Xavier Pennec

Univ. Côte d'Azur and Inria, France

Freely adapted from "Women teaching geometry", in Adelard of Bath translation of Euclid's elements, 1310.

Geometric Statistics

Mathematical foundations and applications in computational anatomy

3/ Advanced Stats: empirical estimation and generalized PCA

ESI semester Infinite-dimensional Geometry: Theory and Applications Week 5, 02/2025

Geometric Statistics: Mathematical foundations and applications in computational anatomy

Intrinsic Statistics on Riemannian Manifolds

Metric and Affine Geometric Settings for Lie Groups

Advances Statistics: CLT & PCA

- Estimation of the empirical Fréchet mean & CLT
- Principal component analysis in manifolds
- Natural subspaces in manifolds: barycentric subspaces
- Rephrasing PCA with flags of subspaces

Several definitions of the mean

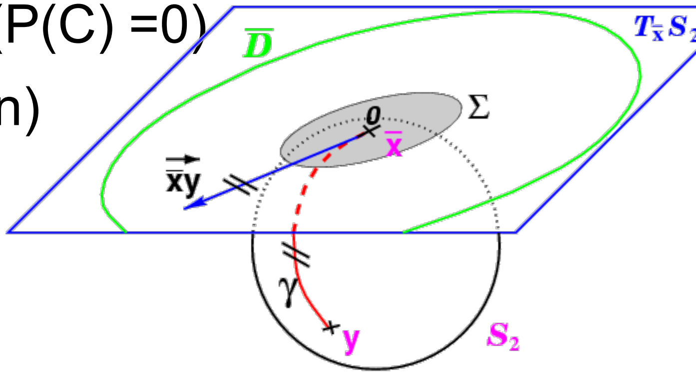
Tensor moments of a random point on M

- $\mathfrak{M}_1(x) = \int_M \overrightarrow{xz}\, dP(z)$ Tangent mean: (0,1) tensor field
- $\mathfrak{M}_2(x) = \int_M \overrightarrow{xz} \otimes \overrightarrow{xz}\, dP(z)$ 2nd moment: (0,2) tensor field
- $\mathfrak{M}_k(x) = \int_M \overrightarrow{xz} \otimes \overrightarrow{xz} \otimes \cdots \otimes \overrightarrow{xz}\, dP(z)$ k-contravariant tensor field
- $\sigma^2(x) = Tr_g\big(\mathfrak{M}_2(x)\big) = \int_M dist^2(x,z)\, dP(z)$ **Mean quadratic deviation**

Mean value = optimum of the variance

- **Frechet mean** [1944] = (global) minima of p-deviation (includes median)
- **Karcher mean** [1977] = local minima
- **Exponential barycenters** = critical points (P(C) =0)

  $\mathfrak{M}_1(\bar{x}) = \int_M \overrightarrow{\bar{x}z}\, dP(z) = 0$ (implicit definition)



Covariance at the mean

- $\Sigma = \mathfrak{M}_2(\bar{x}) = \int_M \overrightarrow{\bar{x}z} \otimes \overrightarrow{\bar{x}z}\, dP(z)$

Asymptotic behavior of the mean

Uniqueness of p-means with convex support

[Karcher 77 / Buser & Karcher 1981 / Kendall 90 / Afsari 10 / Le 11]

- Non-positively curved metric spaces (Aleksandrov): OK [Gromov, Sturm]
- Positive curvature: [Karcher 77 & Kendall 89] concentration conditions:
  Support in a regular geodesic ball of radius $r < r^* = \frac{1}{2}\min(inj(M), \pi/\sqrt{\kappa})$

Bhattacharya-Patrangenaru CLT [BP 2005, B&B 2008]

- Under suitable concentration conditions [KKC], for IID n-samples:
  - $\bar{x}_n \rightarrow \bar{x}$ *(consistency of empirical mean)*
  - $\sqrt{n}\, log_{\bar{x}}(\bar{x}_n) \rightarrow N(0, 4\overline{\boldsymbol{H}}^{-1}\, \boldsymbol{\Sigma}\, \overline{\boldsymbol{H}}^{-1})$ *if* $\overline{H} = \int_M Hess_{\bar{x}}\big(d^2(y, \bar{x})\big)\, \mu(dy)$ *invertible*
- Problems for larger supports [Huckemann & Eltzner, H. Le]

Behavior in high concentration conditions?

- **Interpretation of the mean Hessian?**
- **What happens for a small sample size (non-asymptotic behavior)?**
- **Can we extend results to affine connection spaces?**

Concentration assumptions

- Uniqueness of the mean, support of diameter < ε

Riemannian manifold: Karcher & Kendall Concentr. Cond.

- $\mathrm{Supp}(\mu) \subset B(x,r)$ with $\mathrm{r} < \frac{1}{2}\, inj(x)$
- $\sup_{x \in B(x,r)} \kappa(x) < \pi^2/(4r)^2$

Affine connection spaces: Arnaudon & Li convexity cond.

- $\rho: M \times M \;\to R^+$ separating function
  - Separability: $\rho(x,y) = 0 \Leftrightarrow x = y$
  - Convexity along geodesic: $\rho(\gamma_1(t), \gamma_2(t)): R \to R^+\ convex$
- p-convex geometry: $\mathrm{c\ dist^p(x,y)} \leq \rho(x,y) \leq C\ dist^p(x,y)$
- Uniqueness of exponential barycenter (compact support)

Principle and difficulty

The empirical mean $\bar{x}_n$ of an IID n-sample with population mean $\bar{x}$ is a random variable on M

- Locate $\bar{x}_n$ in a normal coordinate system at x for a given empirical law
- Compute the moments of the empirical mean $\bar{x}_n$at $\bar{x}$:
  - Expectation at the population mean: $\mathbf{Bias}(\bar{x}_n) = \mathbb{E}\big(log_{\bar{x}}(\bar{x}_n)\big)$
  - Covariance matrix $\mathbf{Cov}(\bar{x}_n) = \mathbb{E}\big(log_{\bar{x}}(\bar{x}_n) \otimes log_{\bar{x}}(\bar{x}_n)\big)$
  - Compare with asymptotic BP-CLT for large n

Empirical and population means are exponential barycenters

- n-sample $\mathrm{X_n} = \frac{1}{n}\sum_i \delta_{x_i}$ → tangent mean vector field is $\mathfrak{M}_1(x) = \frac{1}{n}\sum_i \log_x(x_i)$
- Locate the zero $\bar{x}_n$ → **Taylor expansion of $\boldsymbol{log_{x_v}(y)}$ for $\boldsymbol{x_v = exp_x(v)}$?**

Riemannian distance derivatives

How does the (squared) distance (Synge's world function) vary with endpoints?

- First order derivatives is easy

$$D_v\left(\mathrm{dist}^2(x_v, y)\right) = -2\log_{x_v}(y) \text{ with } \mathrm{x_v} = \exp_x(v)$$

- Higher order derivatives begin to be quite involved:
  Taylor expansion in normal coordinates (Grey 1973, Brewin 1998, 2009)

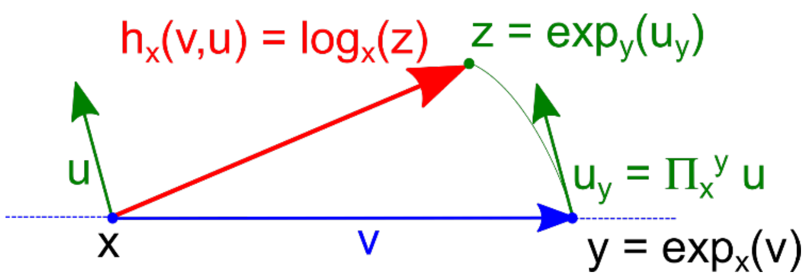
$$\begin{aligned}
D(v) =&\ \mathrm{dist}^2(\exp_x(v), y) = \|y\|_x^2 + D_{,a}v^a + D_{,ab}v^a v^b + D_{,abc}v^a v^b v^c + D_{,abcd}v^a v^b v^c v^d + O(\epsilon^5),\\
D_{,a} =&\ -2y_a\\
D_{,ab} =&\ g_{ab} - \frac{1}{3}y^c y^d R_{acbd} - \frac{1}{12}y^c y^d y^e \nabla_d R_{aebc} - \frac{1}{180}y^c y^d y^e y^f (44R^g_{\ eaf}R_{gcbd} - 3\nabla_{ef}R_{acbd})\\
D_{,abc} =&\ -\frac{1}{12}y^d y^e \nabla_c R_{aebd} + \frac{1}{60}y^d y^e y^f (\nabla_{da}R_{bfce} - 2\nabla_{ad}R_{bfce} + 32R^g_{\ dbe}R_{gacf})\\
&- \frac{1}{108}y^d y^e y^f y^g (8R^h_{\ eaf}\nabla_g R_{hbcd} + 9R^h_{\ eaf}\nabla_h R_{bgcd} + 20R^h_{\ eaf}\nabla_b R_{hgcd} - 6R^h_{\ abe}\nabla_f R_{hgcd})\\
D_{,abcd} =&\ +\frac{1}{180}y^e y^f (8R^g_{\ cdc}R_{gabf} - 9\nabla_{cd}R_{aebf} - 8R^g_{\ dac}R_{gcbf} + 9\nabla_{dc}R_{afbe} - 44R^g_{\ daf}R_{gcbe} - 3\nabla_{db}R_{ceaf})\\
&+ \frac{1}{45}y^c y^f y^g (4R^h_{\ ecf}\nabla_a R_{hbdg} + 4R^h_{\ cae}\nabla_b R_{hfdg} + 4R^h_{\ cae}\nabla_f R_{hbdg} - 3\nabla_{dae}R_{bfcg})\\
&+ \frac{1}{108}y^e y^f y^g (8R^h_{\ eaf}\nabla_d R_{hbcg} + 8R^h_{\ daf}\nabla_g R_{hbce} + 9R^h_{\ eaf}\nabla_h R_{bdcg} + 9R^h_{\ daf}\nabla_h R_{bgce}\\
&\qquad + 20R^h_{\ eaf}\nabla_b R_{hdcg} + 20R^h_{\ daf}\nabla_b R_{hgce} - 6R^h_{\ abe}\nabla_f R_{hdcg} - 6R^h_{\ abe}\nabla_d R_{hgcf})
\end{aligned}$$

- Problem: $\log_{x_v}(y) \in T_{x_v}M$ and not to $\mathrm{T_x}M$: many terms due to $\mathrm{Dexp}_x(v)$

Taylor expansion of geodesic triangles

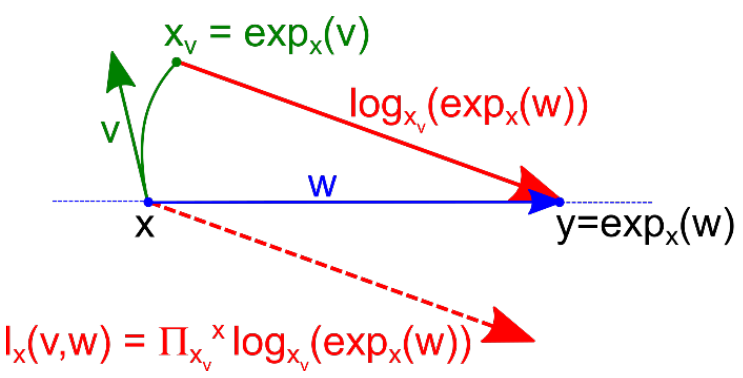
Key idea: use parallel transport rather that normal chart to relate T_xM to $T_{x_v}M$

