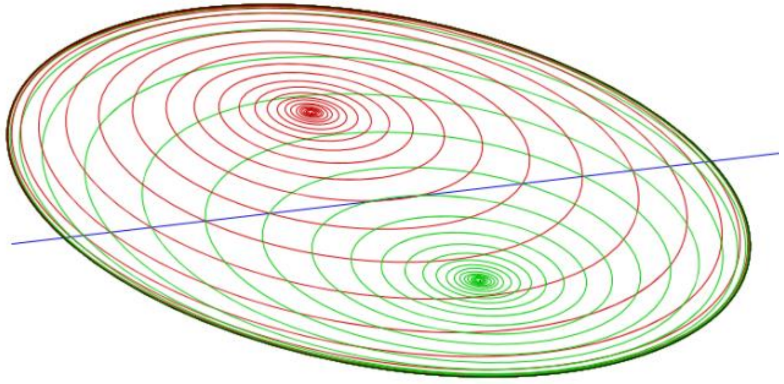
# Hilbert geometry generalizes Klein and Cayley projective geometries

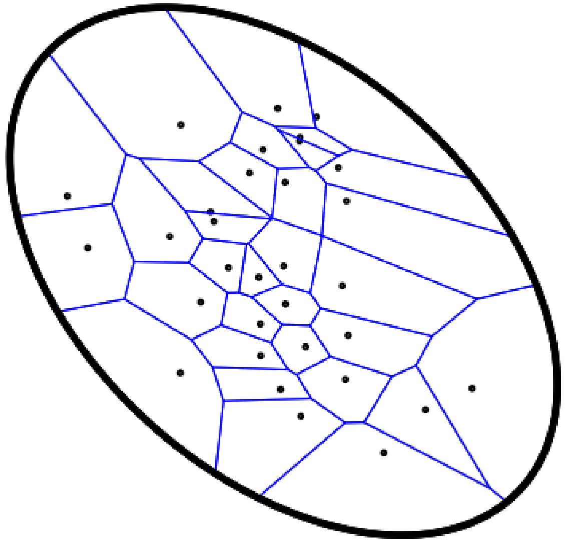



- When Ω=Open disk → Klein geometry
- When Ω=Open ellipsoid → Cayley-Klein geometry
- When Ω=Open convex with smooth boundary → hyperbolic-type Finsler geometry









Klein

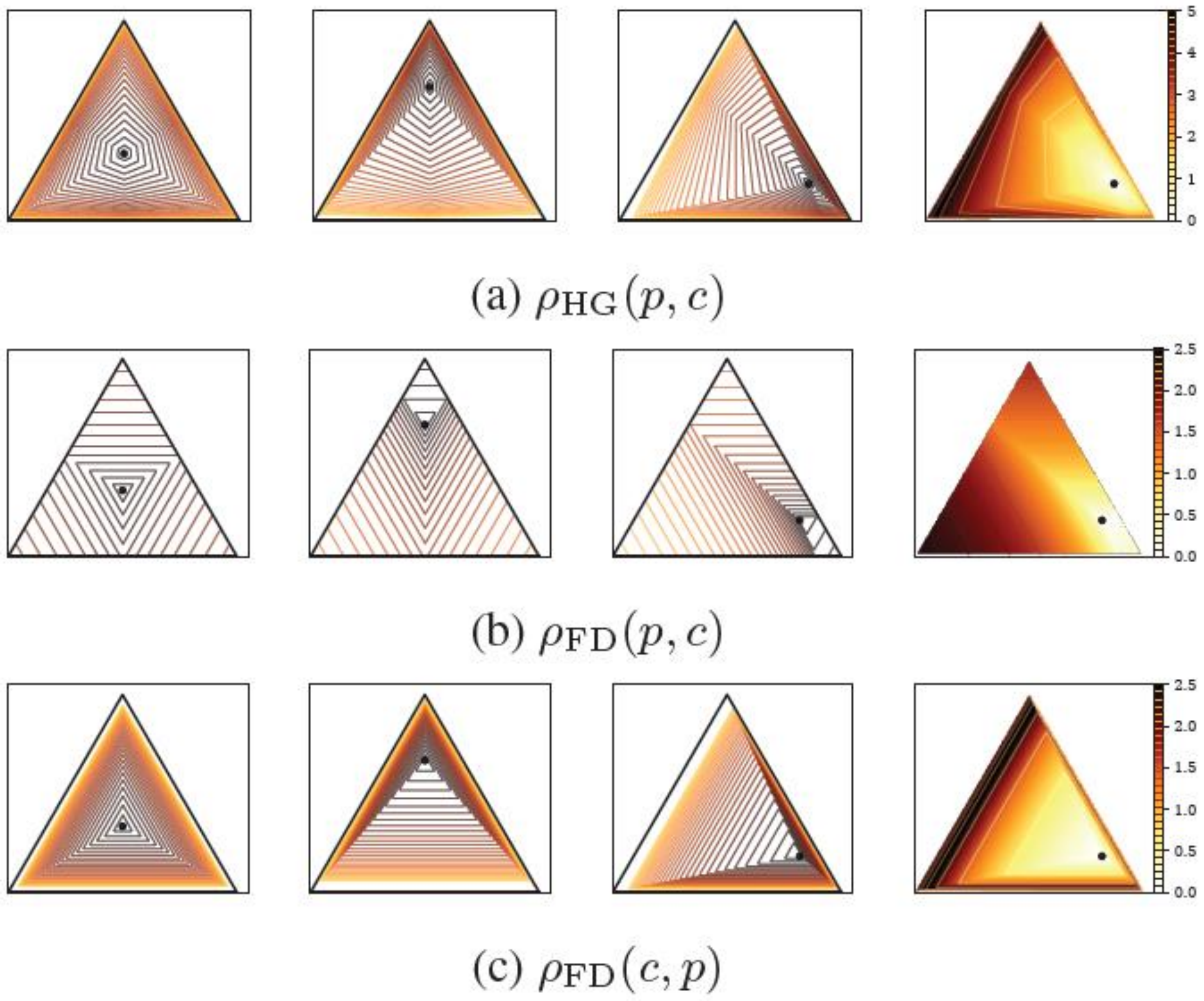
$$\operatorname{arccosh}\left(\frac{1-\langle p,q\rangle}{\sqrt{1-\langle p,p\rangle}\sqrt{1-\langle q,q\rangle}}\right)$$

