Xavier Pennec

Univ. Côte d'Azur and Inria, France

Geometric Statistics for computational anatomy

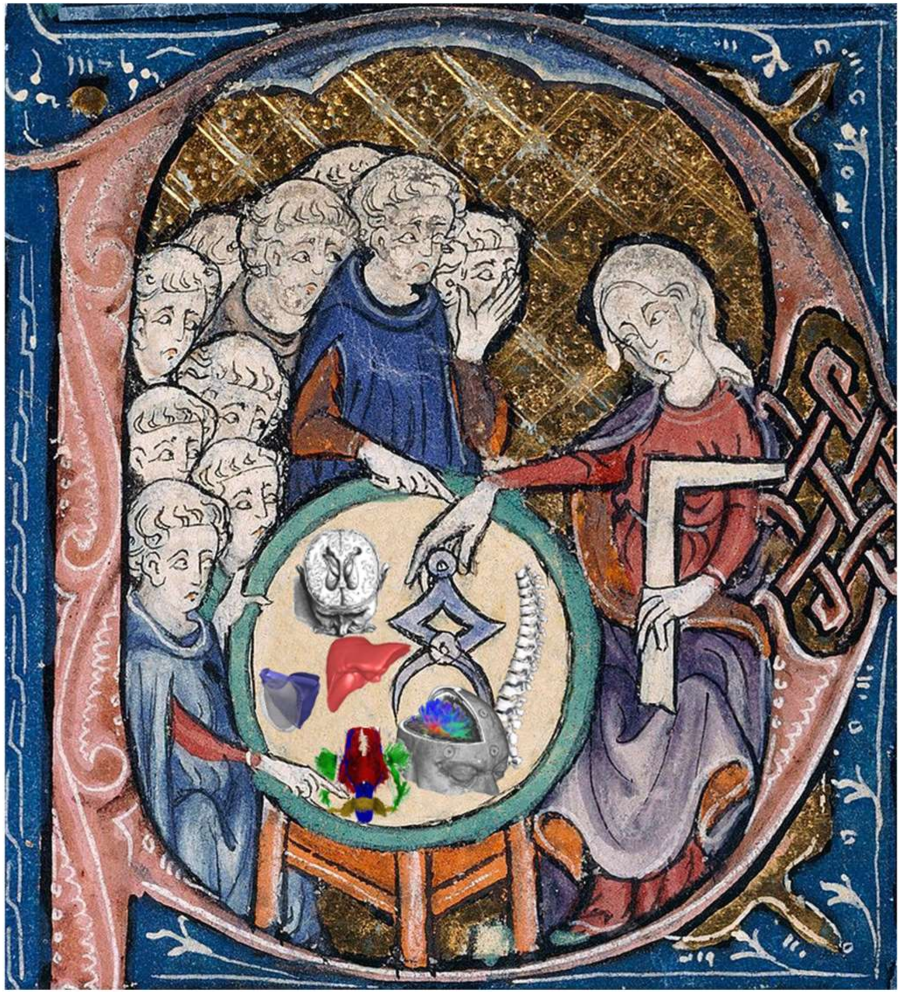
UNIVERSITÉ CÔTE D'AZUR

3iA Côte d'Azur
Interdisciplinary Institute for Artificial Intelligence

Inria Epione
e-patient / e-medicine

erc

ERC AdG 2018-2023 *G-Statistics*

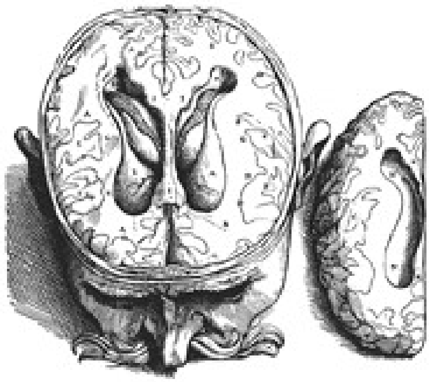


Freely adapted from "Women teaching geometry", in Adelard of Bath translation of Euclid's elements, 1310.

