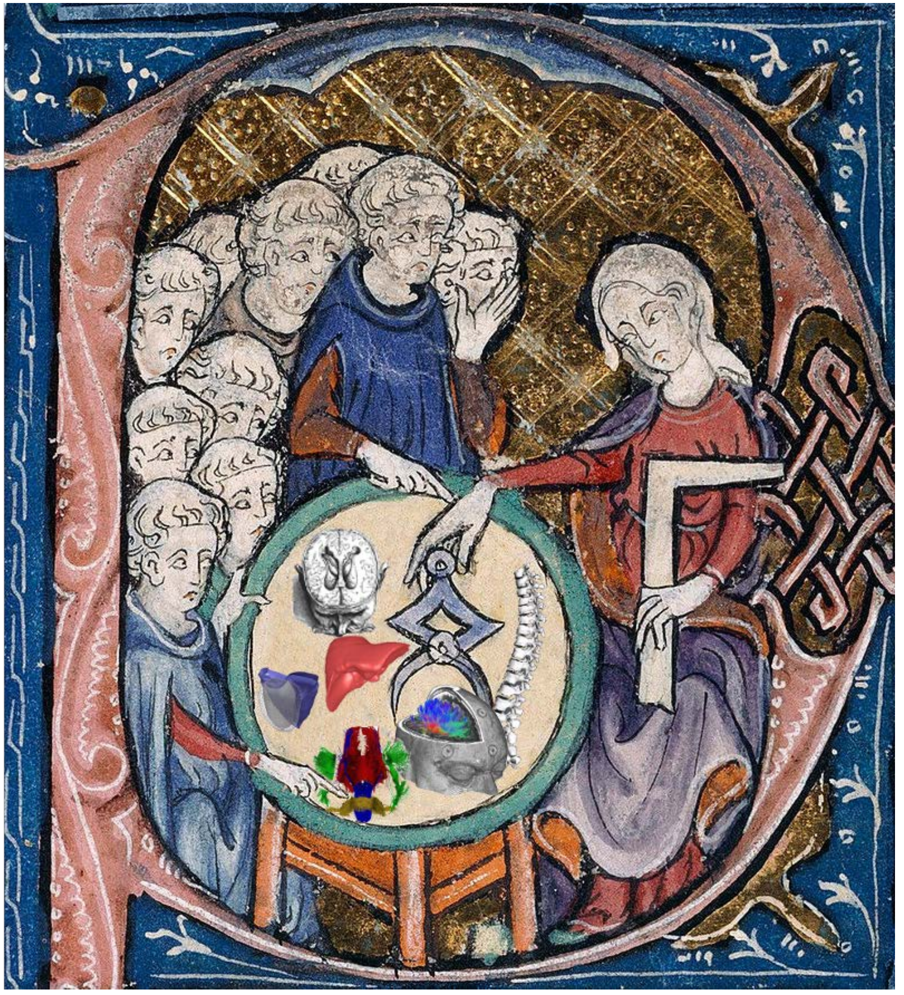
Xavier Pennec

Univ. Côte d’Azur and Inria, France

Geometric Statistics

Mathematical foundations and applications in computational anatomy



Freely adapted from “Women teaching geometry”, in Adelard of Bath translation of Euclid’s elements, 1310.

1/ Intrinsic Statistics on Riemannian Manifolds

Geometric Statistics workshop 09/2019

Epione e-patient / e-medicine

UNIVERSITÉ CÔTE D’AZUR

Inria

Collaborators

Researchers from Epione/Adclepios/Epidaure team

- Maxime Sermesant
- Nicholas Ayache

Former PhD students

- Jonathan Boisvert
- Pierre Fillard
- Vincent Arsigny
- Kristin McLeod
- Nina Miolane
- Loic Devillier
- Marc-Michel Rohé
- Tom Vercauteren
- Stanley Durrleman
- Marco Loreni
- Christof Seiler
- …..

Current PhD students

- Yan Thanwerdas
- Nicolas Guigui
- Shuma Jia

Geometric Statistics

PhDs and Post-docs available

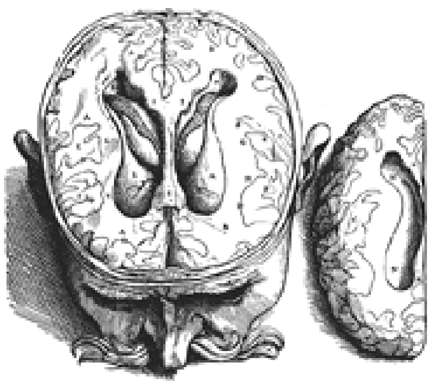
Epione
e-patient / e-medicine

UNIVERSITÉ
CÔTE D'AZUR

Inria

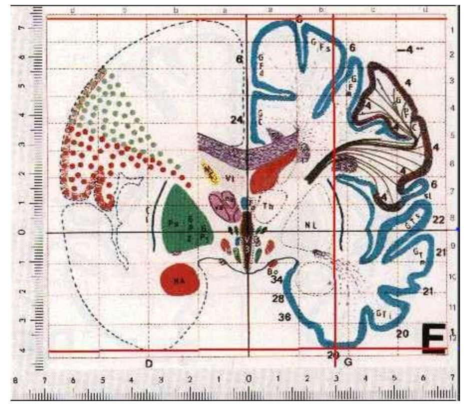
Anatomy

Science that studies the structure and the relationship in space of different organs and tissues in living systems [Hachette Dictionary]

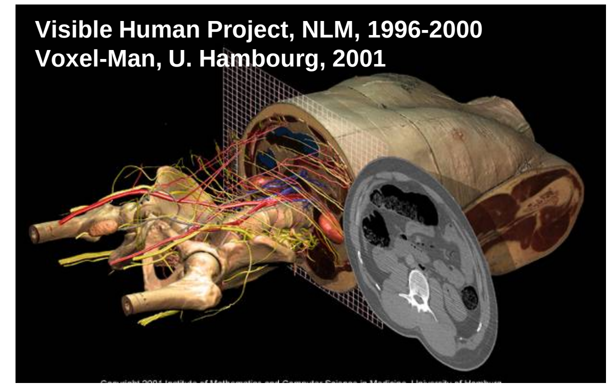


1er cerebral atlas, Vesale, 1543

Paré, 1585



Talairach & Tournoux, 1988

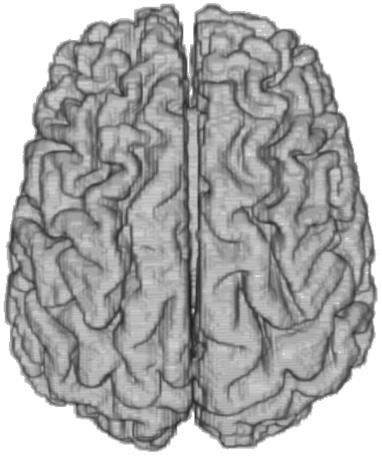
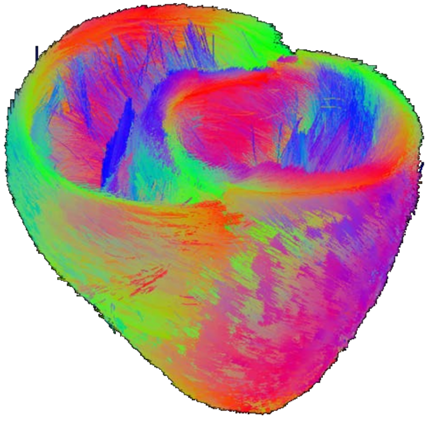
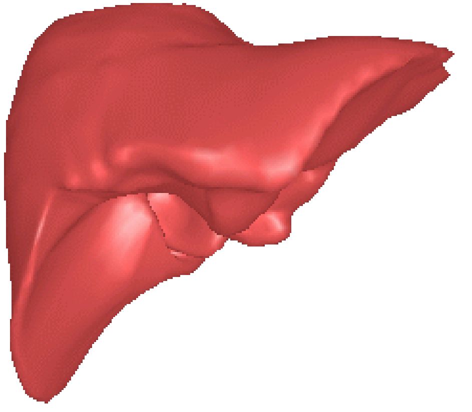


Galien (131-201) | **Vésale (1514-1564)** **Paré (1509-1590)** | **Sylvius (1614-1672)** **Willis (1621-1675)** | **Gall (1758-1828) :** ***Phrenology*** **Talairach (1911-2007)**

Revolution of observation means (~1990):

- From dissection to in-vivo in-situ imaging
- From the description of one representative individual to generative statistical models of the population

Computational Anatomy

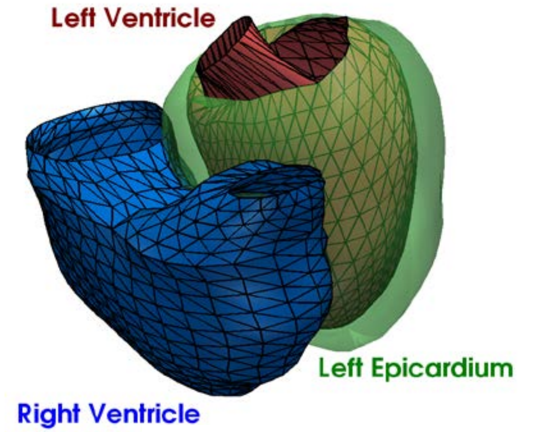
   

Statistics of organ shapes across subjects in species, populations, diseases…

- Mean shape (atlas), subspace of normal vs pathologic shapes
- Shape variability (Covariance)
- Model development across time (growth, ageing, ages…)

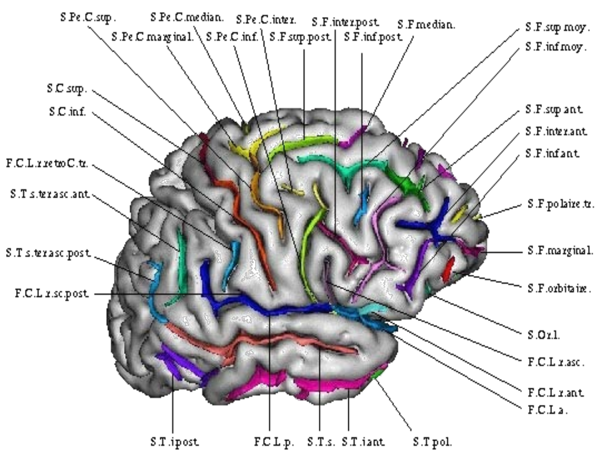
Use for personalized medicine (diagnostic, follow-up, etc)

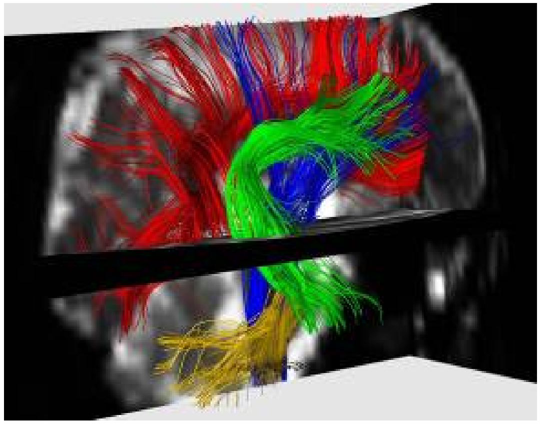
Geometric features in Computational Anatomy

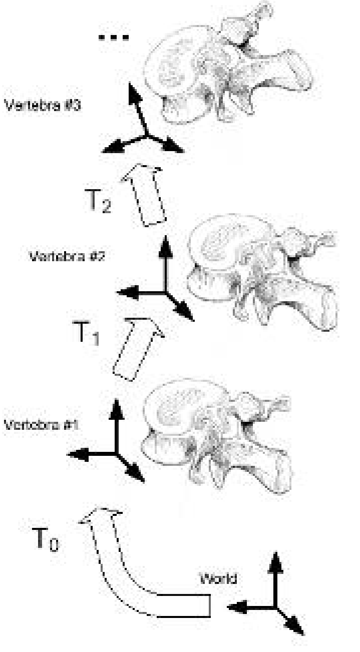


Noisy geometric features

- Curves, sets of curves (fiber tracts)
- Surfaces, SPD matrices
- Transformations







Statistical modeling at the population level

- **Simple Statistics on non-linear manifolds?**
  - Mean, covariance of its estimation, PCA, PLS, ICA
- **GS**: Statistics on manifolds vs **IG**: manifolds of statistical models

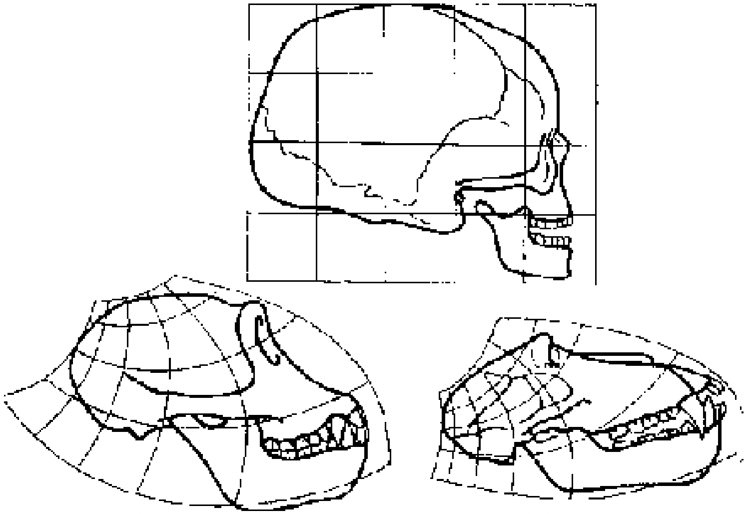
Methods of computational anatomy

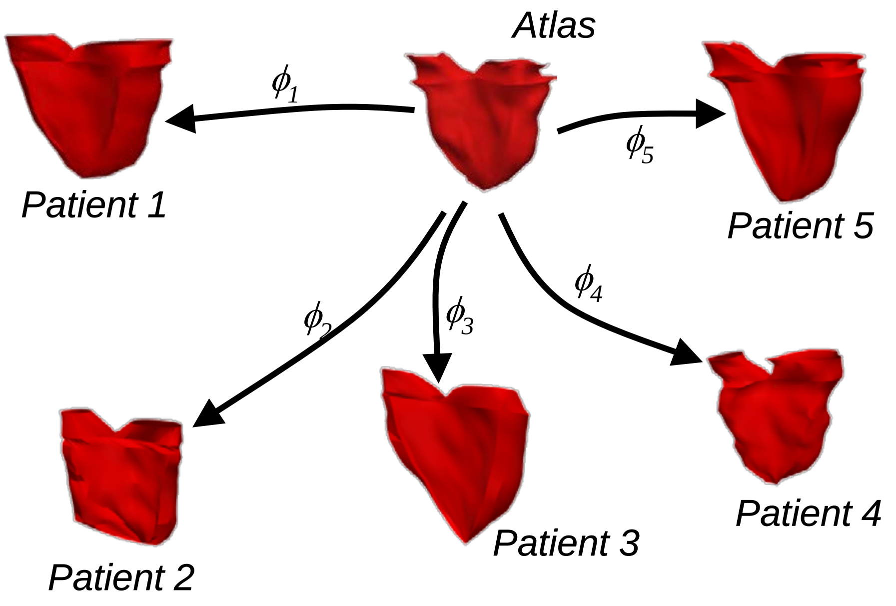
Remodeling of the right ventricle of the heart in tetralogy of Fallot

- Mean shape
- Shape variability
- Correlation with clinical variables
- Predicting remodeling effect

Shape of RV in 18 patients

Morphometry through Deformations





Measure of deformation [D'Arcy Thompson 1917, Grenander & Miller]

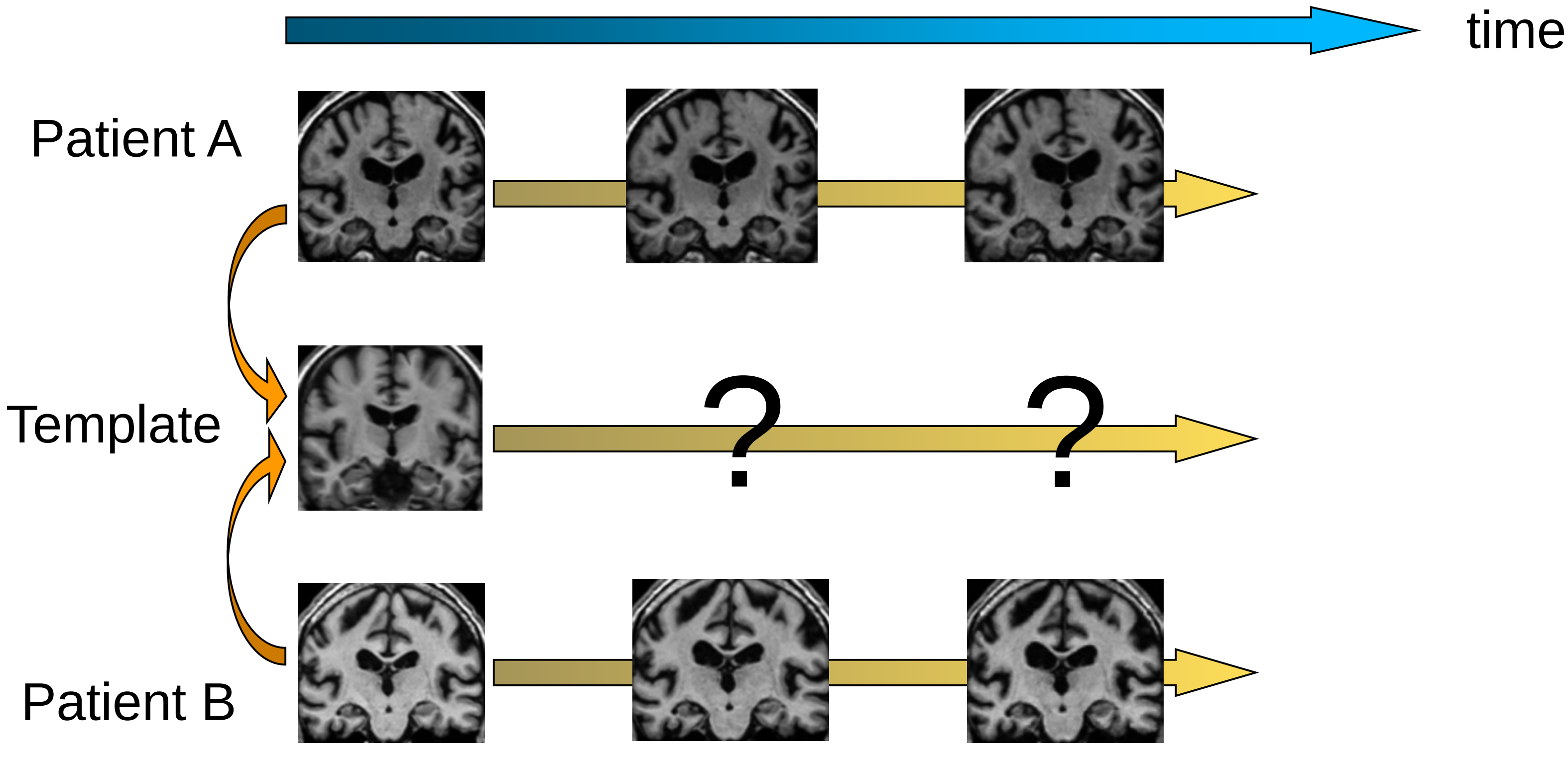
- Observation = "random" deformation of a reference template
- Reference template = Mean (atlas)
- Shape variability encoded by the deformations

Statistics on groups of transformations (Lie groups, diffeomorphism)?

Consistency with group operations (non commutative)?

Longitudinal deformation analysis

Dynamic obervations



How to transport longitudinal deformation across subjects?

What are the convenient mathematical settings?

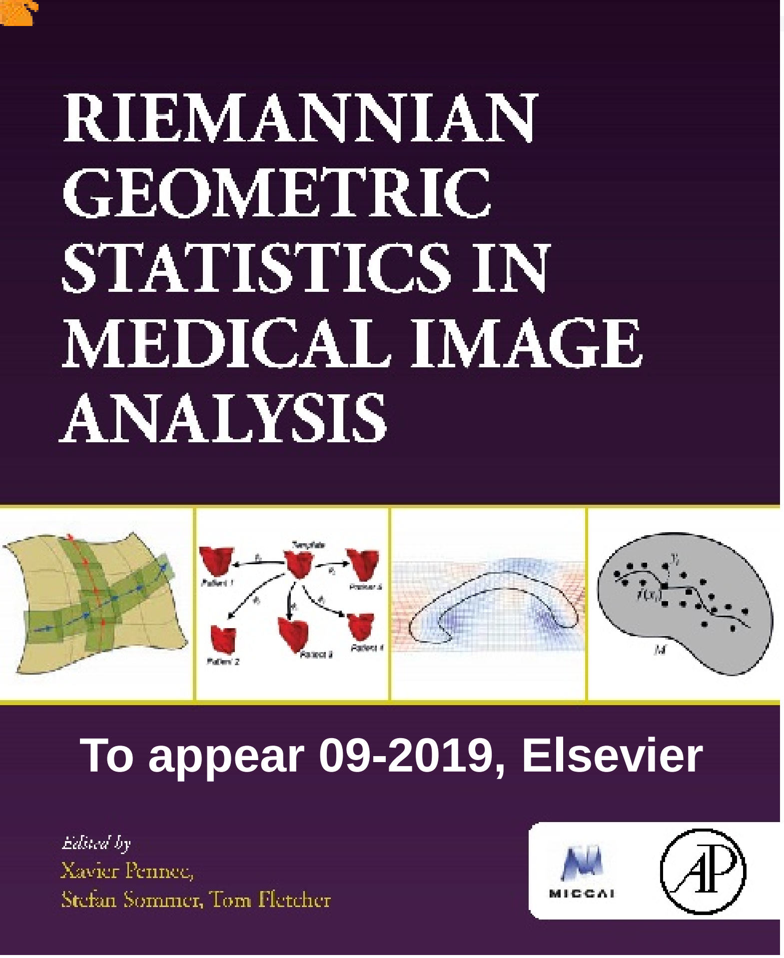
Geometric Statistics: Mathematical foundations and applications in computational anatomy

Intrinsic Statistics on Riemannian Manifolds

Metric and Affine Geometric Settings for Lie Groups

Parallel Transport to Analyze Longitudinal Deformations

Advances Statistics: CLT & PCA



Part 1: Foundations

- 1: Riemannian geometry [Sommer, Fetcher, Pennec]
- 2: Statistics on manifolds [Fletcher]
- 3: Manifold-valued image processing with SPD matrices [Pennec]
- 4: Riemannian Geometry on Shapes and Diffeomorphisms [Marsland, Sommer]
- 5: Beyond Riemannian: the affine connection setting for transformation groups [Pennec, Lorenzi]

Part 2: Statistics on Manifolds and Shape Spaces

- 6: Object Shape Representation via Skeletal Models (s-reps) and Statistical Analysis [Pizer, Maron]
- 7: Inductive Fréchet Mean Computation on S(n) and SO(n) with Applications [Chakraborty, Vemuri]
- 8: Statistics in stratified spaces [Ferage, Nye]
- 9: Bias in quotient space and its correction [Miolane, Devilier,Pennec]
- 10: Probabilistic Approaches to Statistics on Manifolds: Stochastic Processes, Transition Distributions, and Fiber Bundle Geometry [Sommer]
- 11: Elastic Shape Analysis, Square-Root Representations and Their Inverses [Zhang, Klassen, Srivastava]

Part 3: Deformations, Diffeomorphisms and their Applications

- 13: Geometric RKHS models for handling curves and surfaces in Computational Anatomy : currents, varifolds, f-shapes, normal cycles [Charlie, Charon, Glaunes, Gori, Roussillon]
- 14: A Discretize-Optimize Approach for LDDMM Registration [Polzin, Niethammer, Vialad, Modezitski]
- 15: Spatially varying metrics in the LDDMM framework [Vialard, Risser]
- 16: Low-dimensional Shape Analysis In the Space of Diffeomorphisms [Zhang, Fleche, Wells, Golland]
- 17: Diffeomorphic density matching, Bauer, Modin, Joshi]

Supports for the course

http://www-sop.inria.fr/asclepios/cours/Peyresq_2019/

- 1/ Intrinsic Statistics on Riemannian Manifolds
  - Introduction to differential and Riemannian geometry. **Chapter 1**, RGSMIA. Elsevier, 2019.
  - Intrinsic Statistics on Riemannian Manifolds: Basic Tools for Geometric Measurements. JMIV 2006.
- 2/ SPD matrices and manifold-valued image processing
  - Manifold-valued image processing with SPD matrices. **Chapter 3** RGSMIA. Elsevier, 2019.
  - Historical reference: A Riemannian Framework for Tensor Computing. IJCV 2006.
- 3/ Metric and affine geometric settings for Lie groups
  - Beyond Riemannian Geometry The affine connection setting for transformation groups Chapter 5, RGSMIA. Elsevier, 2019.
- 4/ Parallel transport to analyze longitudinal deformations
  - Geodesics, Parallel Transport and One-parameter Subgroups for Diffeomorphic Image Registration. IJCV 105(2), November 2013.
  - Parallel Transport with Pole Ladder: a Third Order Scheme…[arXiv:1805.11436]
- 5/ Advanced statistics: central limit theorem and extension of PCA
  - Curvature effects on the empirical mean in Riemannian and affine Manifolds [arXiv:1906.07418]
  - Barycentric Subspace Analysis on Manifolds. Annals of Statistics. 46(6A):2711-2746, 2018. [arXiv:1607.02833]

Geometric Statistics: Mathematical foundations and applications in computational anatomy

Intrinsic Statistics on Riemannian Manifolds

- Introduction to computational anatomy
- The Riemannian manifold computational structure
- Simple statistics on Riemannian manifolds
- Applications to the spine shape and registration accuracy

Metric and Affine Geometric Settings for Lie Groups

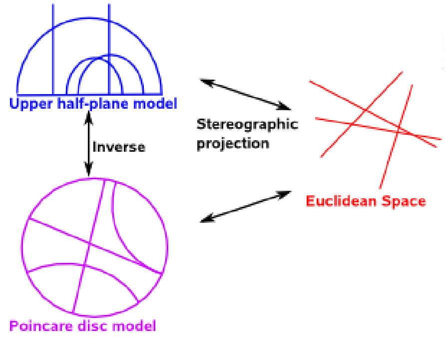
Parallel Transport to Analyze Longitudinal Deformations

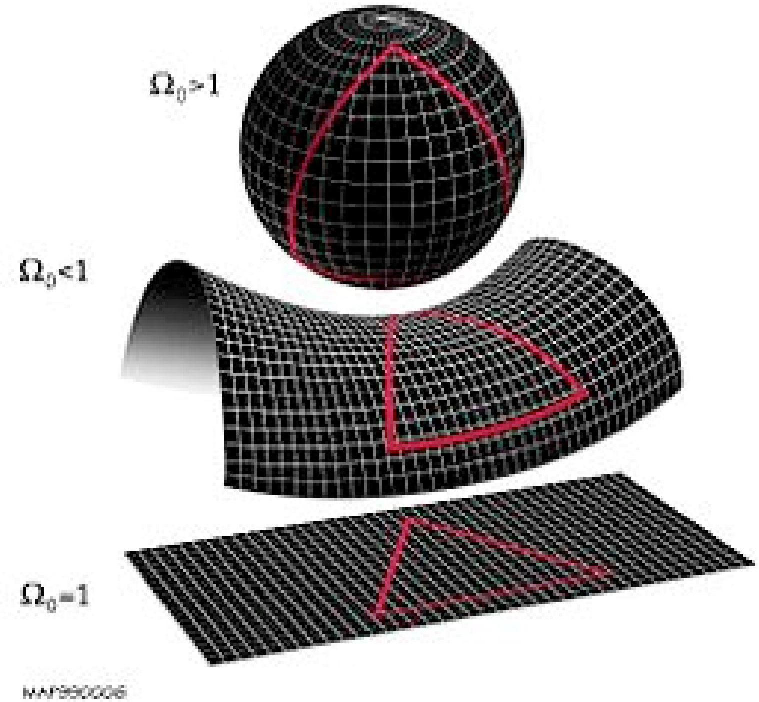
Advances Statistics: CLT & PCA

Which non-linear space?