Gavrilov's double exponential is a tensorial series (2006):



$$
\begin{aligned}
h_x(v,u) &= \log_x(\exp_{x_v}(\Pi_x^{x_v}\, u)) \\
&= v + u + \frac{1}{6}R(u,v)v + \frac{1}{3}R(u,v)u \\
&+ \frac{1}{24}\nabla_v R(u,v)(2v+5u) + \frac{1}{24}\nabla_u R(u,v)(v+2u) + O(5)
\end{aligned}
$$

Neighboring log expansion [XP arXiv:1906.07418, 2019]



$$
\begin{aligned}
l_x(v,w) &= \Pi_{x_v}^{x}\, \log_{x_v}(\exp_x(w)) \\
&= w - v + \frac{1}{6}R(w,v)(v-2w) + \frac{1}{24}\nabla_v R(w,v)(2v-3w) \\
&+ \frac{1}{24}\nabla_w R(w,v)(v-2w) + O(5)
\end{aligned}
$$

Torsion free affine manifolds

Taylor expansion of recentered mean map

$\mathbf{x_v} = \exp_{\boldsymbol{x}}(\boldsymbol{v})$ is an exponential barycenter if $\mathfrak{M}_1(\boldsymbol{x_v}) = \mathbf{0}$

- $\mathfrak{N}_x(v) = \Pi_{x_v}^x \mathfrak{M}_1(x_v) = \int_M \boldsymbol{\Pi_{x_v}^x} \; \boldsymbol{\log_{x_v}(y)} \, \mu(dy)$ has a zero at $\mathrm{v} = \log_x(\bar{x})$
- $\boldsymbol{\mathfrak{M}_1}$ **is a tensor field,** $\boldsymbol{\mathfrak{N}_x}$ **is an analytic endomorphism of** $\boldsymbol{T_x M}$

Taylor expansion with neighboring log:

$$\mathfrak{N}_x(v) = \mathfrak{M}_1 - v + \frac{1}{6} R(\mathfrak{M}_1, v)v - \frac{1}{3} R(*, v) * : \mathfrak{M}_2^{**} + \frac{1}{12} (\nabla_v R)(\mathfrak{M}_1, v)v$$
$$+ \frac{1}{24} (\nabla_* R)(*, v)v : \mathfrak{M}_2^{**} - \frac{1}{8} (\nabla_v R)(*, v) *: \mathfrak{M}_2^{**} - \frac{1}{12} (\nabla_* R)(*, v) *: \mathfrak{M}_3^{***} + O(\varepsilon^5)$$

Solve for the value $\mathrm{v} = \log_x(\bar{x})$ zeroing-out the polynomial

$$\log_x(\bar{x}) = \mathfrak{M}_1 - \frac{1}{3} R(*, \mathfrak{M}_1) * : \mathfrak{M}_2 + \frac{1}{24} (\nabla_* R)(*, \mathfrak{M}_1)\mathfrak{M}_1 : \mathfrak{M}_2^{**}$$
$$- \frac{1}{8} \left(\nabla_{\mathfrak{M}_1} R\right)(*, \mathfrak{M}_1) * : \mathfrak{M}_2^{**} - \frac{1}{12} (\nabla_* R)(*, \mathfrak{M}_1) * : \mathfrak{M}_3^{***} + O(\varepsilon^5)$$

Expectation for a random n-sample

For one empirical n-sample $\mathrm{X_n} = \frac{1}{n}\sum_i \delta_{x_i}$ with moments $\mathfrak{X}_k^n$

- $\log_x(\bar{x}_n) = \mathfrak{X}_1^n - \frac{1}{3} R(*, \mathfrak{X}_1^n) * : \mathfrak{X}_2^n + \frac{1}{24} (\nabla_* R)(*, \mathfrak{X}_1^n) \mathfrak{X}_1^n : \mathfrak{X}_2^n{}^{**}$
  $- \frac{1}{8} \left(\nabla_{\mathfrak{X}_1^n} R\right)(*, \mathfrak{X}_1^n) * : \mathfrak{X}_2^n{}^{**} - \frac{1}{12} (\nabla_* R)(*, \mathfrak{X}_1^n) * : \mathfrak{X}_3^n{}^{***} + O(\varepsilon^5)$

Take expectation for a random IID n-sample

- $\mathbb{E}[\mathfrak{X}_k^n(x)] = \mathfrak{M}_k(x)$
- $\mathbb{E}\left[\mathfrak{X}_p^n \otimes \mathfrak{X}_q^n\right] = \frac{n-1}{n} \mathfrak{M}_{p+q} \otimes \mathfrak{M}_{p+q} + \frac{1}{n} \mathfrak{M}_{p+q}$
- Etc...

Moments of the empirical mean at the population mean:

- $\mathbf{Bias}(\bar{x}_n) = \mathbb{E}[\log_{\bar{x}}(\bar{x}_n)] = \frac{n-1}{6n^2} (\nabla_* R)(*, \circ) \circ : \mathfrak{M}_2^{**} : \mathfrak{M}_2^{\circ\circ} + O(\varepsilon^5)$
- $\mathbf{Cov}(\bar{x}_n) = \mathbb{E}[\log_{\bar{x}}(\bar{x}_n) \otimes \log_{\bar{x}}(\bar{x}_n)]$
  $= \frac{1}{n} \mathfrak{M}_2 - \frac{n-1}{3n^2} \mathfrak{M}_2^{**} : (\circ \otimes R(*, \circ) * + R(*, \circ) * \otimes \circ) : \mathfrak{M}_2^{\circ\circ} + O(\varepsilon^5)$

Asymptotic behavior of empirical Fréchet mean

Moments of the Fréchet mean of a n-sample

- **Surprising Bias** in 1/n on the empirical Fréchet mean (**gradient of curvature**)

$$\mathbf{Bias}(\bar{x}_n) = \mathbb{E}\big(log_{\bar{x}}(\bar{x}_n)\big) = \frac{1}{6n}\,(\mathfrak{M}_2 : \nabla R : \mathfrak{M}_2) + \ O(\epsilon^5, 1/n^2)$$

- **Concentration rate:** term in 1/n modulated by the **curvature**:

$$\mathbf{Cov}(\bar{x}_n) = \mathbb{E}\big(log_{\bar{x}}(\bar{x}_n) \otimes log_{\bar{x}}(\bar{x}_n)\big) = \frac{1}{n}\mathfrak{M}_2 \ + \frac{1}{3n}\mathfrak{M}_2 : R : \mathfrak{M}_2 + \ O(\epsilon^5, 1/n^2)$$

  - **Negative curvature: faster CV than Euclidean**
  - **Positive curvature: slower CV than Euclidean**

Central-limit theorem in manifolds [Bhattacharya & Bhattacharya 2008; Kendall & Le 2011]

- Under Kendall-Karcher concentration conditions:

$$\sqrt{n}\, log_{\bar{x}}(\bar{x}_n) \xrightarrow{D} N(0, H^{-1}\, \Sigma\, H^{-1}) \quad \textit{if}\ \ H = Hess\big(\, MSD(X, \bar{x}_n)\big)\ \textit{invertible}$$

- **Hessian of mean sq. dist:** $\frac{1}{2}\overline{H} = Id + \frac{1}{3}R : \mathfrak{M}_2 + \frac{1}{12}\nabla \mathrm{R} : \mathfrak{M}_3 + O(\epsilon^4, 1/n^2)$

- Same expansion for large n: *modulation of the CV rate by curvature (but our non asymptotic expansion is valid for small data as well)*

[XP, 2018, ARXIV : 1811.01370]

Isotropic distribution in constant curvature spaces

- Symmetric spaces: no bias at order 5
- Modulation of variance w.r.t. Euclidean: $Var(\bar{x}_n) = \alpha \frac{\sigma^2}{n}$

$$\lim_{\kappa\theta^2 = \pi^2/2^2} \alpha = +\infty$$

No CV for uniform distrib on equator

High concentration expansion

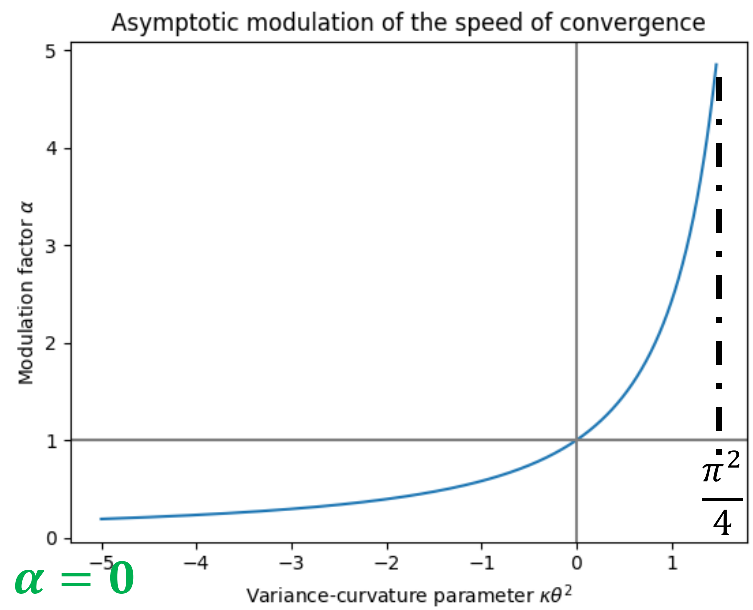
- $\alpha = 1 + \frac{2}{3}\left(1 - \frac{1}{d}\right)\left(1 - \frac{1}{n}\right)\kappa\sigma^2 + O(\epsilon^5)$

Asymptotic BP-CLT expansion

- $\alpha = \left(\frac{1}{d} + \left(1 - \frac{1}{d}\right)\bar{h}\right)^{-2} + O(n^{-2})$

Archetypal modulation factor

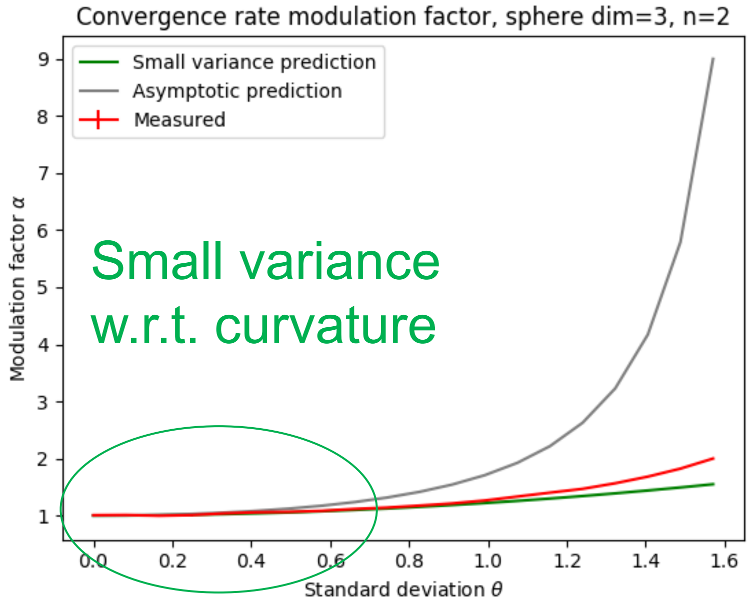
- Uniform distrib on $S(\bar{x}, \theta) \subset M$, large n, large d: $\alpha = \frac{\tan^2(\sqrt{\kappa\theta^2})}{\kappa\theta^2}$