Cayley-Klein

# Funk and Hilbert balls

- Hilbert balls have

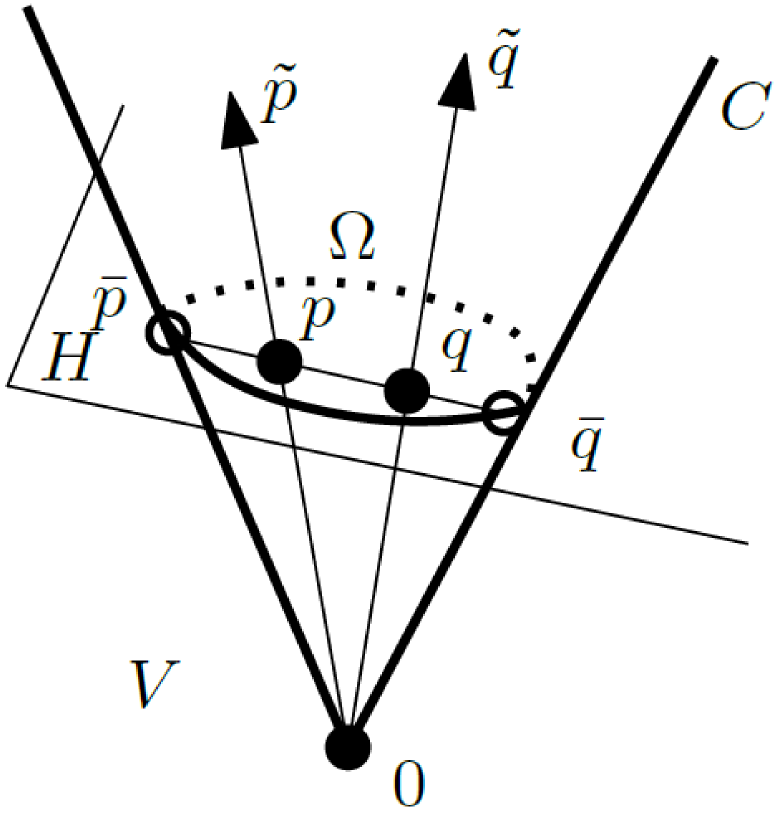
  Euclidean **hexagonal shapes**




**On balls in a Hilbert polygonal geometry, SoCG 2017**

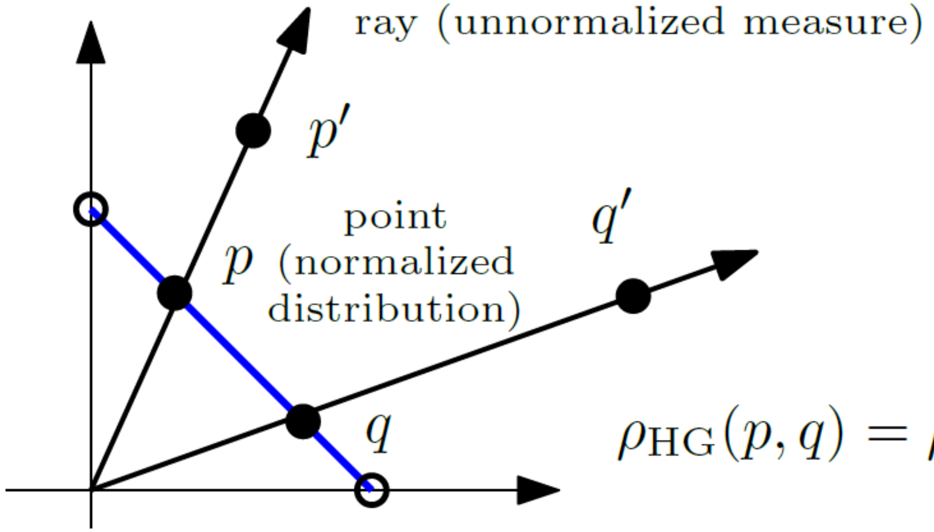
# Hilbert: Metric vs Projective distance

- Metric distance on an open bounded convex set Ω of d-dim space
- Projective distance on the pointed cone C defined by homothets λΩ in dim d
- distances between points vs pseudo-distance between rays




Ω=standard simplex

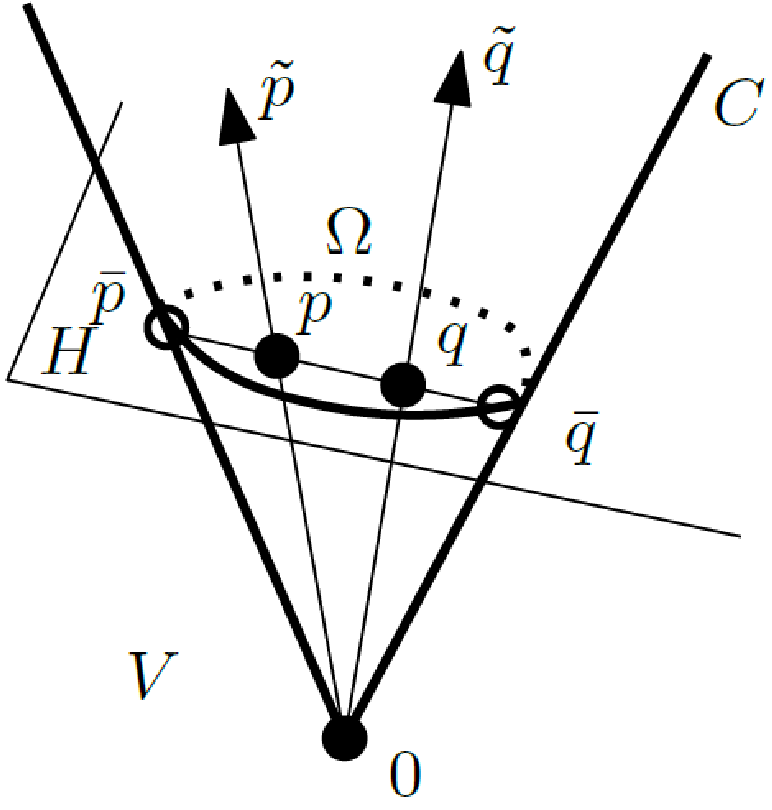
Positive orthant cone $\mathbb{R}^2$




$$\rho_{\mathrm{HG}}(p,q) = \rho_{\mathrm{HG}}(p',q')$$

$$\rho_{\mathrm{HG}}(p,q) = \log \frac{\max_{i\in\{1,\ldots,d\}} \frac{p_i}{q_i}}{\min_{i\in\{1,\ldots,d\}} \frac{p_i}{q_i}}$$

# Hilbert projective distance in a cone




Cone: $C = \{(\lambda, x) \ : \ \lambda \in \mathbb{R}_{>0}, x \in \Omega\}$

Cone defines a partial ordering: $p \preceq_C q \Leftrightarrow q - p \in C$

Projective Hilbert distance: $\rho_{\mathrm{HG}}^{C}(p,q) = \log \frac{M(p,q)}{m(p,q)}$

where $M(p,q) = \inf\{\lambda \in \mathbb{R}_{>0} \ : \ p \preceq_C \lambda q\}$

$m(p,q) = \sup\{\lambda \in \mathbb{R}_{>0} \ : \ \lambda q \preceq_C p\}$

Projective distance: $\rho_{\mathrm{HG}}^{C}(p,q) = \rho_{\mathrm{HG}}^{C}(\alpha p, \beta q), \quad \alpha, \beta > 0$

which becomes a metric distance on Ω: $\rho_{\mathrm{HG}}^{\Omega}(p,q) = \rho_{\mathrm{HG}}^{C}(p,q), \quad \forall p, q \in \Omega$

# Monotone distances

- Let T be a map. A distance is monotone or strictly monotone iff

$$\rho(T(p),T(q)) \leq \rho(p,q), \quad \forall p,q \qquad \rho(T(p),T(q)) < \rho(p,q), \quad \forall p \neq q$$

- In information geometry, the separable monotone "distances" are precisely the f-divergences when d>2.

- Theorem: Funk distance and Hilbert simplex distance are non-separable monotone distances
- Contraction theorem (Birkhoff, 1957): $\rho_{\mathrm{HG}}(Ap, Aq) \leq \tanh\left(\frac{1}{4}\Delta(A)\right) \rho_{\mathrm{HG}}(p,q)$

**Jiao, Information measures: the curious case of the binary alphabet, IEEE Transactions on Information Theory 60.12 (2014)**

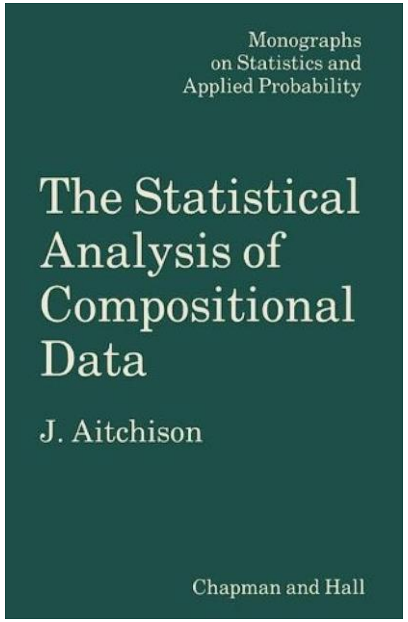
# Aitchison distance:
# Another non-separable monotone distance

- Often used in COmpositional Data Analysis (CODA)

$$\rho_{\mathrm{Aitchison}}(p,q) := \sqrt{\sum_{i=1}^{d}\left(\log\frac{p_i}{G(p)} - \log\frac{q_i}{G(q)}\right)^2}$$

where G is the geometric mean

$$G(p) = \left(\prod_{i=1}^{d} p_i\right)^{\frac{1}{d}} = \exp\left(\frac{1}{d}\sum_{i=1}^{d}\log p_i\right)$$




Erb and Ay, The information-geometric perspective of Compositional Data Analysis
Advances in Compositional Data Analysis: Festschrift in Honour of Vera Pawlowsky-Glahn, 2021.