Constant curvatures spaces

- Sphere,
- Euclidean,
- Hyperbolic





Homogeneous spaces, Lie groups and symmetric spaces

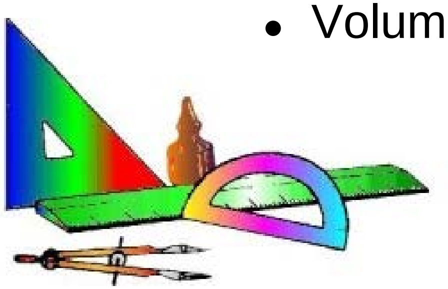
Riemannian or affine connection spaces

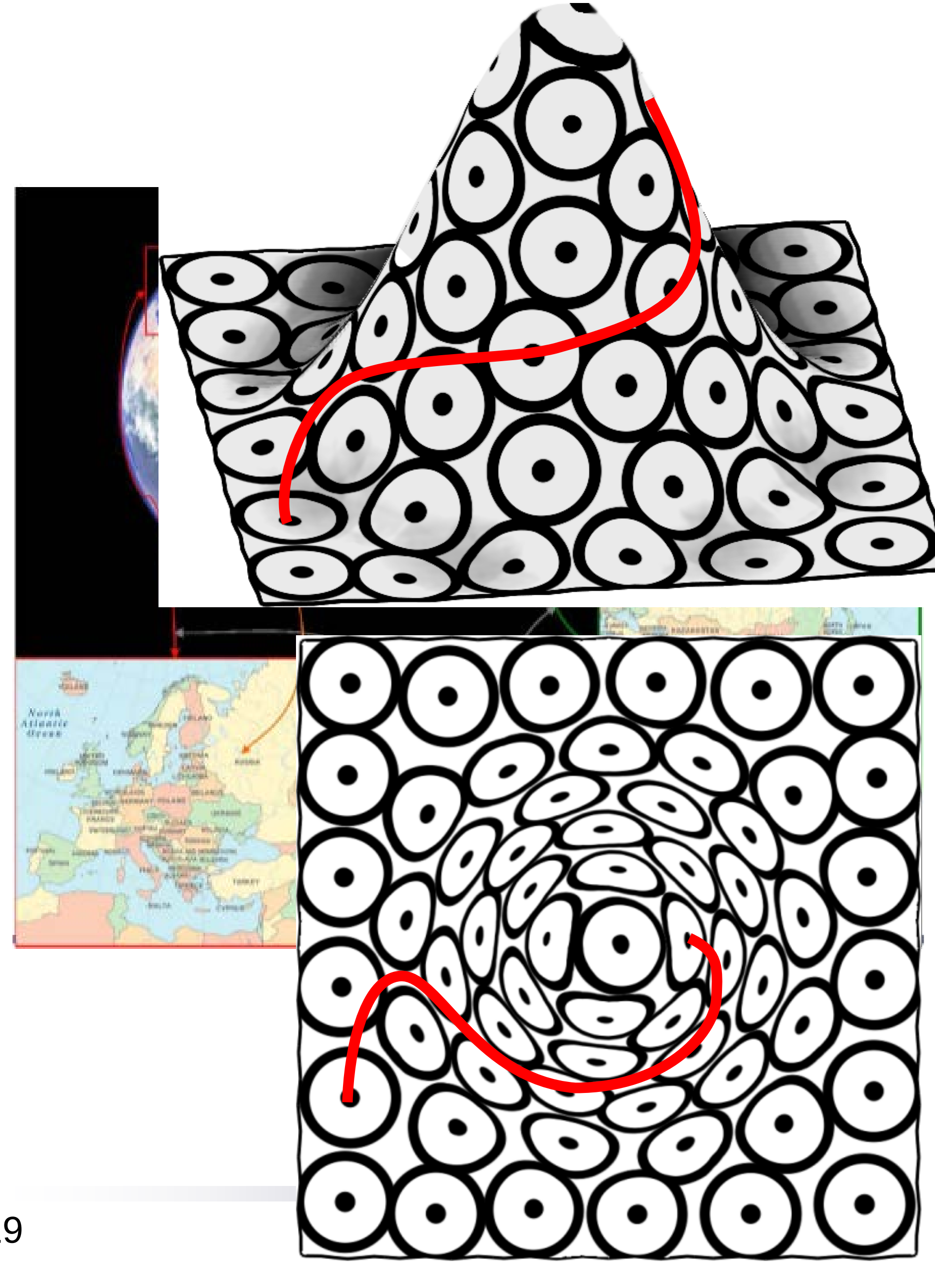
Towards non-smooth quotient and stratified spaces

Differentiable manifolds

Computing on a manifold

- Extrinsic
  - Embedding in $\mathbb{R}^n$
- Intrinsic
  - Coordinates : charts
- Measuring?
  - Lengths
  - Straight lines
  - Volumes

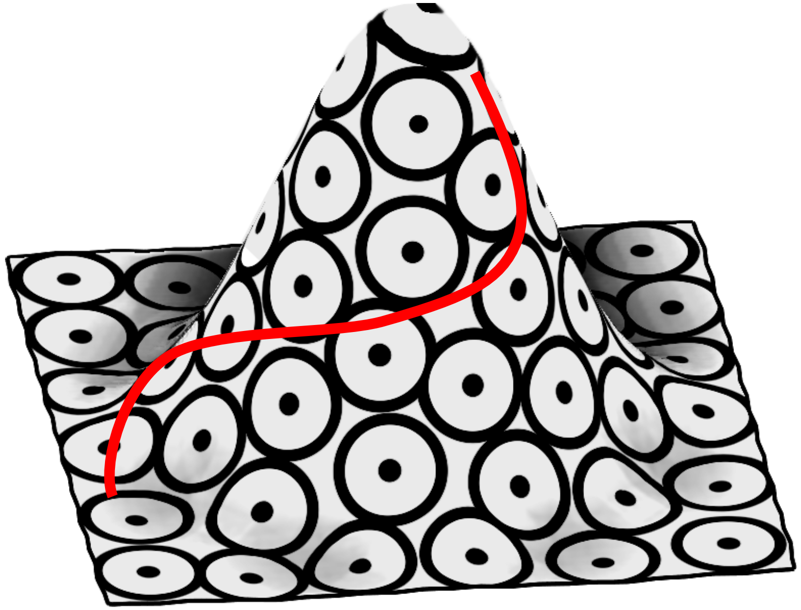




Measuring extrinsic distances

Basic tool: the scalar product

$$< v, w >= v^t w$$

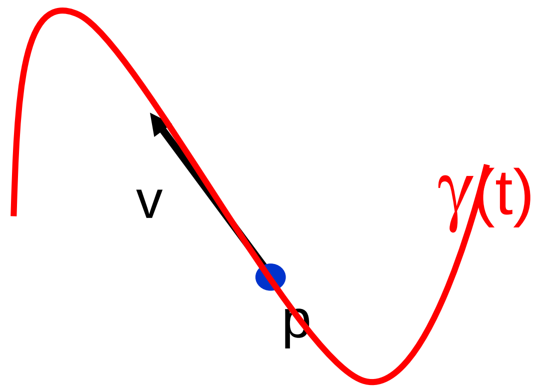


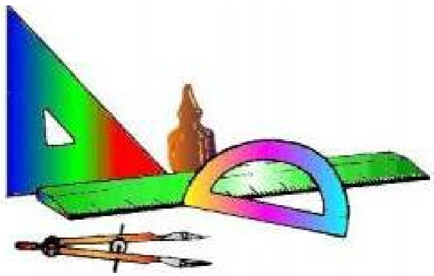
- Norm of a vector

$$\|v\| = \sqrt{< v, v >}$$

- Length of a curve

$$L(\gamma) = \int \|\dot{\gamma}(t)\| \, dt$$



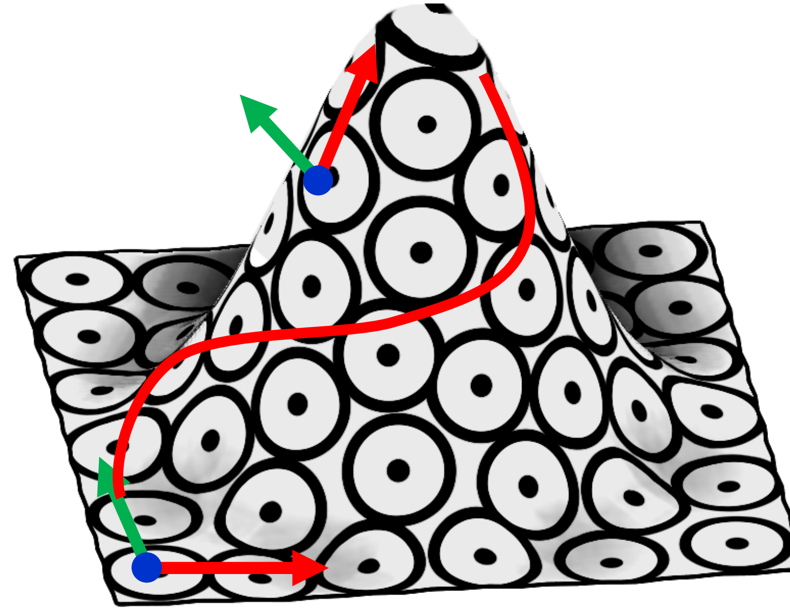


Measuring extrinsic distances

Basic tool: the scalar product

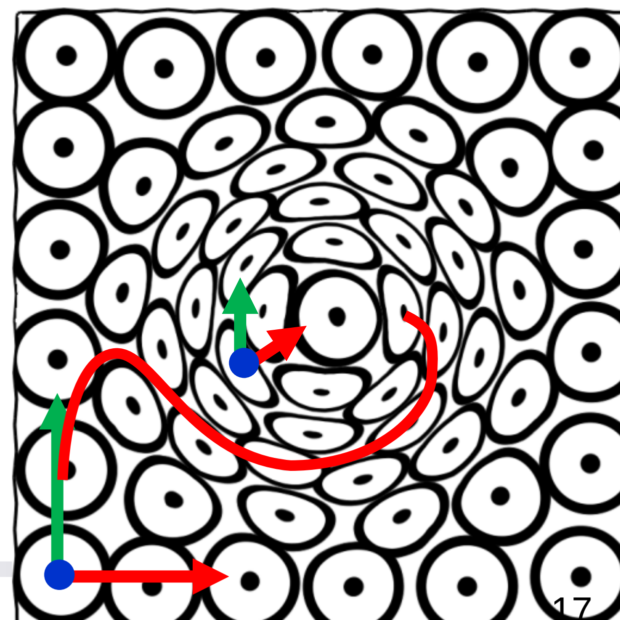
Bernhard Riemann
1826-1866

$$< \mathbf{v},\mathbf{w} >_p = \mathbf{v}^t \mathbf{w} = v^t G(p) w$$



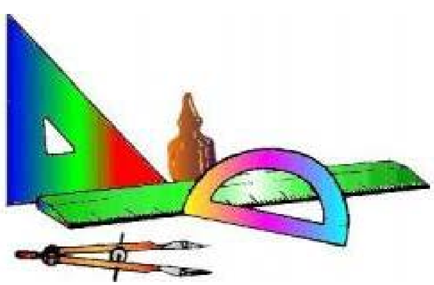
- Norm of a vector

$$\|v\|_p = \sqrt{< v,v >_p}$$



- Length of a curve

$$L(\gamma) = \int \|\dot{\gamma}(t)\| \, dt$$

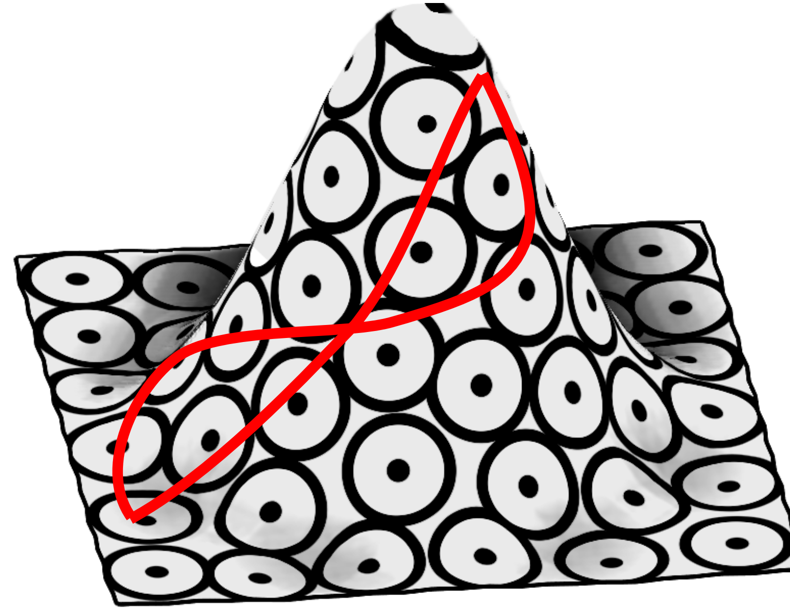


Riemannian manifolds

Basic tool: the scalar product

Bernhard Riemann
1826-1866

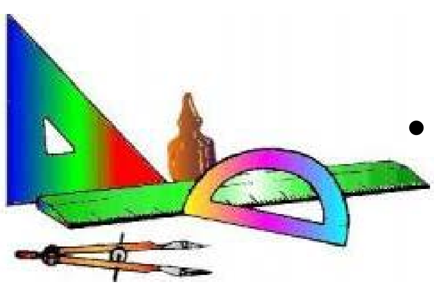
$$< v, w >_p = v^t G(p) w$$

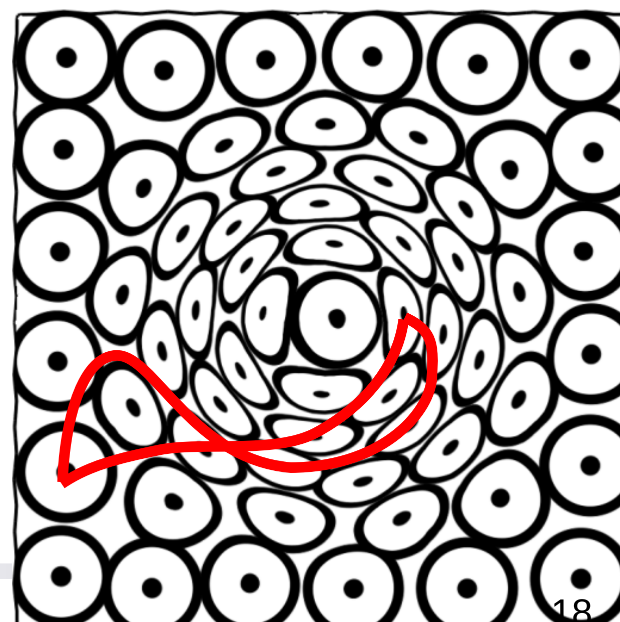


- Geodesics
  - Shortest path between 2 points
  - Calculus of variations (E.L.) : 2^{nd} order differential equation (specifies acceleration)
  - Free parameters: initial speed and starting point

- Length of a curve

$$L(\gamma) = \int \|\dot{\gamma}(t)\| dt$$



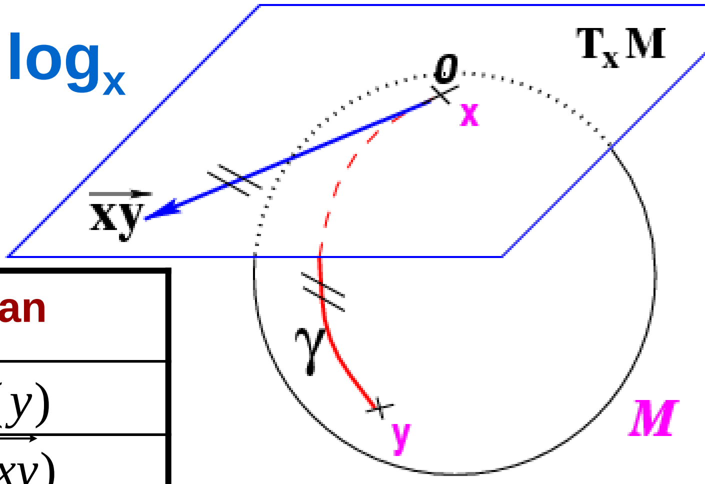


Exponential map (Normal coordinate system):

- **Exp_x** = geodesic shooting parameterized by the initial tangent
- **Log_x** = unfolding the manifold in the tangent space along geodesics
  - Geodesics = straight lines with Euclidean distance
  - Geodesic completeness: covers M \ Cut(x)

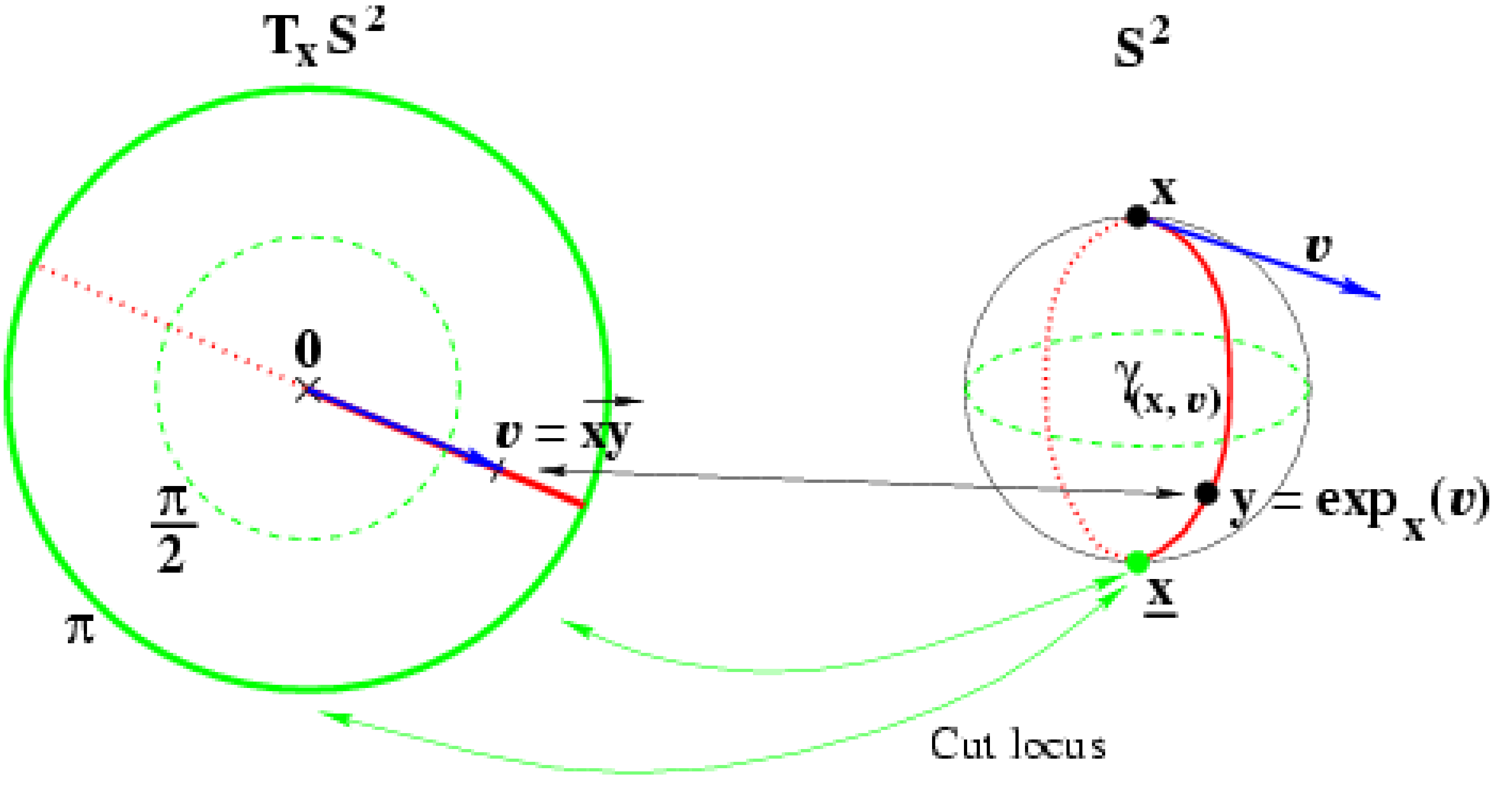
Reformulate algorithms with $\exp_x$ and $\log_x$

Vector -> Bi-point (no more equivalence classes)



Operation	Euclidean space	Riemannian
Subtraction	$\overrightarrow{xy} = y - x$	$\overrightarrow{xy} = Log_x(y)$
Addition	$y = x + \overrightarrow{xy}$	$y = Exp_x(\overrightarrow{xy})$
Distance	$\mathrm{dist}(x,y) = \lVert y - x \rVert$	$\mathrm{dist}(x,y) = \lVert \overrightarrow{xy} \rVert_x$
Gradient descent	$x_{t+\varepsilon} = x_t - \varepsilon \nabla C(x_t)$	$x_{t+\varepsilon} = Exp_{x_t}(-\varepsilon \nabla C(x_t))$

Cut locus



$T_x S^2$
S^2
x
v
0
$v = \overrightarrow{xy}$
$\gamma_{(x,v)}$
$\frac{\pi}{2}$
π
$y = \exp_x(v)$
$\underline{x}$
Cut locus

Geometric Statistics: Mathematical foundations and applications in computational anatomy

Intrinsic Statistics on Riemannian Manifolds

- Introduction to computational anatomy
- The Riemannian manifold computational structure
- Simple statistics on Riemannian manifolds
- Applications to the spine shape and registration accuracy

Metric and Affine Geometric Settings for Lie Groups

Parallel Transport to Analyze Longitudinal Deformations

Advances Statistics: CLT & PCA

Basic probabilities and statistics

Measure: random vector x of pdf $p_{\mathbf{x}}(z)$

Approximation: $\mathbf{x} \sim (\overline{\mathrm{x}}, \mathbf{\Sigma}_{\mathbf{xx}})$

- Mean: $\overline{\mathrm{x}} = \mathrm{E}(\,\mathbf{x}\,) = \int z.p_{\mathbf{x}}(z).dz$
- Covariance: $\Sigma_{\mathbf{xx}} = \mathrm{E}\left[(\mathbf{x}-\overline{\mathrm{x}}).(\mathbf{x}-\overline{\mathrm{x}})^T\right]$

Propagation:

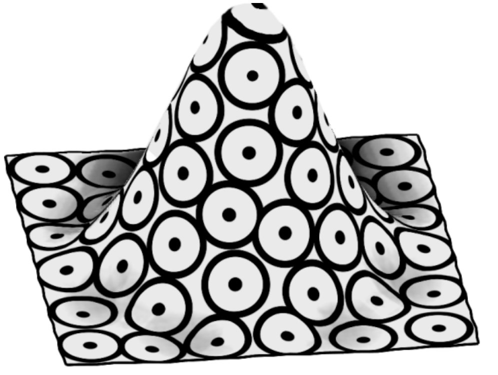
$$\mathbf{y} = h(\mathbf{x}) \sim \left(h(\overline{\mathrm{x}}),\ \frac{\partial h}{\partial \mathrm{x}}.\mathbf{\Sigma}_{\mathbf{xx}}.\frac{\partial h}{\partial \mathrm{x}}^{\mathrm{T}} \right)$$

Noise model: additive, Gaussian...

Principal component analysis

Statistical distance: Mahalanobis and χ^2

Random variable in a Riemannian Manifold



Intrinsic pdf of x

- For every set H

$$P(\mathbf{x} \in H) = \int_H p(y) dM(y)$$

- ~~Lebesgue's measure~~

→ Uniform Riemannian Mesure $dM(y) = \sqrt{\det(G(y))}\, dy$

Expectation of an observable in M

- $\mathbf{E_x}[\phi] = \int_M \phi(y) p(y) dM(y)$
- $\phi = dist^2$ (variance) : $\mathbf{E_x}[dist(.,y)^2] = \int_M dist(y,z)^2 p(z) dM(z)$
- $\phi = \log(p)$ (information) : $\mathbf{E_x}[\log(p)] = \int_M p(y) \log(p(y)) dM(y)$
- ~~$\phi = x$ (mean) : $\mathbf{E_x}[\mathbf{x}] = \int_M y\, p(y) dM(y)$~~

First statistical tools

Moments of a random variable: tensor fields

- $\mathfrak{M}_1(x) = \int_M \overrightarrow{xz}\, P(dz)$ Tangent mean: (0,1) tensor field
- $\mathfrak{M}_2(x) = \int_M \overrightarrow{xz} \otimes \overrightarrow{xz}\, P(dz)$ Covariance: (0,2) tensor field
- $\mathfrak{M}_k(x) = \int_M \overrightarrow{xz} \otimes \overrightarrow{xz} \otimes \cdots \otimes \overrightarrow{xz}\, P(dz)$ k-contravariant tensor field

Fréchet mean set

- Integral only valid in Hilbert/Wiener spaces [Fréchet 44]
- $\sigma^2(x) = Tr_g\big(\mathfrak{M}_2(x)\big) = \int_M dist^2(x,z)\, P(dz)$
- **Fréchet mean** [1948] = global minima
- **Exponential barycenters** [Emery & Mokobodzki 1991]

  $\mathfrak{M}_1(\bar{x}) = \int_M \overrightarrow{\bar{x}z}\, P(dz) = 0$ [critical points if P(C) =0]

Maurice Fréchet
(1878-1973)

Fréchet expectation (1944)

Minimizing the variance

$$\mathrm{E}[\mathbf{x}] = \underset{y \in \mathrm{M}}{\operatorname{argmin}} \left(\mathrm{E}\left[\operatorname{dist}(y, \mathbf{x})^2 \right] \right)$$

Existence

- Finite variance at one point

Characterization as an exponential barycenter (P(C)=0)

$$\operatorname{grad}\left(\sigma_{\mathbf{x}}^2(y)\right) = 0 \quad \Rightarrow \quad \mathrm{E}\left[\overrightarrow{\bar{x}\mathbf{x}}\right] = \int_{\mathrm{M}} \overrightarrow{\bar{x}\mathbf{x}}.p_{\mathbf{x}}(z).d\mathrm{M}(z) = 0$$

Uniqueness **Karcher 77 / Kendall 90 / Afsari 10 / Le 10**

- Unique Karcher mean (thus Fréchet) if distribution has support in a regular geodesic ball with radius $r < r^* = \frac{1}{2}\min(inj(M), \pi/\sqrt{\kappa})$ (k upper bound on sectional curvatures on M)
- Empirical mean: a.s. uniqueness [Arnaudon & Miclo 2013]

Other central primitives

$$\mathrm{E}^{\alpha}[\mathbf{x}] = \underset{y \in \mathrm{M}}{\operatorname{argmin}} \left(\mathrm{E}\left[\operatorname{dist}(y, \mathbf{x})^{\alpha} \right] \right)^{1/\alpha}$$

A gradient descent (Gauss-Newton) algorithm

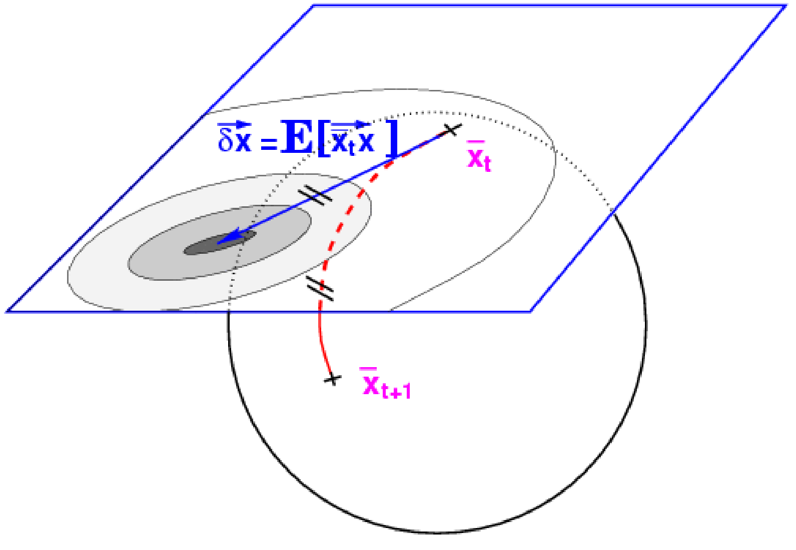
Vector space

$$f(x+v) = f(x) + \nabla f^T.v + \frac{1}{2} v^T.H_f.v$$

$$x_{t+1} = x_t + v \quad \text{with} \quad v = -H_f^{(-1)}.\nabla f$$

Manifold

$$f(\exp_x(v)) = f(x) + \nabla f(v) + \frac{1}{2} H_f(v,v)$$



$$\nabla\left(\sigma_{\mathbf{x}}^2(\mathrm{y})\right) = -2\,\mathrm{E}\left[\overrightarrow{\mathrm{y}\mathbf{x}}\right] = \frac{-2}{n} \sum_i \overrightarrow{\mathrm{y}\mathrm{x}_i}$$

$$H_{\sigma_{\mathbf{x}}^2} \approx 2Id$$

Geodesic marching

$$\bar{\mathrm{x}}_{t+1} = \exp_{\bar{\mathrm{x}}_t}(v) \quad \text{with} \quad v = \mathrm{E}\left[\overrightarrow{\mathrm{y}\mathbf{x}}\right]$$

Example on 3D rotations

Space of rotations SO(3):

- Manifold: $R^T.R=Id$ and $\det(R)=+1$
- Lie group ($R_1 \circ R_2 = R_1.R_2$ & Inversion: $R^{(-1)} = R^T$)

Metrics on SO(3): compact space, there exists a bi-invariant metric

- Left / right invariant / induced by ambient space $\langle X, Y \rangle = Tr(X^T Y)$

Group exponential

- One parameter subgroups = bi-invariant Geodesic starting at Id
  - Matrix exponential and Rodrigue's formula: $R=\exp(X)$ and $X = \log(R)$
- Geodesic everywhere by left (or right) translation

$$Log_R(U) = R \log(R^T U) \qquad Exp_R(X) = R \exp(R^T X)$$

Bi-invariant Riemannian distance

- $d(R,U) = \|\log(R^T U)\| = \theta(R^T U)$

Example with 3D rotations

Principal chart: rotation vector : $r = \theta . n$

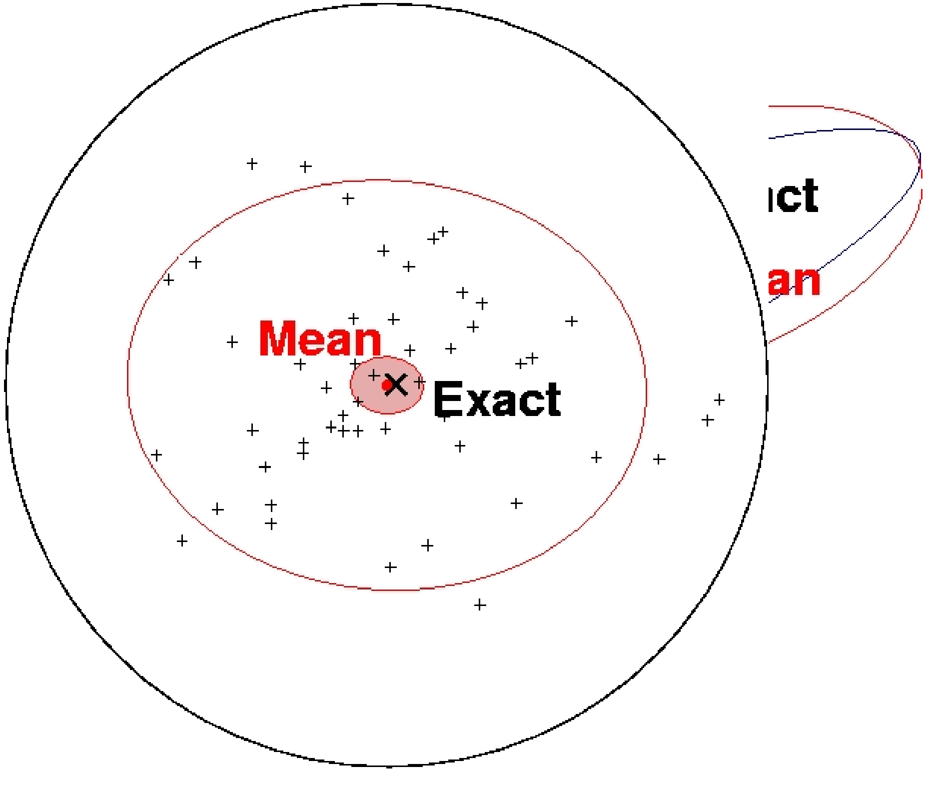
Distance: $\text{dist}(R_1, R_2) = \left\| r_1^{(-1)} \circ r_2 \right\|$