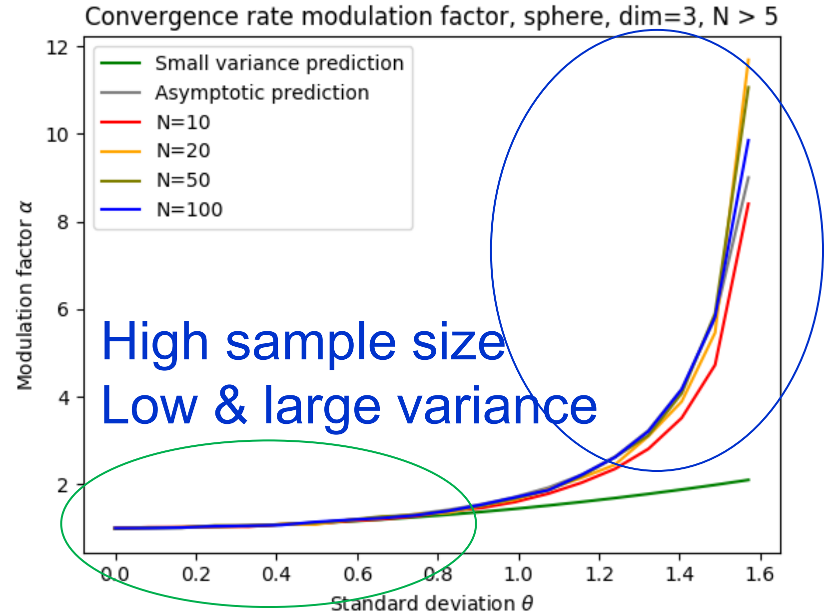


$$\lim_{\kappa\theta^2 = -\infty} \alpha = 0$$

Immediate convergence: sticky mean

Positive curvature

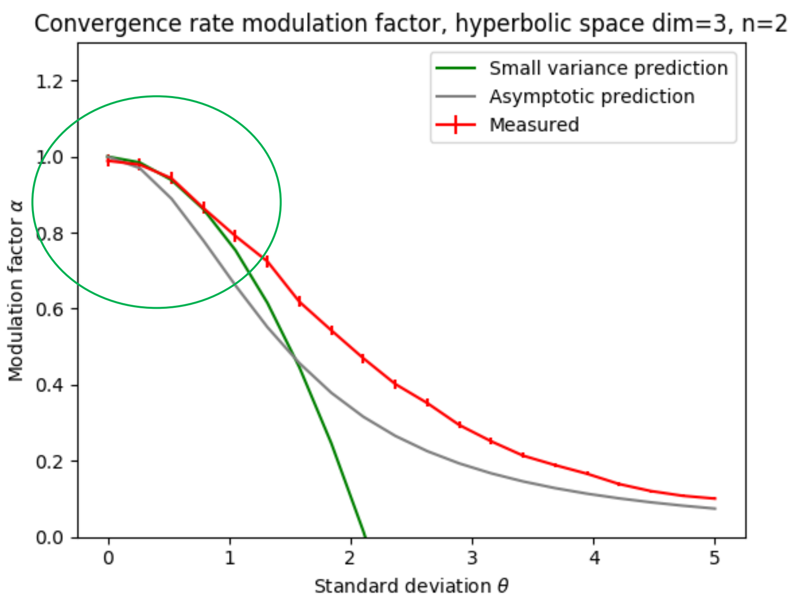


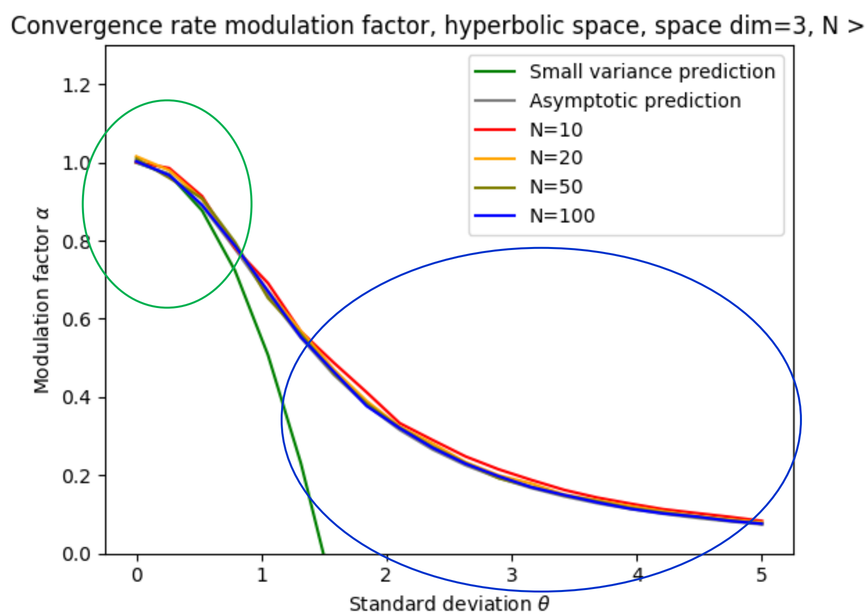


Accurate expansion even with small sample

Accurate asymptotic expansion

Negative curvature

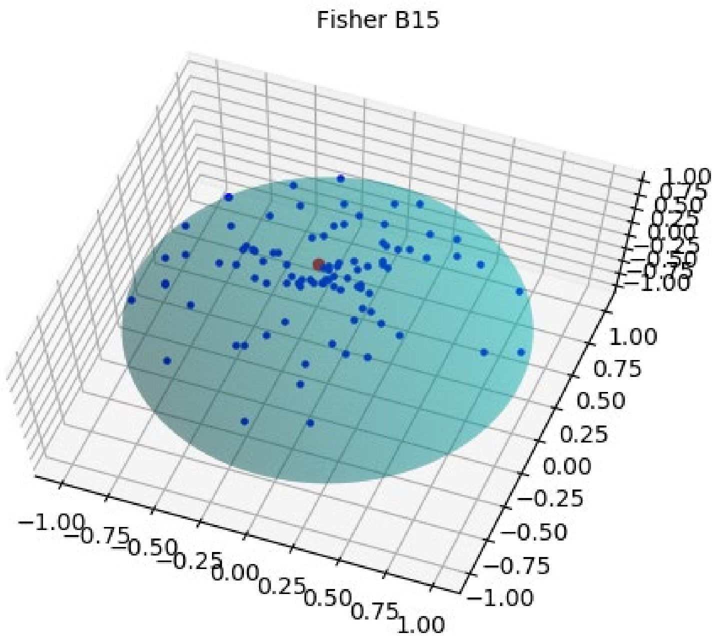


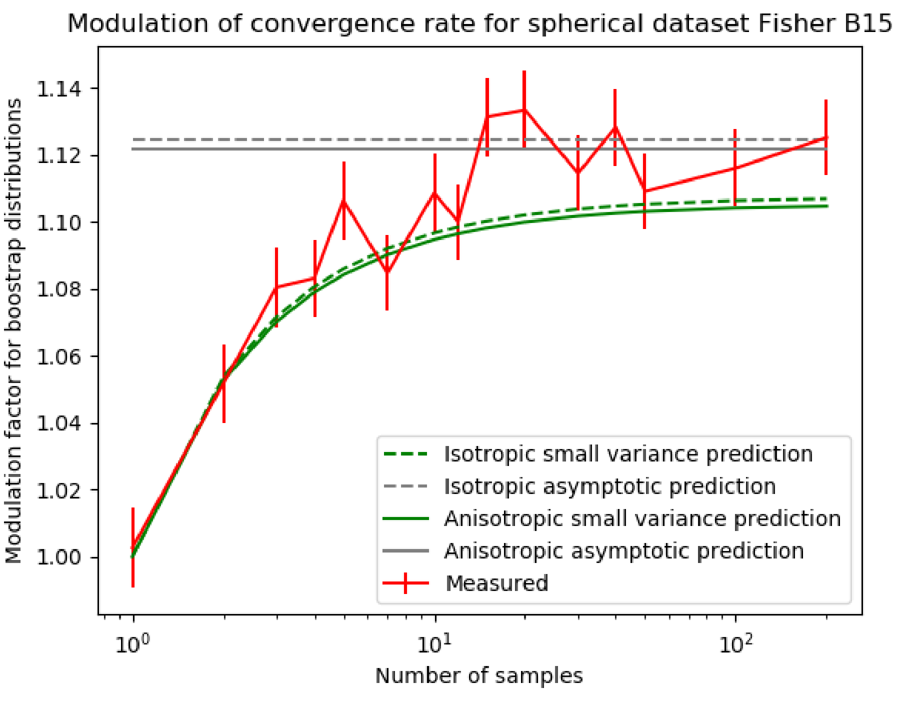


Boostrap on real spherical data from [Fisher, Lewis, Embleton 1987]

B15: high isotropic dispersion (stddev 32°, bbox: 76°x63°)

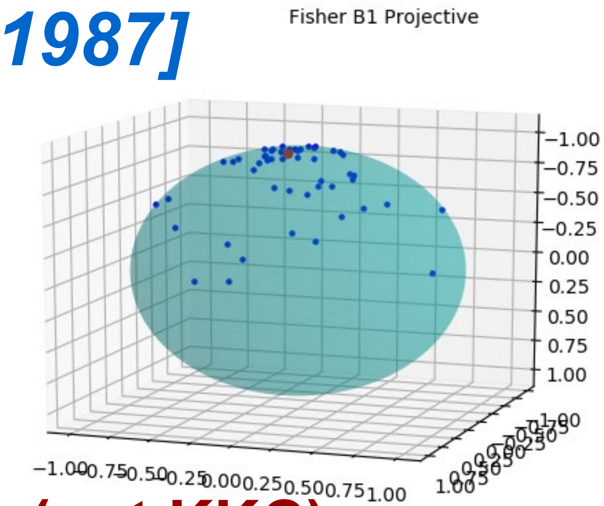
- 94 orientations of dendritic fields in cat's retinas [Keilson et al 1983]
- High dispersion, KKC on the sphere





- Visible modulation (isotropic formulas are good)
- Small sample expansion behavior is well predicted

Boostrap on real projective data from [Fisher, Lewis, Embleton 1987]

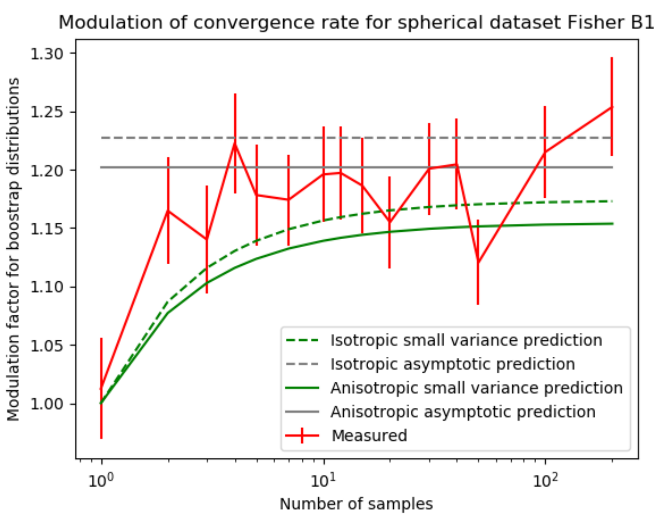


Fisher B1: high dispersion

- 50 pole positions from Paleomagnetic study of new Caledonian laterites (Falvey & Mustgrave)