Geometry & ML workshop 12/07/2021

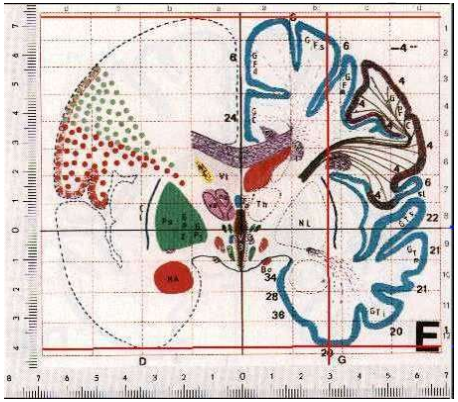
From anatomy…

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

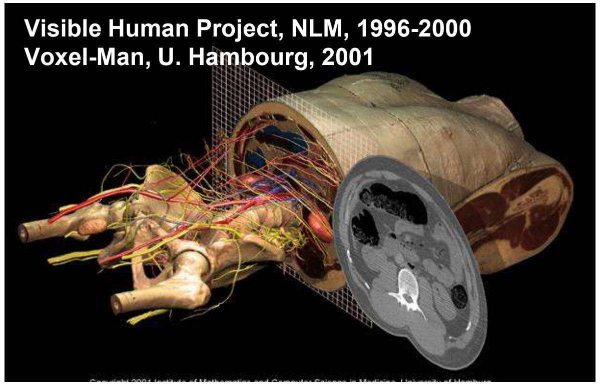


1er cerebral atlas, Vesale, 1543

Paré, 1585



Talairach & Tournoux, 1988



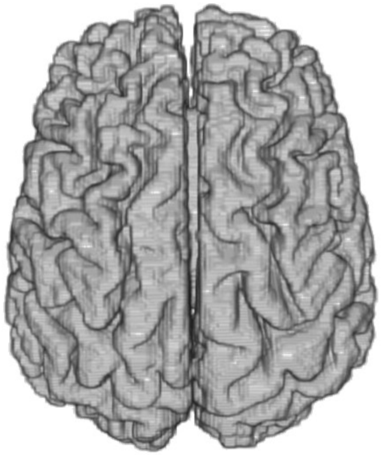
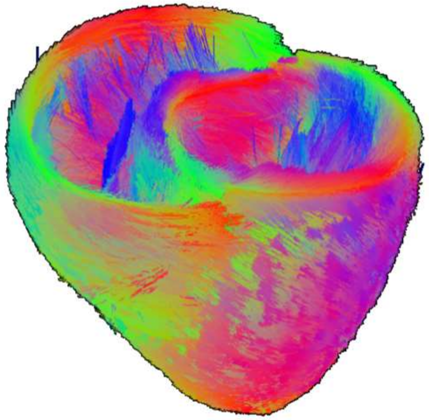
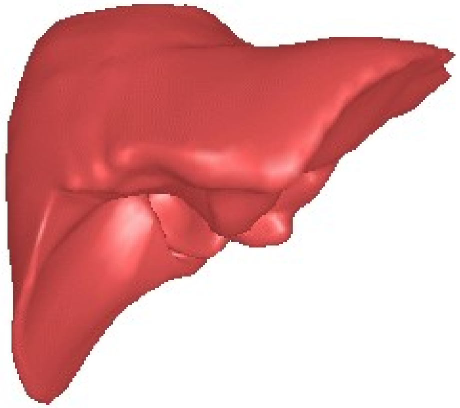
Visible Human Project, NLM, 1996-2000
Voxel-Man, U. Hambourg, 2001

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 (1980-1990):

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

From anatomy… to Computational Anatomy

Methods to compute 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…)

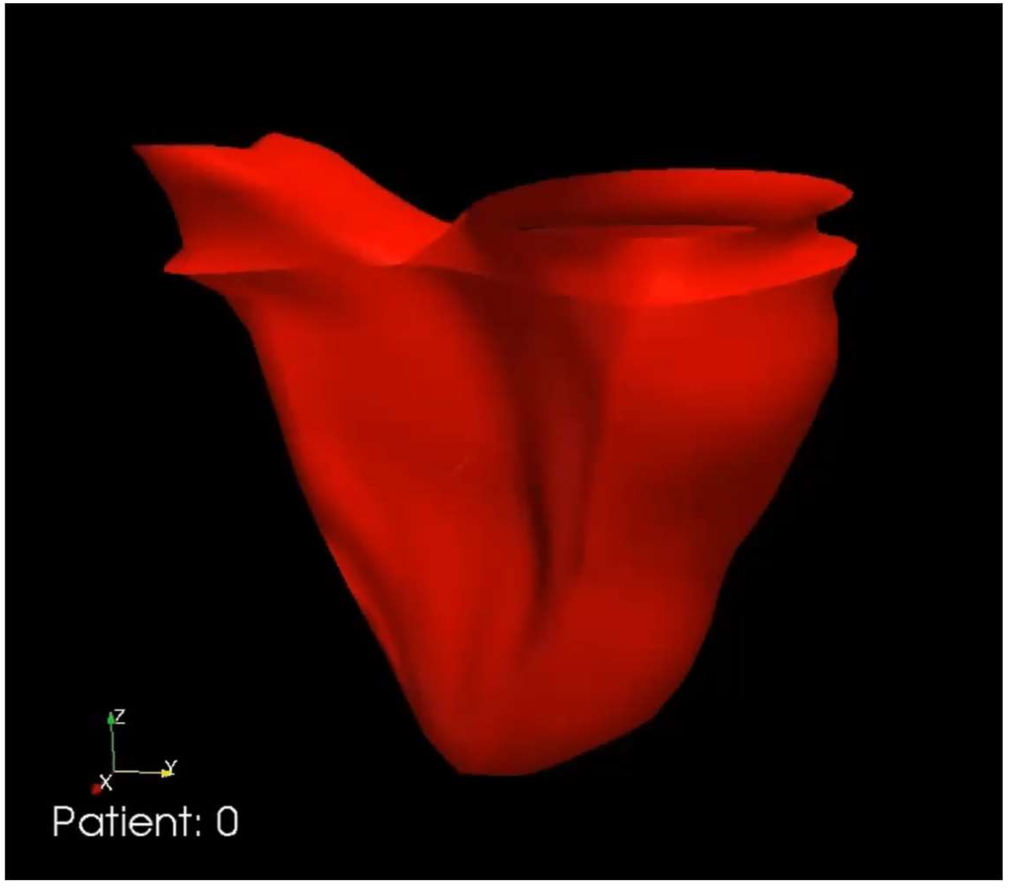
Use for personalized medicine (diagnostic, follow-up, etc)

- Classical use: atlas-based segmentation

Methods of computational anatomy

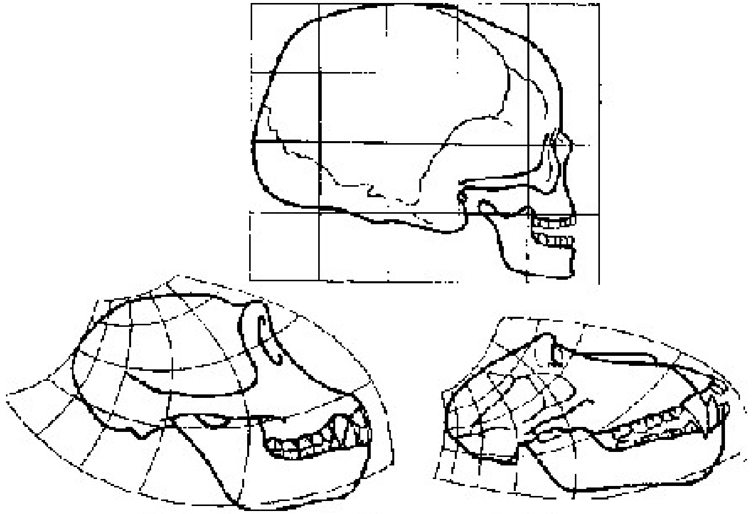
Remodeling of the right ventricle of the heart in tetralogy of Fallot