Frechet mean:

$$\overline{R} = \arg \min_{R \in SO_3} \left(\sum_{\text{i}} \text{dist}(R, R_i) \right)$$

Centered chart:

mean = barycenter



Distributions for parametric tests

Uniform density:

- maximal entropy knowing X $$p_{\mathbf{x}}(z) = \mathrm{Ind}_X(z) / \mathrm{Vol}(X)$$

Generalization of the Gaussian density:

- Stochastic heat kernel p(x,y,t) [complex time dependency]
- Wrapped Gaussian [Infinite series difficult to compute]
- **Maximal entropy knowing the mean and the covariance**

$$N(y) = k.\exp\left(\left(\overrightarrow{\bar{\mathrm{x}}\mathbf{x}}\right)^{\mathrm{T}}.\mathbf{\Gamma}.\left(\overrightarrow{\bar{\mathrm{x}}\mathbf{x}}\right)/\,2\right)$$

$$\mathbf{\Gamma} = \mathbf{\Sigma}^{(-1)} - \tfrac{1}{3}\mathrm{Ric} + O(\sigma) + \varepsilon(\sigma / r)$$

$$k = (2\pi)^{-n/2}.\det(\mathbf{\Sigma})^{-1/2}.\left(1 + O(\sigma^3) + \varepsilon(\sigma / r)\right)$$

Mahalanobis D2 distance / test:

$$\mu_{\mathbf{x}}^2(\mathrm{y}) = \overrightarrow{\bar{\mathrm{x}}\mathrm{y}}^{t}.\Sigma_{\mathbf{xx}}^{(-1)}.\overrightarrow{\bar{\mathrm{x}}\mathrm{y}}$$

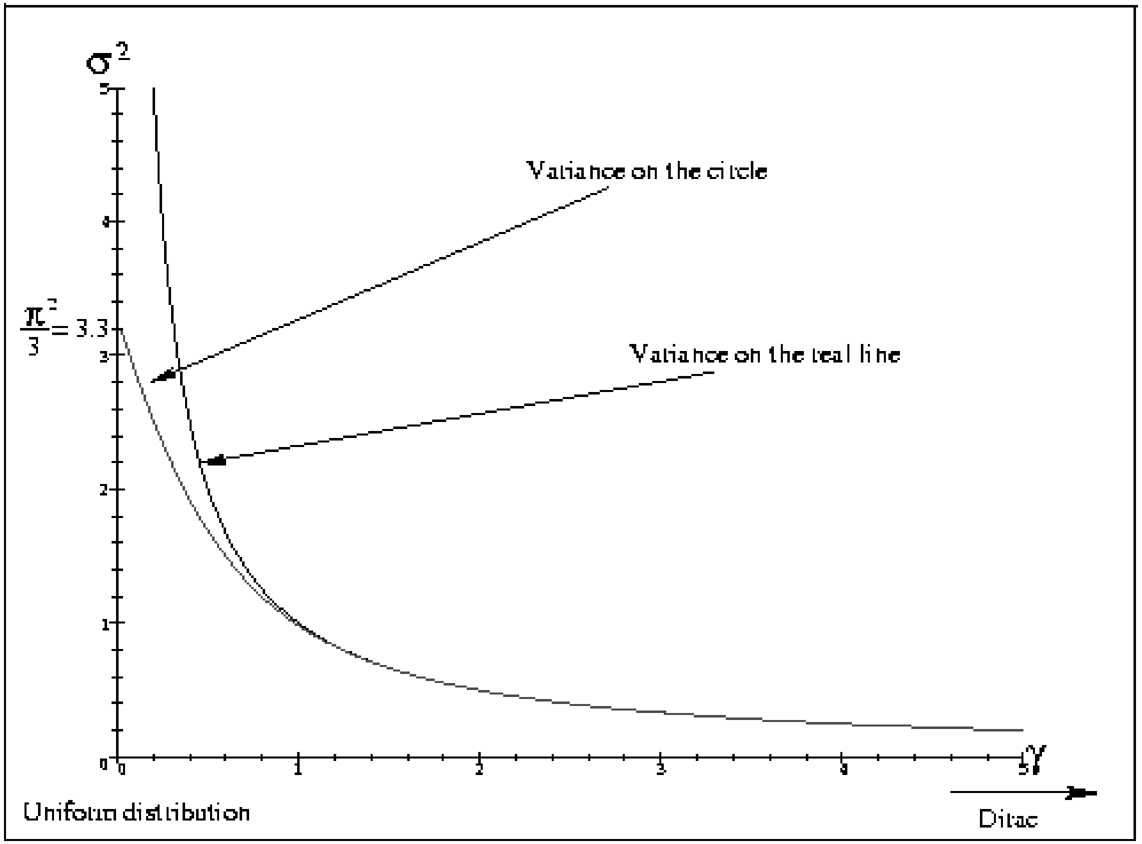
- Any distribution: $$\mathrm{E}\left[\mu_{\mathbf{x}}^2(\mathbf{x})\right] = n$$
- Gaussian: $$\mu_{\mathbf{x}}^2(\mathbf{x}) \propto \chi_n^2 + O(\sigma^3) + \varepsilon(\sigma / r)$$

[Pennec, JMIV06, NSIP'99]

Gaussian on the circle

Exponential chart: $x = r\theta \in \left]-\pi.r\,;\pi.r\right[$

Gaussian: truncated standard Gaussian



$r \to \infty$: standard Gaussian
(Ricci curvature → 0)

$\gamma \to 0$: uniform pdf with

$$\sigma^2 = (\pi.r)^2 / 3$$

(compact manifolds)

$\gamma \to \infty$: Dirac

tPCA vs PGA

tPCA

- Generative model: Gaussian
- Find the subspace that best explains the variance
  - → Maximize the squared distance to the mean

PGA (Fletcher 2004, Sommer 2014)

- Generative model:
  - Implicit uniform distribution within the subspace
  - Gaussian distribution in the vertical space
- Find a low dimensional subspace (geodesic subspaces?) that minimizes the error
  - → Minimize the squared Riemannian distance from the measurements to that sub-manifold (no closed form)

Different models in curved spaces (no Pythagore thm)

Extension to BSA tomorrow

Geometric Statistics: Mathematical foundations and applications in computational anatomy

Intrinsic Statistics on Riemannian Manifolds

- Introduction to computational anatomy
- The Riemannian manifold computational structure
- Simple statistics on Riemannian manifolds
- Applications to the spine shape and registration accuracy

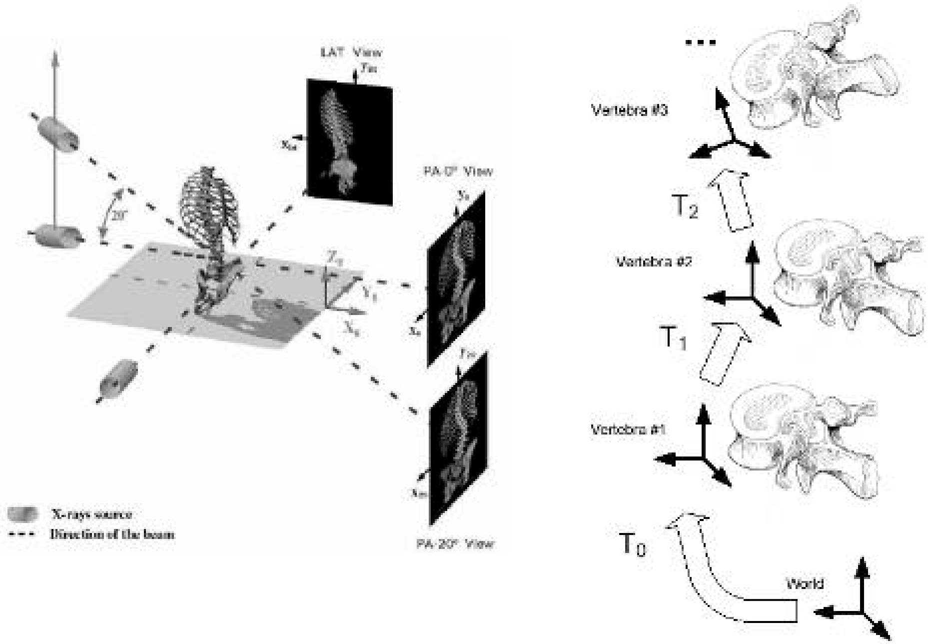
Metric and Affine Geometric Settings for Lie Groups

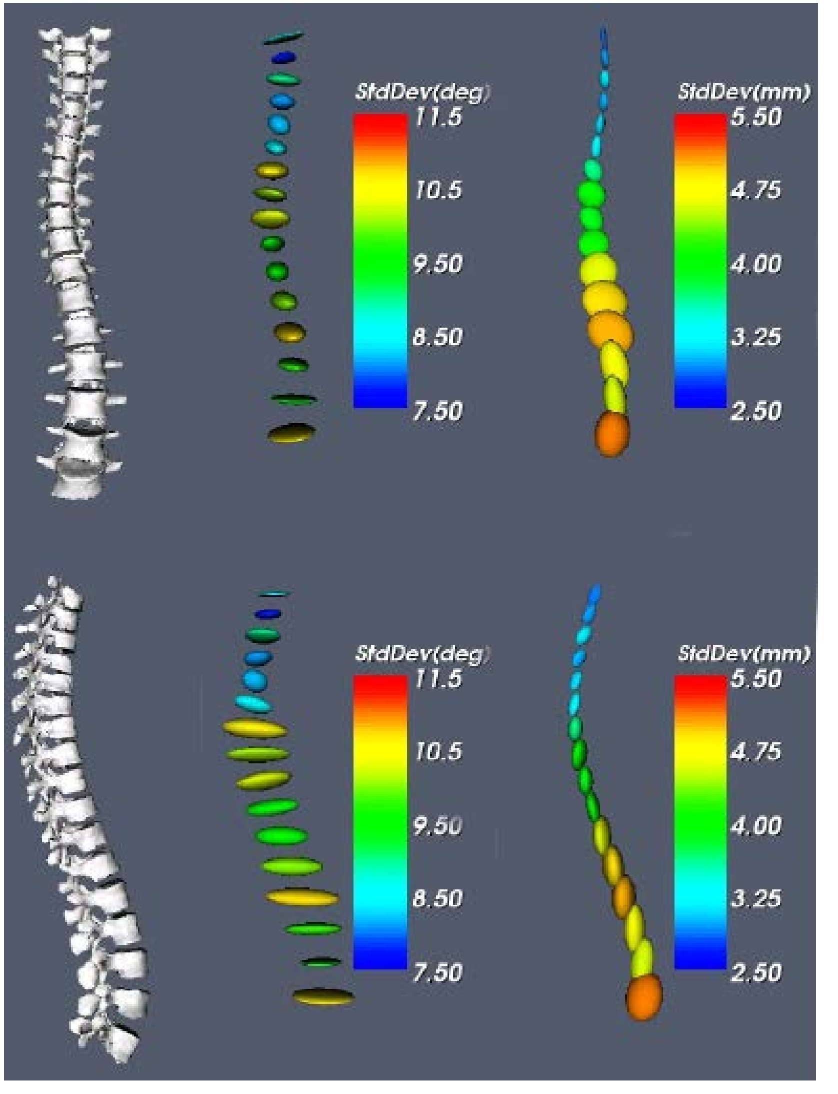
Parallel Transport to Analyze Longitudinal Deformations

Advances Statistics: CLT & PCA

Statistical Analysis of the Scoliotic Spine

[J. Boisvert et al. ISBI’06, AMDO’06 and IEEE TMI 27(4), 2008]





Database

- 307 Scoliotic patients from the Montreal’s Sainte-Justine Hospital.
- 3D Geometry from multi-planar X-rays

Mean

- Main translation variability is axial (growth?)
- Main rot. var. around anterior-posterior axis

Statistical Analysis of the Scoliotic Spine

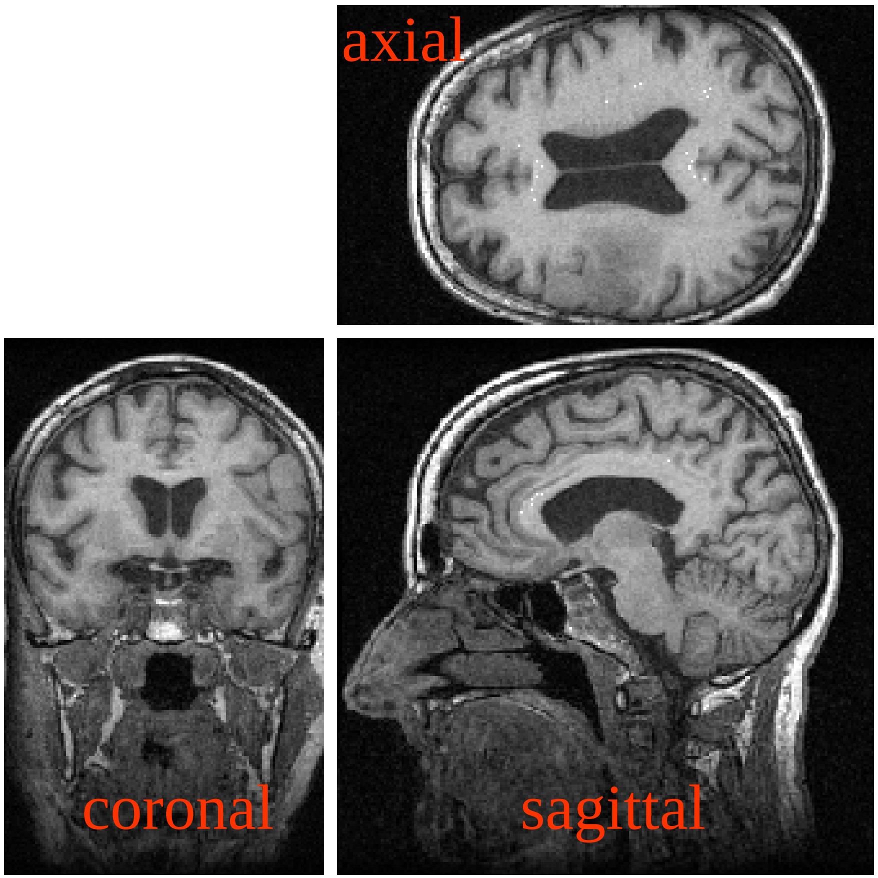
[J. Boisvert et al. ISBI'06, AMDO'06 and IEEE TMI 27(4), 2008]
AMDO'06 best paper award, Best French-Quebec joint PhD 2009

PCA of the Covariance:
4 first variation modes have clinical meaning

- Mode 1: King's class I or III
- Mode 2: King's class I, II, III
- Mode 3: King's class IV + V
- Mode 4: King's class V (+II)

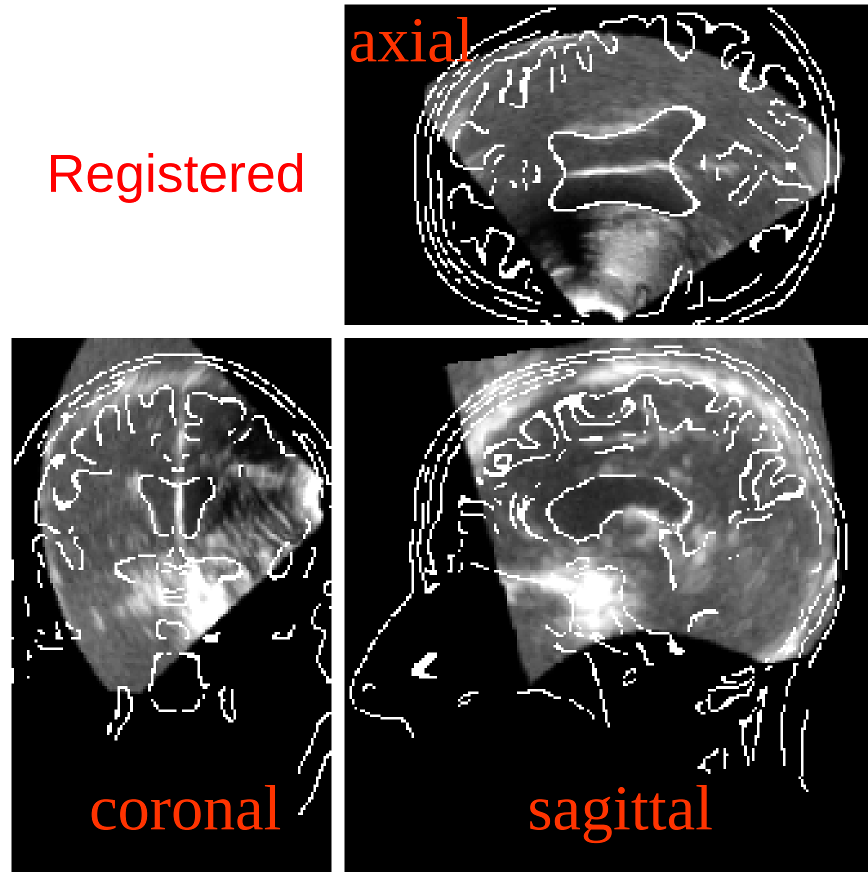
Typical Registration Result with Bivariate Correlation Ratio

Pre - Operative MR Image



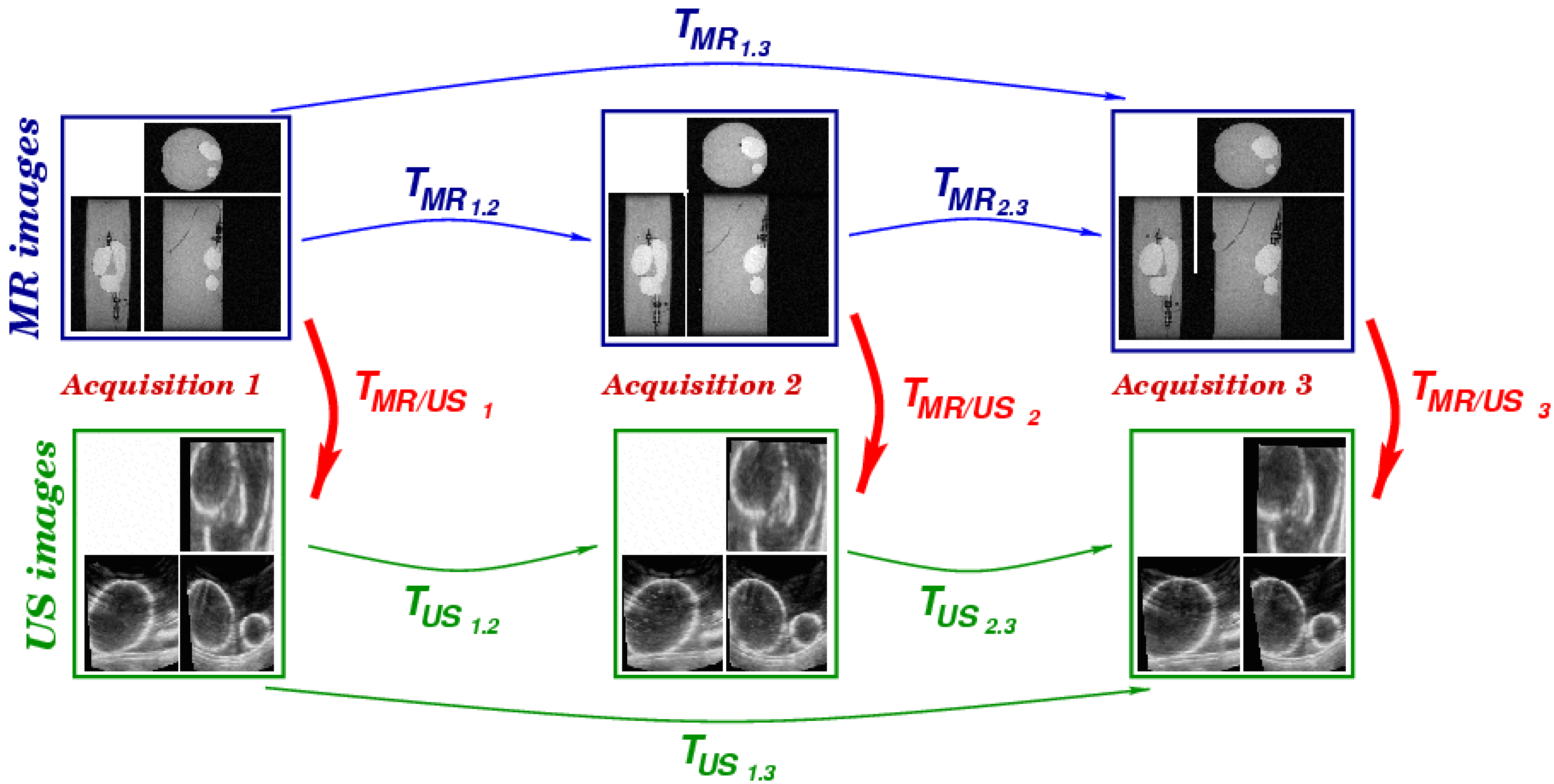
Per - Operative US Image

Registered



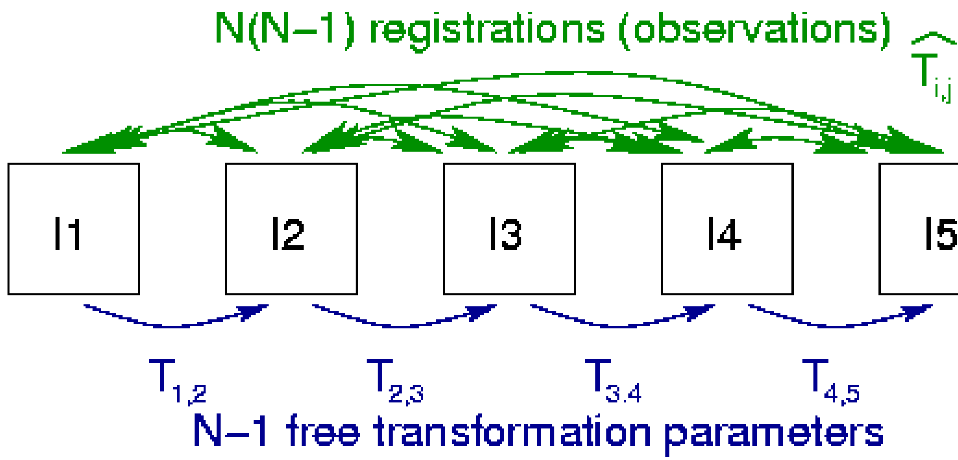
Acquisition of images : L. & D. Auer, M. Rudolf

Accuracy Evaluation (Consistency)



$$\sigma^2_{loop} = 2\sigma^2_{MR/US} + \sigma^2_{MR} + \sigma^2_{US}$$

Bronze Standard Rigid Registration Validation



Best explanation of the observations (ML) : $C = \sum_{ij} d^2(T_{ij}, \hat{T}_{ij})$

- LSQ criterion
- Robust Fréchet mean $d^2(T_1, T_2) = \min\left(\mu^2(T_1, T_2), \chi^2\right)$
- Robust initialization and Newton gradient descent

Result $T_{i,j}, \sigma_{rot}, \sigma_{trans}$

[T. Glatard & al, MICCAI 2006,

Int. Journal of HPC Apps, 2006]

Derive tests on transformations for accuracy / consistency

Results on per-operative patient images

Data (per-operative US)

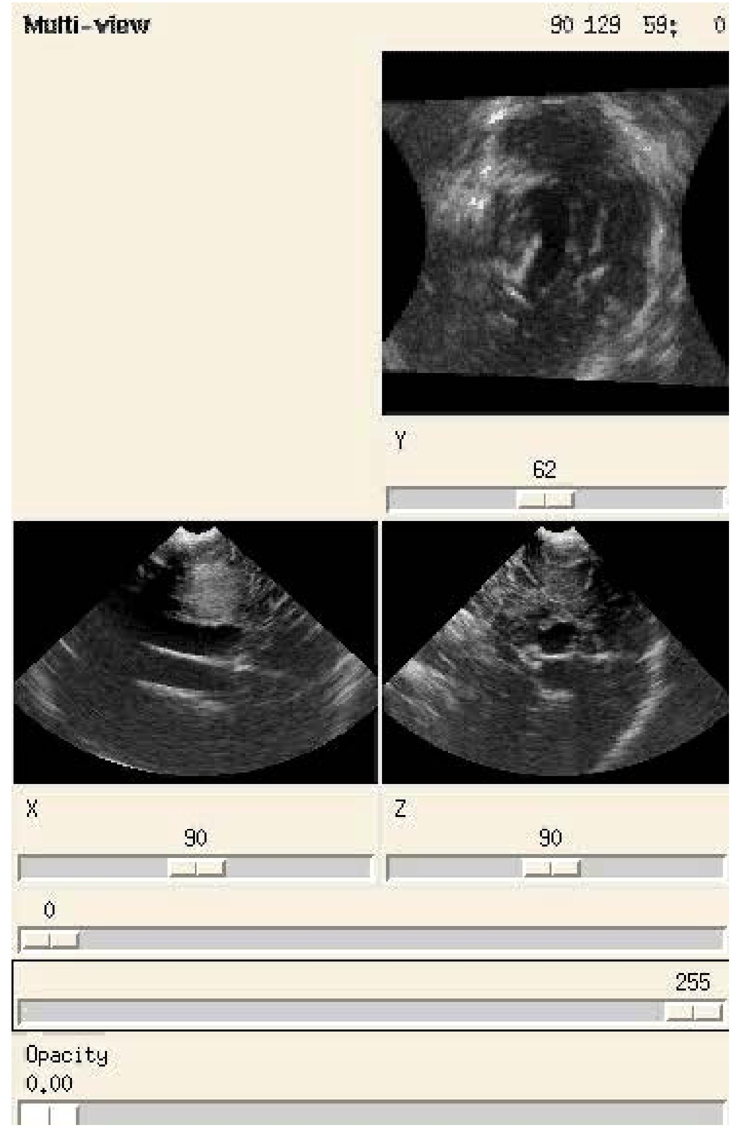
- 2 pre-op MR (0.9 x 0.9 x 1.1 mm)
- 3 per-op US (0.63 and 0.95 mm)
- 3 loops

Robustness and precision

	Success	var rot (deg)	var trans (mm)
MI	29%	0.53	0.25
CR	90%	0.45	0.17
BCR	**85%**	**0.39**	**0.11**

Consistency of BCR

	var rot (deg)	var trans (mm)	var test (mm)
Multiple MR	0.06	0.06	0.10
Loop	2.22	0.82	2.33
MR/US	**1.57**	**0.58**	**1.65**

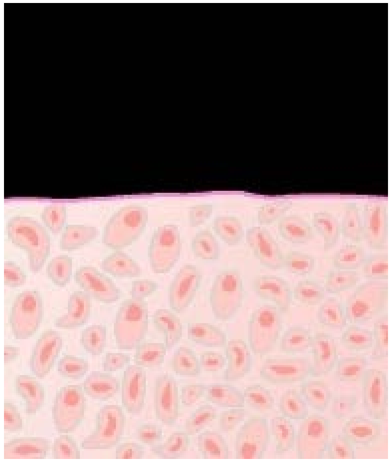


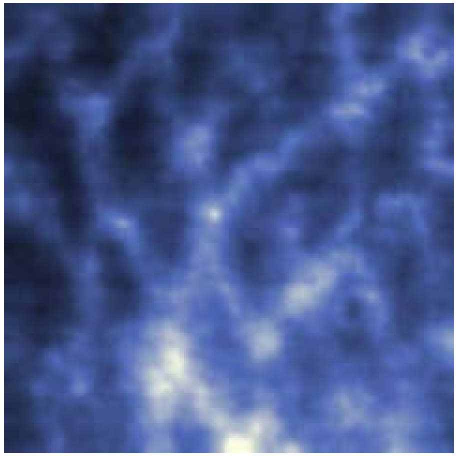
[Roche et al, TMI 20(10), 2001]

[Pennec et al, Multi-Sensor Image Fusion, Chap. 4, CRC Press, 2005]

Mosaicing of Confocal Microscopic in Vivo Video Sequences.

Cellvizio: Fibered confocal fluorescence imaging





FOV 200x200 µm

Courtesy of Mike Booth, MGH, Boston, MA

FOV 2747x638 µm

[T. Vercauteren et al., MICCAI 2005, T.1, p.753-760]

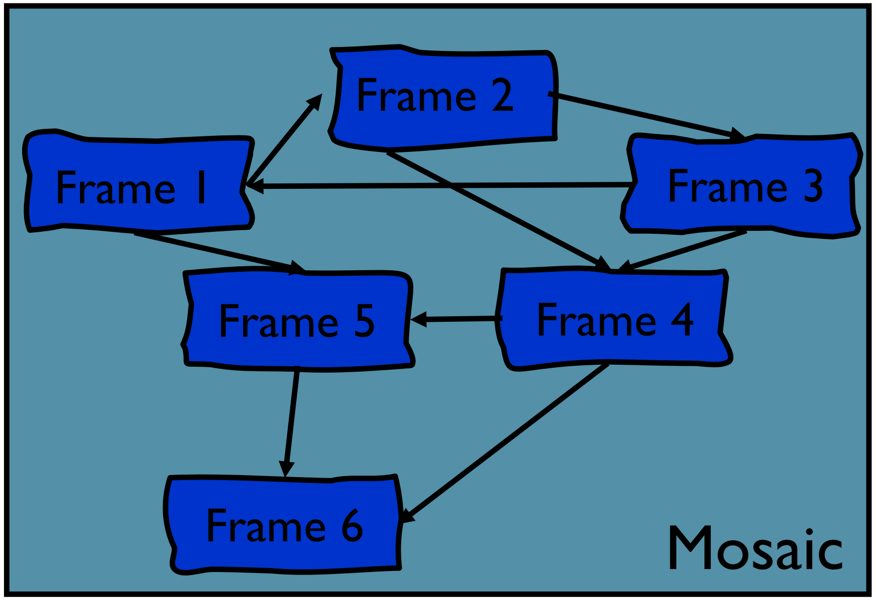
Mosaicing of Confocal Microscopic in Vivo Video Sequences.