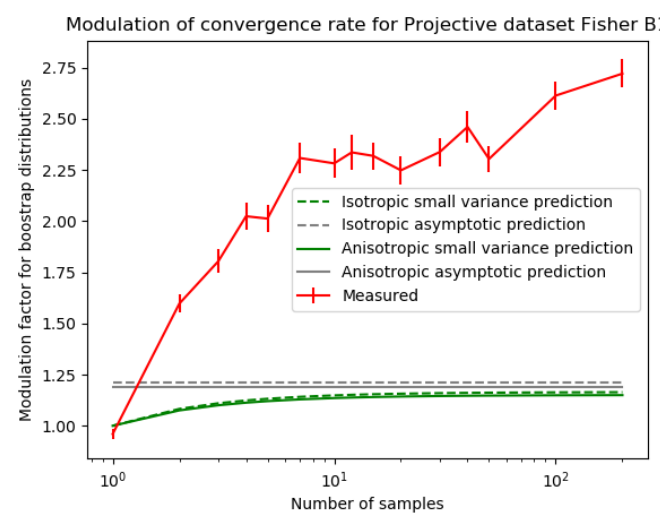
Spherical (not KKC)

- Stddev 41°, bbox: 98° x 67°
- Small var and asymptotic OK



Projective (not KKC)

- Stddev 40°, bbox: 86° x 76°
- Prediction fails: **smeary mean?**



Geometric Statistics: Mathematical foundations and applications in computational anatomy

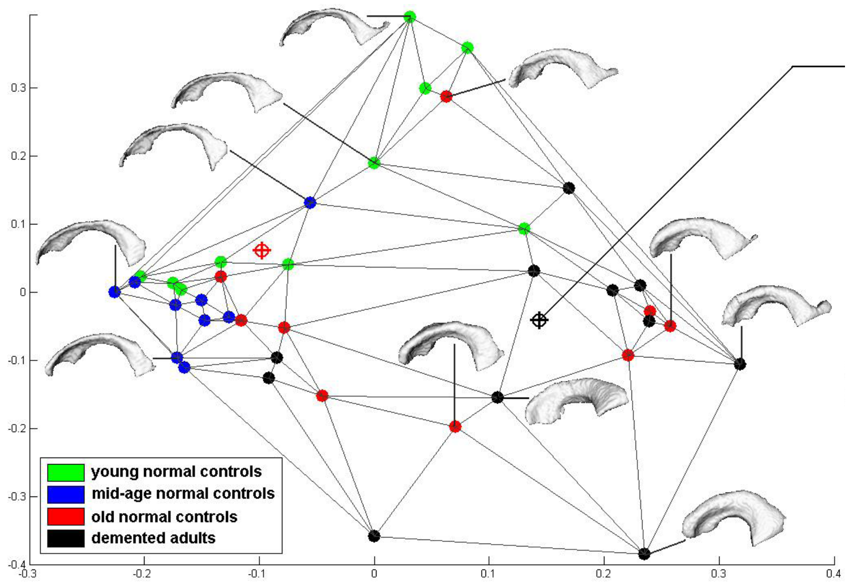
Intrinsic Statistics on Riemannian Manifolds

Metric and Affine Geometric Settings for Lie Groups

Advances Statistics: CLT & PCA

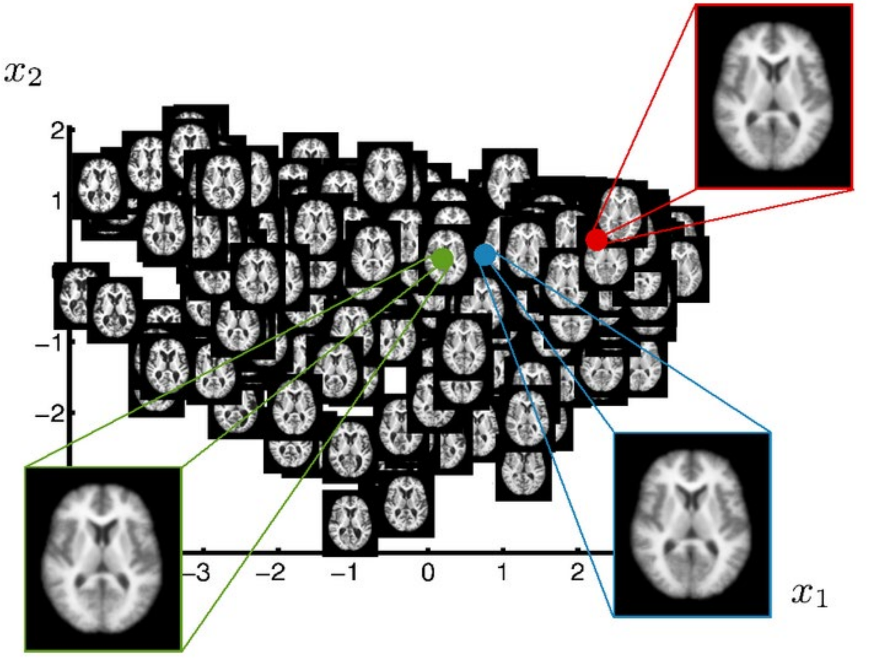
- Estimation of the empirical Fréchet mean & CLT
- Principal component analysis in manifolds
  - Natural subspaces in manifolds: barycentric subspaces
  - Rephrasing PCA with flags of subspaces

Low dimensional subspace approximation?



Manifold of cerebral ventricles

Etyngier, Keriven, Segonne 2007.



Manifold of brain images

S. Gerber et al, Medical Image analysis, 2009.

- Beyond the 0-dim mean → higher dimensional subspaces
- When embedding structure is already manifold (e.g. Riemannian): Not manifold learning (LLE, Isomap,…) but **submanifold learning**
- **Natural subspaces for extending PCA to manifolds?**

Tangent PCA (tPCA)

Maximize the squared distance to the mean (explained variance)

- Algorithm
  - Unfold data on tangent space at the mean
  - Diagonalize covariance at the mean $\Sigma(x) \propto \sum_i \overrightarrow{\bar{x}x_i}\,\overrightarrow{\bar{x}x_i}^t$

- Generative model:
  - Gaussian (large variance) in the horizontal subspace
  - Gaussian (small variance) in the vertical space

- Find the subspace of $T_x M$ that best explains the variance

Principal Geodesic / Geodesic Principal Component Analysis

Minimize the squared Riemannian distance to a low dimensional subspace (unexplained variance)

- **Geodesic Subspace**: $GS(x, w_1, \dots w_k) = \{\exp_x(\sum_i \alpha_i w_i) \ for \ \alpha \in R^k \}$
  - Parametric subspace spanned by geodesic rays from point x
  - Beware: GS have to be restricted to be well posed [XP, AoS 2018]
    - PGA (Fletcher et al., 2004, Sommer 2014)
    - Geodesic PCA (GPCA, Huckeman et al., 2010)

- Generative model:
  - Unknown (uniform ?) distribution within the subspace
  - Gaussian distribution in the vertical space

Asymmetry w.r.t. the base point in $GS(x, w_1, \dots w_k)$

- Totally geodesic at x only

Patching the Problems of tPCA / PGA

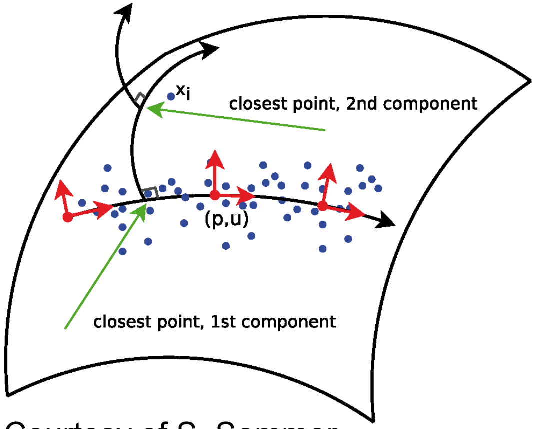
Improve the flexibity of the geodesics

- 1D regression with higher order splines **[Gu, Machado, Leite, Vialard, Singh, Niethammer, Absil,…]**
  - **Control of dimensionality for n-D Polynomials on manifolds?**

Iterated Frame Bundle Development

[HCA, Sommer GSI 2013]

- Iterated construction of subspaces
- Parallel transport in frame bundle
  - **Intrinsic asymmetry between components**



Courtesy of S. Sommer

Nested "algebraic" subspaces

- Principal nested spheres **[Jung, Dryden, Marron 2012]**
- Quotient of Lie group action **[Huckemann, Hotz, Munk, 2010]**
  - **No general semi-direct product space structure in general Riemannian manifolds**

Affine span in Euclidean spaces

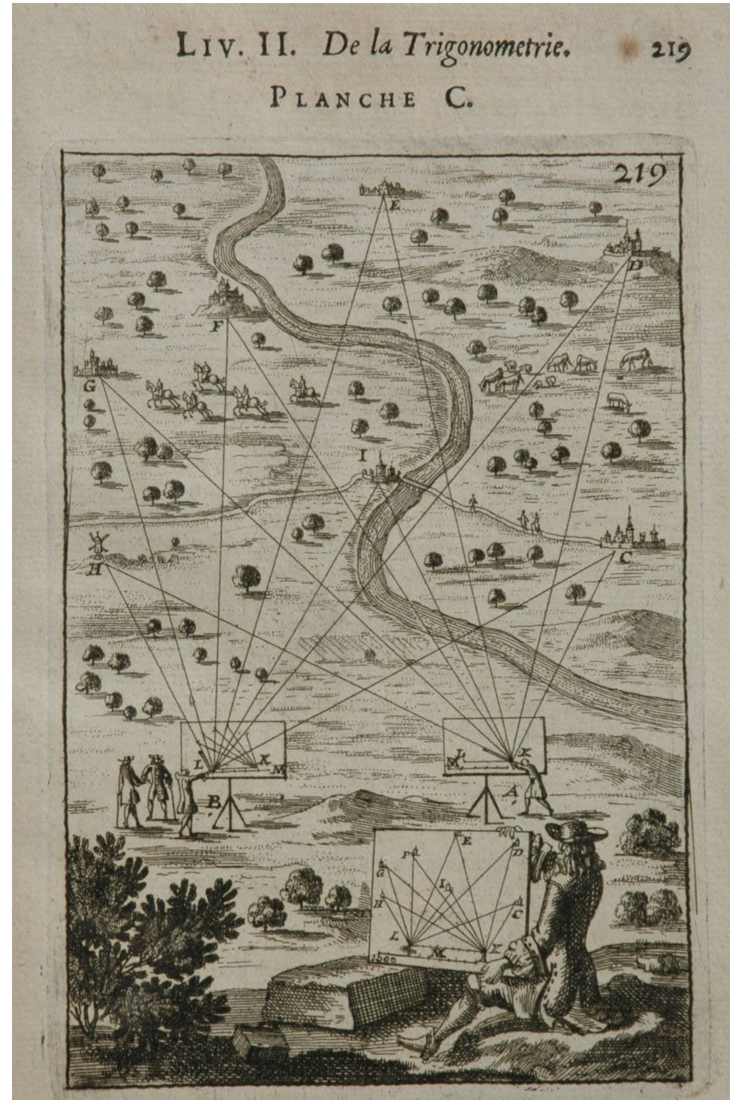
Affine span of (k+1) points: weighted barycentric equation

$$\mathrm{Aff}(\mathrm{x}_0, \mathrm{x}_1, \ldots \mathrm{x}_\mathrm{k}) = \{\mathrm{x} = \textstyle\sum_i \lambda_i\, x_i \;\; with \; \sum_i \lambda_i = 1\}$$
$$= \{\mathrm{x} \in R^n \;\; s.t \; \textstyle\sum_i \lambda_i\, (x_i - x) = 0, \lambda \in P_k^*\}$$

Key ideas:

- ~~tPCA, PGA: Look at data points from the mean (mean has to be unique)~~

- Triangulate from several reference: **locus of weighted means**



A. Manesson-Mallet. La géométrie Pratique, 1702

Barycentric subspaces and Affine span in Riemannian manifolds

Fréchet / Karcher barycentric subspaces (KBS / FBS)

- Normalized weighted variance: $\sigma^2(x,\lambda) = \sum\lambda_i dist^2(x, x_i)/\sum\lambda_i$
- Set of absolute / local minima of the λ-variance
- Works in stratified spaces (may go accross different strata)
  - Non-negative weights: Locus of Fréchet Mean [Weyenberg, Nye]

Exponential barycentric subspace and affine span

- Weighted exponential barycenters: $\mathfrak{M}_1(x,\lambda) = \sum_i \lambda_i \overrightarrow{xx_i} = 0$
- EBS$(x_0, \dots x_k) = \{\, x \in M^*(x_0, \dots x_k) \mid \mathfrak{M}_1(x,\lambda) = 0\}$
- Affine span = closure of EBS in M $\quad Aff(x_0, \dots x_k) = \overline{EBS(x_0, \dots x_k)}$

Questions

- Local structure: local manifold? dimension? stratification?
- Relationship between KBS ⊂ FBS, EBS and affine span?