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

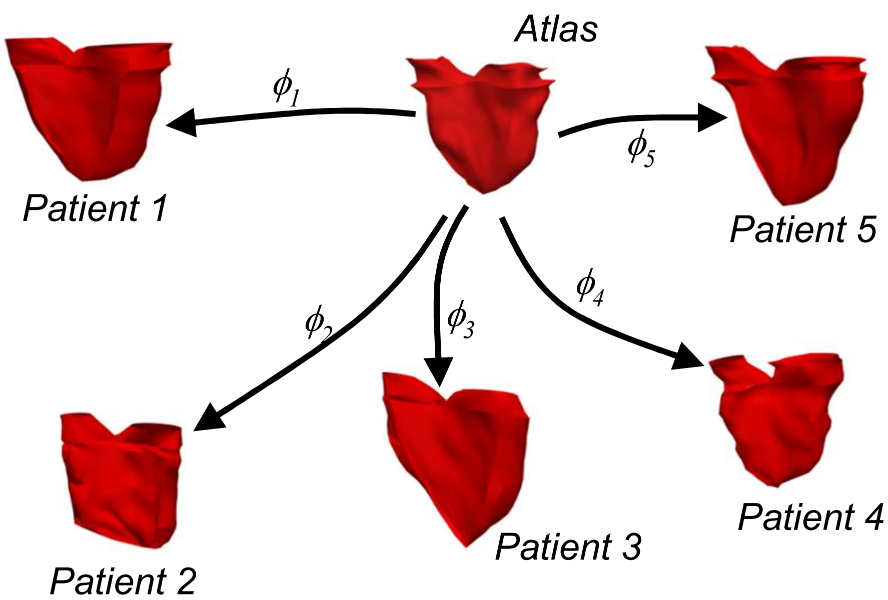


Shape of RV in 18 patients

Diffeomorphometry: Morphometry through Deformations



Skulls of a human, a chimpanzee and a baboon
and transformations between them



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

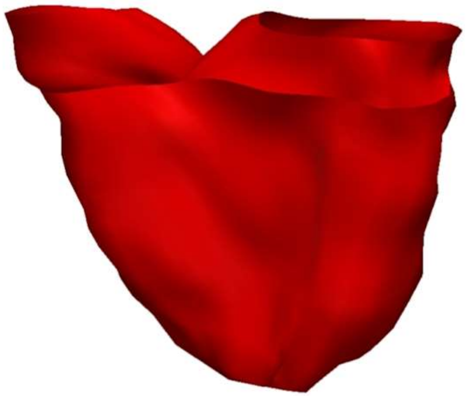
- Observation = "random" deformation of a reference template
- Reference template = Mean (atlas)
- Shape variability encoded by the deformations

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

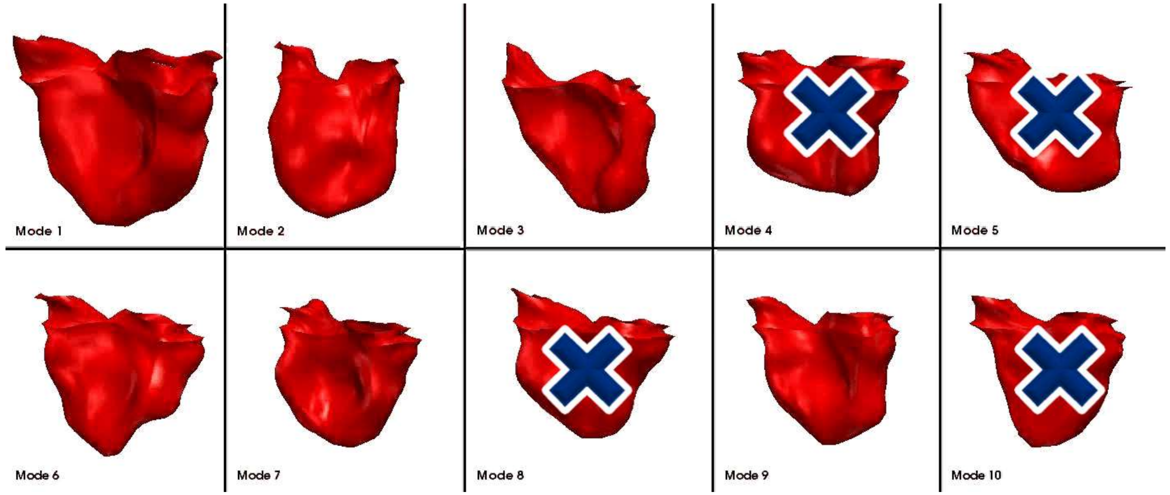
Atlas and Deformations Joint Estimation

Method: LDDMM to compute atlas + PLS on momentum maps

- Find modes that are significantly correlated to clinical variables (body surface area, tricuspid and pulmonary valve regurgitations).
- Create a generative model by regressing shape vs age (BSA)



Average RV anatomy of 18 ToF patients



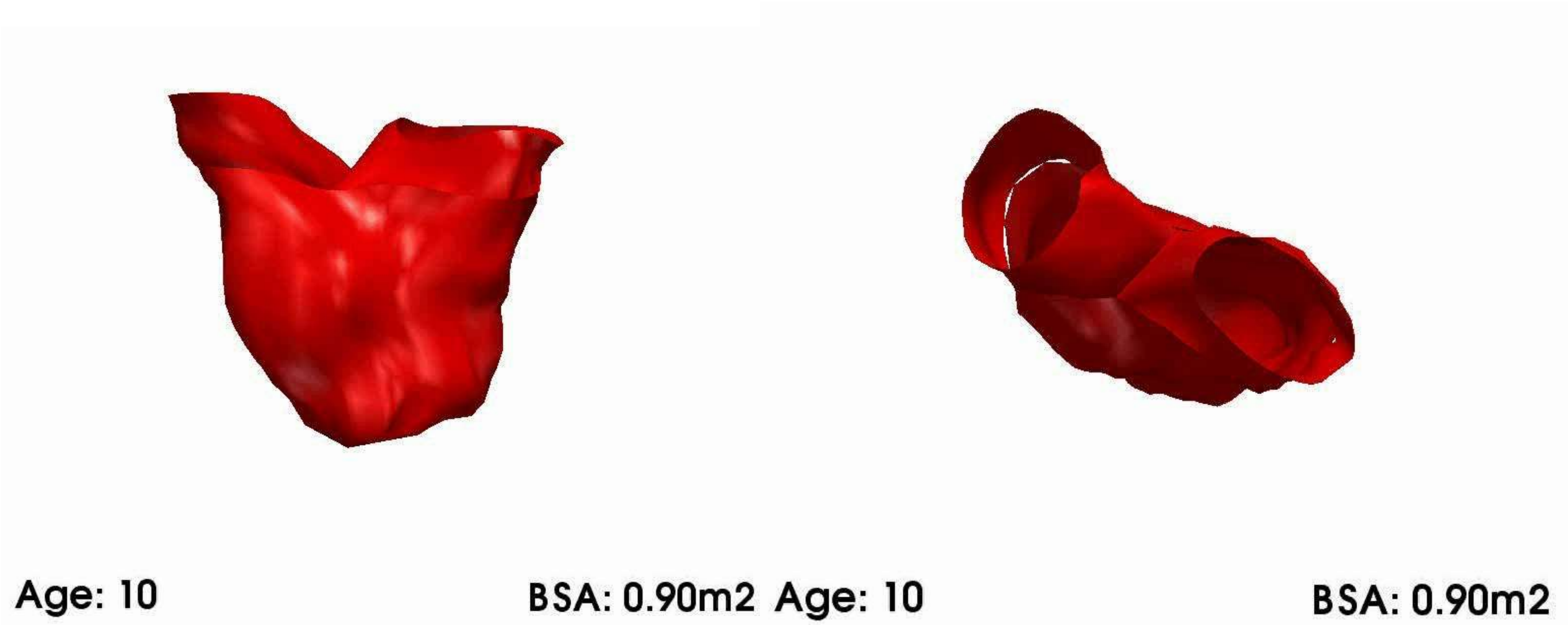
10 Deformation Modes = 90% of spectral energy