Common coordinate system

- **Multiple rigid** registration
- Refine with **non rigid**

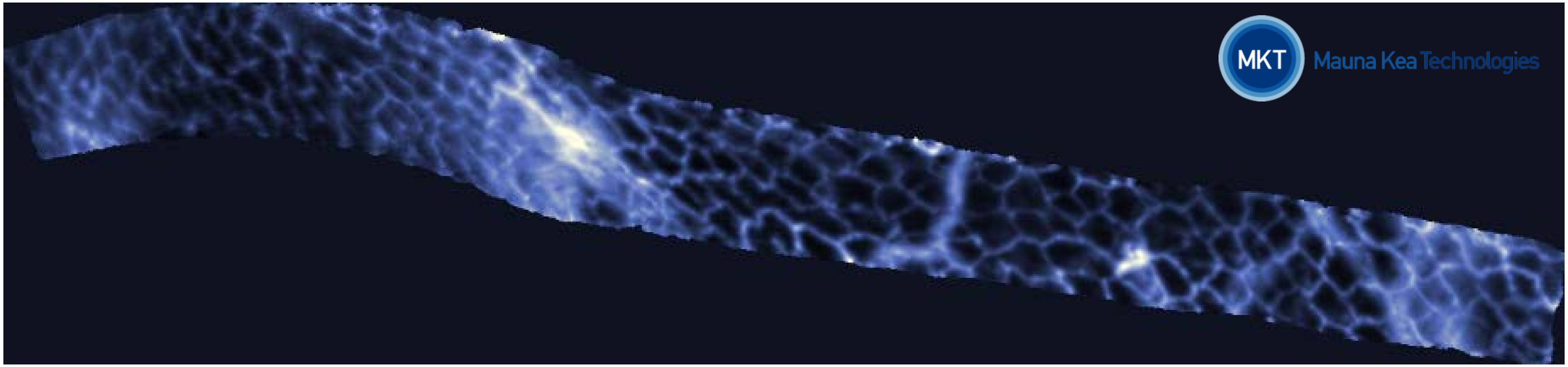
Mosaic image creation

- Interpolation / approximation with **irregular sampling**



Courtesy of Mike Booth, MGH, Boston, MA

FOV 2747x638 µm



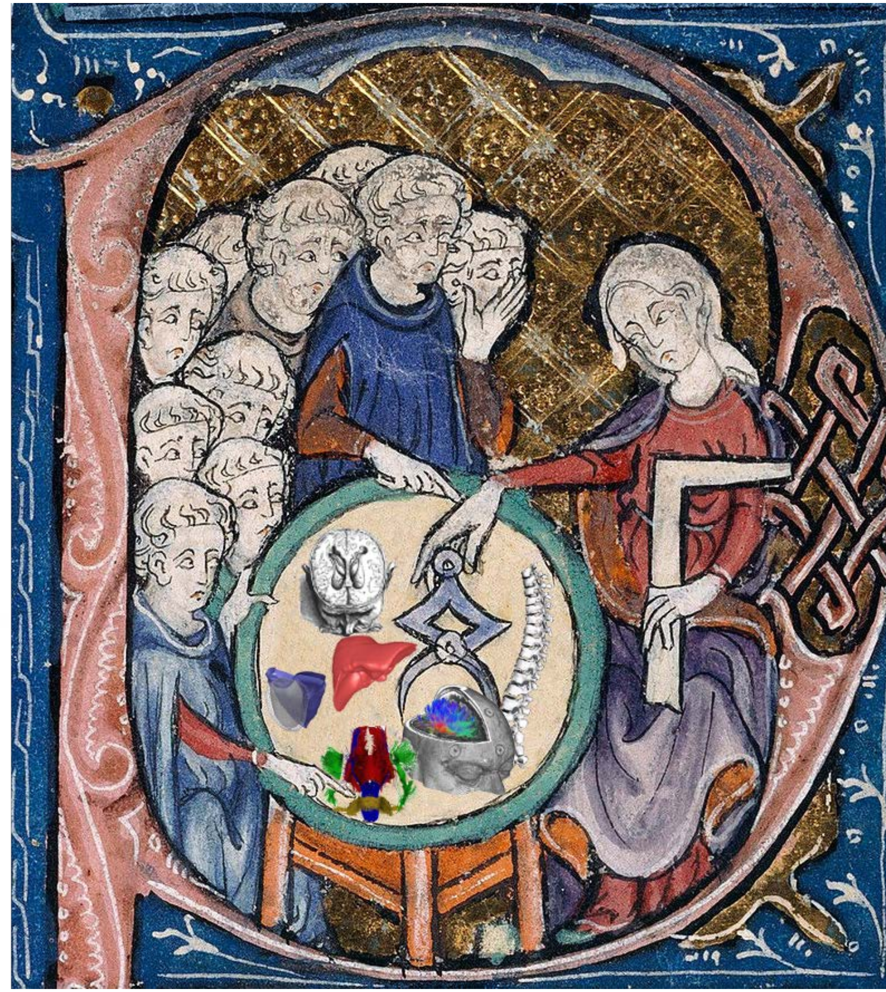
[T. Vercauteren et al., MICCAI 2005, T.1, p.753-760]

Xavier Pennec

Univ. Côte d'Azur and Inria, France

Geometric Statistics

Mathematical foundations and applications in computational anatomy



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

2/ Metric and Affine Geometric Settings for Lie Groups

Geometric Statistics workshop 09/2019

Geometric Statistics: Mathematical foundations and applications in computational anatomy

Intrinsic Statistics on Riemannian Manifolds

Metric and Affine Geometric Settings for Lie Groups

- Riemannian frameworks on Lie groups
- Lie groups as affine connection spaces
- The SVF framework for diffeomorphisms

Advances Statistics: CLT & PCA

Natural Riemannian Metrics on Transformations

Transformation are Lie groups: Smooth manifold G compatible with group structure

- Composition g o h and inversion g^{-1} are smooth
- Left and Right translation $L_g(f) = g \circ f$ $R_g(f) = f \circ g$
- Conjugation $Conj_g(f) = g \circ f \circ g^{-1}$
- Symmetry: $S_g(f) = g \circ f^{-1} \circ g$

Natural Riemannian metric choices

- Chose a metric at Id: $\langle x,y \rangle_{Id}$
- Propagate at each point g using left (or right) translation
  $\langle x,y \rangle_g = \langle DL_{g^{(-1)}}.x , DL_{g^{(-1)}}.y \rangle_{Id}$

Implementation

- Practical computations using left (or right) translations

$$\mathrm{Exp}_f(x) = f \circ \mathrm{Exp}_{Id}(DL_{f^{(-1)}}.x) \qquad \overrightarrow{fg} = \mathrm{Log}_f(g) = DL_f.\mathrm{Log}_{Id}(f^{(-1)} \circ g)$$

General Non-Compact and Non-Commutative case

No Bi-invariant Mean for 2D Rigid Body Transformations

- Metric at Identity: $dist(Id, (\theta; t_1; t_2))^2 = \theta^2 + t_1^2 + t_2^2$

- $T_1 = \left(\frac{\pi}{4}; -\frac{\sqrt{2}}{2}; \frac{\sqrt{2}}{2}\right)$ $T_2 = \left(0; \sqrt{2}; 0\right)$ $T_3 = \left(-\frac{\pi}{4}; -\frac{\sqrt{2}}{2}; -\frac{\sqrt{2}}{2}\right)$

- Left-invariant Fréchet mean: $(0; 0; 0)$
- Right-invariant Fréchet mean: $\left(0; \frac{\sqrt{2}}{3}; 0\right) \simeq (0; 0.4714; 0)$

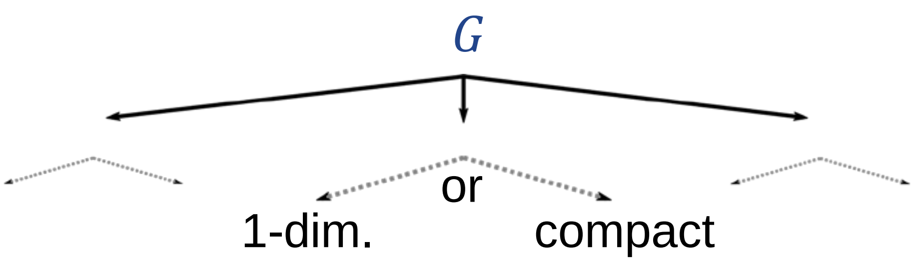
Questions for this talk:

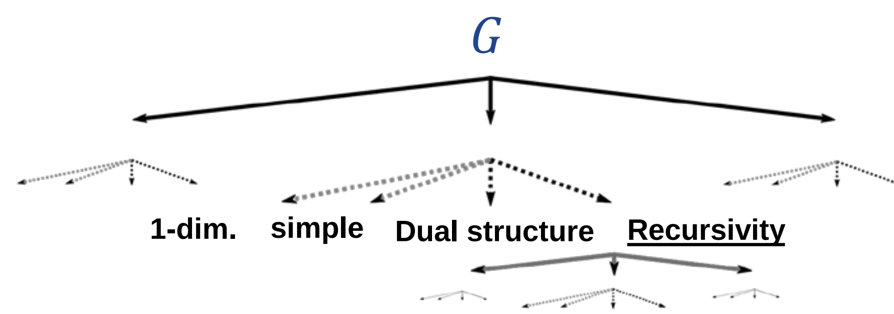
- **Can we design a mean compatible with the group operations?**
- **Is there a more convenient structure for statistics on Lie groups?**

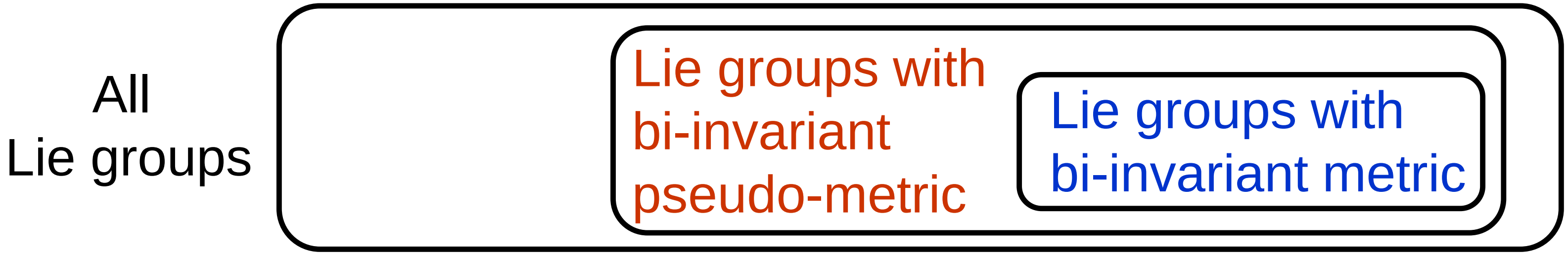
Existence of bi-invariant (pseudo) metrics

[Cartan 50's]:
Bi-invariant metric on G

[Medina, Revoy 80's]:
Bi-invariant pseudo-metric on G







[Miolane, XP, Computing Bi-Invariant Pseudo-Metrics on Lie Groups for Consistent Statistics. Entropy, 17(4):1850-1881, April 2015.]

- Algorithm: decompose the Lie algebra and find a bi-inv. pseudo-metric
- Test on rigid transformations SE(n): bi-inv. ps-metric for n=1 or 3 only

Geometric Statistics: Mathematical foundations and applications in computational anatomy

Intrinsic Statistics on Riemannian Manifolds

Metric and Affine Geometric Settings for Lie Groups

- Riemannian frameworks on Lie groups
- Lie groups as affine connection spaces
- The SVF framework for diffeomorphisms

Advances Statistics: CLT & PCA

Basics of Lie groups

Flow of a left invariant vector field $\tilde{X} = DL.x$ from identity

- $\gamma_x(t)$ exists for all time
- One parameter subgroup: $\gamma_x(s+t) = \gamma_x(s).\,\gamma_x(t)$

Lie group exponential

- Definition: $x \in \mathfrak{g} \rightarrow Exp(x) = \gamma_x(1) \in G$
- Diffeomorphism from a neighborhood of 0 in $\mathfrak{g}$ to a neighborhood of e in G (not true in general for inf. dim)

3 curves parameterized by the same tangent vector

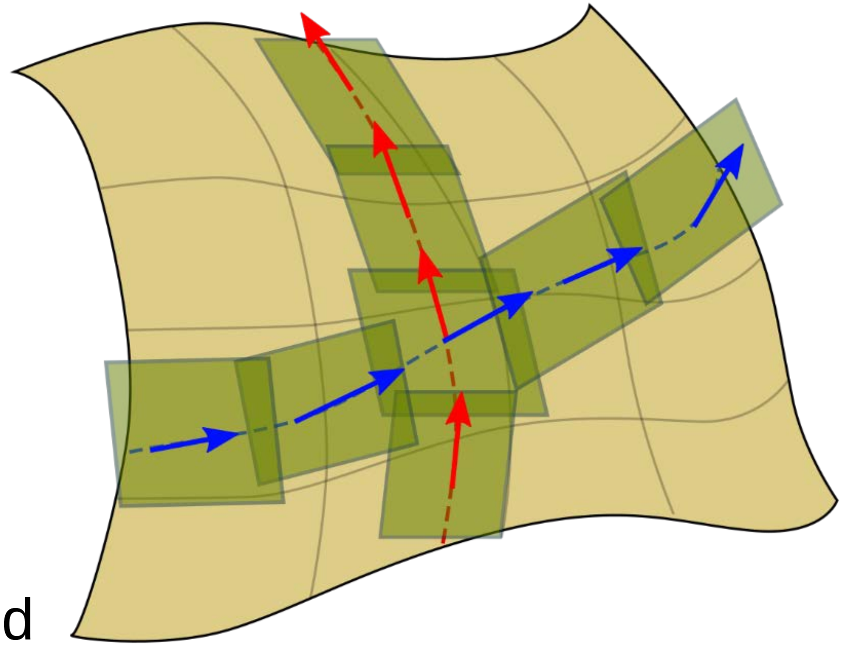
- Left / Right-invariant geodesics, one-parameter subgroups

Question: Can one-parameter subgroups be geodesics?

Affine connection spaces: Drop the metric, use connection to define geodesics

Affine Connection (infinitesimal parallel transport)

- Acceleration = derivative of the tangent vector along a curve
- Projection of a tangent space on a neighboring tangent space



Geodesics = straight lines

- Null acceleration: $\nabla_{\dot{\gamma}}\dot{\gamma} = 0$
- 2nd order differential equation: Normal coordinate system
- Local exp and log maps, well defined in a convex neighborhood

[Lorenzi, Pennec. Geodesics, Parallel Transport & One-parameter Subgroups for Diffeomorphic Image Registration. Int. J. of Computer Vision, 105(2):111-127, 2013.]

Canonical Affine Connections on Lie Groups

A unique Cartan-Schouten connection

- Bi-invariant and symmetric (no torsion)
- Geodesics through Id are one-parameter subgroups (group exponential)
  - Matrices : M(t) = A exp(t.V)
  - Diffeos : **translations of Stationary Velocity Fields (SVFs)**

Levi-Civita connection of a bi-invariant metric (if it exists)

- Continues to exists in the absence of such a metric (e.g. for rigid or affine transformations)

Symmetric space with central symmetry $S_{\psi}(\phi) = \psi\phi^{-1}\psi$

- Matrix geodesic symmetry: $S_A(M(t)) = A\exp(-tV)A^{-1}A = M(-t)$

[Lorenzi, Pennec. Geodesics, Parallel Transport & One-parameter Subgroups for Diffeomorphic Image Registration. Int. J. of Computer Vision, 105(2):111-127, 2013.]

Statistics on an affine connection space

~~Fréchet mean~~: exponential barycenters

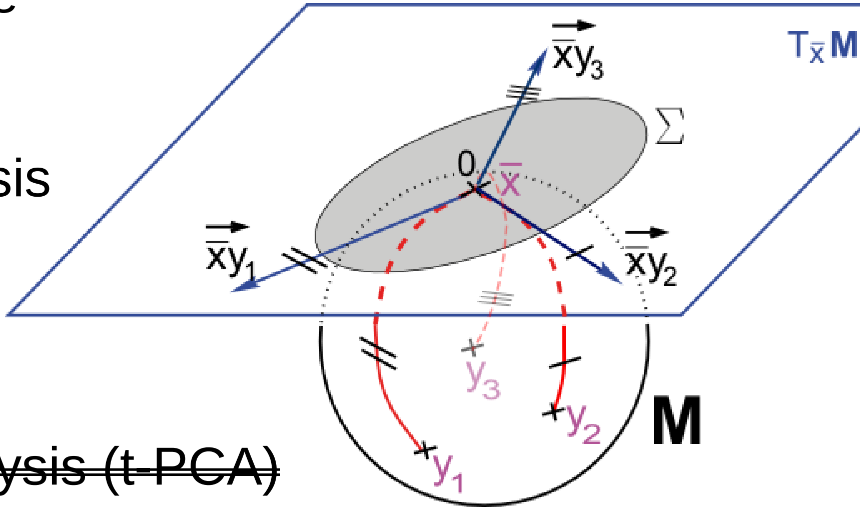
- $\sum_i Log_x(y_i) = 0$ [Emery, Mokobodzki 91, Corcuera, Kendall 99]
- Existence **local uniqueness** if local convexity [Arnaudon & Li, 2005]

Covariance matrix & higher order moments

- Defined as tensors in tangent space

$$\Sigma = \int Log_x(y) \otimes Log_x(y)\, \mu(dy)$$

- Matrix expression changes with basis



Other statistical tools

- Mahalanobis distance, chi^2 test
- ~~Tangent Principal Component Analysis (t-PCA)~~
- Independent Component Analysis (ICA)?

[XP & Arsigny, 2012, XP & Lorenzi, Beyond Riemannian Geometry, 2019]

Statistics on an affine connection space

For Cartan-Schouten connections [Pennec & Arsigny, 2012]

- Locus of points *x* such that $\sum Log(x^{-1}.y_i) = 0$
- Algorithm: fixed point iteration (**local convergence**)

$$x_{t+1} = x_t \circ Exp\left(\frac{1}{n}\sum Log(x_t^{-1}.y_i)\right)$$

- **Mean stable by left / right composition and inversion**

Matrix groups with no bi-invariant metric

- Heisenberg group: bi-invariant mean is unique (conj. ok for solvable)
- Rigid-body transformations: uniqueness if unique mean rotation
- SU(n) and GL(n): log does not always exist (need 2 exp to cover)

[XP and V. Arsigny. Exponential Barycenters of the Canonical Cartan Connection and Invariant Means on Lie Groups. In Matrix Information Geometry. 2012]

Example mean of 2D rigid-body transformation

$$T_1 = \left(\frac{\pi}{4};\ -\frac{\sqrt{2}}{2};\frac{\sqrt{2}}{2}\right) \quad T_2 = (0;\ \sqrt{2};0\ ) \quad T_3 = \left(-\frac{\pi}{4};\ -\frac{\sqrt{2}}{2};-\frac{\sqrt{2}}{2}\right)$$

- Metric at Identity: $dist(Id,(\theta;t_1;t_2))^2 = \theta^2 +\ t_1^2 + t_2^2$
- Left-invariant Fréchet mean: $(0;0;0\ )$
- Log-Euclidean mean: $\left(0;\frac{\sqrt{2}-\pi/4}{3};0\right) \simeq (0;0.2096;0)$
- Bi-invariant mean: $\left(0;\frac{\sqrt{2}-\pi/4}{1+\pi/4(\sqrt{2}+1)};0\right) \simeq (0;0.2171;0)$
- Right-invariant Fréchet mean: $\left(0;\frac{\sqrt{2}}{3};0\right) \simeq (0;0.4714;0)$

Cartan Connections vs Riemannian

What is similar

- Standard differentiable geometric structure [curved space without torsion]
- Normal coordinate system with Exp_x et Log_x [finite dimension]

Limitations of the affine framework

- No metric (but no choice of metric to justify)
- The exponential does always not cover the full group
  - Pathological examples close to identity in finite dimension
  - In practice, similar limitations for the discrete Riemannian framework

What we gain with Cartan-Schouten connection

- A globally invariant structure invariant by composition & inversion
- Simple geodesics, efficient computations (stationarity, group exponential)
- Consistency with any bi-invariant (pseudo)-metric
- The simplest linearization of transformations for statistics on Lie groups?

Geometric Statistics: Mathematical foundations and applications in computational anatomy

Intrinsic Statistics on Riemannian Manifolds

Metric and Affine Geometric Settings for Lie Groups

- Riemannian frameworks on Lie groups
- Lie groups as affine connection spaces
- The SVF framework for diffeomorphisms

Advances Statistics: CLT & PCA

Riemannian Metrics on diffeomorphisms

Space of deformations

- Transformation y=ϕ(x)
- Curves in transformation spaces: ϕ(x,t)
- Tangent vector = speed vector field

$$v_t(x) = \frac{d\phi(x,t)}{dt}$$

Right invariant metric

$$\|v_t\|_{\phi_t} = \left\|v_t \circ \phi_t^{-1}\right\|_{Id}$$

- Eulerian scheme
- Sobolev Norm H_k or H_∞ (RKHS) in LDDMM → diffeomorphisms [Miller, Trouve, Younes, Holm, Dupuis, Beg… 1998 – 2009]

Geodesics determined by optimization of a time-varying vector field

- Distance

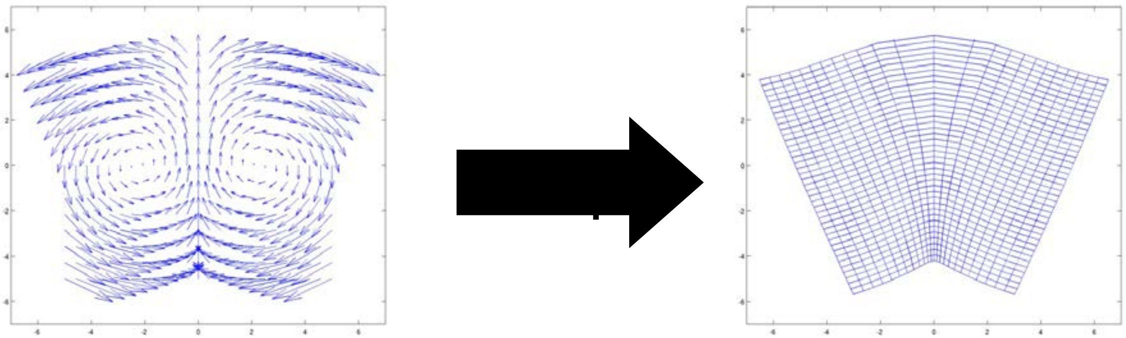
$$d^2(\phi_0, \phi_1) = \arg\min_{v_t}\left(\int_0^1 \|v_t\|_{\phi_t}^2 \, dt\right)$$

- Geodesics characterized by initial velocity / momentum
- Optimization for images is quite tricky (and lenghty)

The SVF framework for Diffeomorphisms

Idea: [Arsigny MICCAI 2006, Bossa MICCAI 2007, Ashburner Neuroimage 2007]

- Exponential of a smooth vector field is a diffeomorphism
- Parameterize deformation by ~~time-varying~~ Stationary Velocity Fields



Stationary velocity field — Diffeomorphism

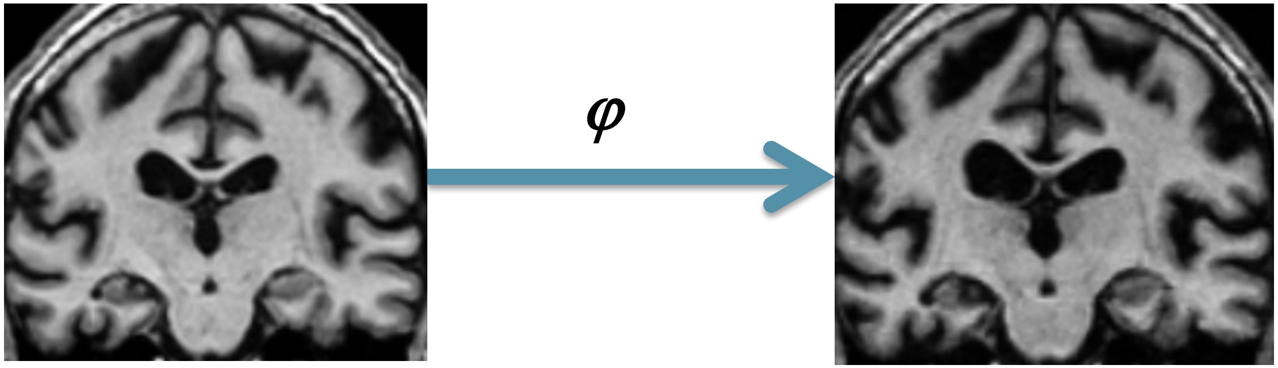
Direct generalization of numerical matrix algorithms

- Computing the deformation: **Scaling and squaring** **[Arsigny MICCAI 2006]**
  recursive use of $\exp(\boldsymbol{v})=\exp(\boldsymbol{v}/2) \circ \exp(\boldsymbol{v}/2)$
- Computing the Jacobian: $D\exp(v) = D\exp(v/2) \circ \exp(v/2) \,.\, D\exp(v/2)$
- Updating the deformation parameters: **BCH formula** **[Bossa MICCAI 2007]**

  $$\exp(\boldsymbol{v}) \circ \exp(\varepsilon \boldsymbol{u}) = \exp(\boldsymbol{v} + \varepsilon \boldsymbol{u} + [\boldsymbol{v},\varepsilon \boldsymbol{u}]/2 + [\boldsymbol{v},[\boldsymbol{v},\varepsilon \boldsymbol{u}]]/12 + \ldots)$$

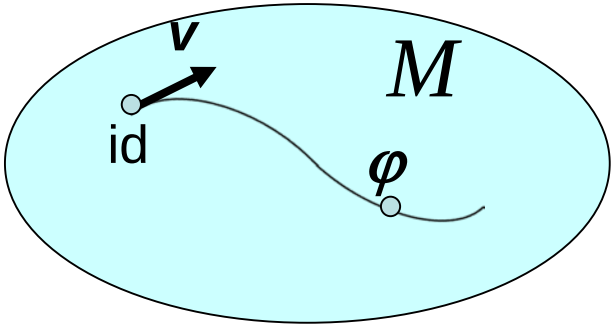
  - Lie bracket $[\boldsymbol{v},\boldsymbol{u}](p) = Jac(\boldsymbol{v})(p).\boldsymbol{u}(p) - Jac(\boldsymbol{u})(p).\boldsymbol{v}(p)$

Parallel transport of deformation trajectories



SVF setting

- ***v stationary*** *velocity field*
- *Lie group Exp(v) non-metric geodesic wrt Cartan connections*



LDDMM setting

- ***v time-varying*** *velocity field*
- *Riemannian* $exp_{id}(v)$ *metric geodesic wrt Levi-Civita connection*
- *Defined by* ***intial momentum***

Transporting trajectories:

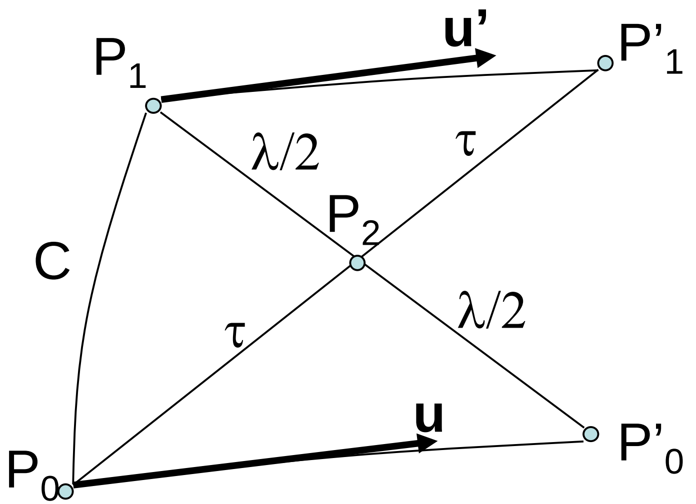
Parallel transport of initial tangent vectors

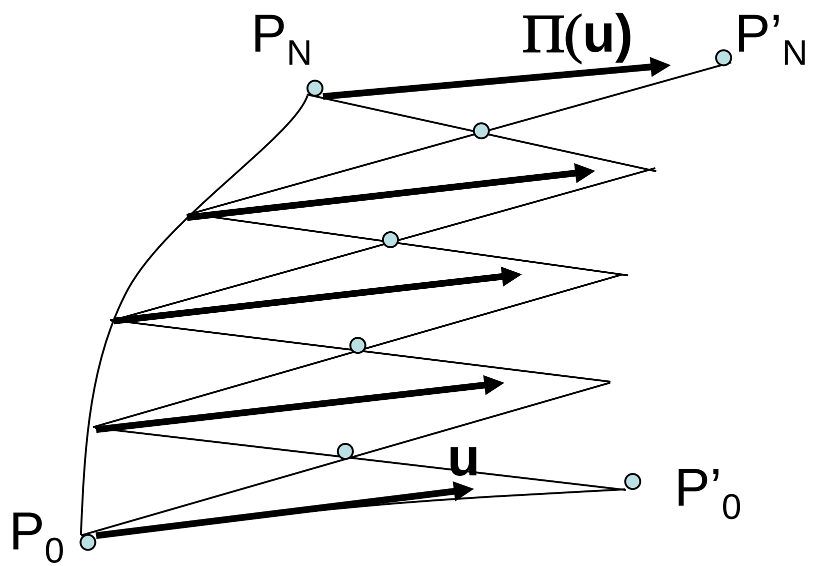
LDDMM: parallel transport along geodesics using Jacobi fields [Younes et al. 2008]

Parallel transport along arbitrary curves

A numerical scheme to integrate for symmetric connections: Schild's Ladder [Elhers et al, 1972]