[X.P. Barycentric Subspace Analysis on Manifolds. Annals of statistics. 2018. arXiv:1607.02833]

Analysis of Barycentric Subspaces

Exp. barycenters are critical points of λ-variance on M*

- $\nabla\sigma^2(\mathrm{x},\lambda) = -2\mathfrak{M}_1(\mathrm{x},\lambda) = 0$ $\quad$ $\boldsymbol{KBS} \cap \boldsymbol{M}^* \subset \boldsymbol{EBS}$

Caractérisation of local minima: Hessian (if non degenerate)

$$H(\mathrm{x},\lambda) = -2\sum_i \lambda_i D_x \log_x(x_i) = \mathbf{Id} - \frac{\mathbf{1}}{\mathbf{3}}\mathbf{Ric}\left(\mathfrak{M}_2(\mathbf{x},\boldsymbol{\lambda})\right) + \mathrm{HOT}$$

Regular and positive pts (non-degenerated critical points)

- $\boldsymbol{EBS}^{\boldsymbol{Reg}}(\boldsymbol{x}_0, \dots \boldsymbol{x}_k) = \{\, \boldsymbol{x} \in \boldsymbol{Aff}(\boldsymbol{x}_0, \dots \boldsymbol{x}_k), \boldsymbol{s.t.}\ \boldsymbol{H}(\boldsymbol{x}, \boldsymbol{\lambda}^*(\boldsymbol{x})) \neq \mathbf{0}\}$
- $\boldsymbol{EBS}^{+}(\boldsymbol{x}_0, \dots \boldsymbol{x}_k) = \{\, \boldsymbol{x} \in \boldsymbol{Aff}(\boldsymbol{x}_0, \dots \boldsymbol{x}_k), \boldsymbol{s.t.}\ \boldsymbol{H}(\boldsymbol{x}, \boldsymbol{\lambda}^*(\boldsymbol{x}))\ \boldsymbol{Pos.def.}\}$

Theorem: EBS partitioned into cells by the index of the Hessian of λ-variance: KBS = EBS^+ on M*

[X.P. Barycentric Subspace Analysis on Manifolds. Annals of statistics. 2018. arXiv:1607.02833]

KBS / FBS with 3 points on the sphere

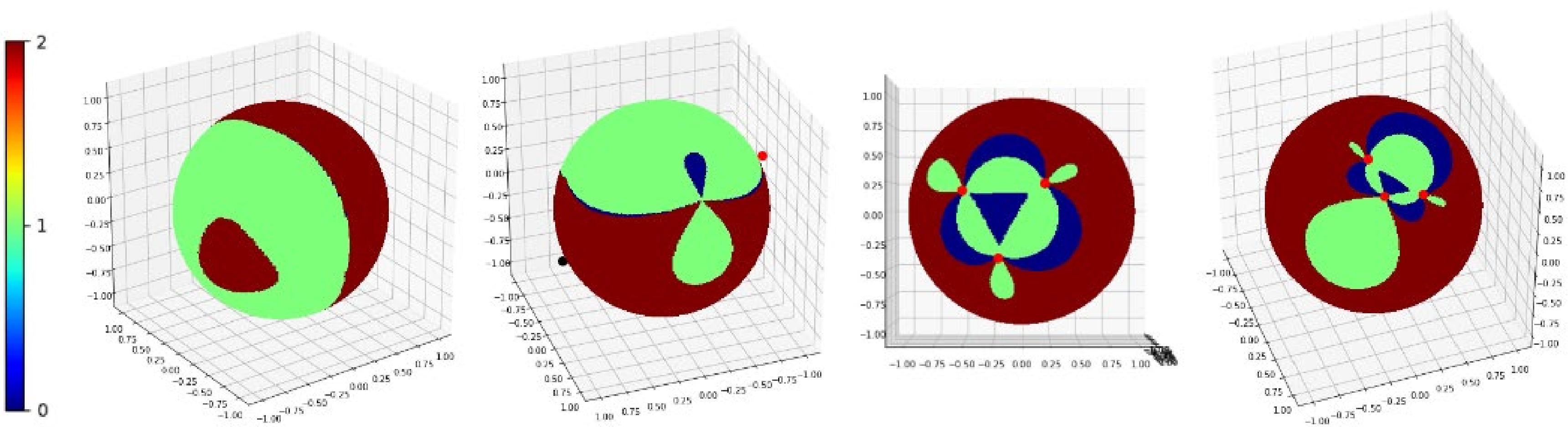
EBS: great subspheres spanned by reference points (mod cut loci)

$$\mathrm{EBS}(x_0, \dots x_k) = Span(X) \cap S_n \setminus Cut(X) \qquad Aff(x_0, \dots x_k) = Span(X) \cap S_n$$

KBS/FBS: look at index of the Hessian of λ-variance

$$\mathrm{H}(\mathrm{x},\lambda) = \sum \lambda_i \theta_i \cot(\theta_i)\,(\mathrm{Id} - \mathrm{xx}^{\mathrm{t}}) + \sum (1 - \lambda_i \theta_i \cot(\theta_i))\, \overrightarrow{xx_i}\, \overrightarrow{xx_i}^{t}$$

- Complex algebric geometry problem [Buss & Fillmore, ACM TG 2001]
- 3 points of the n-sphere: EBS partitioned in cell complex by index of critical point
- **KBS/EBS less interesting than EBS/affine span**



Weighed Hessian index: **brown = -2 (min) = KBS** / green = -1 (saddle) / blue = 0 (max)

Example on the hyperbolic space

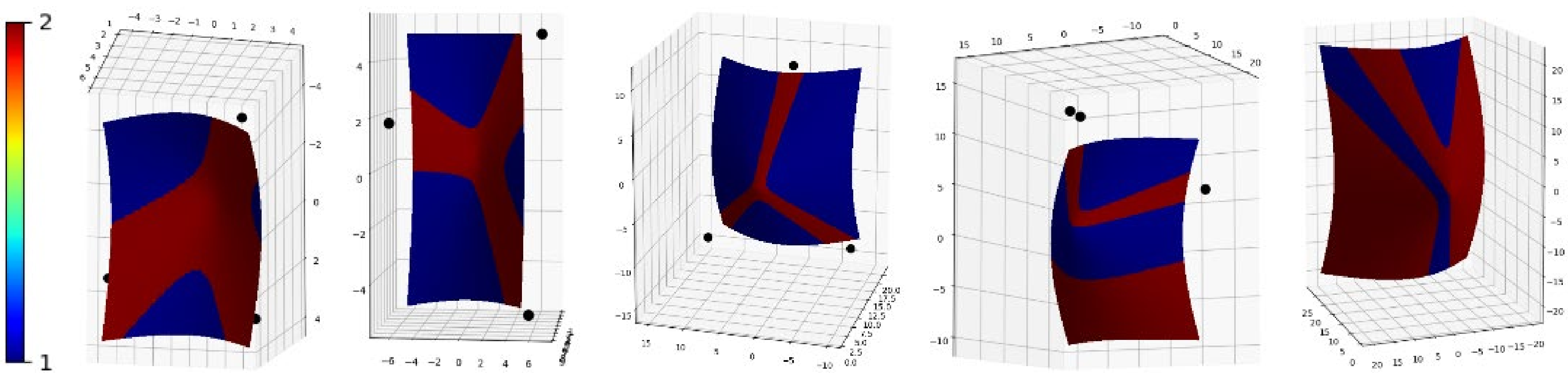
EBS = Affine span: great sub-hyperboloids spanned by reference points

$$\mathrm{EBS}(x_0, \dots x_k) = Aff(x_0, \dots x_k) = Span(X) \cap H_n$$

KBS: locus of maximal index of the Hessian of λ-variance

$$\mathrm{H}(\mathrm{x},\lambda) = \sum \lambda_i \theta_i \coth(J + J\mathrm{x}\mathrm{x}^{\mathrm{t}} J^t) + \sum (1 - \lambda_i \coth(\theta_i)) J\, \overrightarrow{xx_i}\, \overrightarrow{xx_i}^{\,t} J^t$$

- Complex algebric geometry problem
- 3 points on H^n: better than for spheres, but still disconnected components



Weighted Hessian Index: **brown = -2 (min) = KBS** / blue = 1 (saddle)