# HSG: Isometry to a vector space (Normed Hilbert)

- The only Hilbert geometries isometric to vector spaces are obtained with open simplex domains: $(\Delta_d, \rho_{\mathrm{HG}}) \cong (V_d, \|\cdot\|_{\mathrm{NH}})$
- Slanted hyperplane: $V_d = \{v \in \mathbb{R}^d \ : \ \sum_{i=1}^{d} v_i = 1\} \subset \mathbb{R}^d$
- Isometric mapping: $v : p = (p_1, \ldots, p_d) \in \Delta_d \rightarrow v(p) = (v_1, \ldots, v_d) \in V_d$

$$v_i = \frac{1}{d}\left((d-1)\log p_i - \sum_{j \neq i} \log p_j\right) = \log p_i - \frac{1}{d}\sum_{j=1}^{d} \log \ p_j = \log \frac{p_i}{G(p)}$$

- Norm/unit ball in vector space:

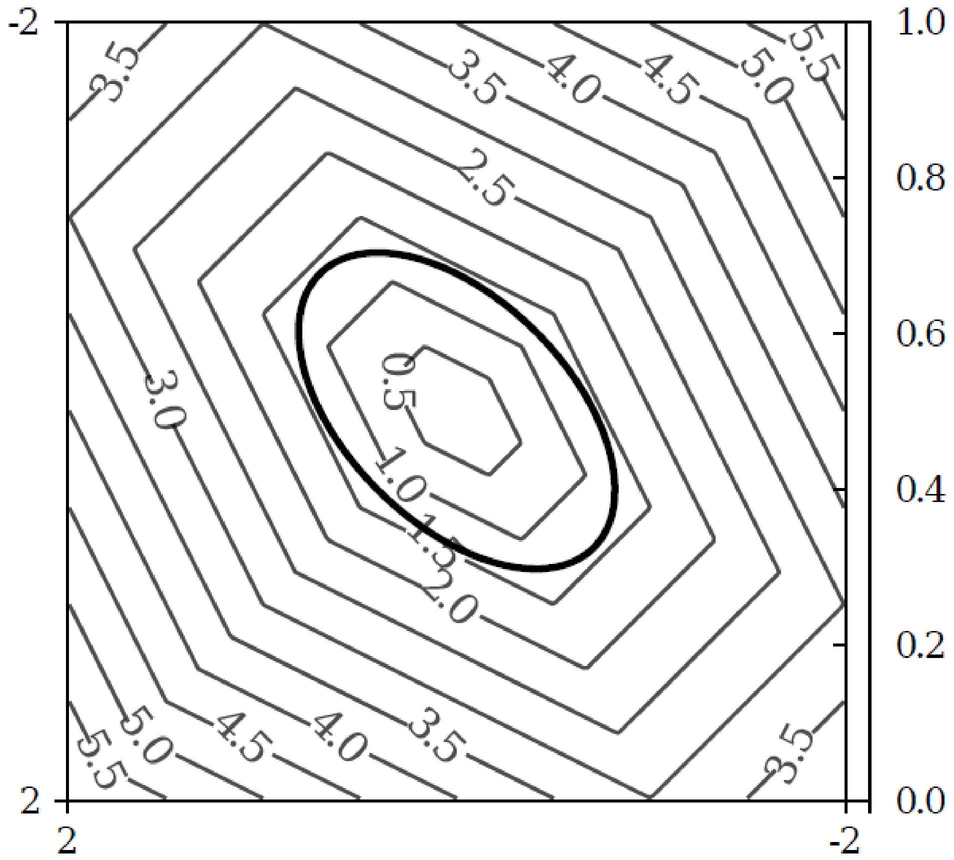
$$B_V = \{v \in V_d \ : \ |v_i - v_j| \leq 1, \forall i \neq j\} = \{v \in V_d \ : \ \|v\|_{\mathrm{NH}} \leq 1\}$$

- Distance: $\rho_V(v, v') = \|v - v'\|_{\mathrm{NH}} = \inf\{\tau \in \mathbb{R}_{>0} \ : \ v' \in \tau(B_V \oplus \{v\})\}$

Minkowski sum of sets: $A \oplus B = \{a + b \ : \ a \in A, b \in B\}$

Foertsch and Karlsson, Hilbert metrics and Minkowski norms, Journal of Geometry 83.1-2 (2005)

# Symmetric polytope norm NH



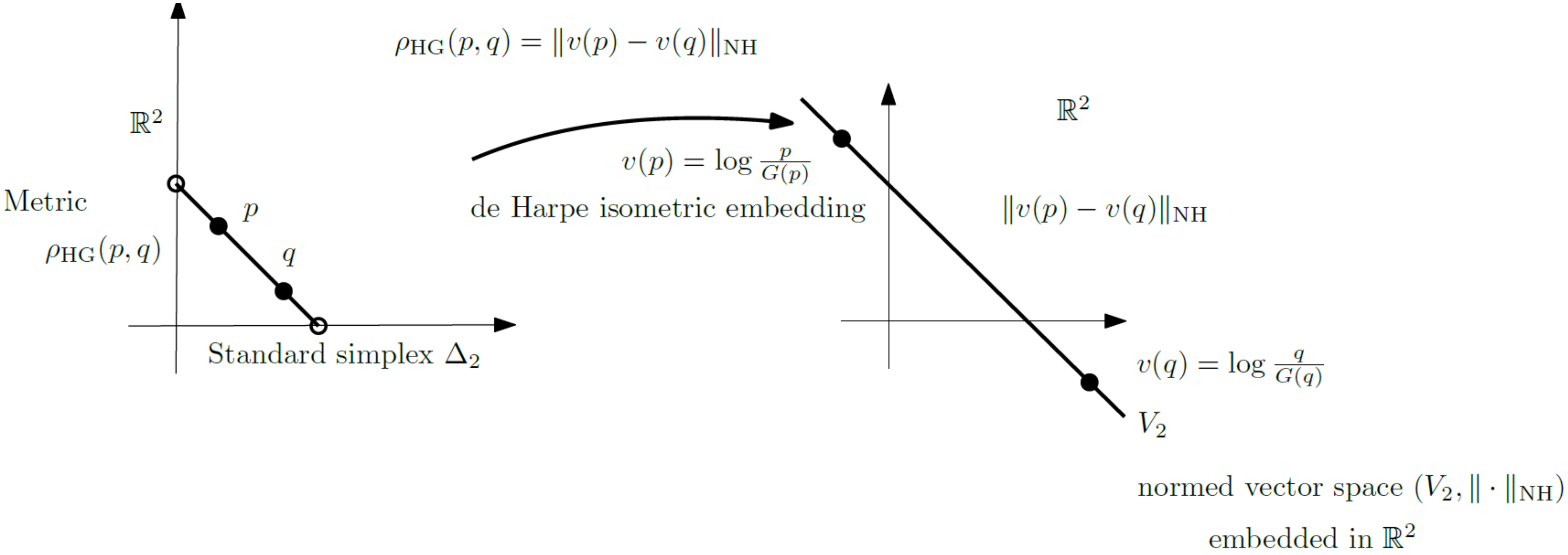

Visualized on the slanted plane $V_2$:

$$V_d = \{v \in \mathbb{R}^d \;:\; \sum_{i=1}^{d} v_i = 1\} \subset \mathbb{R}^d$$

$$B_V = \{v \in V_d \;:\; |v_i - v_j| \leq 1, \forall i \neq j\} = \{v \in V_d \;:\; \|v\|_{\mathrm{NH}} \leq 1\}$$

**de la Harpe, P. On Hilbert's metric for simplices.**
**In Geometric Group Theory, volume 1, 1991**

# HSG isometry to a normed space




$$v_i = \frac{1}{d}\left((d-1)\log p_i - \sum_{j\neq i} \log p_j\right) = \log p_i - \frac{1}{d}\sum_{j=1}^{d} \log\ p_j = \log \frac{p_i}{G(p)}$$

# Logarithmic mapping and variation semi-norm

Logarithmic representation: $l(p) = (\log p_1, \ldots, \log p_d) \in \mathbb{R}^d$

Hilbert distance as normed distance: $\rho_{\mathrm{HG}}(p,q) = \|v(p) - v(q)\|_{\mathrm{NH}} = \left\|\log \frac{p}{G(p)} - \log \frac{q}{G(q)}\right\|_{\mathrm{NH}}$

Hilbert projective distance as semi-normed variation distance:

$$
\begin{aligned}
\rho_{\mathrm{HG}}(p,q) &= \max_i \log \frac{p_i}{q_i} - \min_i \log \frac{p_i}{q_i} \\
&= \left\|\log \frac{p}{q}\right\|_{\mathrm{var}} \\
&= \|l(p) - l(q)\|_{\mathrm{var}} \\
&= \|l(\alpha p) - l(\beta q)\|_{\mathrm{var}}, \forall \alpha, \beta \\
&= \|l(\alpha p) - l(\beta q)\|_{\mathrm{var}}, \alpha = \frac{1}{G(p)}, \beta = \frac{1}{G(q)} \\
&= \left\|\log \frac{p}{G(p)} - \log \frac{q}{G(q)}\right\|_{\mathrm{var}}
\end{aligned}
$$