- Build geodesic parallelogrammoid
- Iterate along the curve





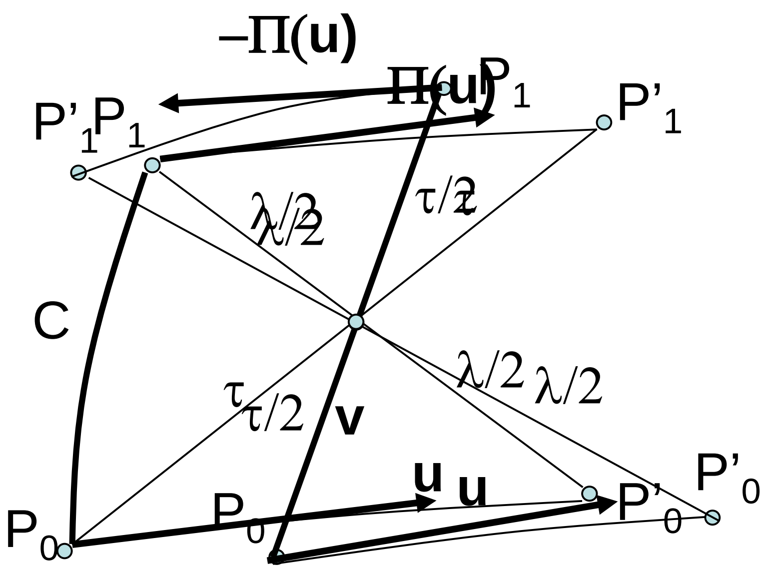
[Lorenzi, Pennec: Efficient Parallel Transport of Deformations in Time Series of Images: from Schild's to pole Ladder, JMIV 50(1-2):5-17, 2013]

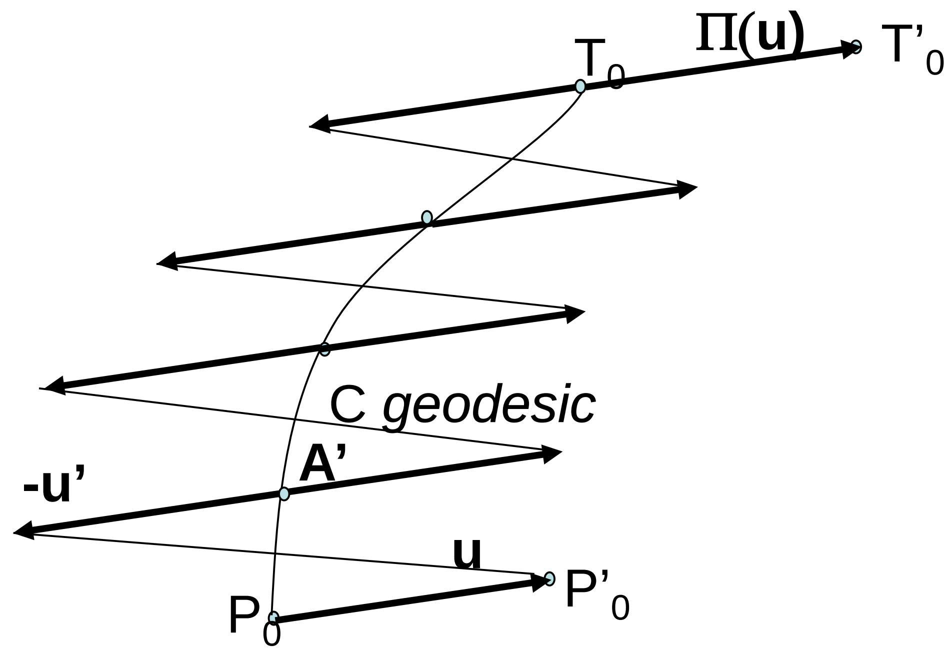
Parallel transport along geodesics

Simpler scheme along geodesics: Pole Ladder

$$\mathrm{Exp}(\Pi(u)) = \mathrm{Exp}(v/2) \circ \mathrm{Exp}(u) \circ \mathrm{Exp}(-v/2)$$

$$\Pi_{BCH}(u) = u + [v,u] + \frac{1}{2}[v[v,u]]$$

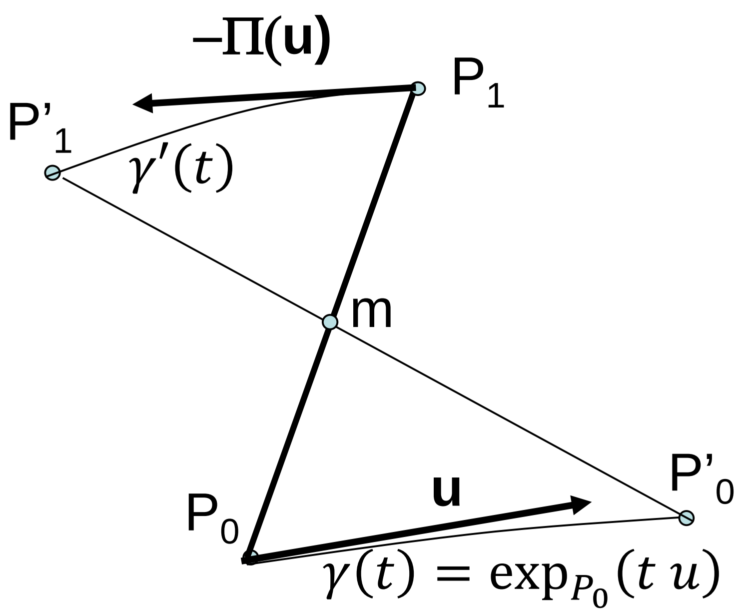




[Lorenzi, Pennec: Efficient Parallel Transport of Deformations in Time Series of Images: from Schild's to pole Ladder, JMIV 50(1-2):5-17, 2013]

Parallel transport along geodesics

Simpler scheme along geodesics: Pole Ladder



Pole ladder is exact in 1 step in symmetric space

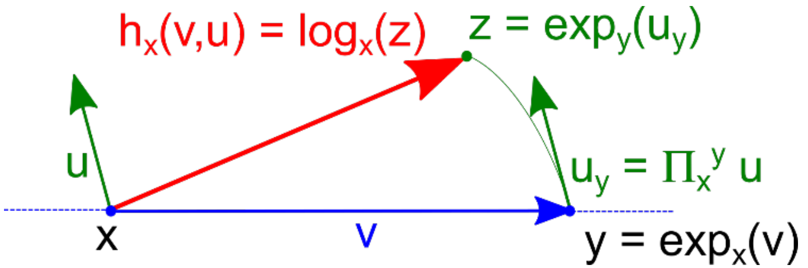
- Symmetry preserves geodesics:
  $$S_m\big(\gamma(t)\big) = \gamma'(t)$$
- Parallel transport is differential of symmetry
  $$\gamma'(t) = \exp_{P_1}(-\Pi(u))$$

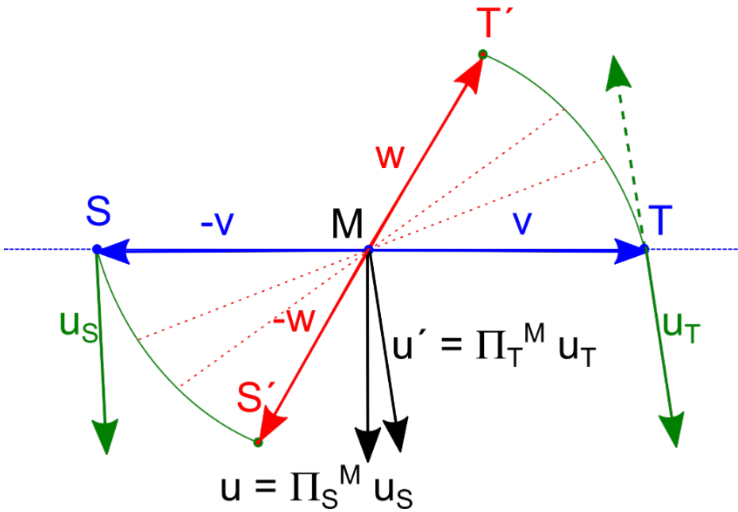
[XP. Parallel Transport with Pole Ladder: a Third Order Scheme in Affine Connection Spaces which is Exact in Affine Symmetric Spaces. Arxiv 1805.11436]

Accuracy of pole ladder

Gavrilov's double exponential series (2006):

$$h_x(v,u) = \log_x(\Pi_x^{\exp_x(v)}\, u)$$
$$= v + u + \frac{1}{6}R(u,v)v + \frac{1}{3}R(u,v)u + \frac{1}{24}\nabla_v R(u,v)(2v+5u) + \frac{1}{24}\nabla_u R(u,v)(v+2u) + O(5)$$





Find u' that satisfies:

$$h_M(v,-u') + h_M(-v,u) = 0$$

$$u' = u + \frac{1}{12}\nabla_v R(u,v)(5u-2v) + \frac{1}{12}\nabla_u R(u,v)(v-2u) + O(5)$$

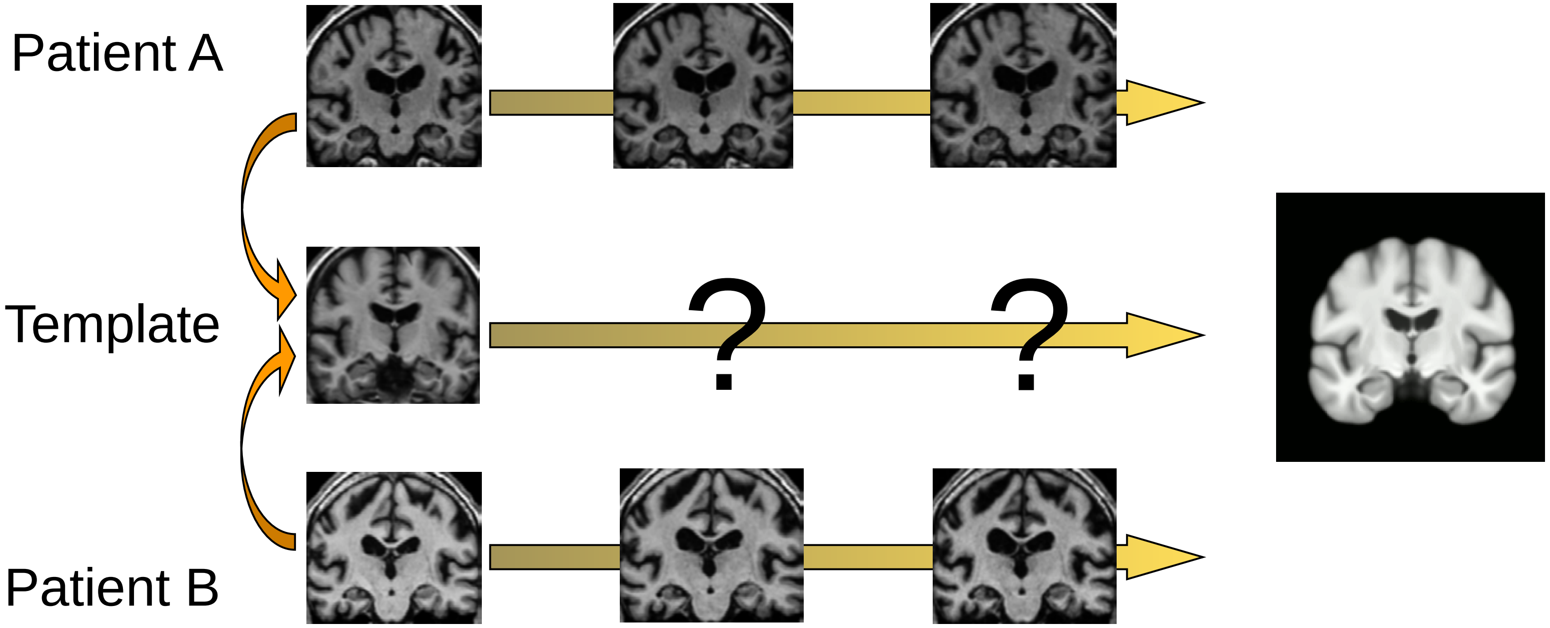
- Error term is of order 4 in general affine manifolds
- Error is even zero for symmetric spaces: pole ladder is exact in one step!

[XP. Parallel Transport with Pole Ladder: a Third Order Scheme in Affine Connection Spaces which is Exact in Affine Symmetric Spaces. Arxiv 1805.11436]

The Stationnary Velocity Fields (SVF) framework for diffeomorphisms

- SVF framework for diffeomorphisms is algorithmically simple
- Compatible with “inverse-consistency” **[Lorenzi, XP. IJCV, 2013]**
- Vector statistics directly generalized to diffeomorphisms.
- **Exact parallel transport** using one step of pole ladder **[XP arxiv 1805.11436 2018]**

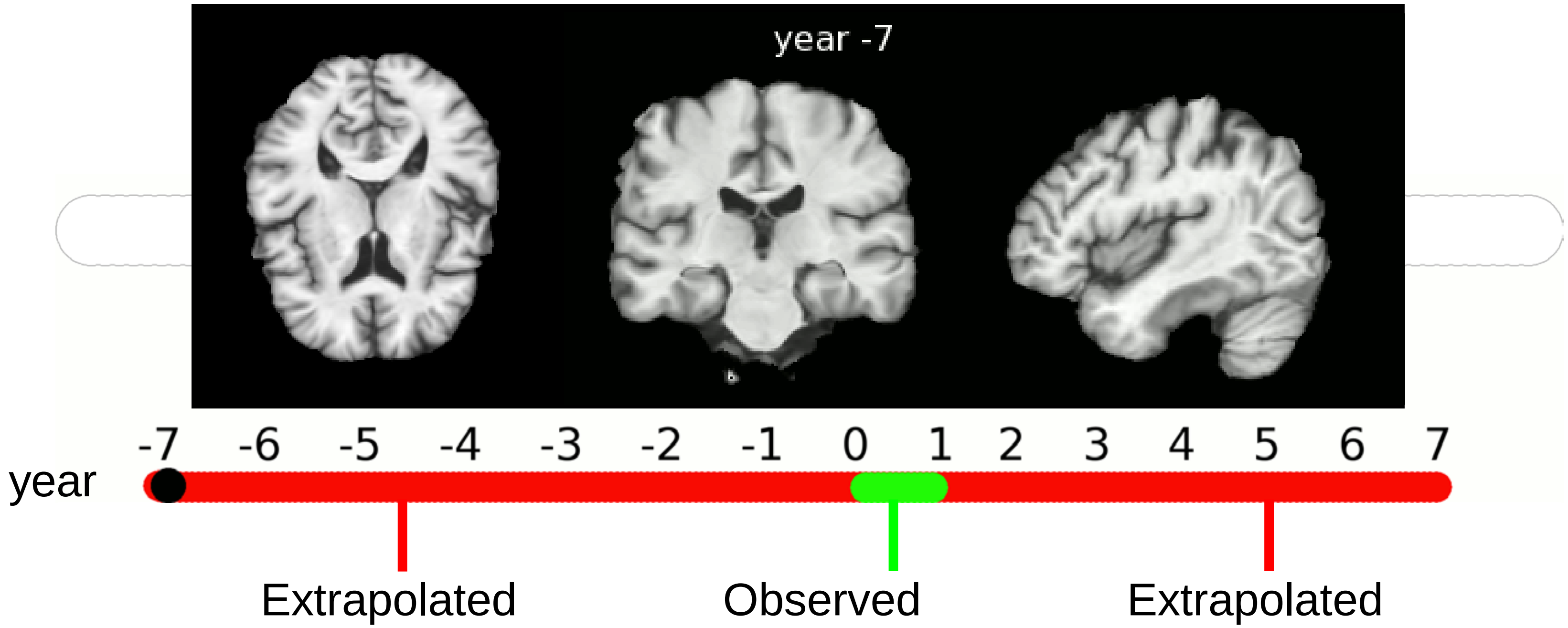
Longitudinal modeling of AD: 70 subjects extrapolated from 1 to 15 years



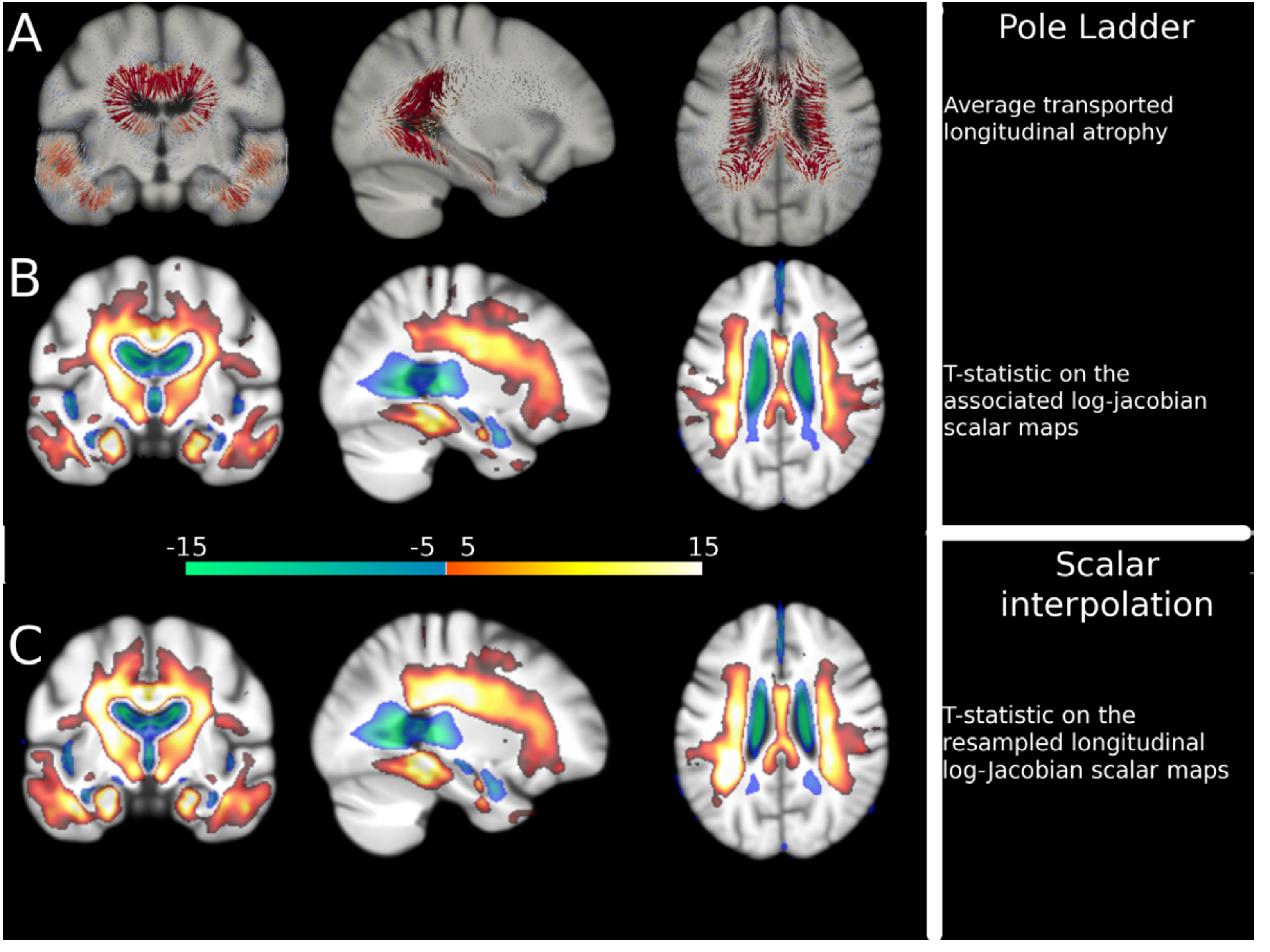
The Stationnary Velocity Fields (SVF) framework for diffeomorphisms

- SVF framework for diffeomorphisms is algorithmically simple
- Compatible with “inverse-consistency” **[Lorenzi, XP. IJCV, 2013]**
- Vector statistics directly generalized to diffeomorphisms.
- **Exact parallel transport** using one step of pole ladder **[XP arxiv 1805.11436 2018]**

Longitudinal modeling of AD: 70 subjects extrapolated from 1 to 15 years

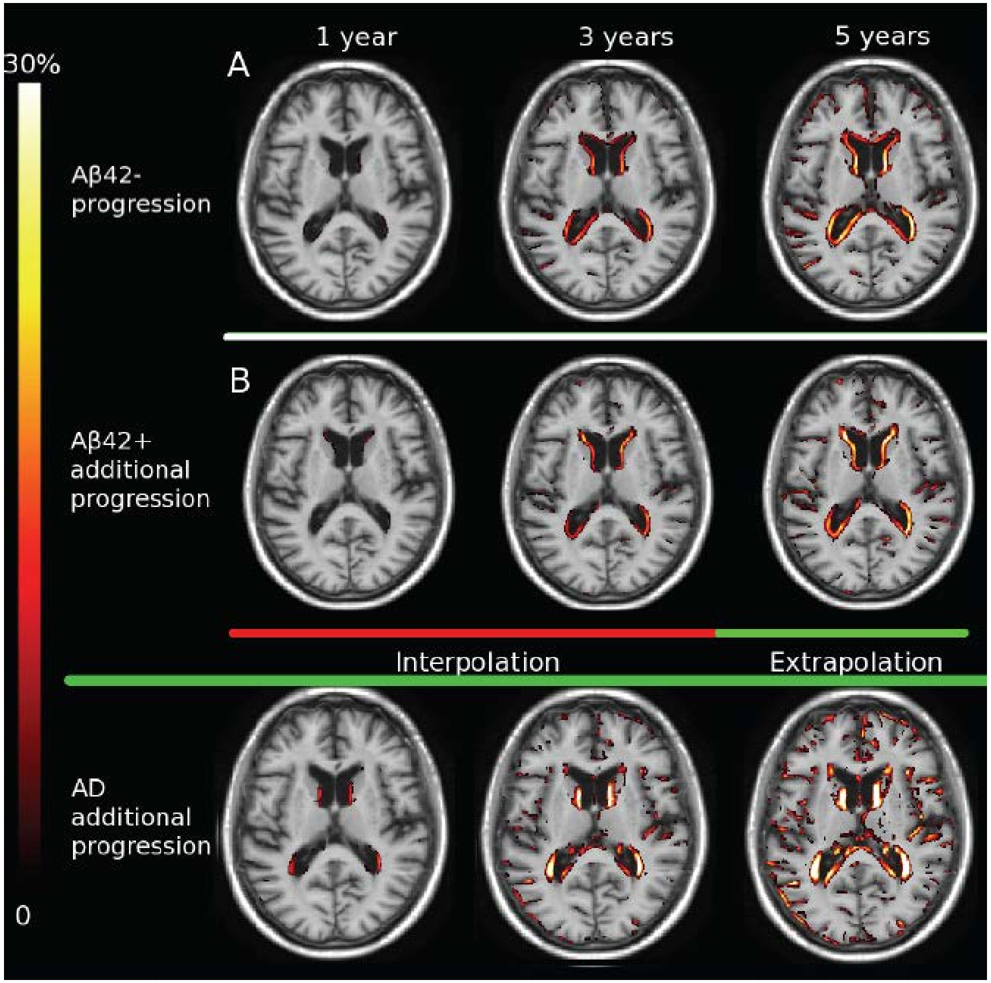


Modeling longitudinal atrophy in AD from images

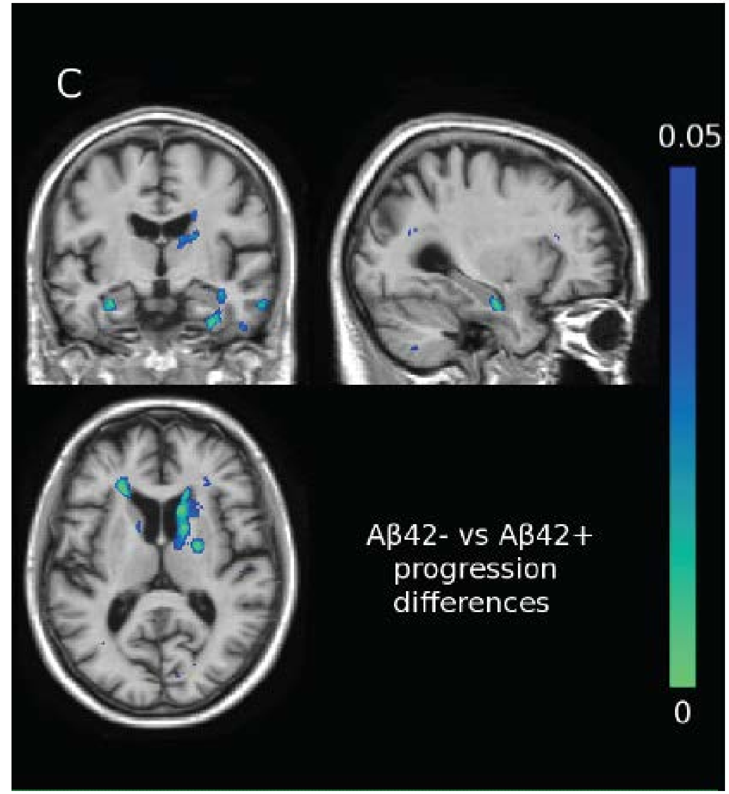


Study of prodromal Alzheimer's disease

Linear regression of the SVF over time: interpolation + prediction



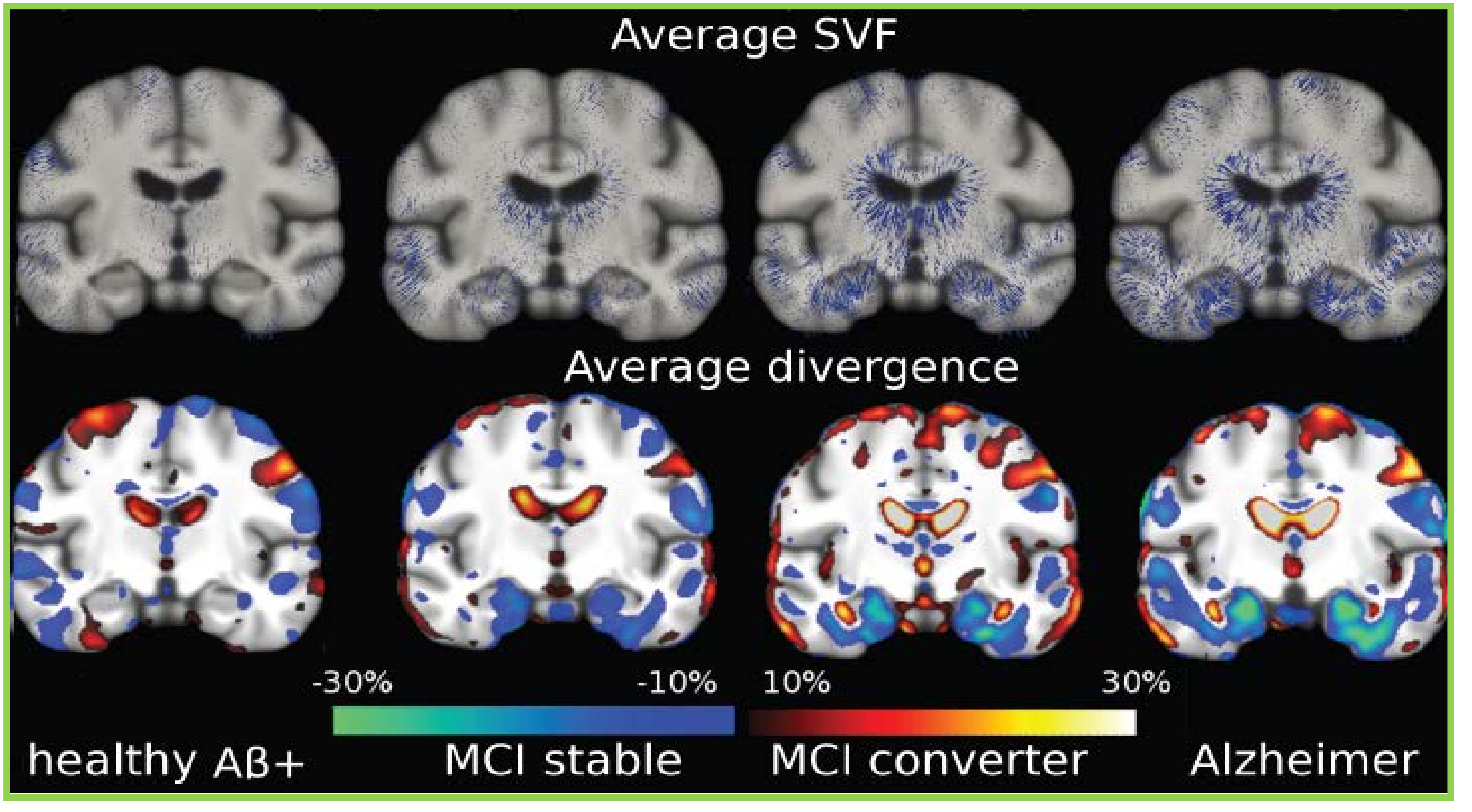
$$T(t) = Exp(\tilde{v}(t)) * T_0$$



Multivariate group-wise comparison of the transported SVFs shows statistically significant differences (nothing significant on log(det))

[Lorenzi, Ayache, Frisoni, Pennec, in Proc. of MICCAI 2011]

Mean deformation / atrophy per group



M Lorenzi, N Ayache, X Pennec G B. Frisoni, for ADNI. Disentangling the normal aging from the pathological Alzheimer's disease progression on structural MR images. 5th Clinical Trials in Alzheimer's Disease (CTAD'12), Monte Carlo, October 2012. (see also MICCAI 2012)

References for Statistics on Manifolds and Lie Groups

Statistics on Riemannnian manifolds

- Xavier Pennec. Intrinsic Statistics on Riemannian Manifolds: Basic Tools for Geometric Measurements. Journal of Mathematical Imaging and Vision, 25(1):127-154, July 2006. http://www.inria.fr/sophia/asclepios/Publications/Xavier.Pennec/Pennec.JMIV06.pdf

Invariant metric on SPD matrices and of Frechet mean to define manifold-valued image processing algorithms

- Xavier Pennec, Pierre Fillard, and Nicholas Ayache. A Riemannian Framework for Tensor Computing. International Journal of Computer Vision, 66(1):41-66, Jan. 2006. http://www.inria.fr/sophia/asclepios/Publications/Xavier.Pennec/Pennec.IJCV05.pdf

Bi-invariant means with Cartan connections on Lie groups

- Xavier Pennec and Vincent Arsigny. Exponential Barycenters of the Canonical Cartan Connection and Invariant Means on Lie Groups. In Frederic Barbaresco, Amit Mishra, and Frank Nielsen, editors, Matrix Information Geometry, pages 123-166. Springer, May 2012. http://hal.inria.fr/hal-00699361/PDF/Bi-Invar-Means.pdf

Cartan connexion for diffeomorphisms:

- Marco Lorenzi and Xavier Pennec. Geodesics, Parallel Transport & One-parameter Subgroups for Diffeomorphic Image Registration. International Journal of Computer Vision, 105(2), November 2013 https://hal.inria.fr/hal-00813835/document

Xavier Pennec

Univ. Côte d'Azur and Inria, France

Geometric Statistics

Mathematical foundations and applications in computational anatomy

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

3/ Advanced Stats: empirical estimation and generalized PCA

Geometric Statistics workshop 09/2019

Epione e-patient / e-medicine UNIVERSITÉ CÔTE D'AZUR Inria

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

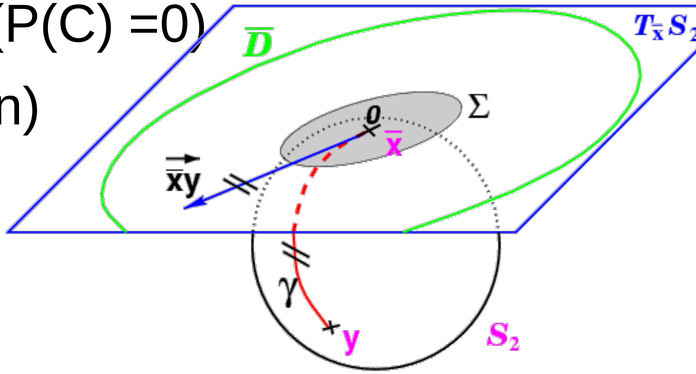
Tensor moments of a random point on M

- $\mathfrak{M}_1(x) = \int_M \overrightarrow{xz}\, dP(z)$ 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
- $\sigma^2(x) = Tr_g\left(\mathfrak{M}_2(x)\right) = \int_M dist^2(x,z)\, dP(z)$ **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

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

Algorithms to compute the mean

Karcher flow (gradient descent)

$$\bar{x}_{t+1} = \exp_{\bar{x}_t}(\epsilon_t \, v_t) \;\; with \;\; v_t = \mathrm{E}(\overrightarrow{\mathrm{y}\mathbf{x}}) = \frac{1}{\mathrm{n}} \sum_i \log_{\bar{x}_t}(x_i)$$

- Usual algorithm with $\epsilon_t = 1$ can diverge on SPD matrices
  [Bini & Iannazzo, Linear Algebra Appl., 438:4, 2013]
- Convergence for non-negative curvature (p-means)
  [Afsari, Tron and Vidal, SICON 2013]

Inductive / incremental weighted means

- $\bar{x}_{k+1} = \exp_{\bar{x}_k}\left(\frac{1}{k} \, v_k\right) \;\; with \;\; v_k = \log_{\bar{x}_k}(x_{k+1})$
- On negatively curved spaces [Sturm 2003],
  BHV centroid [Billera, Holmes, Vogtmann, 2001]
- On non-positive spaces [G. Cheng, J. Ho, H. Salehian, B. C. Vemuri 2016]

Stochastic algorithm

- [Bonnabel IEE TAC 58(9) 2013]
- [Arnaudon & Miclo, Stoch. Proc. & App. 124, 2014]

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)*
  - $\sqrt{n}\, log_{\bar{x}}(\bar{x}_n) \rightarrow N(0, 4\bar{H}^{-1}\, \Sigma\, \bar{H}^{-1})$ *if* $\bar{H} = \int_M Hess_{\bar{x}}\left(d^2(y, \bar{x})\right) \mu(dy)$ *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)

Taylor expansion in manifolds

The mean is an exponential barycenter

- The zero of the tangent mean field (Brewin Taylor expansion)

$$\mathfrak{M}_1(x) = \int_M \log_x(z)\, \mu(dz) \text{ has a zero at } \bar{x}.$$

Lots of additional terms in higher order derivatives since the vector field expression at $x_v = \exp_x(v)$ in a normal coordinates at x is modulated by $\mathrm{Dexp}_x(v)$.