Geodesic subspaces are limit cases of affine span

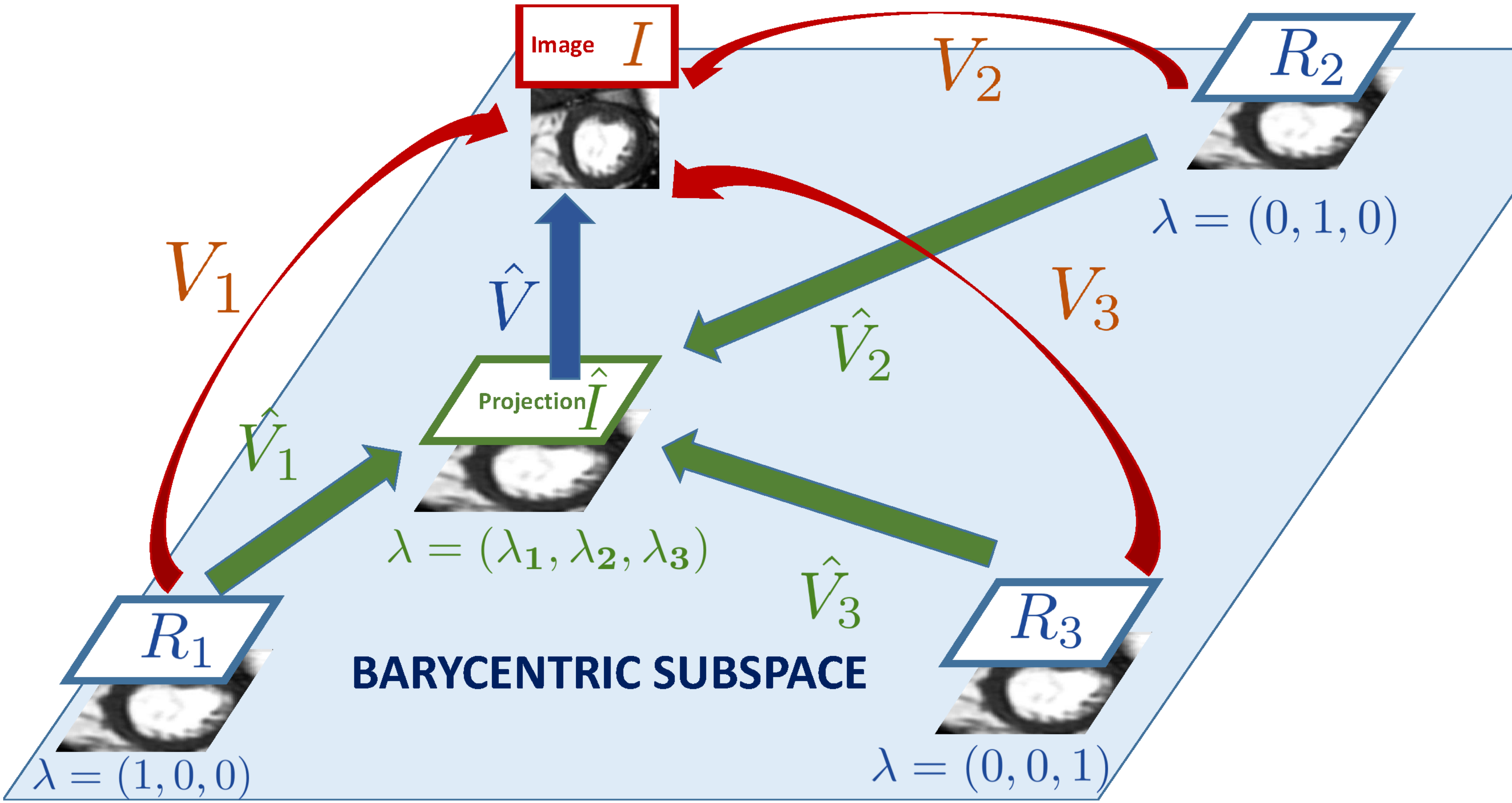
Theorem

- $GS(x, w_1, \ldots w_k) = \{\exp_x(\sum_i \alpha_i w_i) \ for\ \alpha \in R^k \}$ is the limit of $Aff\left(x_0, \exp_{x_o}(\epsilon\, w_1), \ldots \exp_{x_o}(\epsilon\, w_k)\right)$ when $\epsilon \to 0$.
- Reference points converge to a 1st order (k,n)-jet
  - PGA [Fletcher et al. 2004, Sommer et al. 2014]
  - GPGA [Huckemann et al. 2010]

Conjecture

- This can be generalized to higher order derivatives
  - Quadratic, cubic splines [Vialard, Singh, Niethammer]
  - Principle nested spheres [Jung, Dryden, Marron 2012]
  - Quotient of Lie group action [Huckemann, Hotz, Munk, 2010]

Application in Cardiac motion analysis

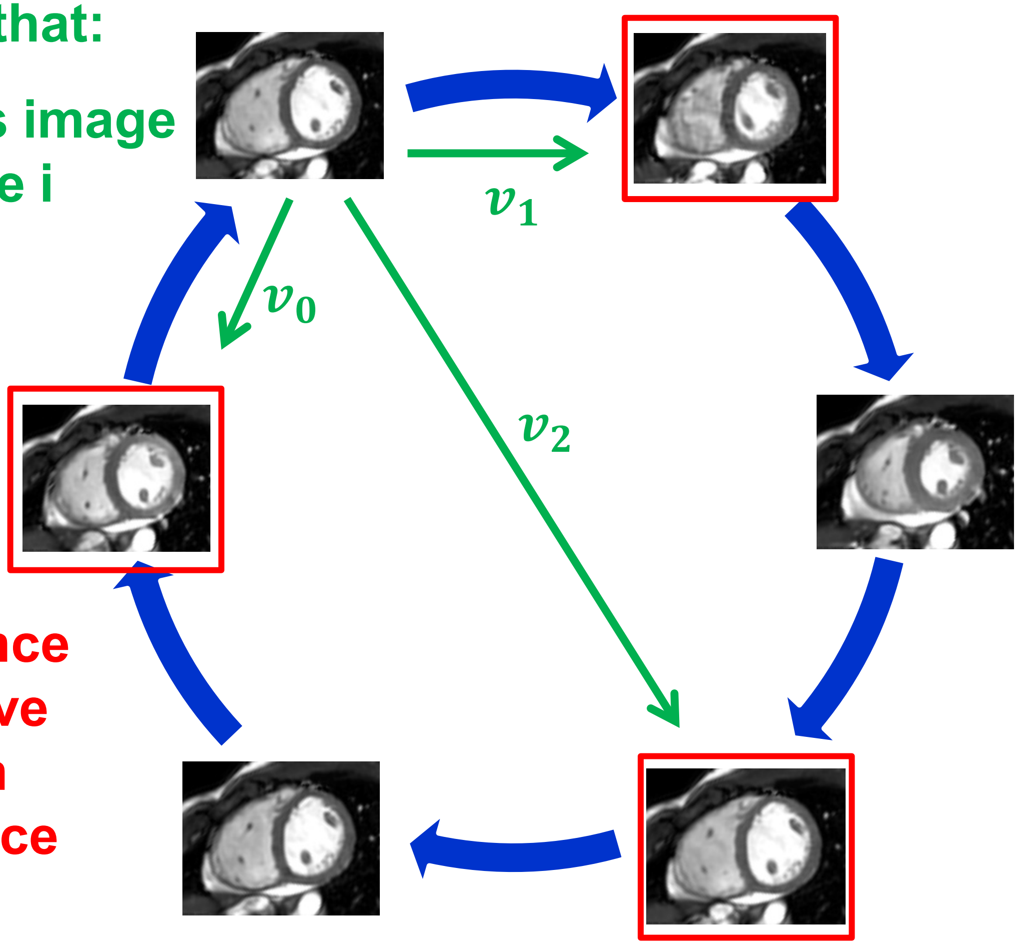


[Marc-Michel Rohé et al., MICCAI 2016, MedIA 45:1-12, 2018]

Application in Cardiac motion analysis

Find weights λ_i and SVFs v_i such that:

- v_i **registers image to reference i**
- $\sum_i \lambda_i v_i = 0$

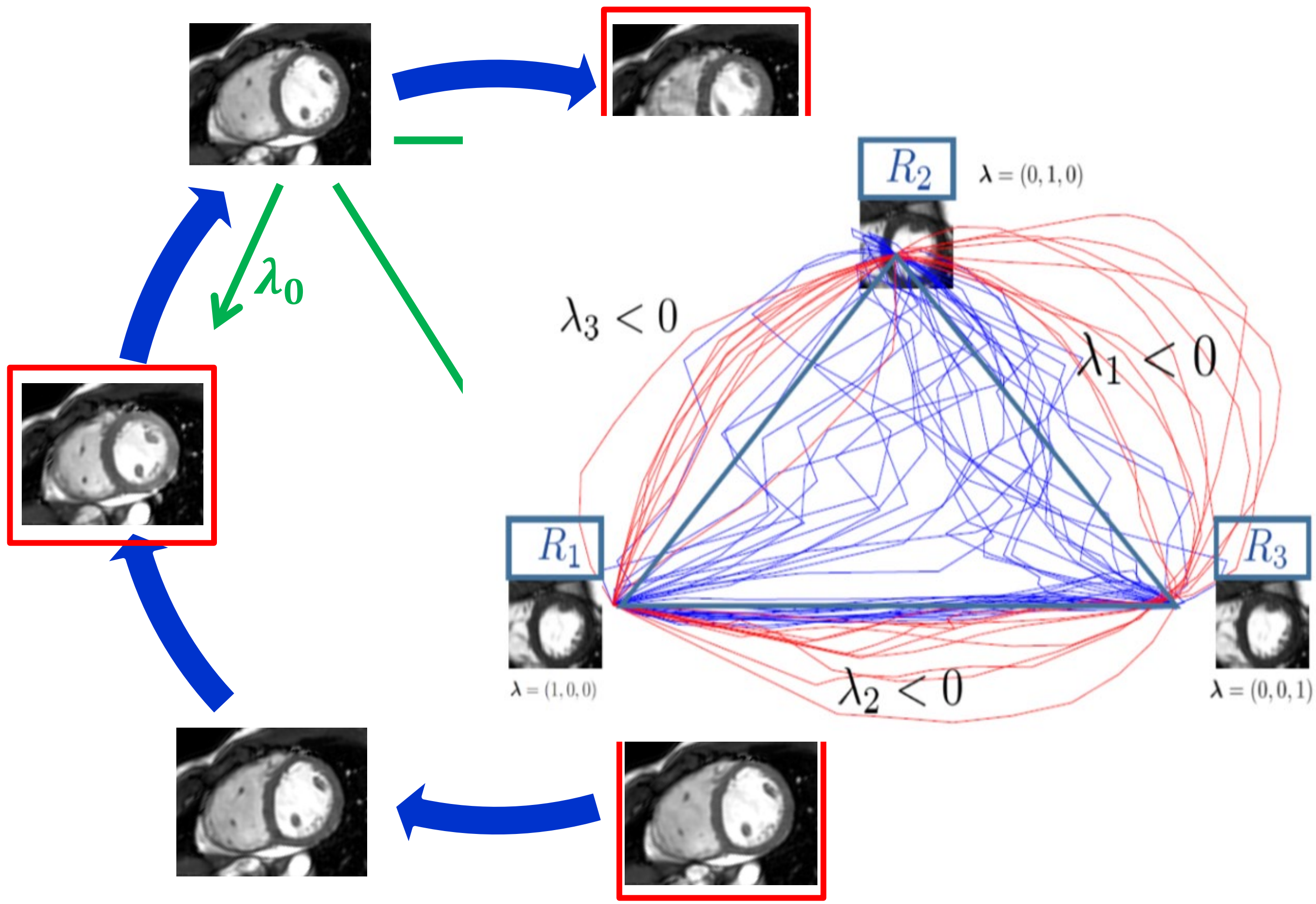


Optimize reference images to achieve best registration over the sequence

[Marc-Michel Rohé et al., MICCAI 2016, MedIA 45:1-12, 2018]

Application in Cardiac motion analysis

Optimal Reference Frames **Barycentric coefficients curves**

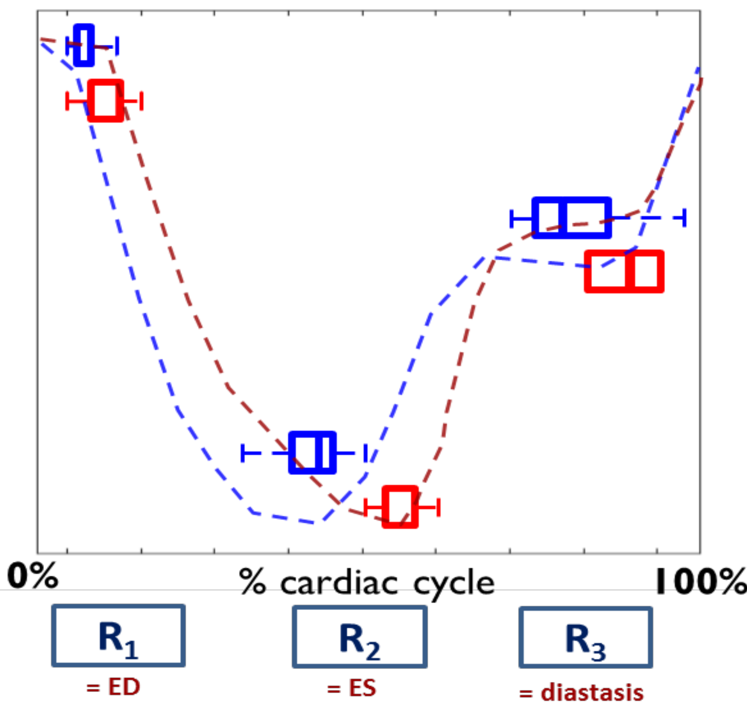


[Marc-Michel Rohé et al., MICCAI 2016, MedIA 45:1-12, 2018]