6 modes significantly correlated to BSA

[Mansi et al, MICCAI 2009, TMI 2011]

Statistical Remodeling of RV in Tetralogy of Fallot

[Mansi et al, MICCAI 2009, TMI 2011]

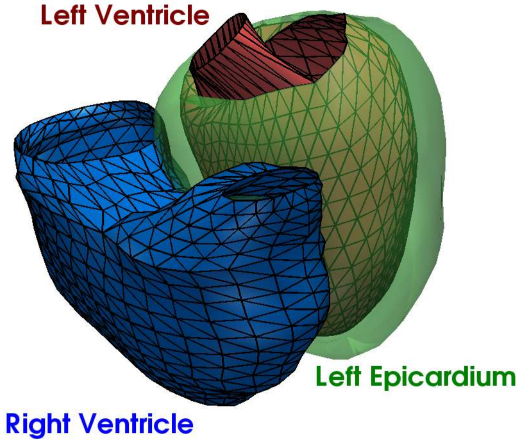


Predicted remodeling effect … has a clinical interpretation

Volume increases → Valve annuli deform → Pulmonary stenosis reduces → RV pressure decreases → Septum pushed inwards → RV free-wall outwards

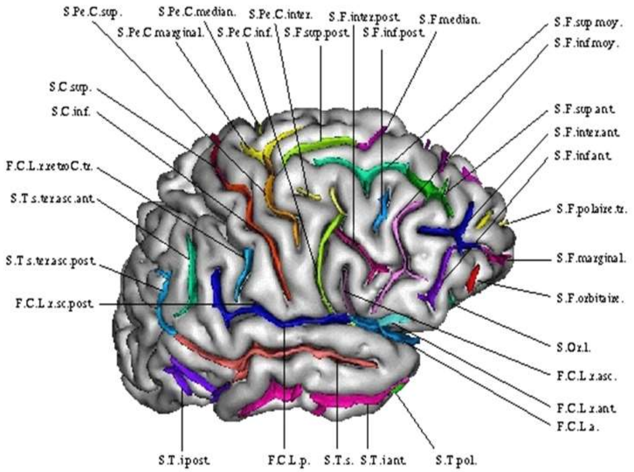
[Mansi et al, MICCAI 2009, TMI 2011]

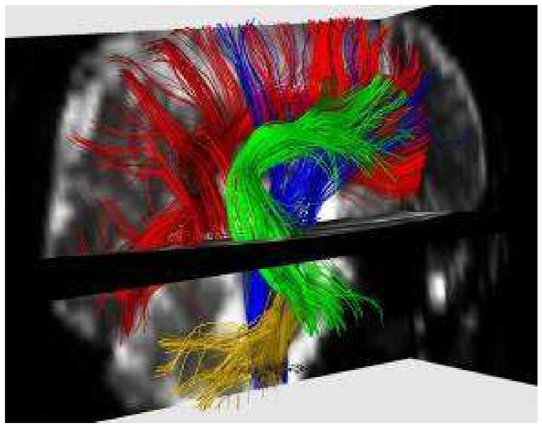
Geometric features in Computational Anatomy

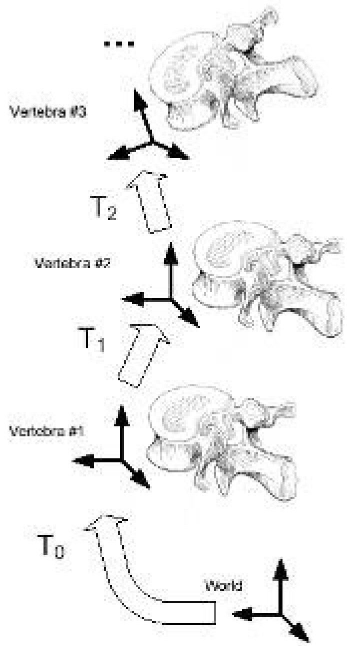


Non-Euclidean geometric features

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







Modeling statistical variability at the group level

- **Simple Statistics on non-linear manifolds?**
  - Mean, covariance of its estimation, PCA, PLS, ICA

Motivations

Simple statistics on Riemannian manifolds

- Bases for computing
- Extending statistics
- Manifold-valued image processing

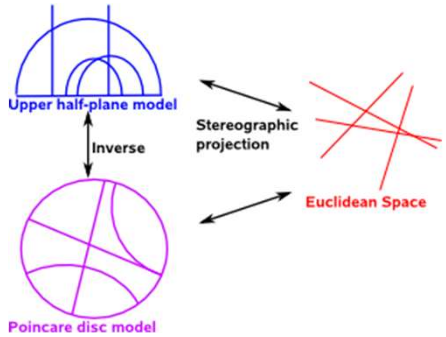
Extension to transformation groups with affine spaces