- The zero of a mapping of vector spaces: the recenterd mean field

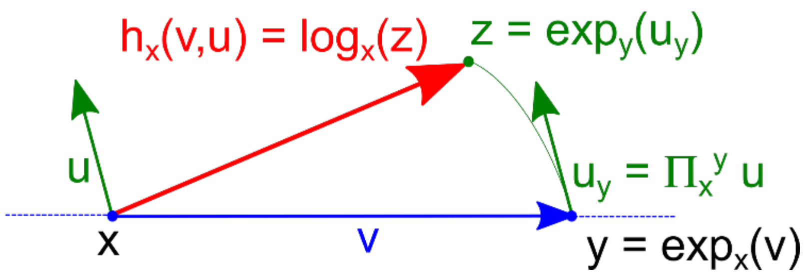
$$N_x(v) = \Pi_{x_v}^{x} \mathfrak{M}_1(\exp_x(v)) = \int_M \Pi_{x_v}^{x} \log_{x_v}(y)\, \mu(dy)$$

$\bar{x}$ is a Fréchet mean iff $N_x(v)$ has a zero at $\mathrm{v} = \log_x(\bar{x})$

Goal: compute a series expansion w.r.t. v

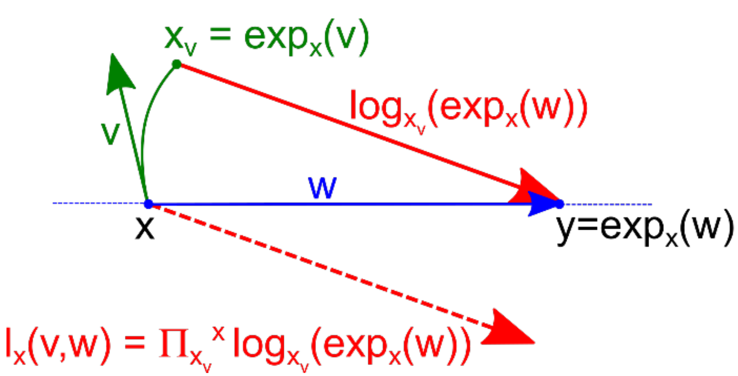
Taylor expansion in manifolds

Gavrilov's double exponential series (2006):



$$h_x(v,u) = \log_x(\Pi_x^{\exp_x(v)} u)$$
$$= v + u + \frac{1}{6} R(u,v)v + \frac{1}{3} R(u,v)u$$
$$+ \frac{1}{24} \nabla_v R(u,v)(2v + 5u)$$
$$+ \frac{1}{24} \nabla_u R(u,v)(v + 2u) + O(5)$$

Neighboring log expansion (new)



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

Taylor expansion of recentered mean map

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

Solving for the value $v = \log_x(\bar{x})$ that zeros the polynomial

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

For an empirical an n-sample $\mathbf{X_n} = \frac{1}{n}\sum_i \delta_{x_i}$

$$\log_x(\bar{x}_n) = \mathfrak{x}_1^n - \frac{1}{3}R(\bullet, \mathfrak{x}_1^n)\bullet : \mathfrak{x}_2^n + \frac{1}{24}\nabla_\bullet R(\bullet, \mathfrak{x}_1^n)\mathfrak{x}_1^n : \mathfrak{x}_2^n - \frac{1}{8}\nabla_{\mathfrak{x}_1^n} R(\bullet, \mathfrak{x}_1^n)\bullet : \mathfrak{x}_2^n - \frac{1}{12}\nabla_\bullet R(\bullet, \mathfrak{x}_1^n)\bullet : \mathfrak{x}_3^n + O(\epsilon^5).$$

Compute the expectation for a random n-sample?

Non-Asymptotic behavior of empirical means

Expectation of product of empirical moments

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

First Moment of the empirical mean

$$\begin{aligned}\mathrm{E}\left[\,\log_x(\bar{x}_n)\,\right] = \mathfrak{M}_1 - \frac{n-1}{3n} R(\bullet, \mathfrak{M}_1)\bullet : \mathfrak{M}_2 \\ + \frac{(n-1)(n-2)}{24n^2}\left(\nabla_\bullet R(\bullet, \mathfrak{M}_1)\mathfrak{M}_1 : \mathfrak{M}_2 - 3\nabla_{\mathfrak{M}_1} R(\bullet, \mathfrak{M}_1)\bullet : \mathfrak{M}_2\right) \\ + \frac{(n-1)}{12n^2}\left(2\nabla_\circ R(\circ, \bullet)\bullet : \mathfrak{M}_2 \mathbin{\dot\vdots} \mathfrak{M}_2 - (1+n)\nabla_\bullet R(\bullet, \mathfrak{M}_1)\bullet : \mathfrak{M}_3\right) + O(\epsilon^5).\end{aligned}$$

At the population mean:

$$\mathrm{E}\left[\,\log_{\bar{x}}(\bar{x}_n)\,\right] = \frac{(n-1)}{6n^2}\nabla_\bullet R(\bullet, \circ)\circ : \mathfrak{M}_2 \mathbin{\dot\vdots} \mathfrak{M}_2 + O(\epsilon^5).$$

Non-Asymptotic behavior of empirical means

Second Moment of the empirical mean a the pop. mean:

$$\mathbf{E}\left[\log_{\bar{x}}(\bar{x}_n)\otimes\log_{\bar{x}}(\bar{x}_n)\right] = \frac{1}{n}\mathfrak{M}_2 - \frac{(n-1)}{3n^2}\mathfrak{M}_2 \,\overset{\circ}{\circ}\, (\circ\otimes R(\bullet,\circ)\bullet + R(\bullet,\circ)\bullet\otimes\circ) \,\overset{\bullet}{\bullet}\, \mathfrak{M}_2 + O(\epsilon^5).$$

In coordinates:

$$\begin{aligned} Bias(\bar{x}_n)^a &= \frac{1}{6n}\left(1-\frac{1}{n}\right)\nabla_b R^a_{cde}\mathfrak{M}_2^{ce}\mathfrak{M}_2^{bd} + O\left(\epsilon^5\right) \\ Cov(\bar{x}_n)^{ab} &= \frac{1}{n}\left(\mathfrak{M}_2^{ab} - \frac{1}{3}\left(1-\frac{1}{n}\right)\mathfrak{M}_2^{cd}\left(\mathfrak{M}_2^{ae}R^b_{cde} + R^a_{cde}\mathfrak{M}_2^{be}\right)\right) + O\left(\epsilon^5\right). \end{aligned}$$

Non-Asymptotic behavior of empirical means

Moments of the Fréchet mean of a n-sample

- **Unexpected bias** in 1/n on empirical mean (**gradient of curvature-cov.**)

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

- **Concentration rate** modulated by the **curvature-covariance**:

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

  - **Asymptotically infinitely fast** CV for negative curvature
  - **No convergence (LLN fails)** at the limit of KKC condition

[XP, Curvature effects on the empirical mean in Manifolds 2019, arXiv:1906.07418]

Comparison with the BP CLT

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

- Under suitable concentration conditions, for IID n-samples:
  - $\bar{x}_n \rightarrow \bar{x}$ *(consistency of empirical mean)*
  - $\sqrt{n}\, log_{\bar{x}}(\bar{x}_n) \rightarrow N(0, 4\bar{H}^{-1}\, \Sigma\, \bar{H}^{-1})$ *if* $\bar{H} = \int_M Hess_{\bar{x}}\left(d^2(y, \bar{x})\right) \mu(dy)$ *invertible*

Hessian: $\frac{1}{2}\bar{H} = Id + \frac{1}{3} R : \mathfrak{M}_2 + \frac{1}{12} \nabla \mathrm{R} : \mathfrak{M}_3 + O(\epsilon^4, 1/n^2)$

$$4[\bar{H}^{(-1)}\mathfrak{M}_2\bar{H}^{(-1)}]^{ab} = 4[\bar{H}^{(-1)}]^a_c[\mathfrak{M}_2]^{cd}[\bar{H}^{(-1)}]^b_d = \mathfrak{M}_2^{ab} - \frac{1}{3}\mathfrak{M}_2^{ef}\left(R^a_{efc}\mathfrak{M}_2^{cb} + \mathfrak{M}_2^{ad}R^b_{efd}\right) + O(\varepsilon^5),$$

Same limiting expansion for large n

Isotropic distribution in constant curvature spaces

- Symmetric spaces: no bias
- Variance is modulated w.r.t. Euclidean: $Var(\bar{x}_n) = \alpha \frac{\sigma^2}{n}$

High concentration expansion

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

Closed form for asymptotic BP-CLT expansion

$$\frac{1}{2} H_x(y) = uu^\top + h(\kappa\theta^2)(\mathrm{Id} - uu^\top) \qquad \text{with} \qquad h(t) = \sqrt{t}\cot(\sqrt{t})$$

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

$$\overline{h} = \mathbf{E}\left[\, h(\kappa\,\mathrm{dist}(\bar{x}, .)^2)\,\right] = \int_{\mathcal{M}} h(\kappa(\,\mathrm{dist}(\bar{x}, y)^2)\,\mu(dy).$$

Isotropic distribution in constant curvature spaces

- Variance is modulated w.r.t. Euclidean: $Var(\bar{x}_n) = \alpha \frac{\sigma^2}{n}$

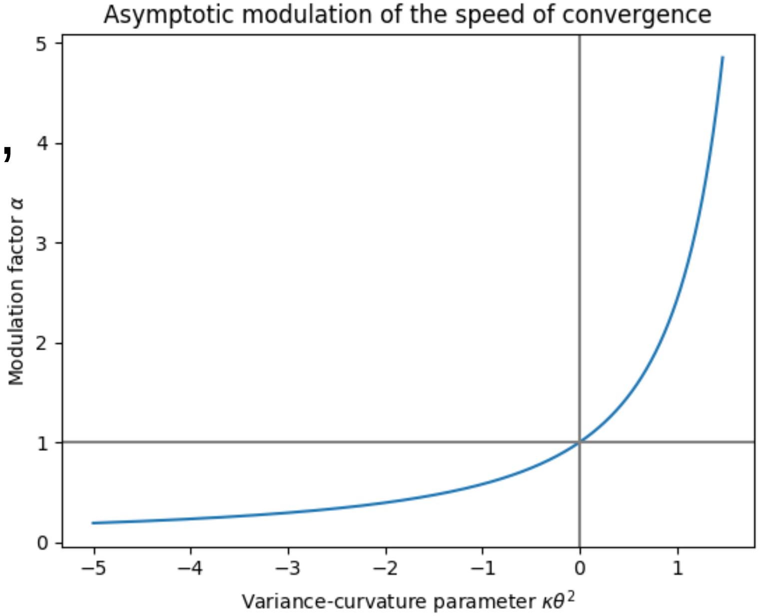
)

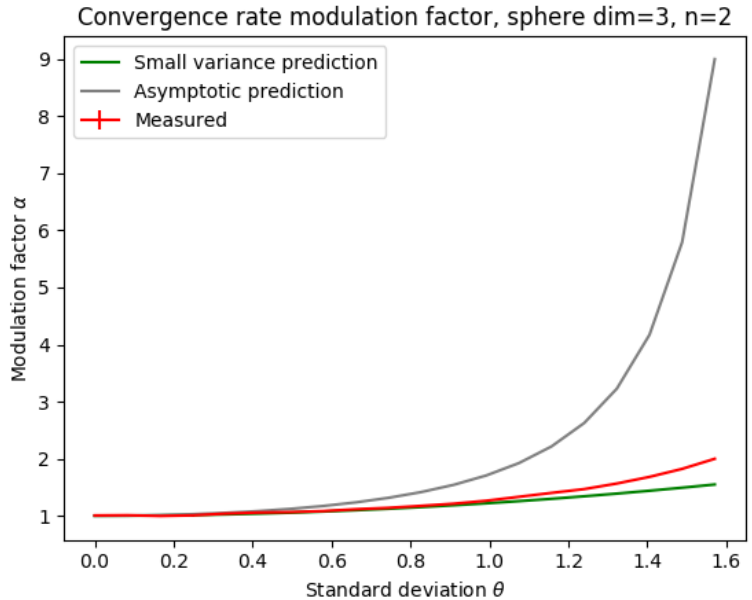
Asymptotic BP-CLT expansion

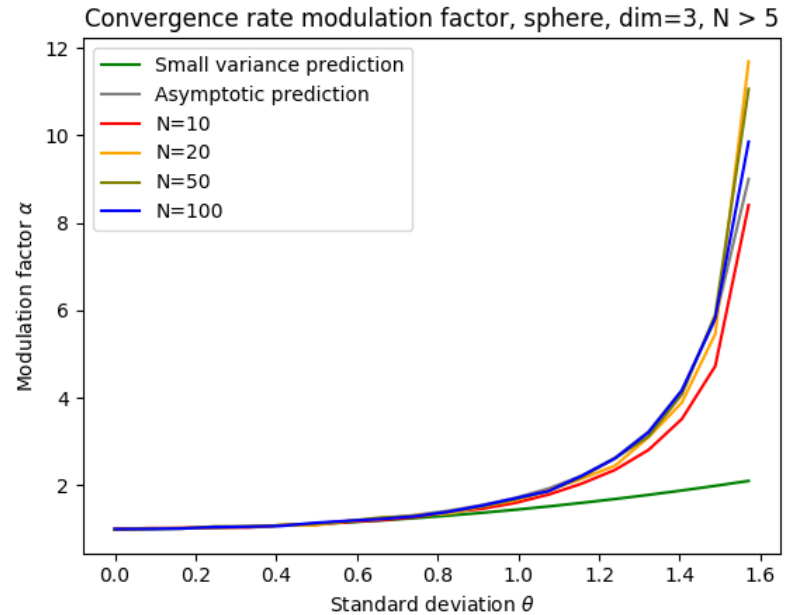
- $\alpha = \left(\frac{1}{d} + \left(1 - \frac{1}{d}\right)\bar{h}\right)^{-2} + O(n^{-2})$

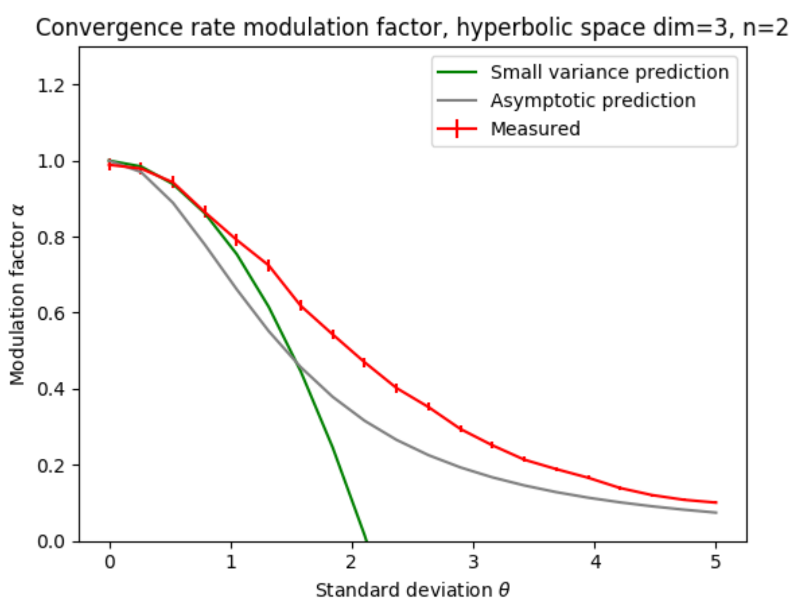
Archetypal modulation factor

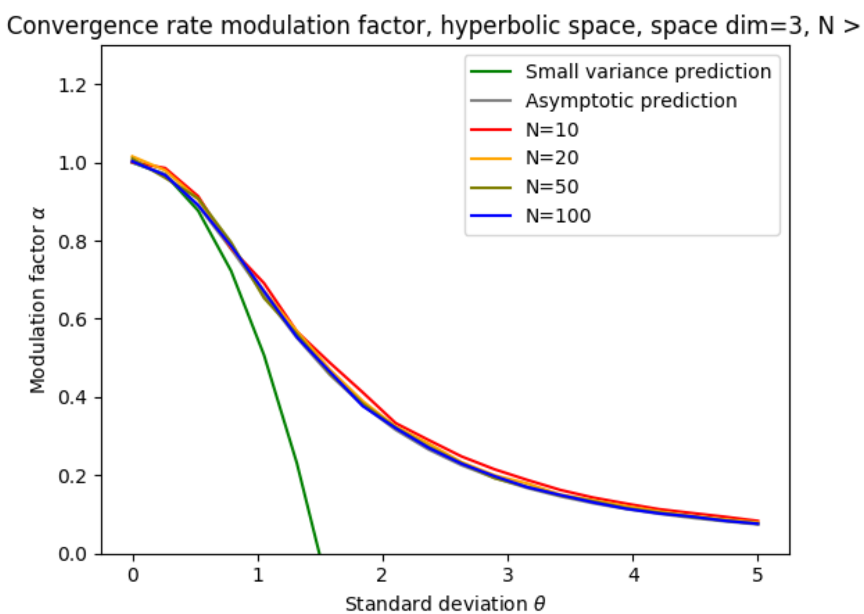
- Uniform distrib on $S(\bar{x}, \theta) \subset M$, large n, large d
- $\alpha = \frac{\tan^2\left(\sqrt{\kappa\theta^2}\right)}{\kappa\theta^2}$




Convergence rate modulation factor, sphere dim=3, n=2
Small variance prediction
Asymptotic prediction
Measured
Modulation factor α
Standard deviation θ


Convergence rate modulation factor, sphere, dim=3, N > 5
Small variance prediction
Asymptotic prediction
N=10
N=20
N=50
N=100
Modulation factor α
Standard deviation θ


Convergence rate modulation factor, hyperbolic space dim=3, n=2
Small variance prediction
Asymptotic prediction
Measured
Modulation factor α
Standard deviation θ


Convergence rate modulation factor, hyperbolic space, space dim=3, N > 5
Small variance prediction
Asymptotic prediction
N=10
N=20
N=50
N=100
Modulation factor α
Standard deviation θ

Conclusions

High concertation expansion very accurate for low theta

Asymptotic expansion very accurate for n> 10

Main variable controlling the modulation is variance-curvature tensor

$$R(\blacksquare, ^\circ)\blacksquare : \mathfrak{M}_2$$

Main variable controling the bias

$$\mathfrak{M}_2 : \nabla \circ R(^\circ, \blacksquare)\blacksquare : \mathfrak{M}_2$$

Geometric Statistics: Mathematical foundations and applications in computational anatomy

Intrinsic Statistics on Riemannian Manifolds

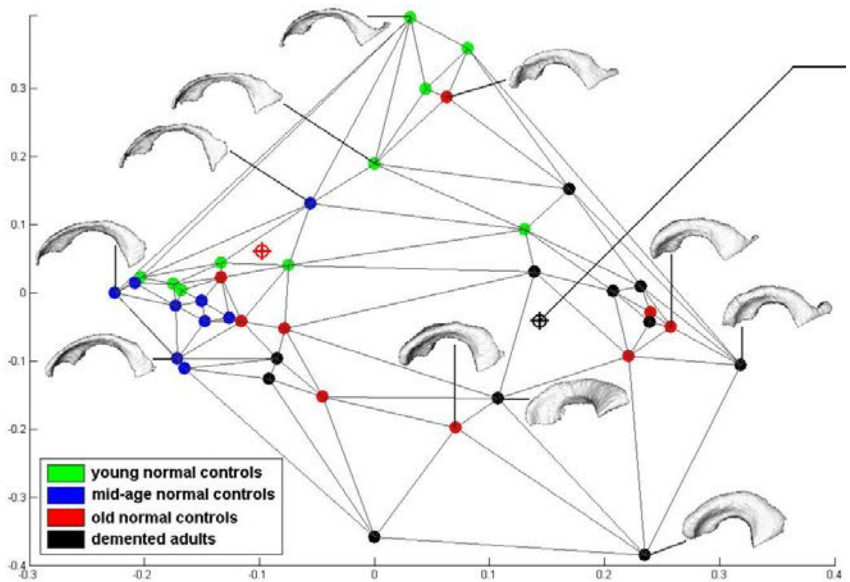
Metric and Affine Geometric Settings for Lie Groups

Parallel Transport to Analyze Longitudinal Deformations

Advances Statistics: CLT & PCA

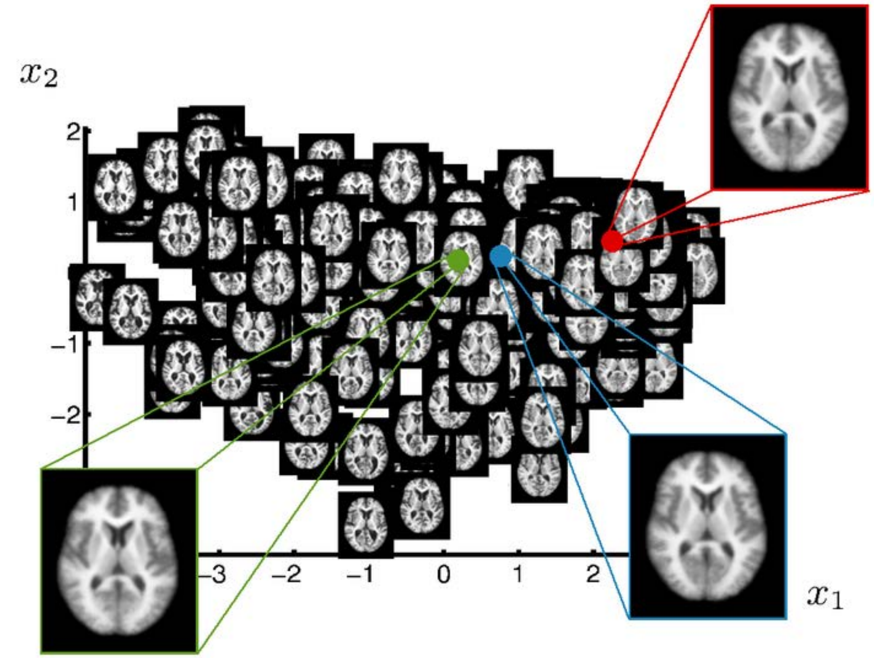
- Estimation of the empirical Fréchet mean & CLT
- Principal component analysis in manifolds
- Natural subspaces in manifolds: barycentric subspaces
- Rephrasing PCA with flags of subspaces

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 $\Sigma(x) \propto \sum_i \overrightarrow{\bar{x}x_i}\, \overrightarrow{\bar{x}x_i}^t$

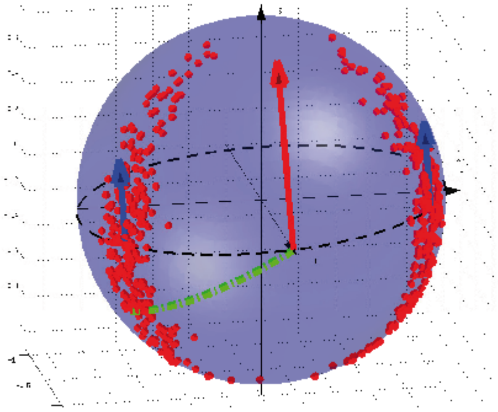
- Generative model:
  - Gaussian (large variance) in the horizontal subspace
  - Gaussian (small variance) in the vertical space

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

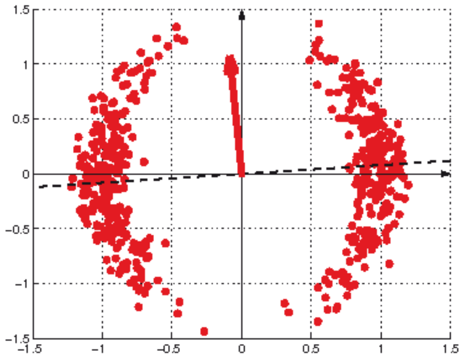
Problems of tPCA

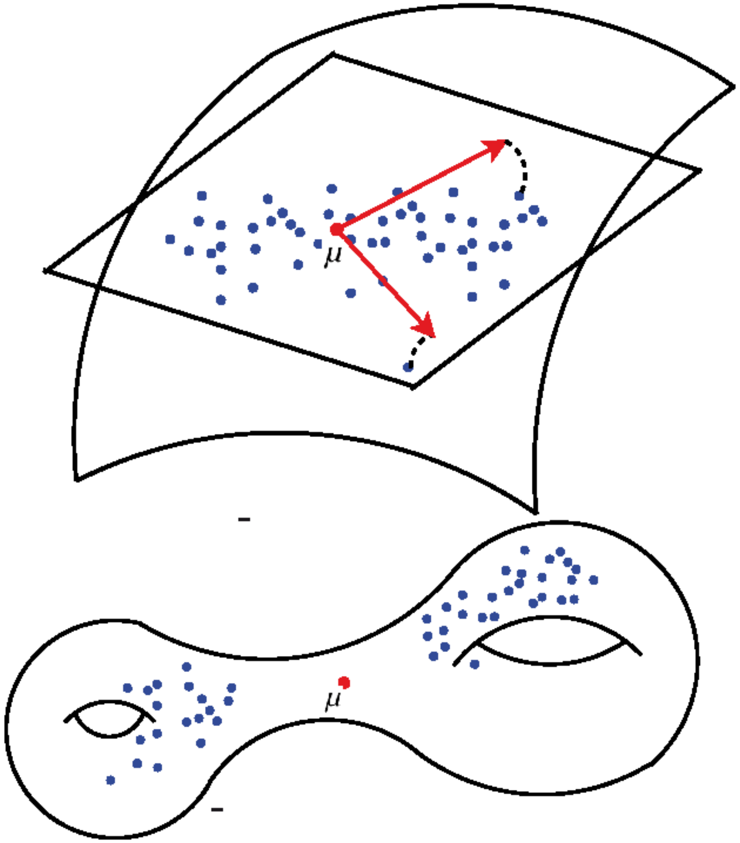
Analysis is done relative to the mean

- What if the mean is a poor description of the data?
  - Multimodal distributions
  - Uniform distribution on subspaces
  - Large variance w.r.t curvature



Bimodal distribution on S2





Images courtesy of S. Sommer

Principal Geodesic / Geodesic Principal Component Analysis

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

- **Geodesic Subspace**: $GS(x, w_1, \ldots w_k) = \{\exp_x(\sum_i \alpha_i w_i) \;\; for\; \alpha \in R^k\}$
  - 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 $GS(x, w_1, \ldots w_k)$