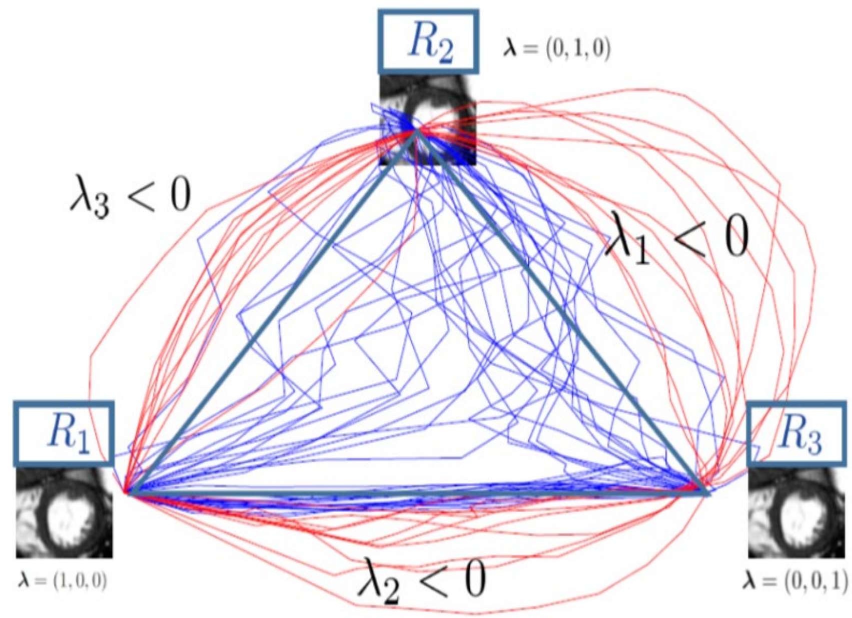
Cardiac Motion Signature

Low-dimensional representation of motion using:

Optimal Reference Frames



Barycentric coefficients curves



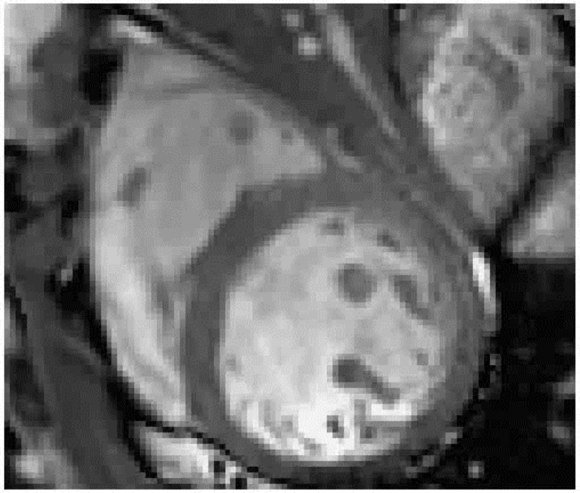
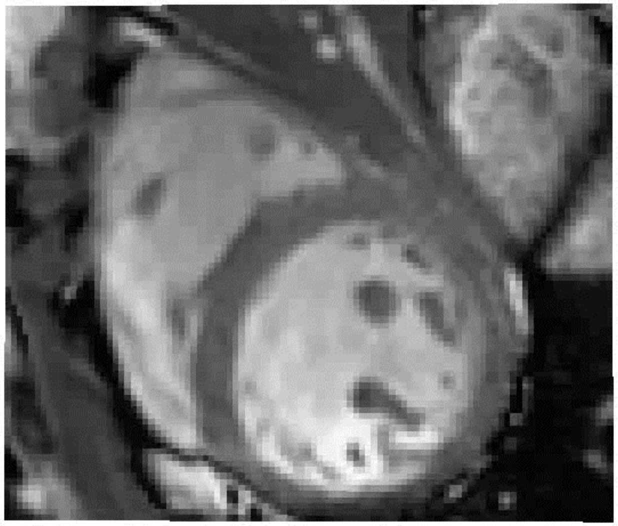
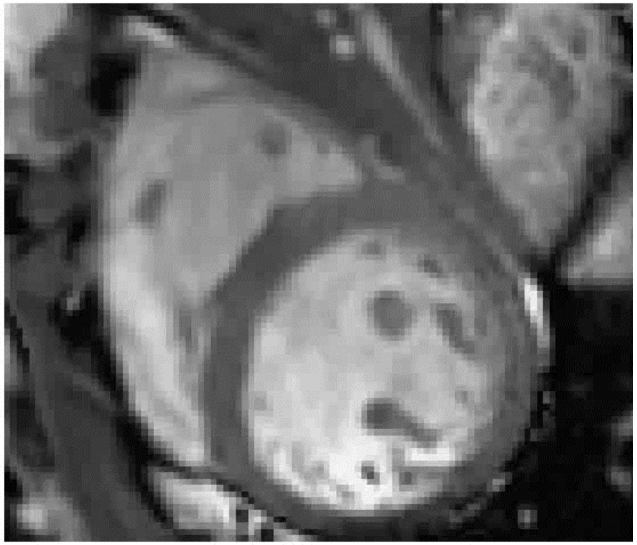
Dimension reduction from **+10M voxels** to **3 reference** frames + **60 coefficients**

Tested on **10 controls** [1] and **16 Tetralogy of Fallot** patients [2]

[1] Tobon-Gomez, C., et al.: Benchmarking framework for myocardial tracking and deformation algorithms: an open access database. Medical Image Analysis (2013)

[2] Mcleod K., et al.: Spatio-Temporal Tensor Decomposition of a Polyaffine Motion Model for a Better Analysis of Pathological Left Ventricular Dynamics. IEEE TMI (2015)

Cardiac motion synthesis

Original Sequence	Barycentric Reconstruction (3 images)	PCA Reconstruction (2 modes)
		
30 images	3 images + 2 coeff.	1 image + 2 SVF + 2 coeff.
	Reconstr. error: 18.75 Compression ratio: 1/10	Reconstr. error: 26.32 (+40%) Compression ratio: 1/4

[Marc-Michel Rohé et al., MICCAI 2016, MedIA 45:1-12, 2018]

Geometric Statistics: Mathematical foundations and applications in computational anatomy

Intrinsic Statistics on Riemannian Manifolds

Metric and Affine Geometric Settings for Lie Groups

Advances Statistics: CLT & PCA

- Estimation of the empirical Fréchet mean & CLT
- Principal component analysis in manifolds
  - Natural subspaces in manifolds: barycentric subspaces
  - Rephrasing PCA with flags of subspaces

The natural object for PCA: Flags of subspaces in manifolds

Subspace approximations with variable dimension

- Optimal unexplained variance → non nested subspaces
- Nested forward / backward procedures → not optimal
- Optimize first, decide dimension later → Nestedness required
  [Principal nested relations: Damon, Marron, JMIV 2014]

Flags of affine spans in manifolds: $FL(x_0 \prec x_1 \prec \cdots \prec x_n)$

- Sequence of nested subspaces

$$Aff(x_0) \subset Aff(x_0, x_1) \subset \cdots Aff(x_0, \ldots x_i) \subset \cdots Aff(x_0, \ldots x_n) = M$$

Barycentric subspace analysis (BSA):

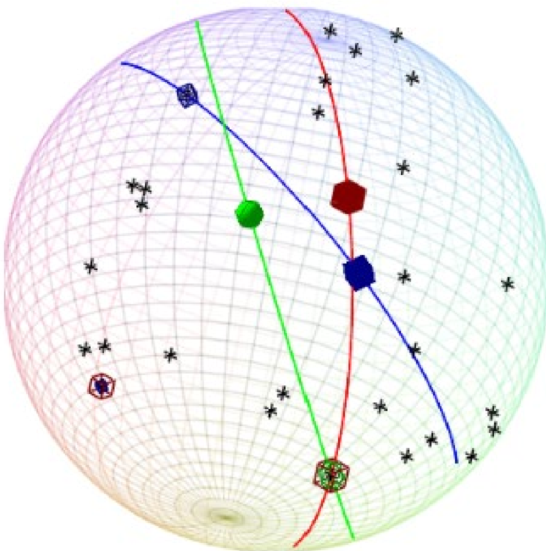
- Energy on flags: Accumulated Unexplained Variance
  → **optimal flags of subspaces in Euclidean spaces = PCA**

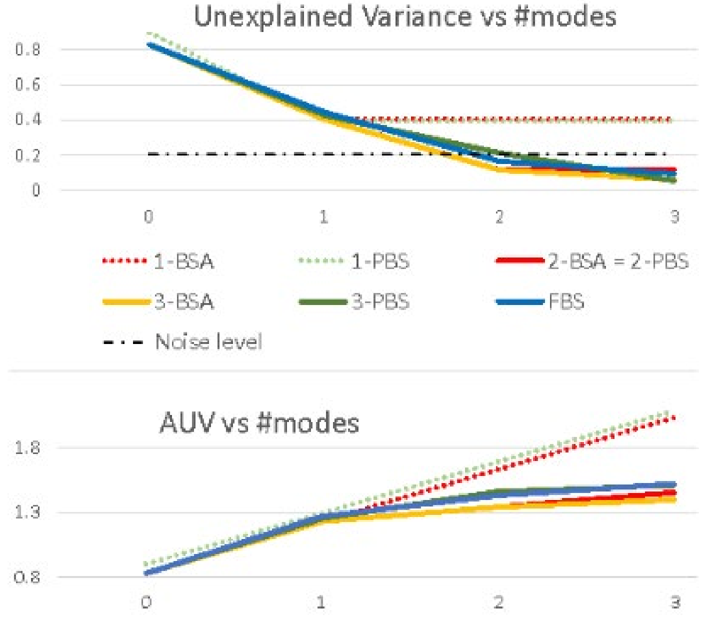
[X.P. Barycentric Subspace Analysis on Manifolds, Annals of Statistics 2018]

Sample-limited barycentric subspace inference

Restrict the inference to data points only

- Fréchet mean / template [Lepore et al 2008]
- First geodesic mode [Feragen et al. 2013, Zhai et al 2016]
- Higher orders: challenging with PGA… but not with BSA





- **FBS: Forward Barycentric Subspace**
- **k-PBS: Pure Barycentric Subspace with backward ordering**
- **k-BSA: Barycentric Subspace Analysis up to order k**

Robustness with L_p norms

Affine spans is stable to p-norms

- $\sigma^p(x,\lambda) = \frac{1}{p}\sum\lambda_i dist^p(x, x_i) / \sum\lambda_i$
- Critical points of $\sigma^p(x,\lambda)$ are also critical points of $\sigma^2(x,\lambda')$ with $\lambda'_i = \lambda_i\, dist^{p-2}(x, x_i)$ (non-linear reparameterization of affine span)