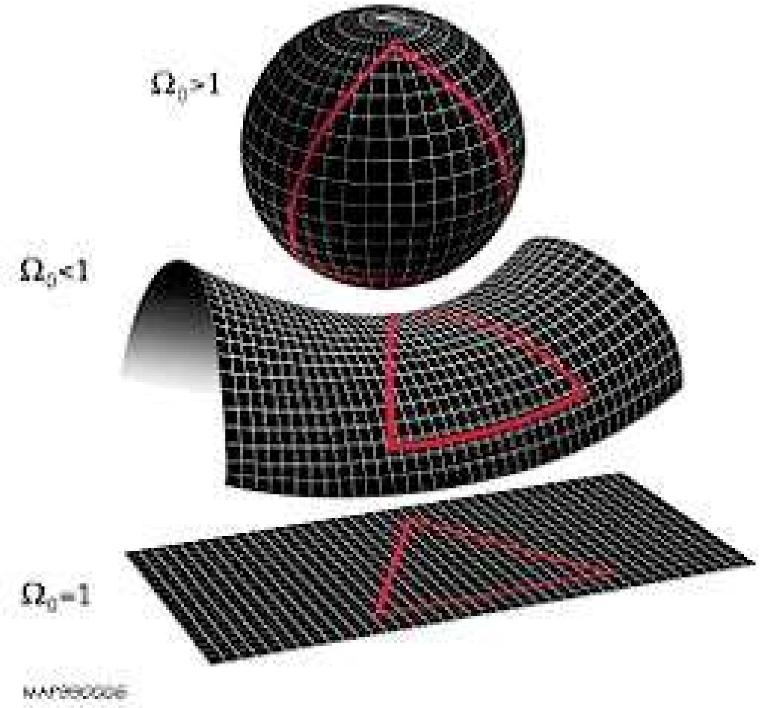
Perspectives, open problems

Which non-linear space?

Constant curvatures spaces

- Sphere,
- Euclidean,
- Hyperbolic





Homogeneous spaces, Lie groups and symmetric spaces

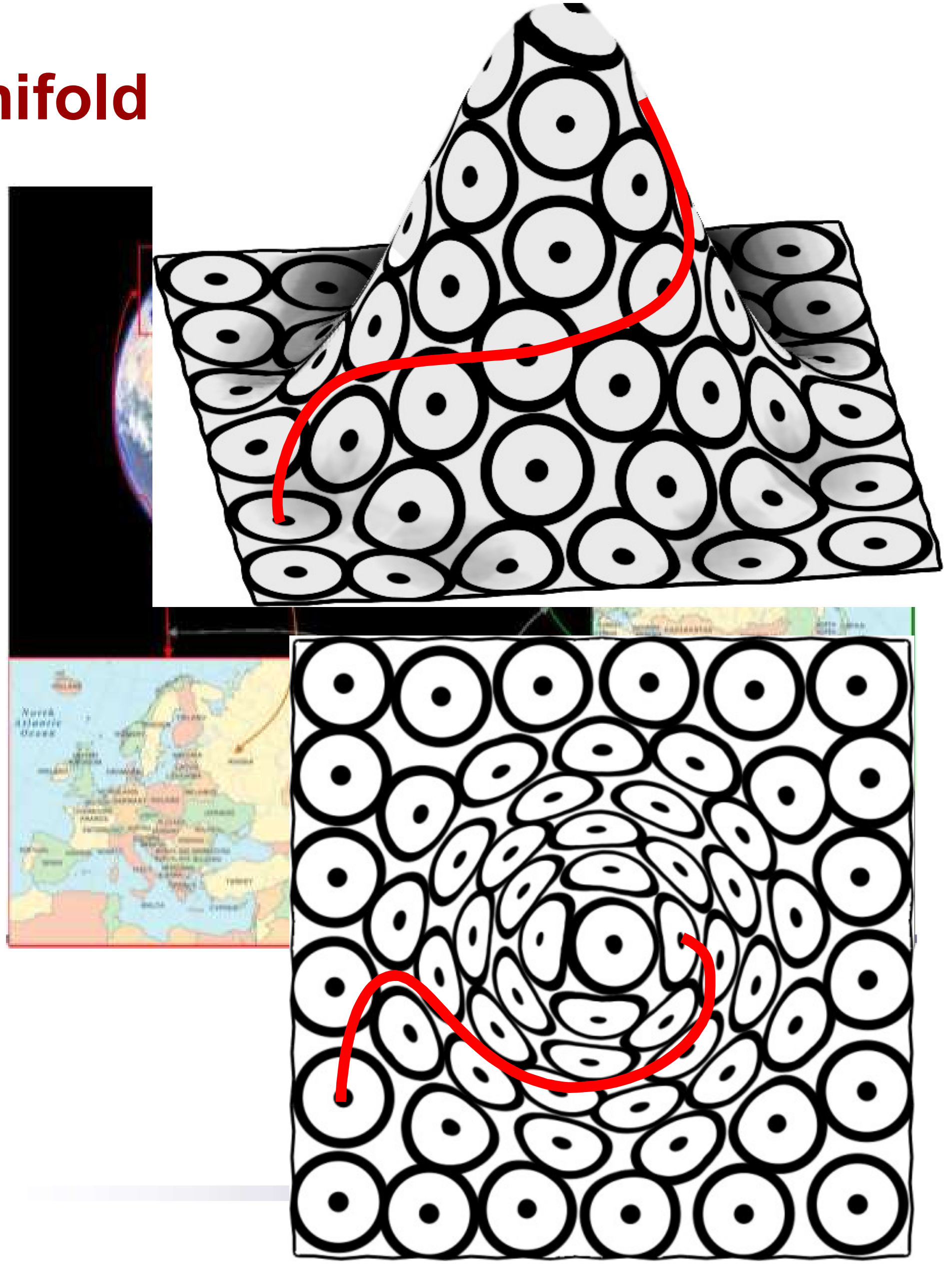
Riemannian and affine connection spaces

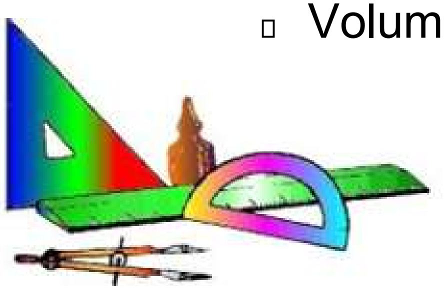
Towards non-smooth quotient and stratified spaces