- Totally geodesic at x only

Patching the Problems of tPCA / PGA

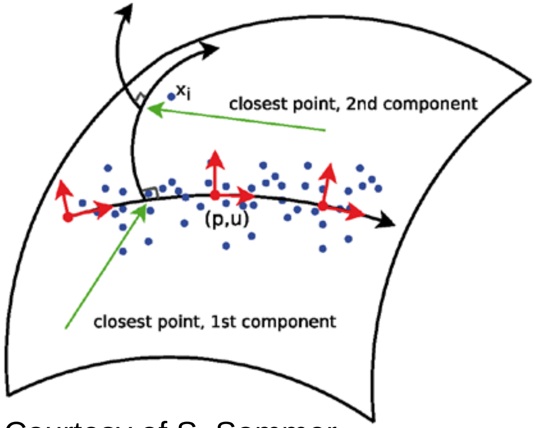
Improve the flexibity of the geodesics

- 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**

Geometric Statistics: Mathematical foundations and applications in computational anatomy

Intrinsic Statistics on Riemannian Manifolds

Metric and Affine Geometric Settings for Lie Groups

Parallel Transport to Analyze Longitudinal Deformations

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

Affine span in Euclidean spaces

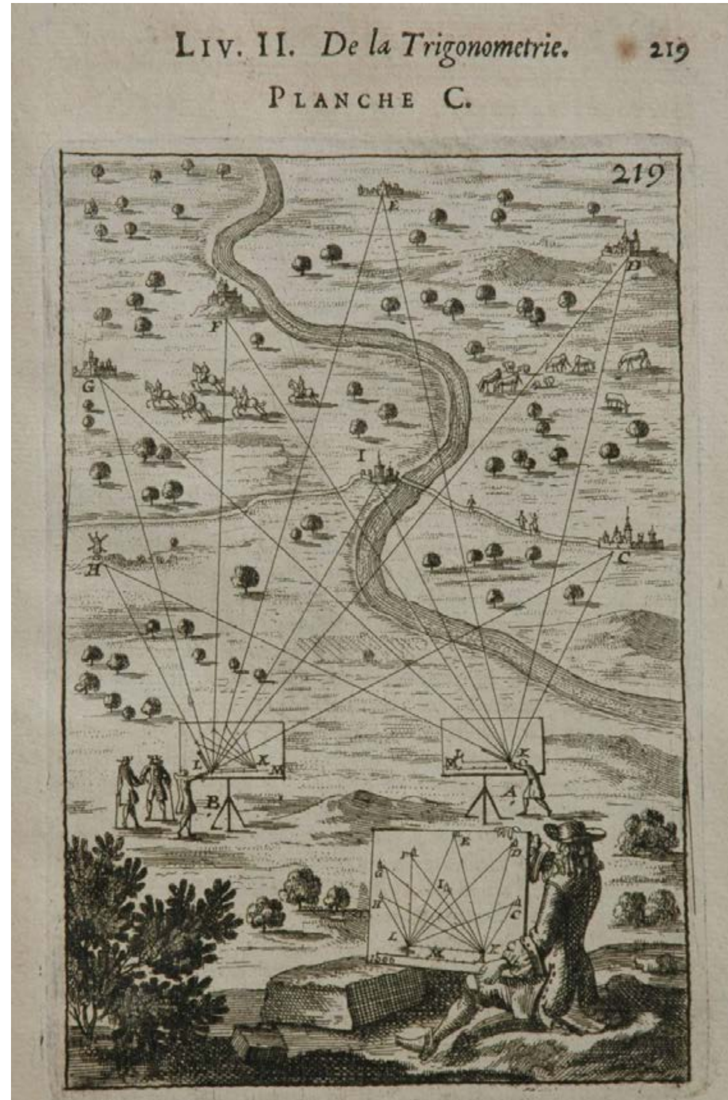
Affine span of (k+1) points: weighted barycentric equation

$$\mathrm{Aff}(x_0, x_1, \ldots x_k) = \{x = \textstyle\sum_i \lambda_i x_i \;\; with \; \sum_i \lambda_i = 1\}$$
$$= \{x \in R^n \;\; s.t \; \textstyle\sum_i \lambda_i (x_i - x) = 0, \lambda \in P_k^*\}$$

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$(x_0, \ldots x_k) = \{\, x \in M^*(x_0, \ldots x_k) \,|\mathfrak{M}_1(x,\lambda) = 0\}$
- Affine span = closure of EBS in M $\quad Aff(x_0, \ldots x_k) = \overline{EBS(x_0, \ldots x_k)}$

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

Assumptions:

- Restrict to the **punctured manifold** $M^*(x_0, \dots x_k) = M \,/ \cup C(x_i)$
  - $dist^2(x, x_i)$, $\log_x(x_i)$ are smooth but M^* may be split in pieces
- Affinely independent points:
  $\{\overrightarrow{x_i x_j}\}_{0 \le i \neq j \le k}$ exist and are linearly independent for all i

Local well posedness for the barycentric simplex:

- EBS / KBS are well defined in a neighborhood of reference points
- For reference points in a sufficiently small ball and positive weights: unique Frechet = Karcher = Exp Barycenter in that ball: smooth graph of a k-dim function [proof using Buser & Karcher 81]

SVD characterization of EBS: $\mathfrak{M}_1(\mathrm{x}, \lambda) = Z(x)\lambda = 0$

- SVD: $Z(x) = [\, \overrightarrow{x x_0}, \dots \overrightarrow{x x_k}] = U(x) S(x) V^t(x)$
  - $EBS(x_0, \dots x_k)$ = Zero level-set of l>0 singular values of Z(x)
  - Stratification on the number of vanishing singular values

[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*

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

Caractérisation of local minima: Hessian (if non degenerate)

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

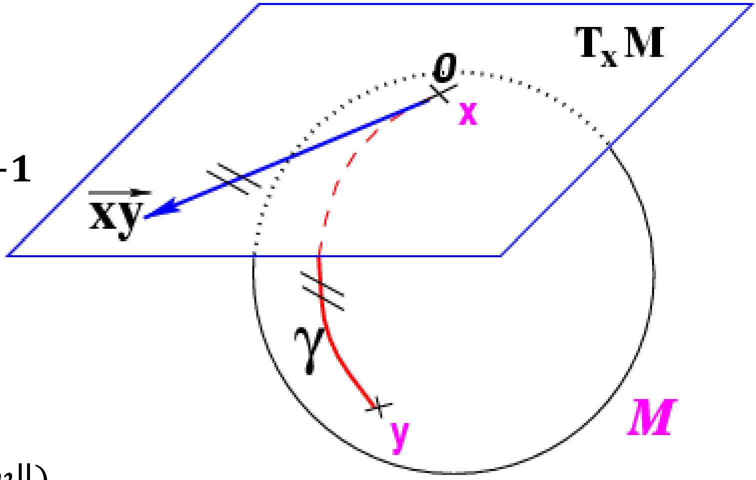
Regular and positive pts (non-degenerated critical points)

- $\boldsymbol{EBS^{Reg}(x_0, \dots x_k) = \{ x \in Aff(x_0, \dots x_k), s.t.\ H(x, \lambda^*(x)) \neq 0\}}$
- $\boldsymbol{EBS^{+}(x_0, \dots x_k) = \{ x \in Aff(x_0, \dots x_k), s.t.\ H(x, \lambda^*(x))\ Pos.def.\}}$

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]

Example on the sphere



Manifold

- Unit sphere $\mathcal{M} = \boldsymbol{S_n}$ embedded in $\mathbb{R}^{n+1}$
- ||x|| =1

Exp and log map

$$\exp_x(v) = \cos(\|v\|)\, x + \frac{\sin(\|v\|)}{\|v\|} v$$

$$\log_x(y) = f(\theta)(y - \cos(\theta)) \quad \text{with} \quad \theta = \arccos(x^t y)$$

Distance

$$dist(x,y) = \|\log_x(y)\| = \theta$$

(k+1)-pointed & punctured Sphere

- $X = [x_0, x_1, \dots, x_k] \in (S_n)^k$
- Punctured sphere: exclude antipodal points: $S_n^* = S_n / -X$

KBS / FBS with 3 points on the sphere

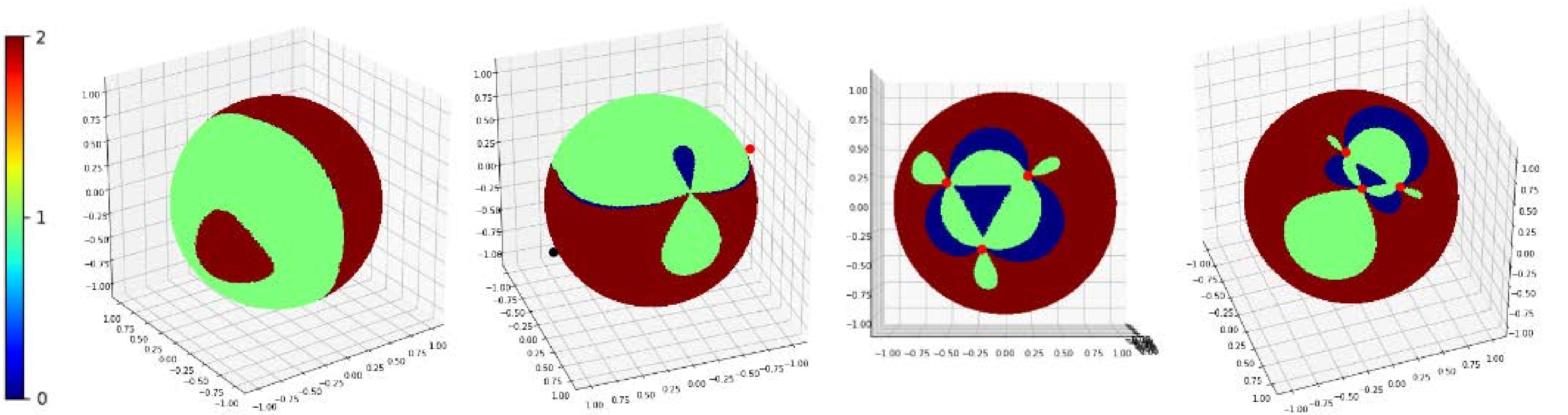
EBS: great subspheres spanned by reference points (mod cut loci)

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

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

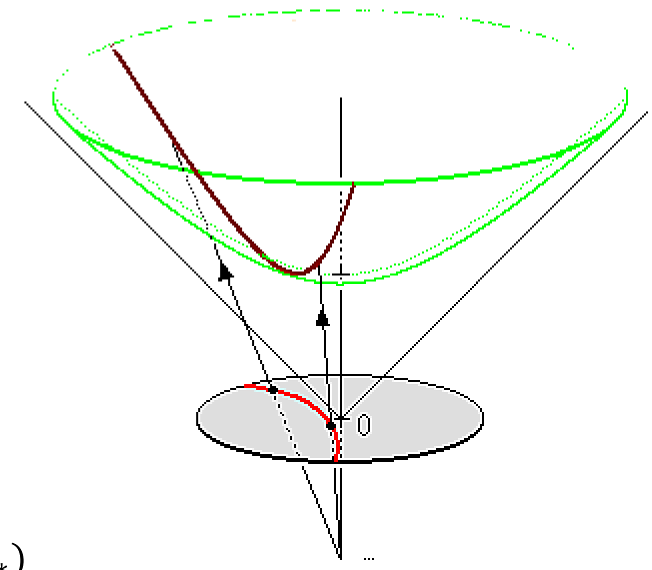
$$\mathrm{H}(\mathrm{x},\lambda) = \sum \lambda_i \theta_i \cot(\theta_i)\,(\mathrm{Id} - \mathrm{xx}^{\mathrm{t}}) + \sum (1 - \lambda_i \theta_i \cot(\theta_i))\,\overrightarrow{xx_i}\,\overrightarrow{xx_i}^t$$

- 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



Manifold

- Unit pseudo-sphere $\mathcal{M} = \boldsymbol{H_n}$ embedded in Minkowski space $\mathbb{R}^{\mathbf{1},\boldsymbol{n}}$
- $\|x\|_*^2 = -x_0^2 + x_1^2 + \cdots x_n^2 = -1$

Exp and log map

$$\exp_x(v) = \cosh(\|v\|_*)\, x + \frac{\sinh(\|v\|_*)}{\|v\|_*} v$$

$$\log_x(y) = f_*(\theta)(y - \cosh(\theta)) \quad \text{with} \quad \theta = \operatorname{arcosh}(-\langle x|y\rangle_*)$$

Distance

$$dist(x,y) = \|\log_x(y)\|_* = \theta$$

Punctured hyperbolic space: no cut locus to exclude

Example on the hyperbolic space

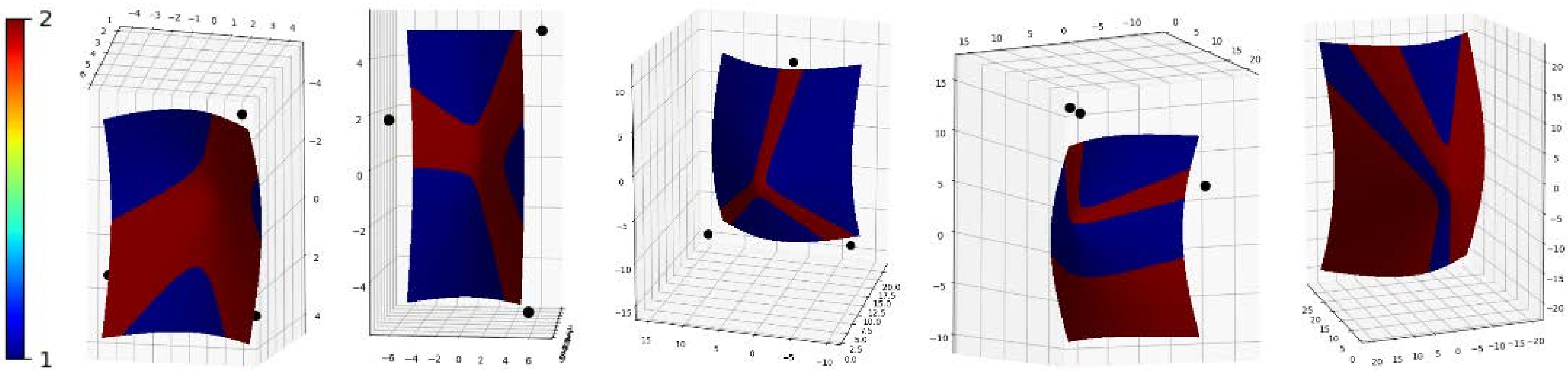
EBS = Affine span: great sub-hyperboloids spanned by reference points

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

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

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

- 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

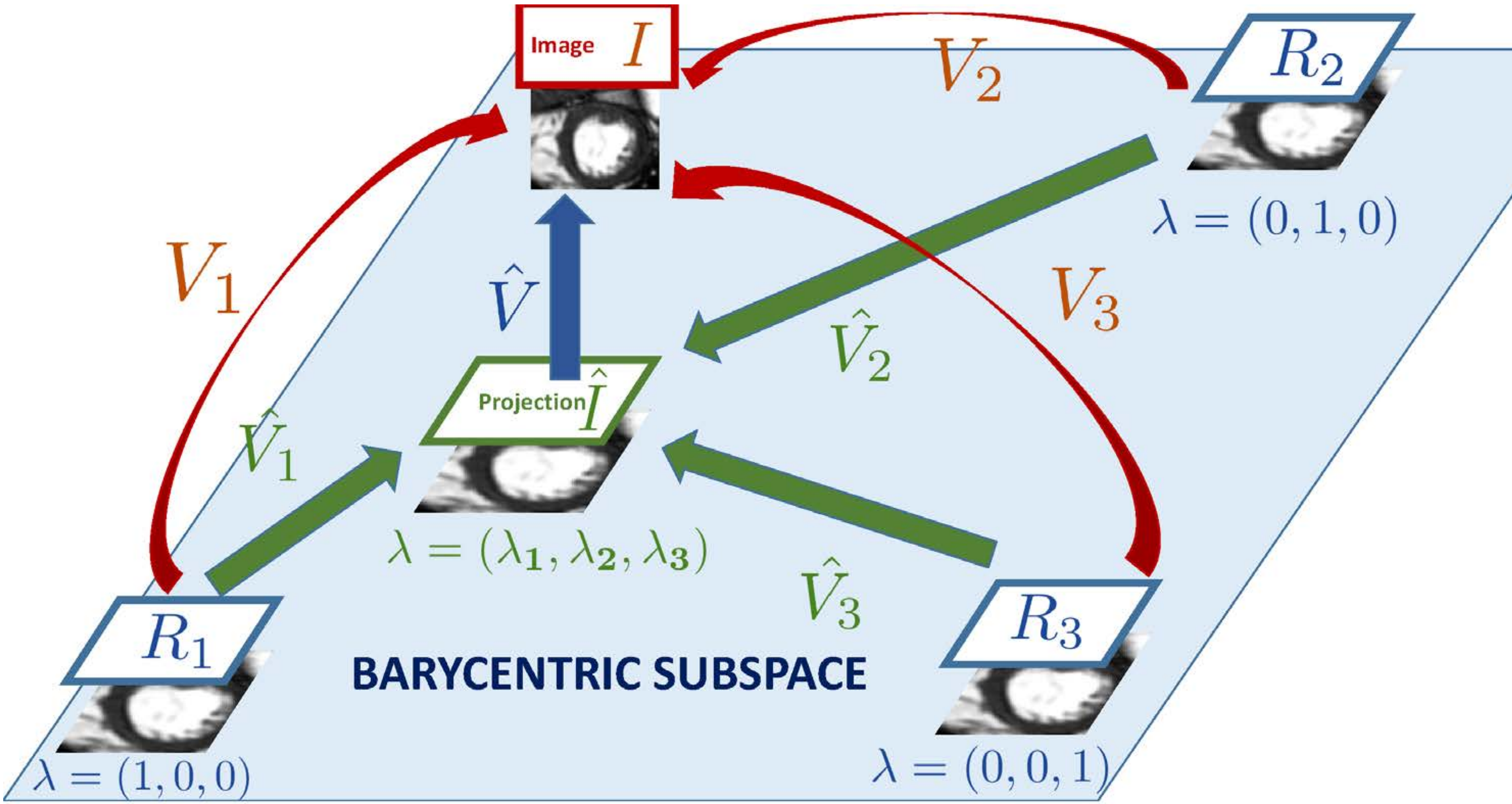
Theorem

- $GS(x, w_1, \ldots w_k) = \{\exp_x(\sum_i \alpha_i w_i) \;\; for\; \alpha \in R^k\}$ 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

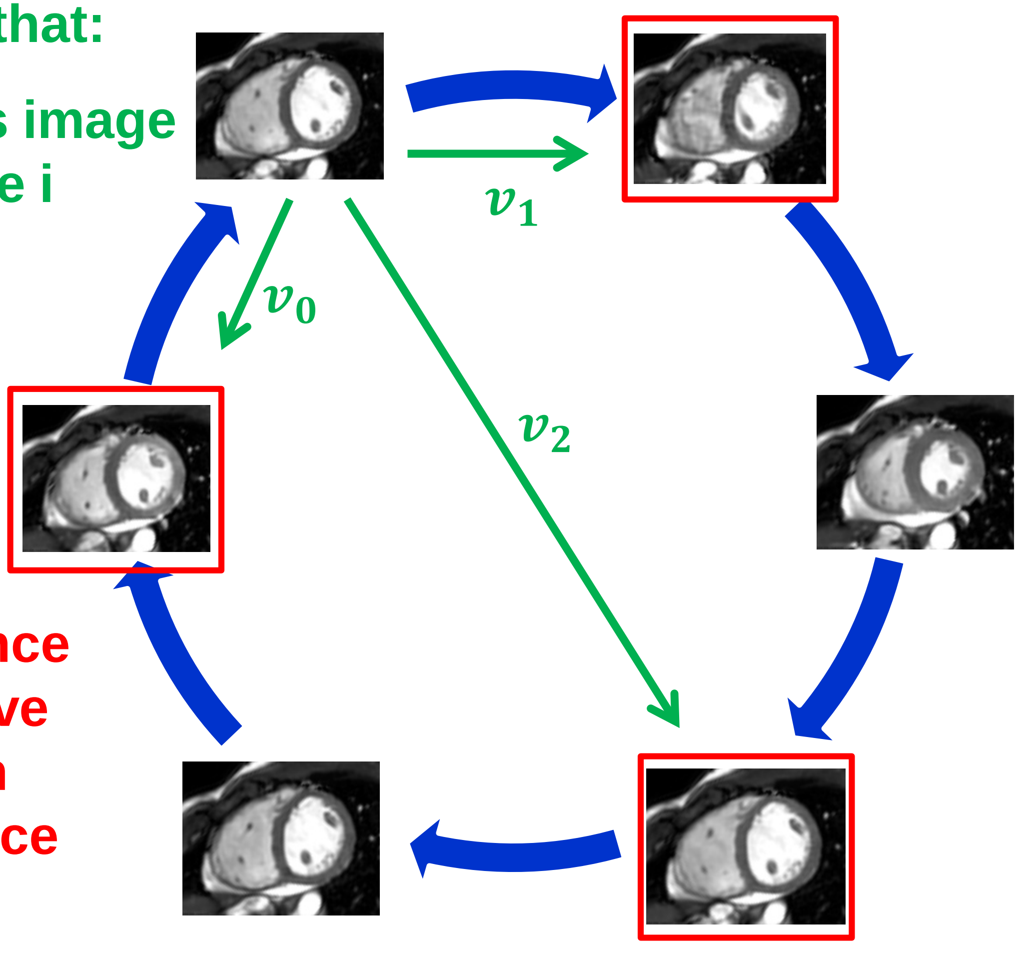


[Marc-Michel Rohé et al., MICCAI 2016, MedIA 45:1-12, 2018]

Application in Cardiac motion analysis

Find weights λ_i and SVFs v_i such that:

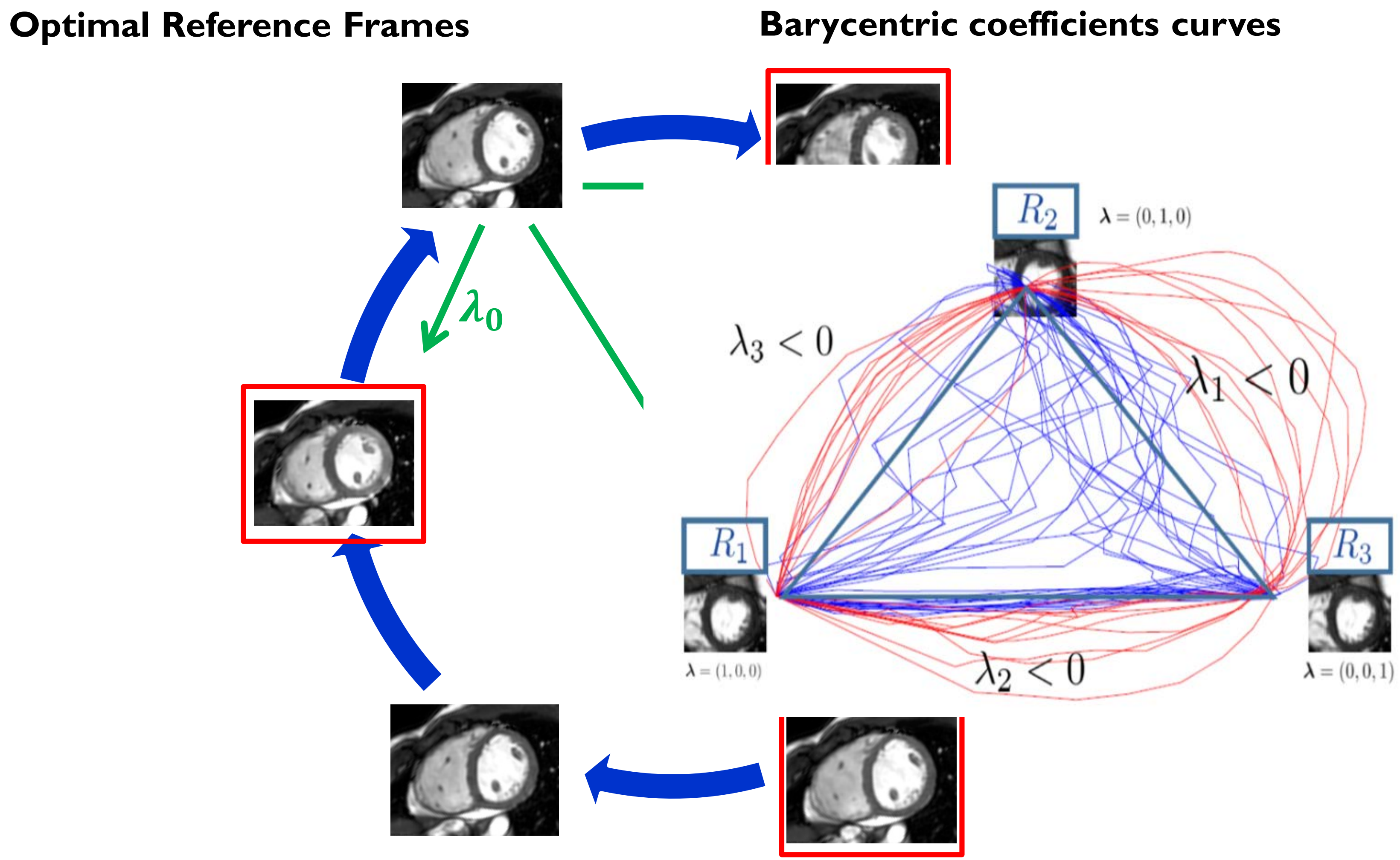
- v_i **registers image to reference i**
- $\sum_i \lambda_i v_i = 0$



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

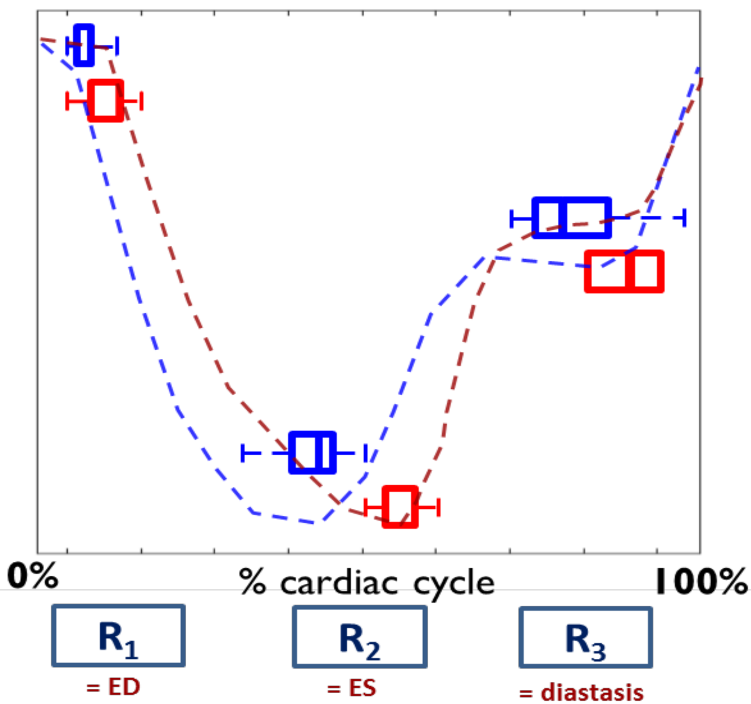


[Marc-Michel Rohé et al., MICCAI 2016, MedIA 45:1-12, 2018]

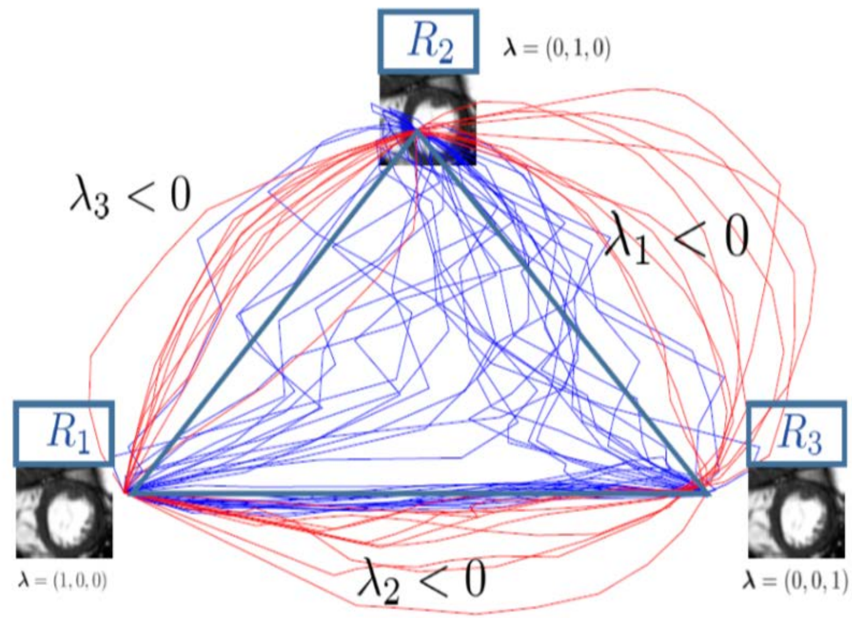
Cardiac Motion Signature

Low-dimensional representation of motion using:

Optimal Reference Frames



Barycentric coefficients curves



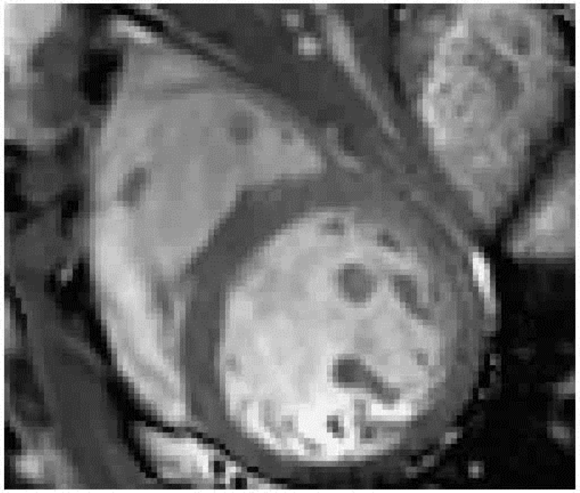
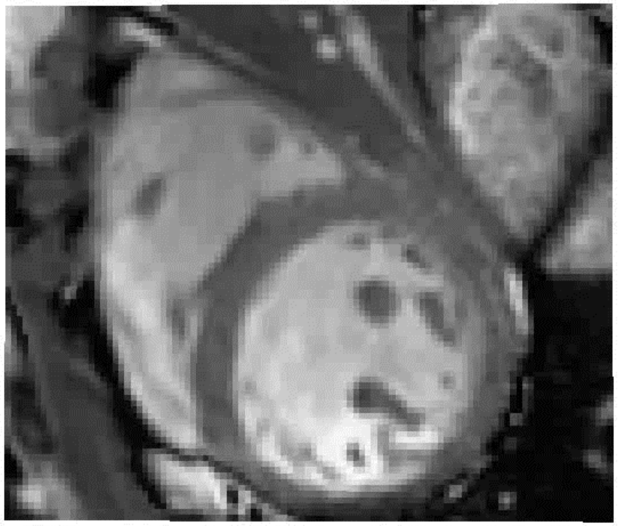
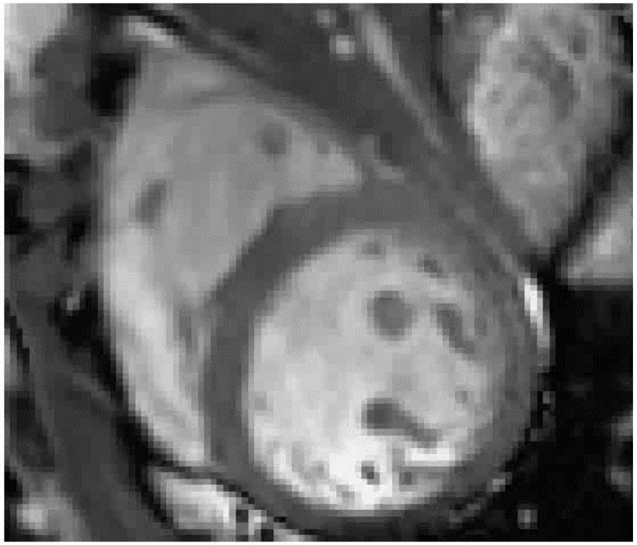
Dimension reduction from **+10M voxels** to **3 reference** frames + **60 coefficients**

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

Manifold-Valued Image Processing

Metric and Affine Geometric Settings for Lie Groups

Parallel Transport to Analyze Longitudinal Deformations

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

Forward, Backward and Nested Analysis

Forward Barycentric Subspace (k-FBS) decomposition

- Iteratively add points x_j from j=0 to k
- $x_0 = Mean(y_j)$, $x_1 = argmin_x(\ \sigma^2_{out}(x_0, x))$ … PGA-like
- Start with 2 points: $(x_0, x_1) = \text{argmin}_{(x,y)}(\ \sigma^2_{out}(x, y))$ GPGA-like

Backward analysis: Pure Barycentric Subspace (k-PBS)

- Find $Aff(x_0, \dots x_k)$ minimizing the unexplained variance:

$$\sigma^2_{out}(x_0, \dots x_k) = \sum_j \ dist^2\left(y_j,\ Proj_{Aff(x_0 \dots x_k)}(y_j)\right)$$

- Iteratively remove one point from $(x_0, \dots x_j)$ from j=0 to k
- One optimization only for k+1 points and discrete backward reordering

From greedy to global optimization?