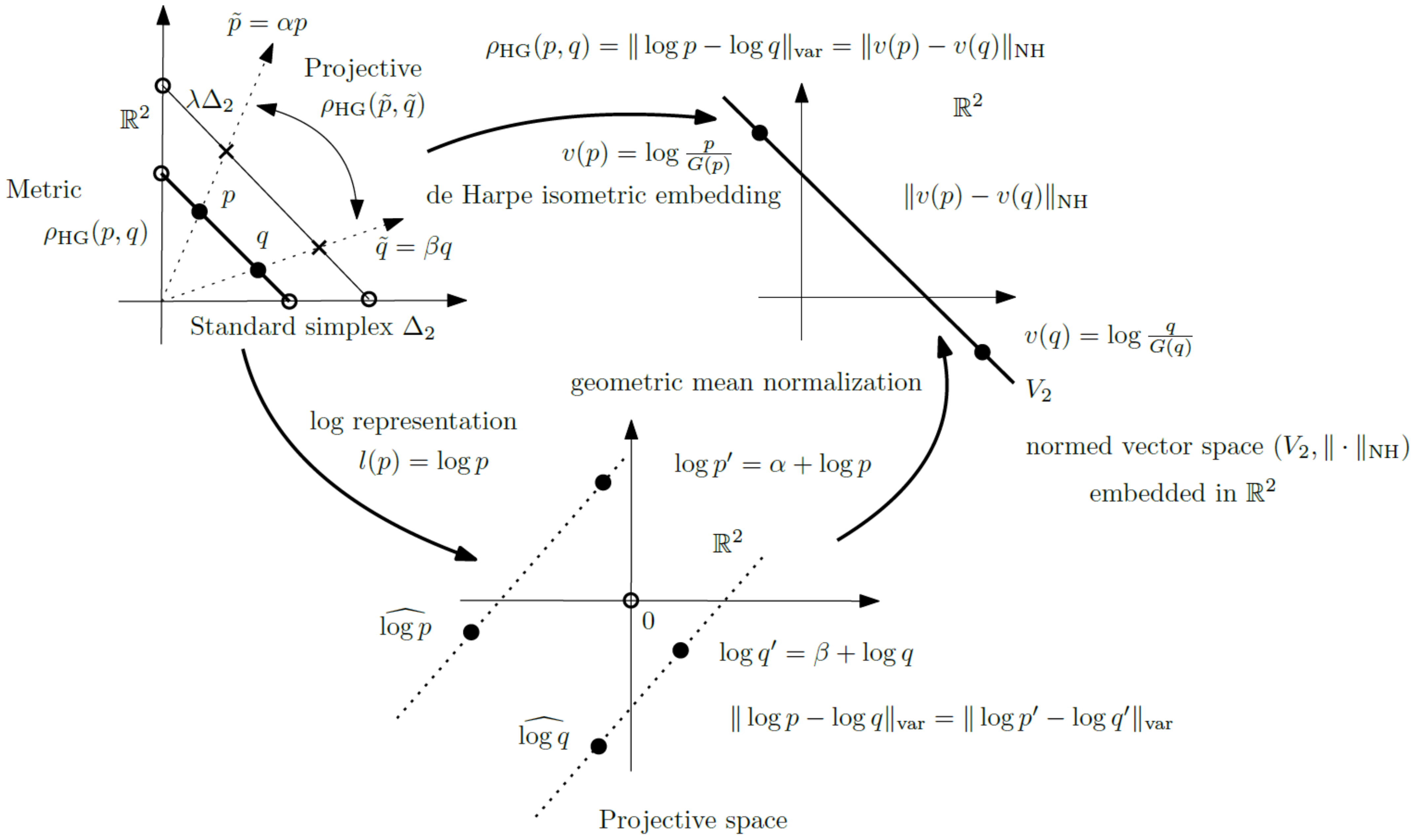
(log monotone increasing: log max=max log, log min=min log)

variation semi-norm:

$$\|x\|_{\mathrm{var}} := \max_i x_i - \min_i x_i = \|x\|_{+\infty} - \|x\|_{-\infty}$$

$$\|(\lambda, \ldots, \lambda)\|_{\mathrm{var}} = 0, \quad \forall \lambda \in \mathbb{R}$$

# Metric/projective Hilbert simplex/pos. orthant distance



# Relationship between HSG and Aitchison

Logarithm mapping normalized by homogeneous geometric mean:

$$\forall \lambda > 0, \quad \log \frac{p}{G(p)} = \log \frac{\lambda p}{G(\lambda p)}$$

Hilbert distance:

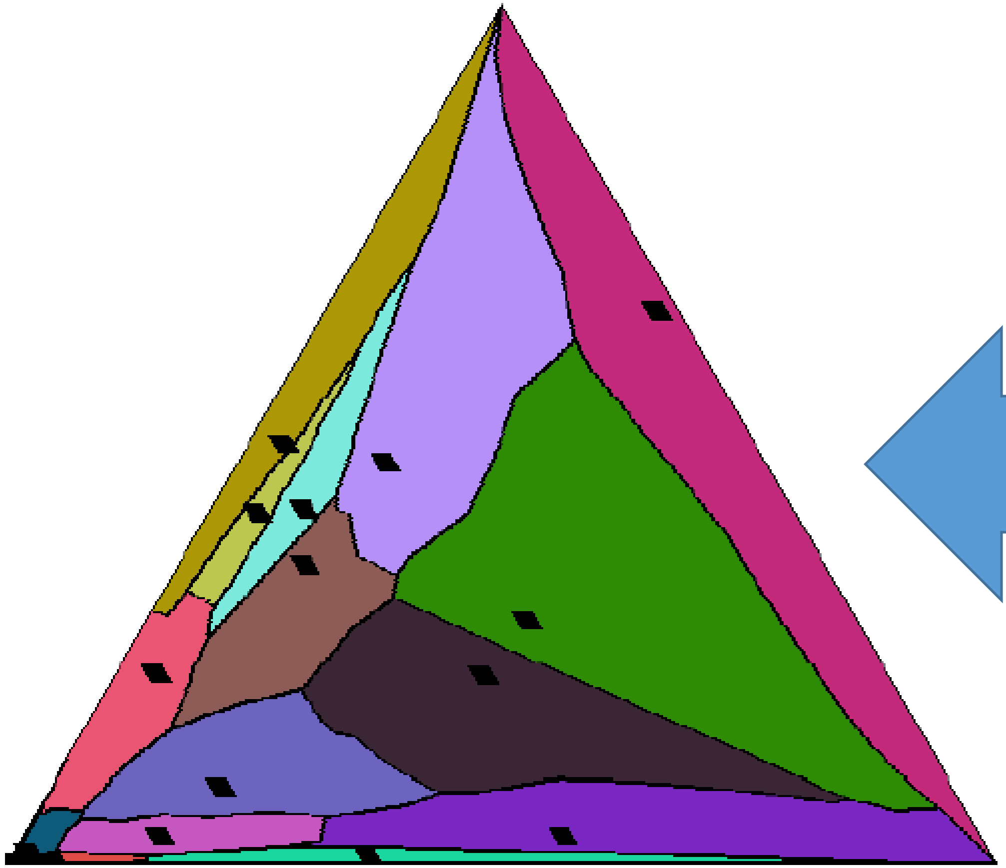
$$\begin{aligned} \rho_{\mathrm{HG}}(p,q) &= \|v(p) - v(q)\|_{\mathrm{NH}} \\ &= \left\| \log \frac{p}{G(p)} - \log \frac{q}{G(q)} \right\|_{\mathrm{NH}} \\ &= \left\| \log \frac{p}{G(p)} - \log \frac{q}{G(q)} \right\|_{\mathrm{var}} \end{aligned}$$

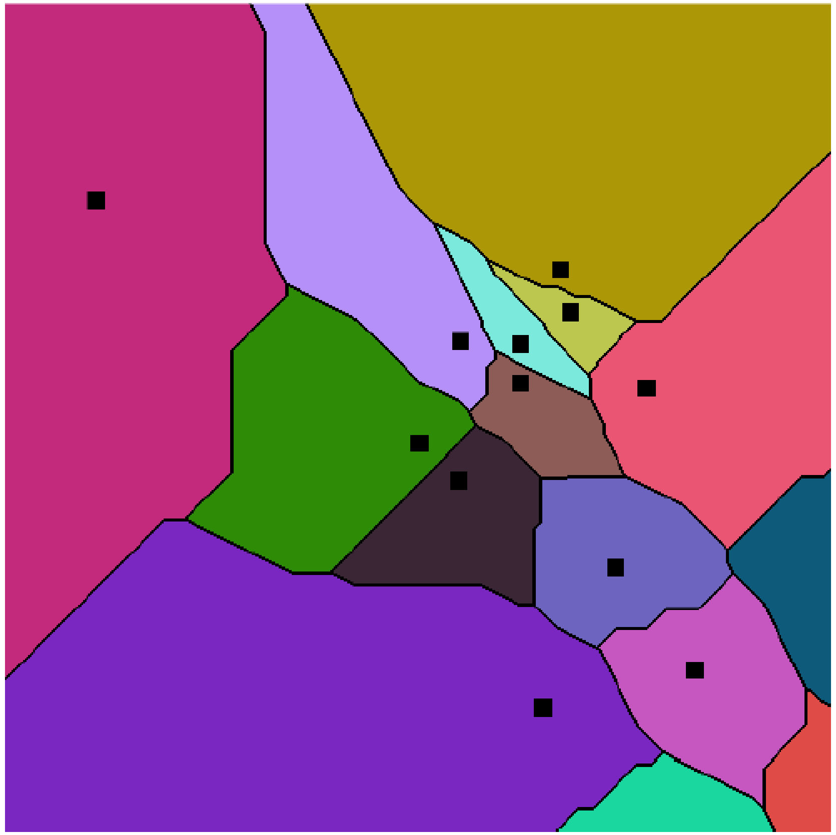
Aitchison distance:

$$\rho_{\mathrm{Aitchison}}(p,q) = \left\| \log \frac{p}{G(p)} - \log \frac{q}{G(q)} \right\|_2$$

$$\rho_{\mathrm{Aitchison}}(p,q) := \sqrt{\sum_{i=1}^{d} \left( \log \frac{p_i}{G(p)} - \log \frac{q_i}{G(q)} \right)^2}$$