Differentiable manifolds

Computing on a smooth manifold

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

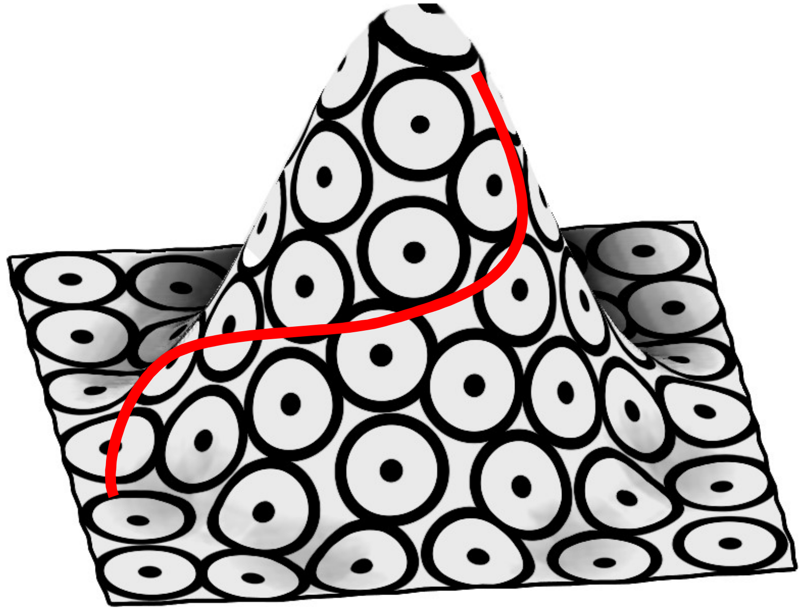




Measuring length

Basic tool: the scalar product

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

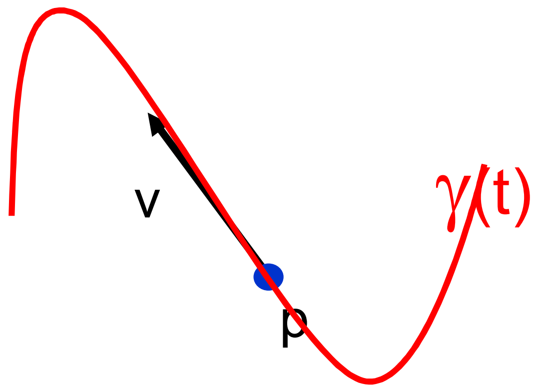


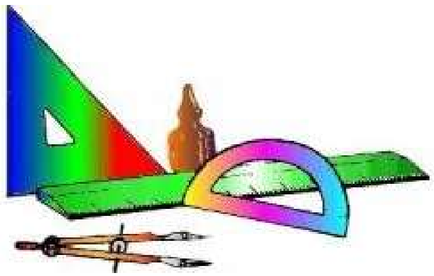
- Norm of a vector

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

- Length of a curve

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





Measuring length

Basic tool: the scalar product

Bernhard Riemann
1826-1866

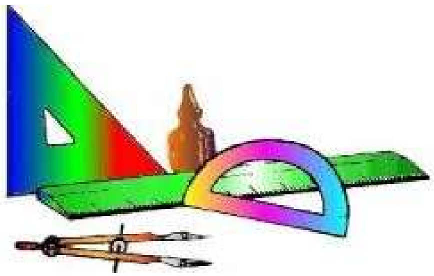
$$< v, w >_p = v^t G(p) w$$

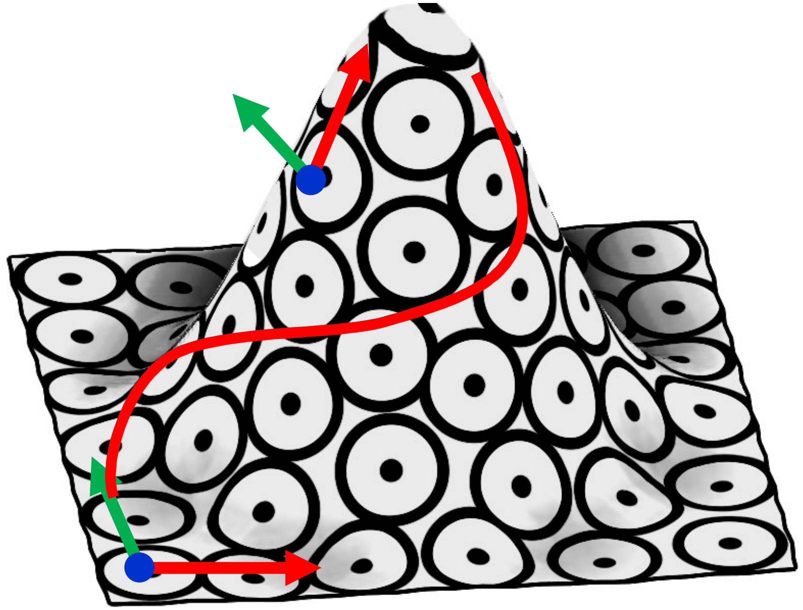
- Norm of a vector

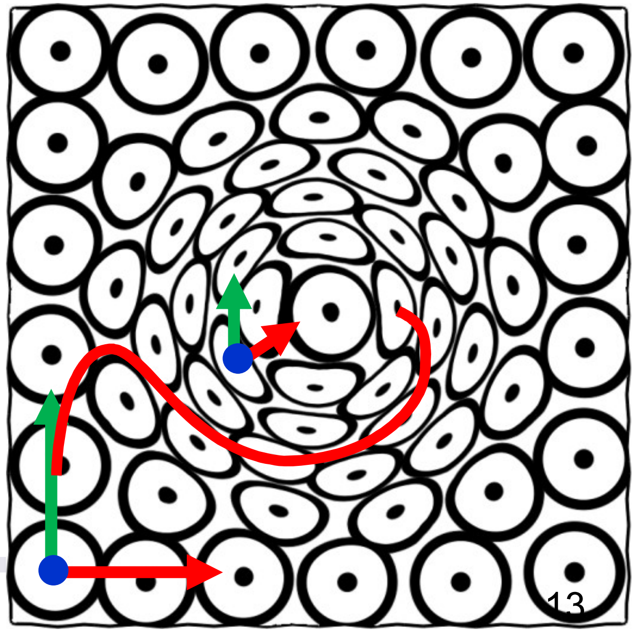
$$\|v\|_p = \sqrt{< v, v >_p}$$

- Length of a curve

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





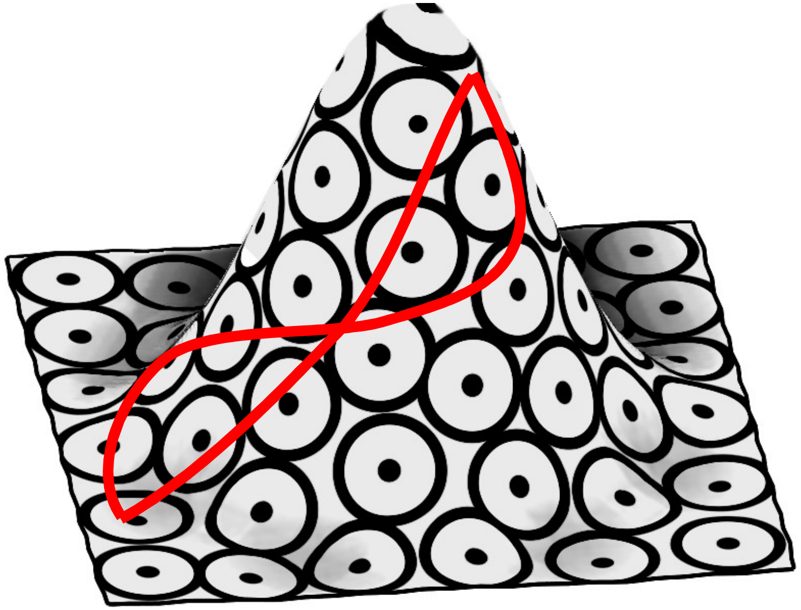


Riemannian manifolds

Basic tool: the scalar product

Bernhard Riemann
1826-1866

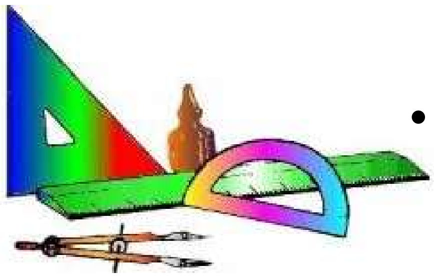
$$< v, w >_p = v^t G(p) w$$

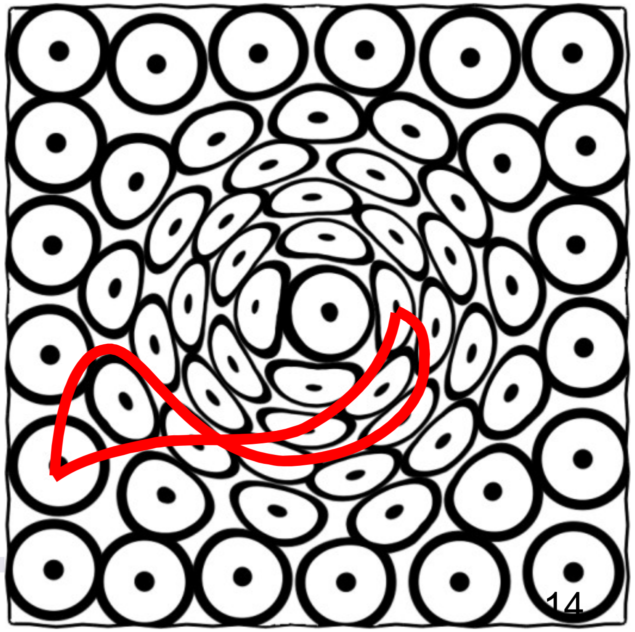


- Length of a curve

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