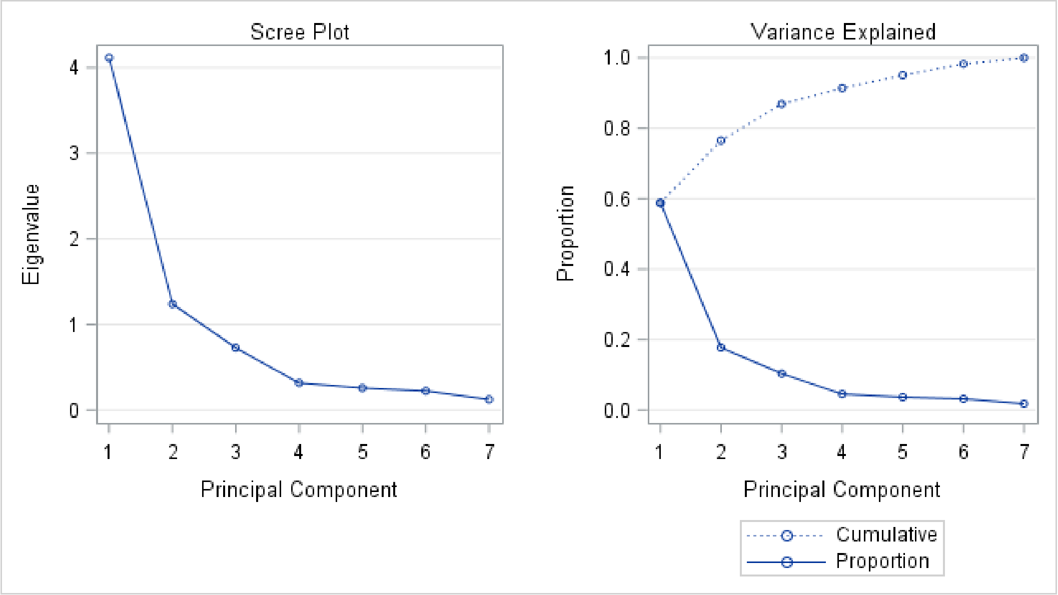
- Optimal unexplained variance → non nested subspaces
- Nested forward / backward procedures → not optimal
- Optimize first, decide dimension later → Nestedness required
  [Principal nested relations: Damon, Marron, JMIV 2014]

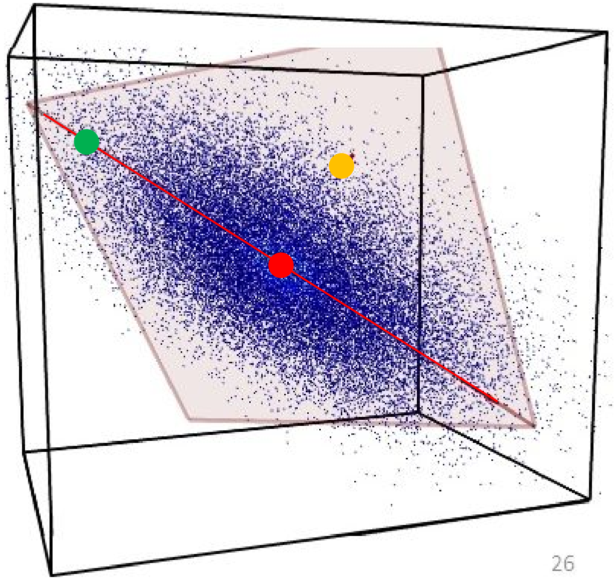
Barycentric Subspace Analysis (k-BSA)

The natural object for PCA: Flags of subspaces in manifolds

- $x_0 \prec x_1 \prec \cdots \prec x_k$ are k +1 n distinct ordered points of M.
- $FL(x_0 \prec x_1 \prec \cdots \prec x_k)$ is the sequence of properly nested subspaces $FL_{i(x_0 \prec x_1 \prec \cdots \prec x_k)} = Aff(x_0, \ldots x_i)$

$$Aff(x_0) = \{x_0\} \subset \ \ldots Aff(x_0, \ldots x_k) \ \ldots \subset Aff(x_0, \ldots x_n) = M$$
$$\sigma^2_{out}(x_0) \geq \ \ldots \geq \ \sigma^2_{out}(x_0, \ldots x_k) \geq \ \ldots \geq \ \sigma^2_{out}(x_0, \ldots x_n) = 0$$





Adapted from 3DM slides by Marc van Kreveld

Barycentric Subspace Analysis (k-BSA)

Accumulated unexplained variance (area under the curve)

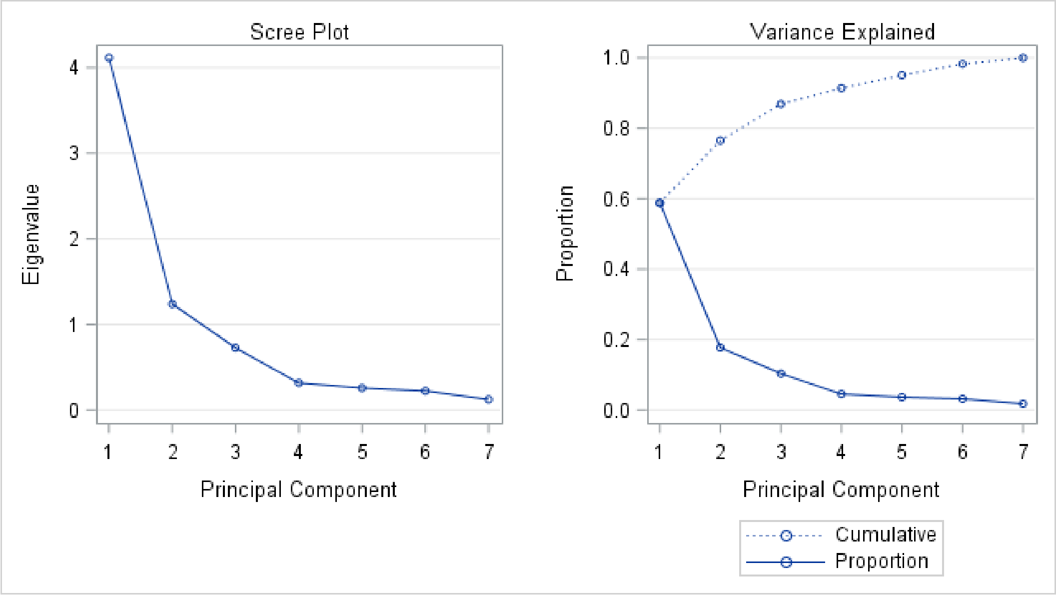
- k-BSA optimizes: $AUV(k) = \sum_{i=0}^{k} \sigma_{out}^2(x_0, \dots x_i)$
- In a Euclidean space with Gaussian $N(x_0, \Sigma = diag(\sigma_1^2, \dots \sigma_n^2))$

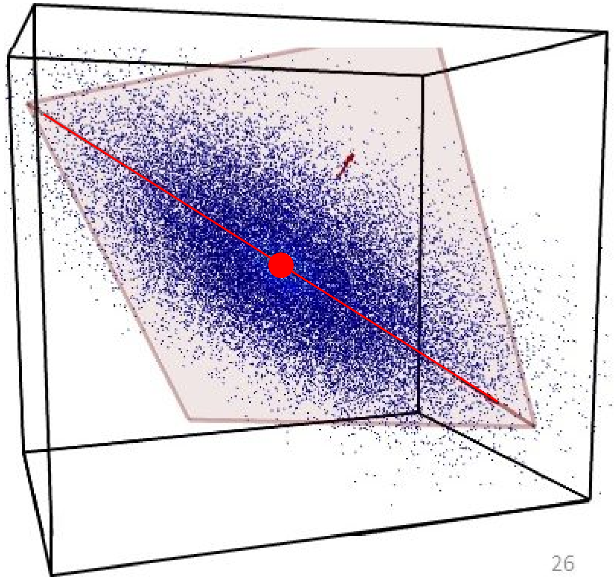
  $\sigma_{out}^2(x_0, \dots x_i) = \sigma_{i+1}^2 + \cdots \sigma_n^2 \rightarrow AUV(k) = \sum_{i=0}^{k} i\, \sigma_i^2 + (k+1) \sum_{i=k+1}^{n} \sigma_i^2$

$\rightarrow$ minimal for ordered eigenmodes of Σ with $\sigma_1 \geq \sigma_2 \; \dots \geq \sigma_n$

\

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



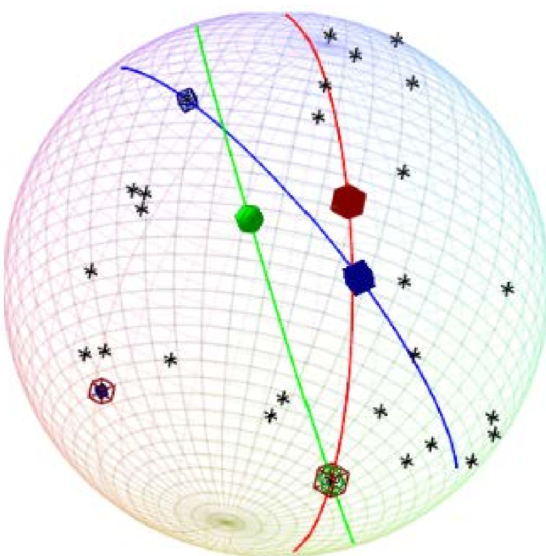


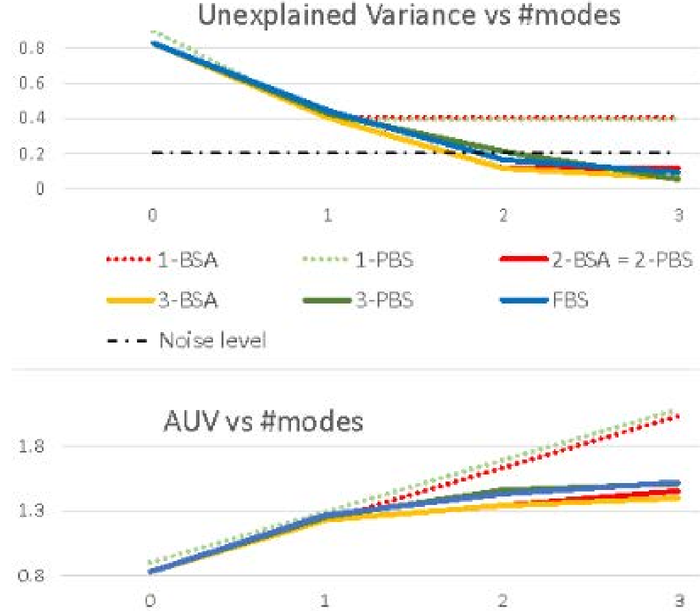
Adapted from 3DM slides by Marc van Kreveld

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)

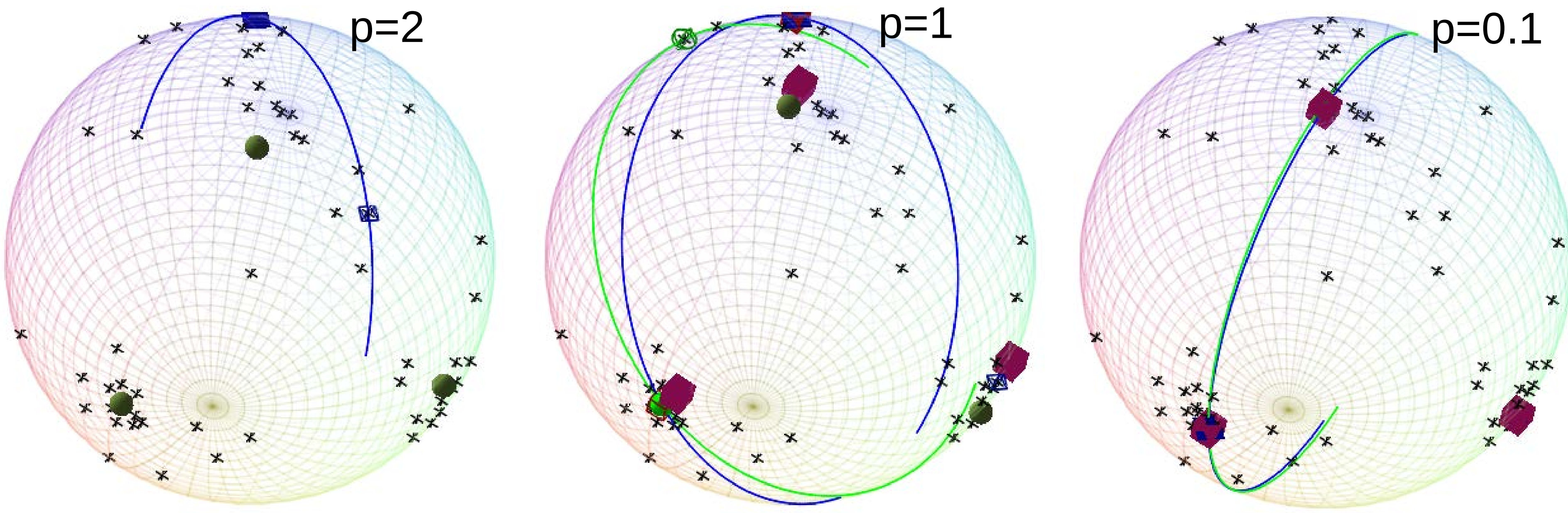
Unexplained p-variance of residuals

- $2 < p \rightarrow +\infty$: 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

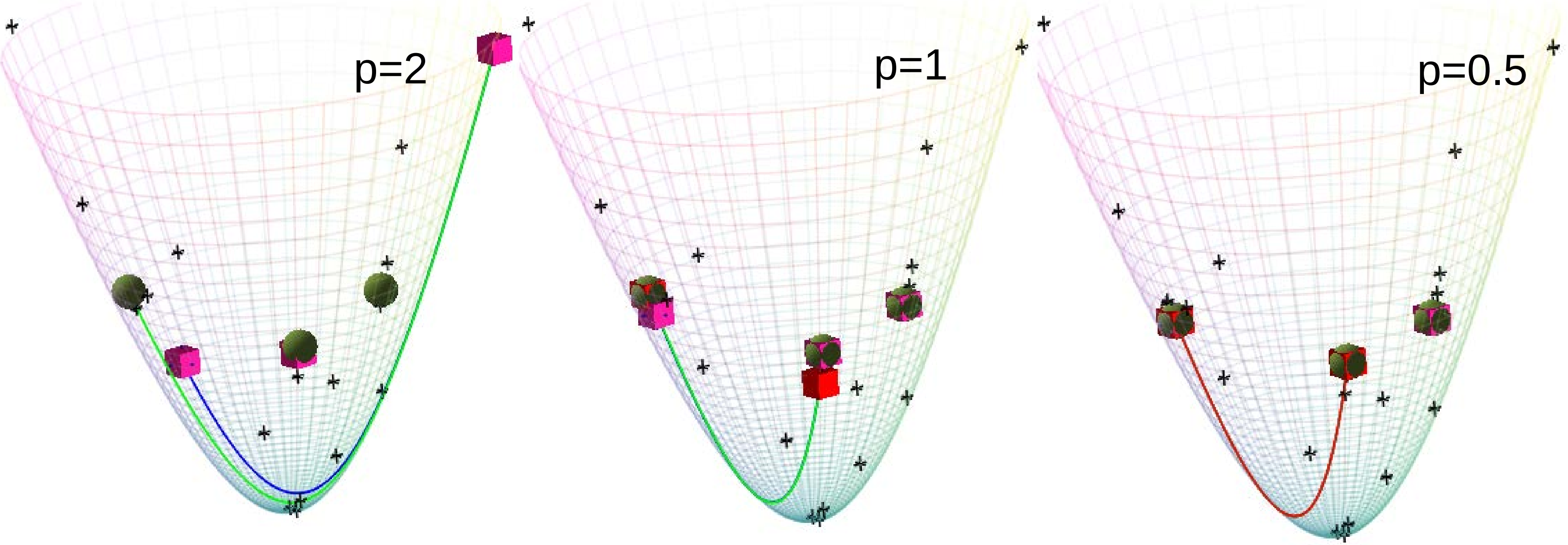


- **FBS: Forward Barycentric Subspace: mean and median not in clusters**
- **1-PBS / 2-PBS: Pure Barycentric Subspace with backward ordering: ok for k=2 only**
- **1-BSA / 2-BSA: Barycentric Subspace Analysis up to order k: less sensitive to p & k**

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 $p \leq 0.5$**
- **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

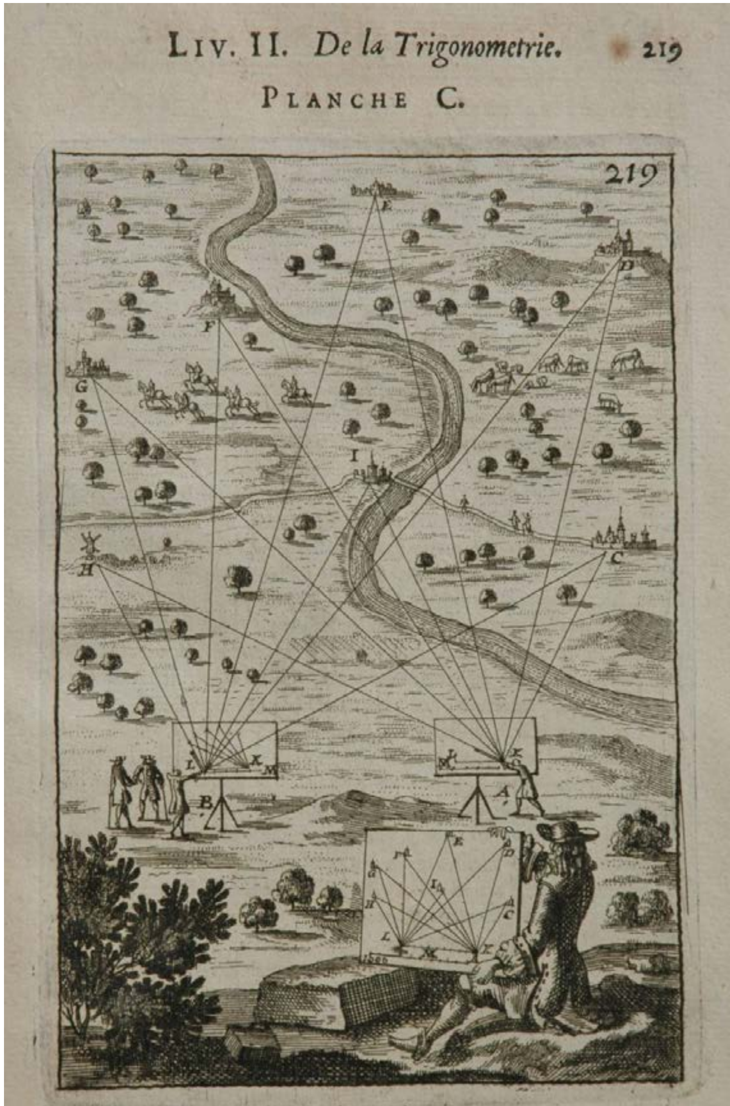
- PGA & Godesic subspaces:
  look at data points from the (unique) mean
- Barycentric subspaces:
  « triangulate » several reference points
  - Justification of multi-atlases?

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

Open research avenues

Other iterative least squares methods?

- ICA, PLS
- Manifold learning → Submanifold learning

Modulate BSA to account for within subspace distribution

- Gaussian: central points
- Clusters: mixtures of modes
- Extremal references: archetypal analysis

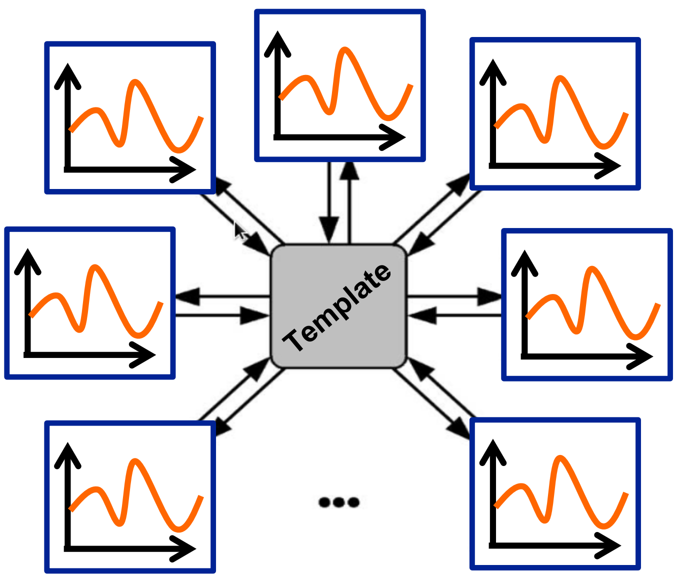
And applications

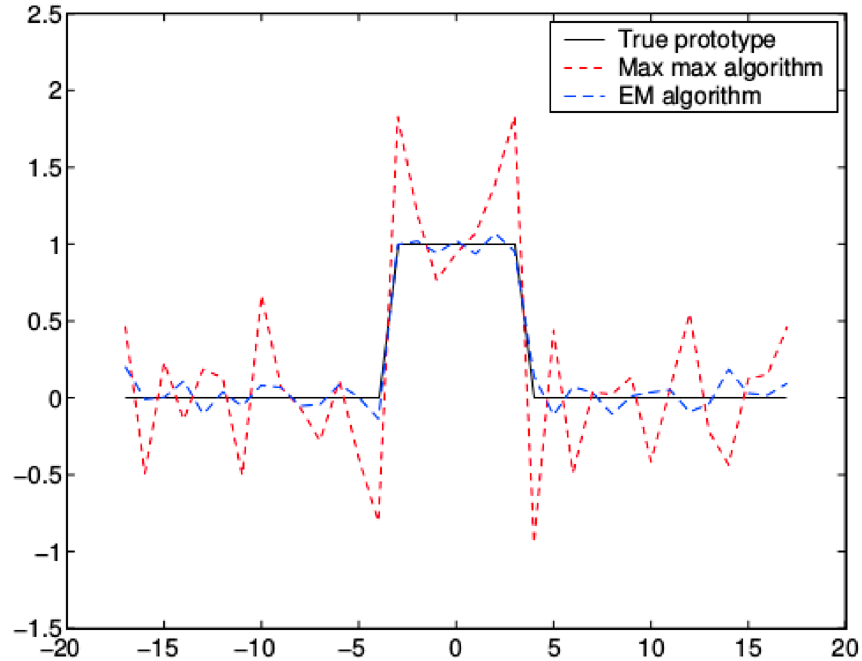
- Multi-atlases (brains, heart motion image sequences)
- SPD matrices (BCI)

Quotient spaces

Functions/Images modulo time/space parameterization

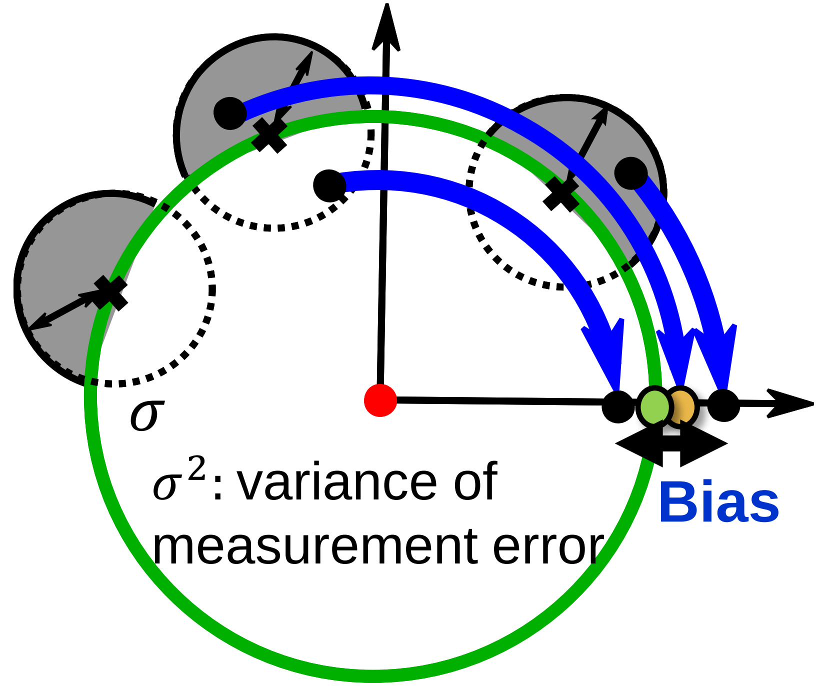
- Amplitude and phase discrimination problem





[Allassoniere, Amit, Trouvé, 2005],
Example by Loic Devillier, IPMI 2017

Noise in top space = Bias in quotient spaces



The curvature of the **template shape's orbit and presence of noise** creates a repulsive bias

Theorem [Miolane et al. (2016)]: Bias of estimator $\hat{T}$ of the template T

$$\text{Bias}(\hat{T}, T) = \frac{\sigma^2}{2} \boldsymbol{H}(T) + \mathcal{O}(\sigma^4)$$

where $\mathbf{H}(T)$: **mean curvature vector of template's orbit**

Extension to Hilbert of ∞-dim: bias for $\sigma > 0$, asymptotic for $\sigma \to \infty$,
[Devilliers, Allasonnière, Trouvé and XP. SIIMS 2017, Entropy, 2017]

→ Estimated atlas is topologically more complex than should be

References on Barycentric Subpsace Analysis

- **Barycentric Subspace Analysis on Manifolds**
  X. P. Annals of Statistics. 46(6A):2711-2746, 2018. [arXiv:1607.02833]
  - **Barycentric Subspaces and Affine Spans in Manifolds** Geometric Science of Information GSI'2015, Oct 2015, Palaiseau, France. LNCS 9389, pp.12-21, 2015.
    Warning: change of denomination since this paper: EBS →affine span
  - **Barycentric Subspaces Analysis on Spheres** Mathematical Foundations of Computational Anatomy (MFCA'15), Oct 2015, Munich, Germany. pp.71-82, 2015. https://hal.inria.fr/hal-01203815
- **Sample-limited L p Barycentric Subspace Analysis on Constant Curvature Spaces.** X.P. Geometric Sciences of Information (GSI 2017), Nov 2017, Paris, France. LNCS 10589, pp.20-28, 2017.
- **Low-Dimensional Representation of Cardiac Motion Using Barycentric Subspaces: a New Group-Wise Paradigm for Estimation, Analysis, and Reconstruction.** M.M Rohé, M. Sermesant and X.P. Medical Image Analysis vol 45, Elsevier, April 2018, 45, pp.1-12.
  - **Barycentric subspace analysis: a new symmetric group-wise paradigm for cardiac motion tracking.** M.M Rohé, M. Sermesant and X.P. Proc of MICCAI 2016, Athens, LNCS 9902, p.300-307, Oct 2016.