# Hilbert simplex Voronoi diagrams





$$\rho_{\mathrm{HG}}(p,q) = \log \frac{\max_{i\in\{1,\ldots,d\}} \frac{p_i}{q_i}}{\min_{i\in\{1,\ldots,d\}} \frac{p_i}{q_i}}$$

$$\begin{aligned}\rho_{\mathrm{HG}}(p,q) &= \|v(p)-v(q)\|_{\mathrm{NH}}\\ &= \left\|\log\frac{p}{G(p)} - \log\frac{q}{G(q)}\right\|_{\mathrm{NH}}\\ &= \left\|\log\frac{p}{G(p)} - \log\frac{q}{G(q)}\right\|_{\mathrm{var}}\end{aligned}$$

# Differentiable approximation of Hilbert simplex distance

- max and min operations in Hilbert simplex distance are not differentiable
- **log-sum-exp** (LSE) commonly used in ML to approximate max operator

- We approximate Hilbert simplex distance by **differentiable function**:

$$\tilde{\rho}_{\mathrm{LSE}^T}(p,q) = \frac{1}{T}\log\left(\sum_i \left(\frac{p_i}{q_i}\right)^T\right)\left(\sum_i \left(\frac{q_i}{p_i}\right)^T\right)$$

with guarantees

$$\rho_{\mathrm{HG}}(p,q) + 2\epsilon_1(r,T) \leq \tilde{\rho}_{\mathrm{LSE}^T}(p,q) \leq \rho_{\mathrm{HG}}(p,q) + 2\epsilon_2(r,T)$$

where

$$\varepsilon_1(x,T) := \frac{1}{T}\log\left[1 + (d-1)\exp(-T\|x\|_{\mathrm{var}})\right]$$

$$\varepsilon_2(x,T) := \frac{1}{T}\log\left[d - 1 + \exp(-T\|x\|_{\mathrm{var}})\right]$$

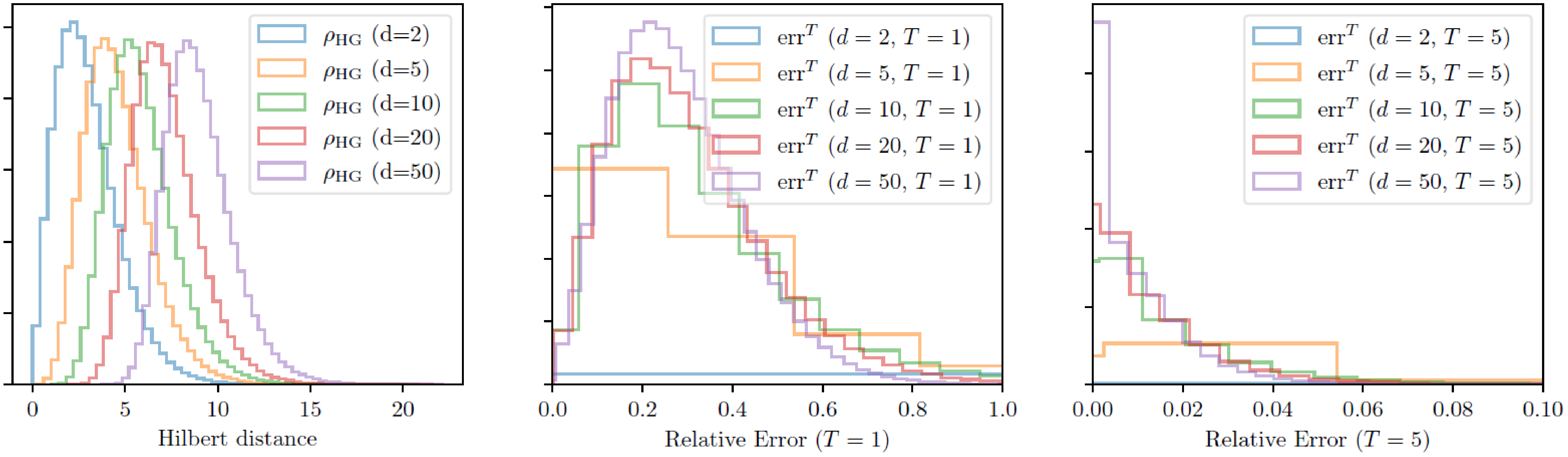
The larger T, the better the approximation: $0 < \varepsilon_1(x,T) \leq \varepsilon_2(x,T) \leq \frac{1}{T}\log d$

**Guaranteed bounds on information-theoretic measures of univariate mixtures using piecewise log-sum-exp inequalities.** ***Entropy*** **18.12 (2016)**

# Differentiable approximation: Experiments

$10^6$ pairs of points randomly sampled inside the d-dimensional standard simplex

We measure: $$\mathrm{err}^T(p,q) := \frac{\tilde{\rho}_{\mathrm{LSE}^T}(p,q) - \rho_{\mathrm{HG}}(p,q)}{\rho_{\mathrm{HG}}(p,q)}$$



# Non-linear embeddings: Evaluation metrics

Remark: When $(M_1,\rho_1)$ isometric to $(M_2,\rho_2)$ $\leftrightarrow$ same representation power

Loss associated to a **distance matrix** $[D_{ij}]$ (e.g., calculated from weighted graphs):

$$\ell(\mathcal{D}, \mathcal{M}^d) := \inf_{\boldsymbol{Y} \in (\mathcal{M}^d)^n} \frac{1}{n^2} \sum_{i=1}^{n} \sum_{j=1}^{n} \left(\mathcal{D}_{ij} - \rho_{\mathcal{M}}(\boldsymbol{y}_i, \boldsymbol{y}_j)\right)^2$$

or use **row-stochastic probability matrix** $[P_{ij}]$ (loss in manifold learning/graph embedding):

$$\ell(\mathcal{P}, \mathcal{M}^d) := \inf_{\boldsymbol{Y} \in (\mathcal{M}^d)^n} \frac{1}{n} \sum_{i=1}^{n} \sum_{j:j \neq i} \mathcal{P}_{ij} \log \frac{\mathcal{P}_{ij}}{q_{ij}(\boldsymbol{Y})}$$

empirical average of the KL divergence between pmf $P_i$ and $q_i$

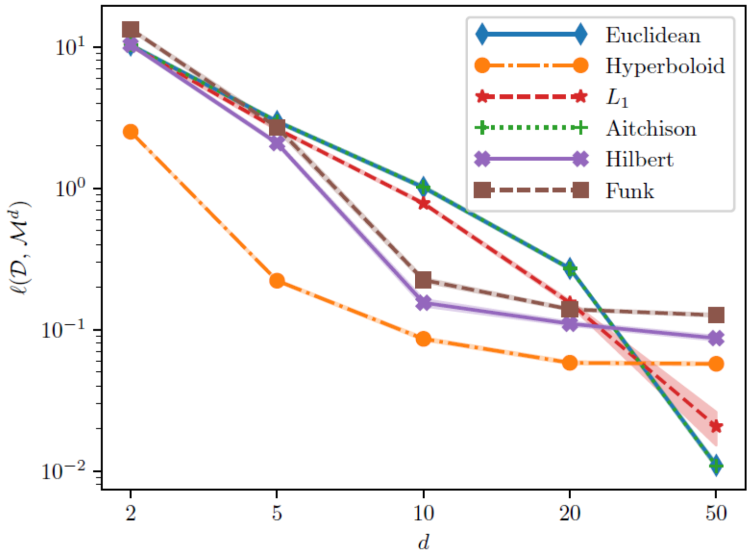
$$q_{ij}(\boldsymbol{Y}) := \frac{\exp(-\rho^2_{\mathcal{M}}(\boldsymbol{y}_i, \boldsymbol{y}_j))}{\sum_{j:j \neq i} \exp(-\rho^2_{\mathcal{M}}(\boldsymbol{y}_i, \boldsymbol{y}_j))}$$