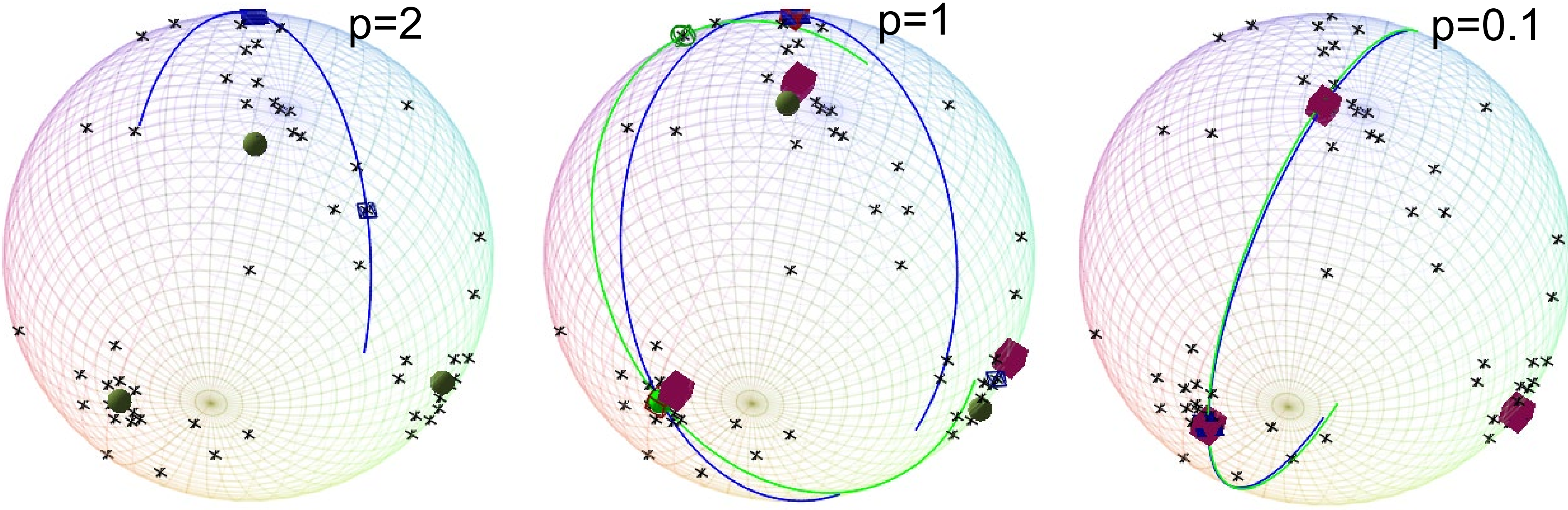
Unexplained p-variance of residuals

- $2 < p \to +\infty$: more weight on the tail,
  at the limit: penalizes the maximal distance to subspace
- $0 < p < 2$: less weight on the tail of the residual errors: statistically robust estimation
  - Non-convex for p<1 even in Euclidean space
  - But sample-limited algorithms do not need gradient information

Experiments on the sphere

3 clusters on a 5D sphere

- 10, 9 and 8 points (stddev 6 deg) around three orthogonal axes plus 30 points uniformly samples on 5D sphere

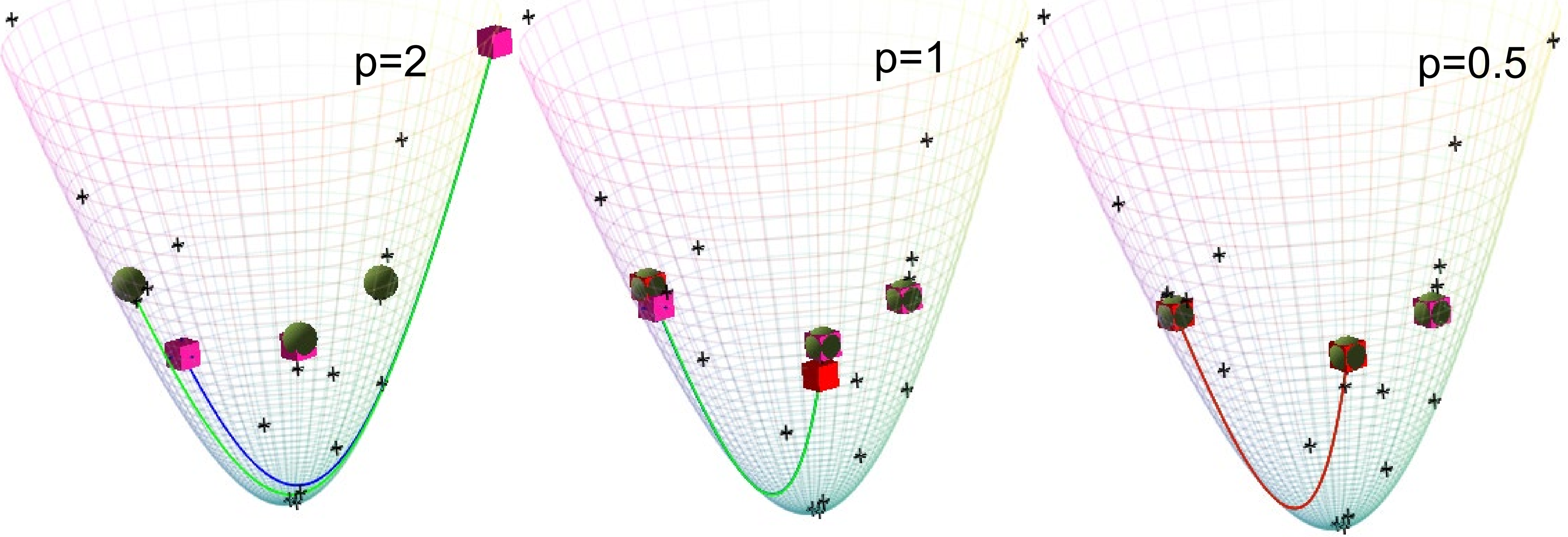


- **FBS: Forward Barycentric Subspace: mean and median not in clusters**
- **1-PBS / 2-PBS: Pure Barycentric Subspace with backward ordering: ok for k=2 only**
- **1-BSA / 2-BSA: Barycentric Subspace Analysis up to order k: less sensitive to p & k**

Experiments on the hyperbolic space

3 clusters on a 5D hyperboloid (50% outliers)

- 15 random points (stddev 0.015) around an equilateral triangle of length 1.57 plus 15 points of stddev 1.0 (truncated at max 1.5)



- **FBS: Forward Barycentric Subspace: ok for p≤0.5**
- **1-PBS / 2-PBS: Pure Barycentric Subspace with backward ordering: ok for k=2 only**
- **1-BSA / 2-BSA: Barycentric Subspace Analysis up to order k: ok for $p \leq 1$**

Take home messages

Natural subspaces in manifolds

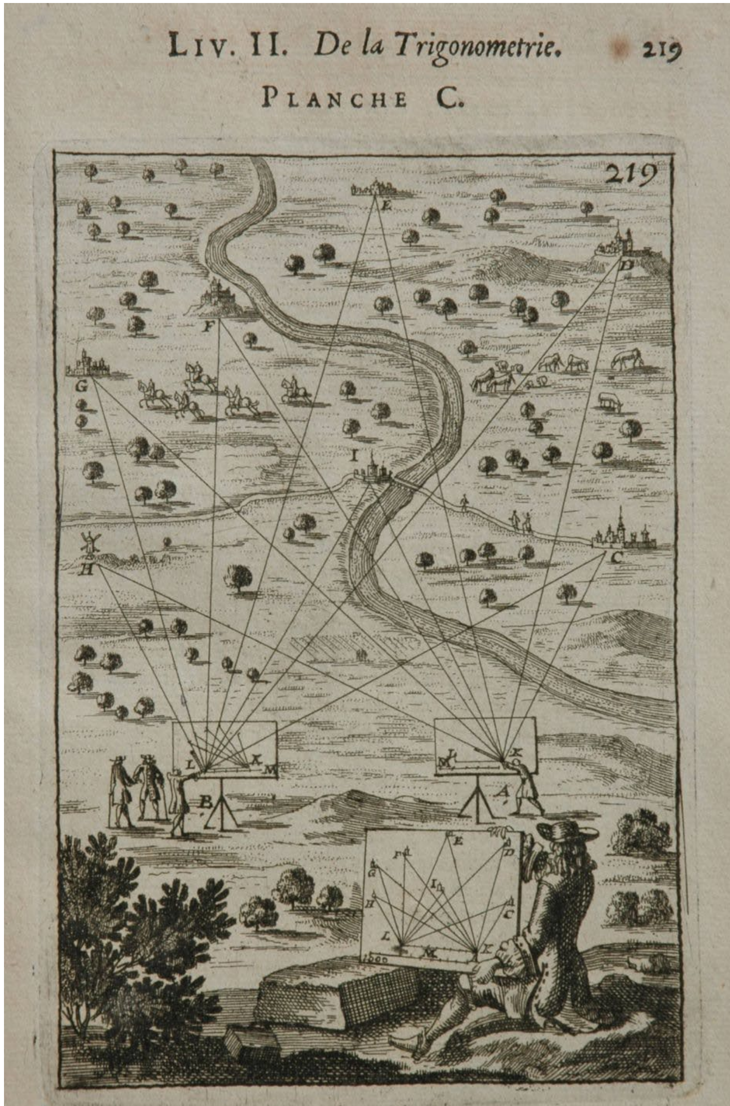
- PGA & Godesic subspaces:
  look at data points from the (unique) mean
- Barycentric subspaces:
  « triangulate » several reference points
  - Justification of multi-atlases?

Critical points (affine span) rather than minima (FBS/KBS)

- Barycentric coordinates need not be positive (convexity is a problem)
- Affine notion (more general than metric)
  - Generalization to Lie groups (SVFs)?

Natural flag structure for PCA

- Hierarchically embedded approximation subspaces to summarize / describe data



A. Manesson-Mallet. La géométrie Pratique, 1702

References

1/ Intrinsic Statistics on Riemannian Manifolds

- XP. **Intrinsic Statistics on Riemannian Manifolds: Basic Tools for Geometric Measurements**. *J. Math. Imaging & Vision* 25 (1), 2006, pp.127-154. [DOI: 10.1007/s10851-006-6228-4][Preprint]
- XP, P. Fillard, N. Ayache. **A Riemannian Framework for Tensor Computing**. *Int. J. of Computer Vision*, 66(1):41-66, Jan. 2006. [DOI: 10.1007/s11263-005-3222-z]. [Preprint]
- XP. **Manifold-valued image processing with SPD matrices**. *In Riemannian Geometric Statistics in Medical Image Analysis*, Chap. 3, pp.75-134, Elsevier, 2020. [DOI: 10.1016/B978-0-12-814725-2.00010-8] [Preprint]

2/ Metric and affine geometric settings for Lie groups

- XP, M. Lorenzi. **Beyond Riemannian Geometry - The affine connection setting for transformation groups**. *In Riemannian Geometric Statistics in Medical Image Analysis*, Chap. 5, pp.169-229 RGSMIA. Elsevier, 2020. [DOI: 10.1016/B978-0-12-814725-2.00012-1] [Preprint] .
- XP, M. Lorenzi. **Geodesics, Parallel Transport and One-parameter Subgroups for Diffeomorphic Image Registration**. *Int. J. of Computer Vision* 105(2):111-127. 2013. [DOI: 10.1007/s11263-012-0598-4] [Preprint]
- XP. **Parallel Transport with Pole Ladder: a Third Order Scheme in Affine Connection Spaces which is Exact in Affine Symmetric Spaces**. 2018. [arXiv:1805.11436]
- N. Guigui, XP. **Numerical Accuracy of Ladder Schemes for Parallel Transport on Manifolds**. *Foundations of Computational Mathematics*, 22:757-790, 2022. [DOI: 10.1007/s10208-021-09515-x]

3/ Advanced statistics: central limit theorem and extension of PCA

- XP. **Curvature effects on the empirical mean in Riemannian and affine Manifolds**. 2019. [arXiv:1906.07418]
- XP. **Barycentric Subspace Analysis on Manifolds**. *Annals of Statistics.*46(6A):2711-2746, 2018. [DOI: 10.1214/17-AOS1636]