- 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





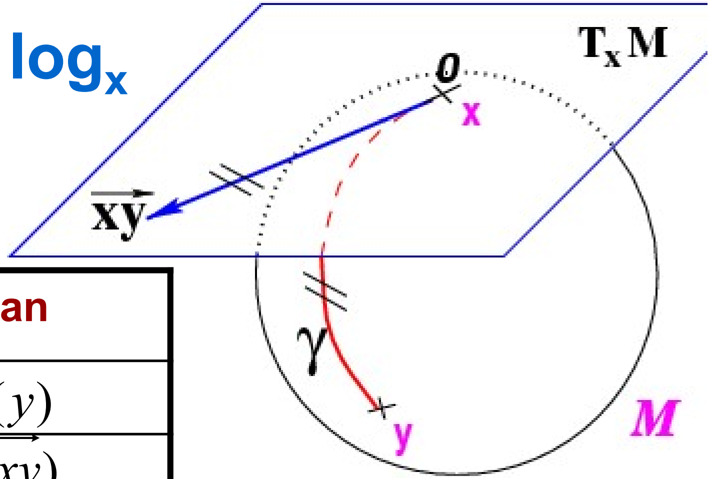
Bases of Algorithms in Riemannian Manifolds

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	$x_{t+\varepsilon} = x_t - \varepsilon \nabla C(x_t)$	$x_{t+\varepsilon} = Exp_{x_t}(-\varepsilon \nabla C(x_t))$

Geometric Statistics for Computational Anatomy

Motivations

Simple statistics on Riemannian manifolds

- Bases for computing
- Extending statistics
- Manifold-valued image processing

Extension to transformation groups with affine spaces

Perspectives, open problems

First statistical tools

Maurice Fréchet (1878-1973)

Fréchet mean set

- Integral only valid in Hilbert/Wiener spaces [Fréchet 44]
- $\sigma^2(x) = Tr_g\left(\mathfrak{M}_2(x)\right) = \int_M dist^2(x,z)\, P(dz)$
- **Fréchet mean** [1948] = global minima of Mean Sq. Dev.
- **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]

Moments of a random variable: tensor fields

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

Estimation of Fréchet 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 $r < r^* = \frac{1}{2}\min(inj(M), \pi/\sqrt{\kappa})$