heat kernel

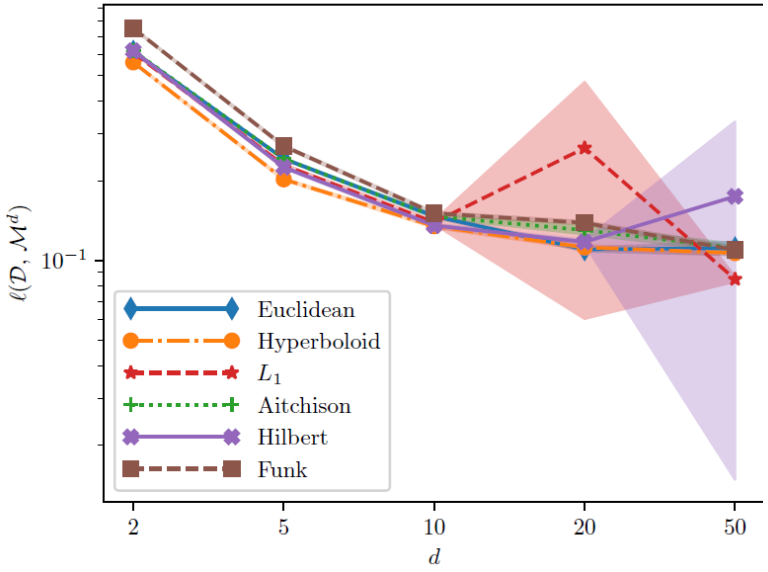
# Non-linear embeddings: Results (MSE)

Use Adam local optimizer [Kingma & Ba, 2015]

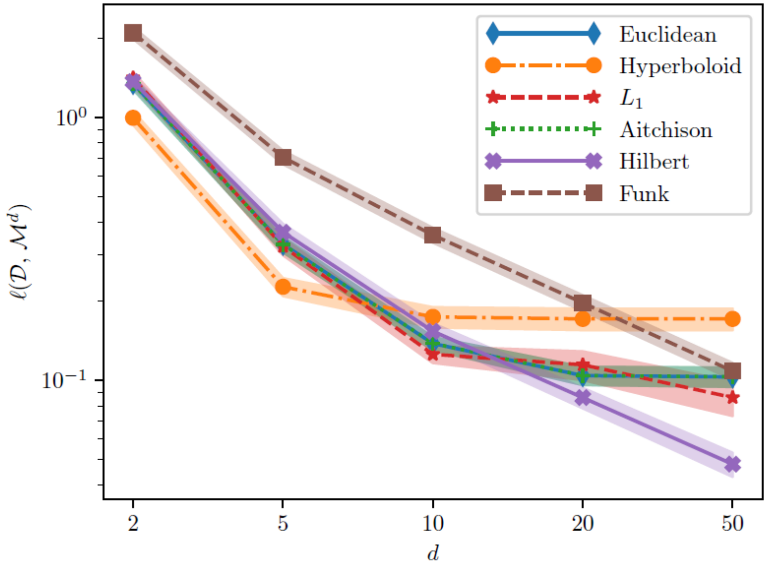
Repeat 10 different instances to get standard deviation shown in color bands




(a) 100 random points in $\mathbb{R}^{100}$




(b) Erdős–Rényi graphs $G(n, p)$ $(n = 200, p = 0.2)$



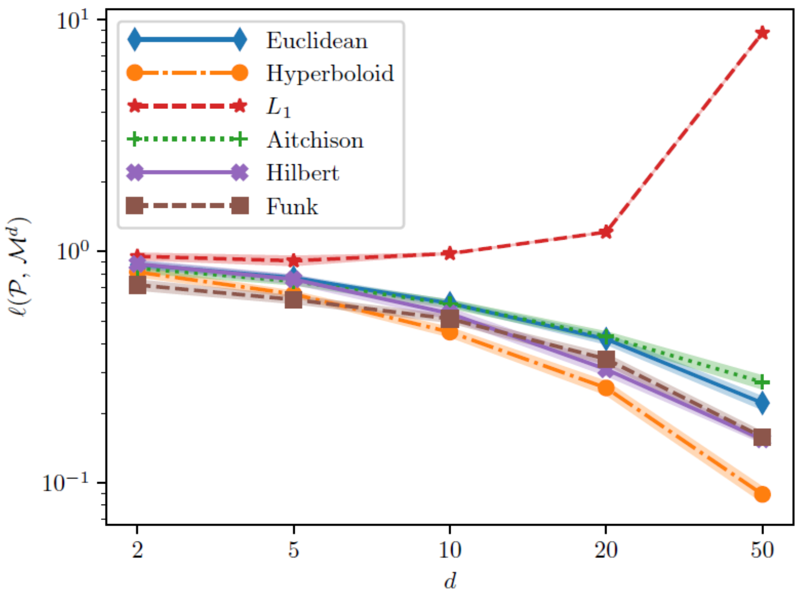

(c) Barabási–Albert graphs $G(n, m)$ $(n = 200, m = 2)$

Hilbert simplex geometry
is the winner
for discrete graphs

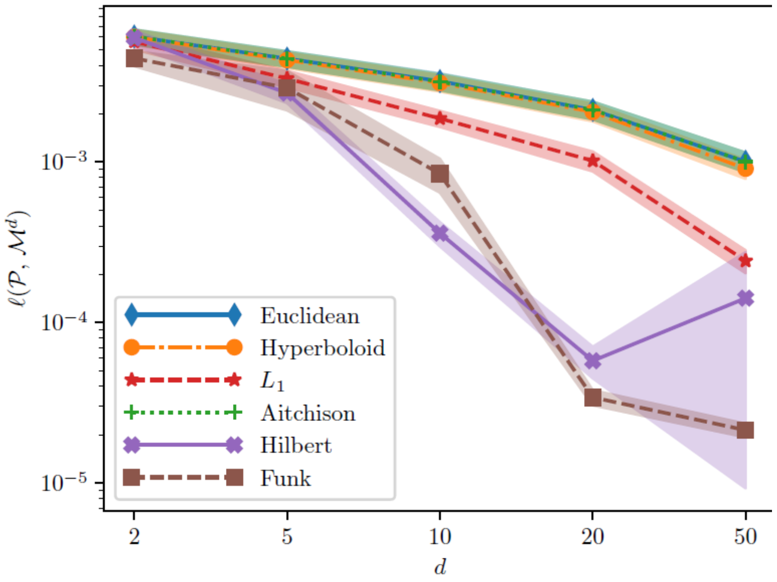
The larger the embedding dimension, the better!

**Hilbert & hyperbolic hyberboloid geometries experimentally performed best**

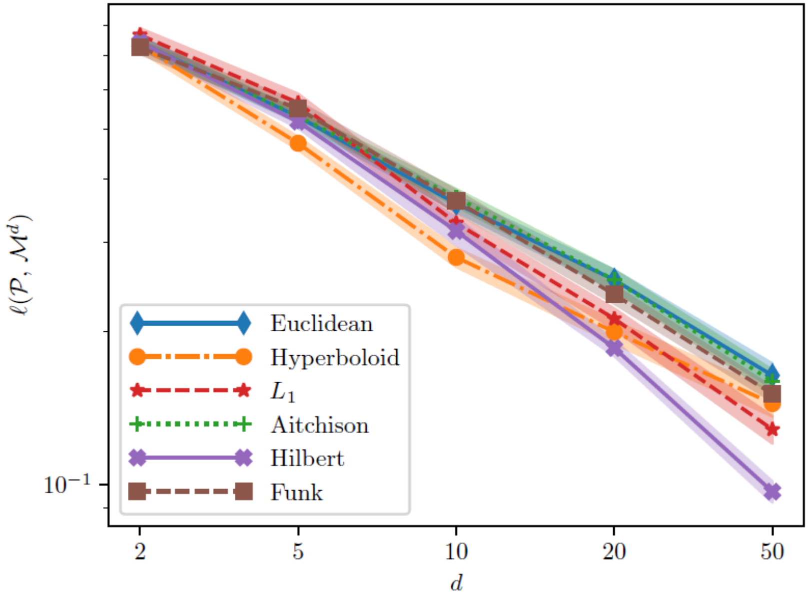
# Non-linear embeddings: Results (empirical KLD)



(a) 100 random points in $\mathbb{R}^{100}$



(b) Erdős–Rényi graphs $G(n, p)$ ($n = 200$, $p = 0.2$)



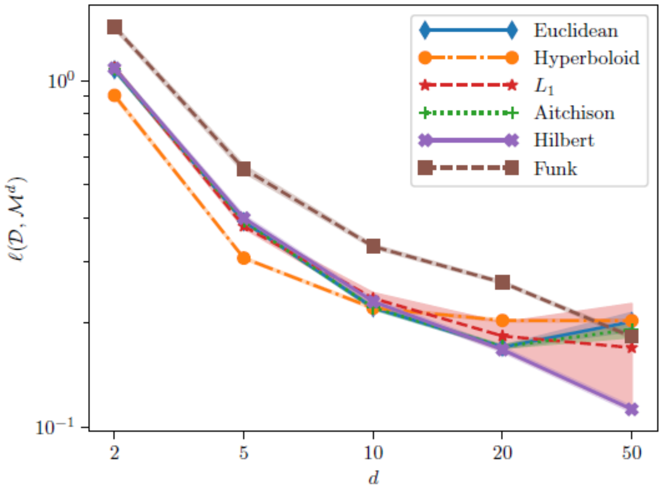
(c) Barabási–Albert graphs $G(n, m)$ ($n = 200$, $m = 2$)

hyperbolic geometry
is the winner for continuous data

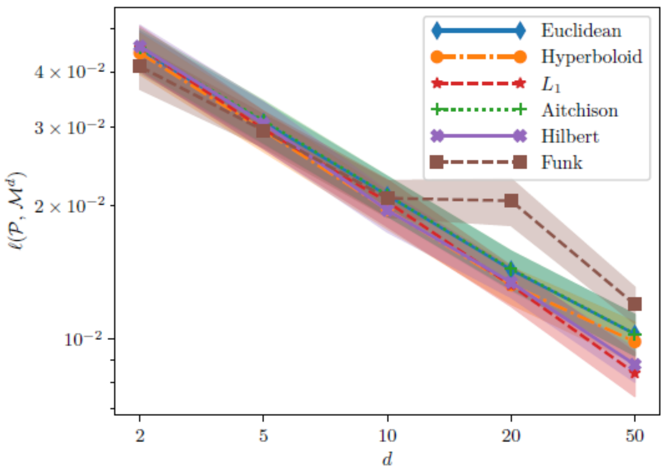
Funk geometry also
good for embedding!

Hilbert simplex geometry
is the winner
for discrete graphs

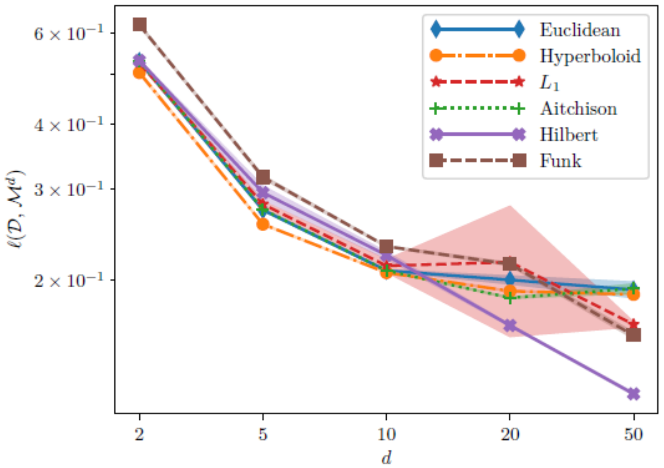
# Non-linear embeddings: Comparative results



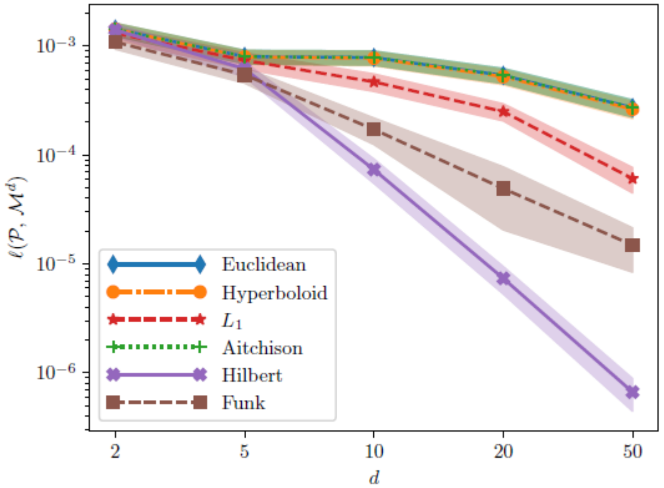

(a) $\ell(\mathcal{D}, \mathcal{M}^d)$ $(n = 200, p = 0.05)$




(b) $\ell(\mathcal{P}, \mathcal{M}^d)$ $(n = 200, p = 0.05)$




(c) $\ell(\mathcal{D}, \mathcal{M}^d)$ $(n = 200, p = 0.5)$




(d) $\ell(\mathcal{P}, \mathcal{M}^d)$ $(n = 200, p = 0.5)$

Embedding losses against $d$ (Erdős–Rényi random graph $G(n, p)$).

# Summary

- Proposed differentiable approximation of Hilbert distance
- Presented Hilbert simplex geometry for graph embeddings via distance matrices using Adam optimizer
- Results for non-linear embeddings: Hilbert/Funk simplex geometry is experimentally fast, robust, and competitive compared to $L_1$, $L_2$, Aitchison and hyperbolic hyperboloid embeddings

- Proved the monotonicity of Funk and Hilbert distances
- Shown a connection between Hilbert distance and Aitchison distance via

the normalized logarithmic representation of standard simplex points

# Erdős–Rényi random graph datasets G(n,p)

Graph with n nodes constructed by connecting nodes randomly
An edge Eij is included in the graph with probability p
(independently from other edges)

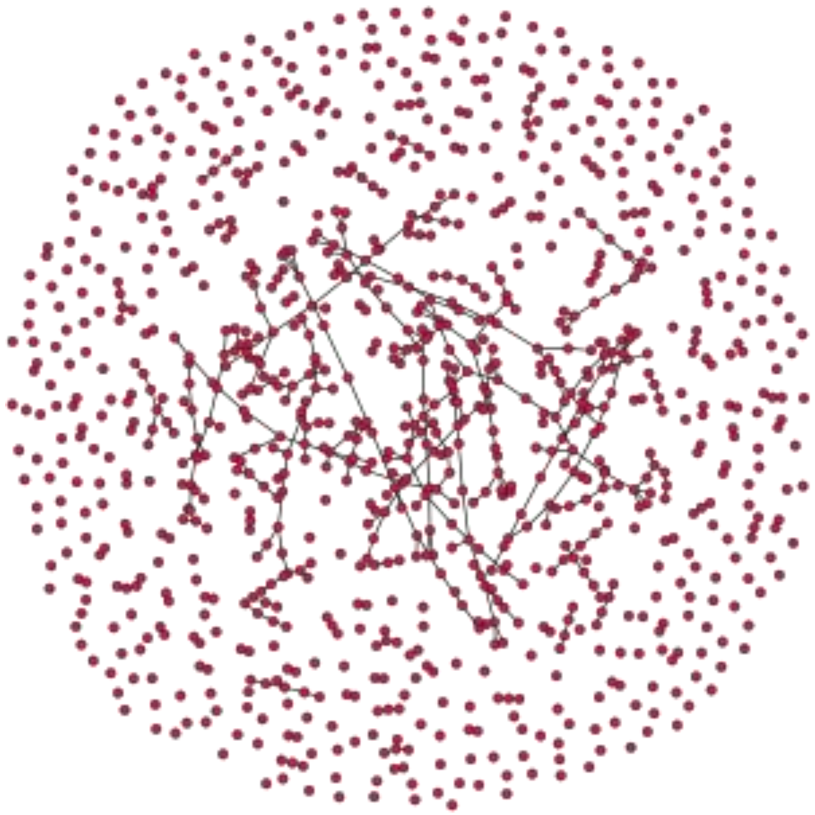


Image courtesy of Wikipedia

Caption:
"An Erdős–Rényi–Gilbert graph with 1000 vertices at the critical edge probability p=1/(n−1) showing a large component and many small ones"

# Barabási–Albert random graph datasets G(n,m)

- Generating random scale-free networks  with power-law degree distributions
- Preferential attachment m:

  the more connected a node is, the more likely it is to receive new links.
- Begins with an initial connected network of $m_0$ vertices.
- Add vertices incrementally: A new vertex v is connecto to already existing vertex $v_i$ with probability $deg(v_i)/\Sigma_j deg(v_j)$

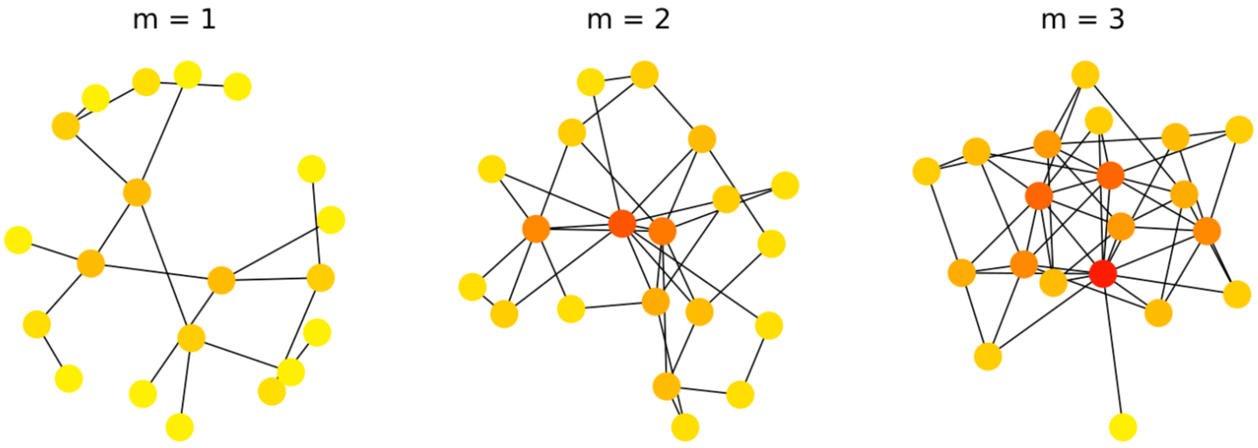



Image courtesy of Wikipedia

Caption:

"Display of three graphs generated with the Barabasi-Albert (BA) model. Each has 20 nodes and a parameter of attachment m as specified. The color of each node is dependent upon its degree (same scale for each graph)."

# References


- **Clustering in Hilbert's projective geometry: The case studies of the probability simplex and the elliptope of correlation matrices**, *Geometric structures of information (2019)*

- **Non-linear Embeddings in Hilbert Simplex Geometry**, *Topology, Algebra, and Geometry in Machine Learning Workshop (ICML TAGML'23), arXiv:2203.11434 (2022)*

- *Home page:* **https://franknielsen.github.io/HSG/**

# Non-linear Embeddings in Hilbert Simplex Geometry

- Hilbert distance on the simplex:

$$\rho_{\mathrm{HG}}(p,q) = \log \frac{\max_{i\in\{1,\ldots,d\}} \frac{p_i}{q_i}}{\min_{i\in\{1,\ldots,d\}} \frac{p_i}{q_i}}$$

HD is projective distance on the positive orthant cone.

- Differentiable approximation of the Hilbert distance:

$$\tilde{\rho}_{\mathrm{LSE}^T}(p,q) = \frac{1}{T}\log\left(\sum_i \left(\frac{p_i}{q_i}\right)^T\right)\left(\sum_i \left(\frac{q_i}{p_i}\right)^T\right)$$

- Loss functions for embedding distance matrices:

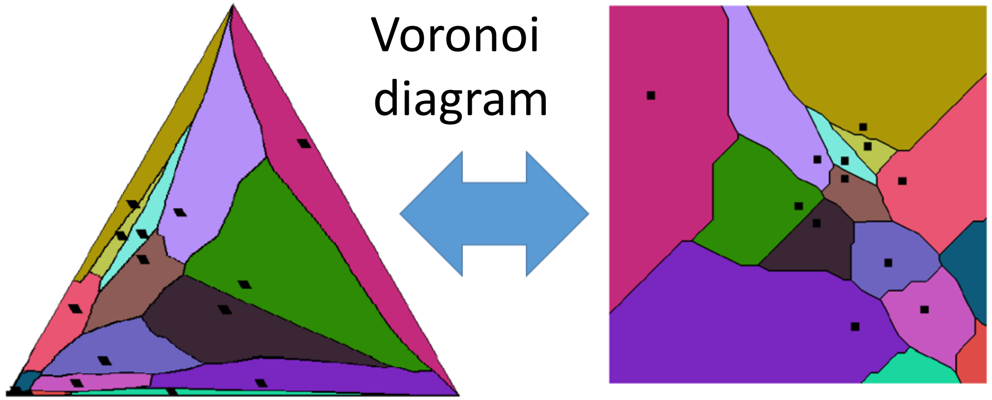
$$\ell(\mathcal{D},\mathcal{M}^d) := \inf_{\boldsymbol{Y}\in(\mathcal{M}^d)^n} \frac{1}{n^2}\sum_{i=1}^n\sum_{j=1}^n \left(\mathcal{D}_{ij} - \rho_{\mathcal{M}}(\boldsymbol{y}_i,\boldsymbol{y}_j)\right)^2$$

or empirical average Kullback-Leibler divergence:

$$\ell(\mathcal{P},\mathcal{M}^d) := \inf_{\boldsymbol{Y}\in(\mathcal{M}^d)^n} \frac{1}{n}\sum_{i=1}^n\sum_{j:j\neq i} \mathcal{P}_{ij}\log\frac{\mathcal{P}_{ij}}{q_{ij}(\boldsymbol{Y})}$$

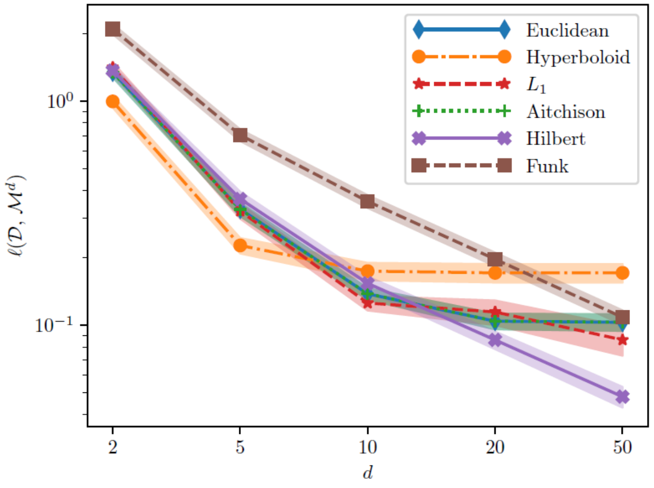
$$q_{ij}(\boldsymbol{Y}) := \frac{\exp(-\rho^2_{\mathcal{M}}(\boldsymbol{y}_i,\boldsymbol{y}_j))}{\sum_{j:j\neq i}\exp(-\rho^2_{\mathcal{M}}(\boldsymbol{y}_i,\boldsymbol{y}_j))}$$

- Hilbert simplex geometry isometric to normed vector space:

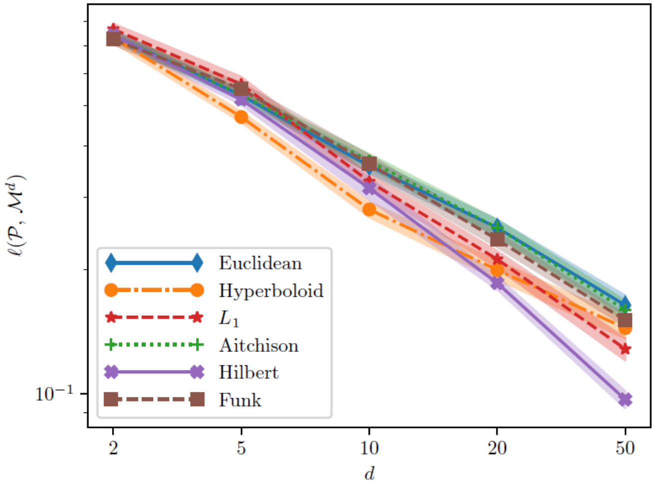



$$\rho_{\mathrm{HG}}(p,q) = \log \frac{\max_{i\in\{1,\ldots,d\}} \frac{p_i}{q_i}}{\min_{i\in\{1,\ldots,d\}} \frac{p_i}{q_i}}$$

$$\begin{aligned}\rho_{\mathrm{HG}}(p,q) &= \|v(p)-v(q)\|_{\mathrm{NH}} \\ &= \left\|\log\frac{p}{G(p)} - \log\frac{q}{G(q)}\right\|_{\mathrm{NH}} \\ &= \left\|\log\frac{p}{G(p)} - \log\frac{q}{G(q)}\right\|_{\mathrm{var}}\end{aligned}$$




(c) Barabási–Albert graphs $G(n,m)$ ($n=200$, $m=2$)




(c) Barabási–Albert graphs $G(n,m)$ ($n=200$, $m=2$)