Law of large numbers and CLT in manifolds

- Under Kendall-Karcher concentration conditions: EFM is a consistent estimator

$$\sqrt{n}\, log_{\bar{x}}(\bar{x}_n) \;\rightarrow\; N(0, \bar{H}^{-1}\, \Sigma\, \bar{H}^{-1}) \quad \textit{if } \bar{H} = Hess(\sigma^2(X, \bar{x}_n)) \textit{ invertible}$$

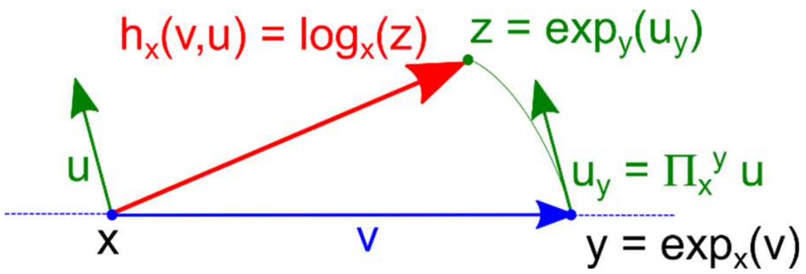
[Bhattacharya & Patrangenaru 2005, Bhatt. & Bhatt. 2008; Kendall & Le 2011]

- **Expression for Hessian? interpretation of covariance modulation?**
- **What happens for a small sample size?**

Taylor expansion in affine connection manifolds

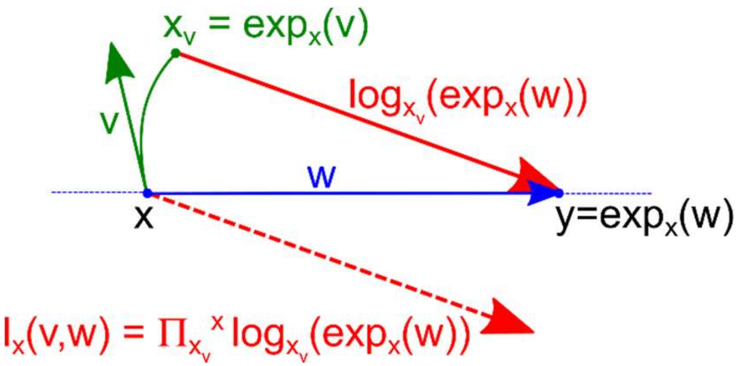
Key idea: use parallel transport instead of the chart to relate $T_x M$ to $T_{x_v} M$

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



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

Neighboring log expansion (new, XP 2019) [XP, arXiv:1906.07418]



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

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**)

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

  - No bias in symmetric spaces (covariantly constant curvature)

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

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

  - Negative curvature: faster CV than Euclidean, prelude to stickiness
  - Positive curvature: slower CV than Euclidean, prelude to smeariness

LLN / CLT in manifolds [Bhattacharya & Bhattacharya 2008; Kendall & Le 2011]

- Under Kendall-Karcher concentration conditions:

$$\sqrt{n}\, log_{\bar{x}}(\bar{x}_n) \xrightarrow{D} N(0, H^{-1}\,\Sigma\, H^{-1}) \quad \textit{if}\;\; H = Hess\big(\,\sigma^2(X, \bar{x}_n)\big) \;\textit{invertible}$$

- Hessian: $\frac{1}{2}\bar{H} = Id + \frac{1}{3}R : \mathfrak{M}_2 + \frac{1}{12}\nabla \mathrm{R} : \mathfrak{M}_3 + O(\epsilon^4, 1/n^2)$
- Same expansion for large n: *modulation of the CV rate by curvature*

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

Beyond the mean: principal components?

Maximize the explained variance

- Tangent PCA (tPCA): eigenvectors of covariance in $T_{\bar{x}}M$ generate a geodesic subspace $GS(\bar{x}, v_1, v_2, \dots v_k)$

Minimize the sum of squared residuals to a subspace

- PGA, GPCA: Geodesic subspace $GS(\bar{x}, v_1, v_2, \dots v_k)$
  [Fletcher et al., 2004, Sommer et al 2014, Huckeman et al., 2010]

- BSA: **Affine span** $\mathrm{Aff}(x_0, x_1, x_2, \dots x_k)$
  Locus of weighted exponential barycenters (geodesic simplex for positive weights)

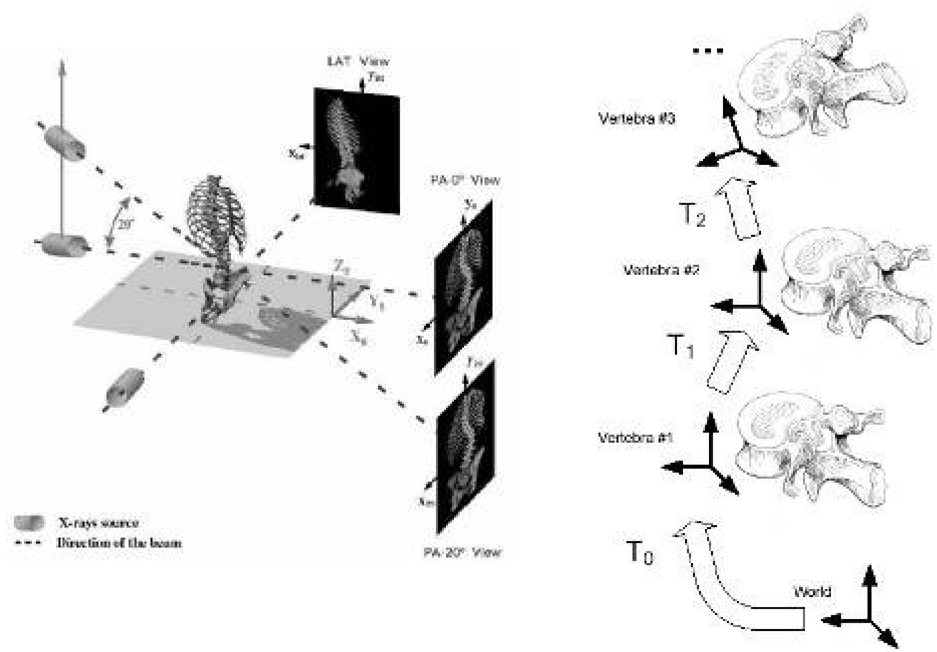
Sequence of properly embedded subspaces (flags)

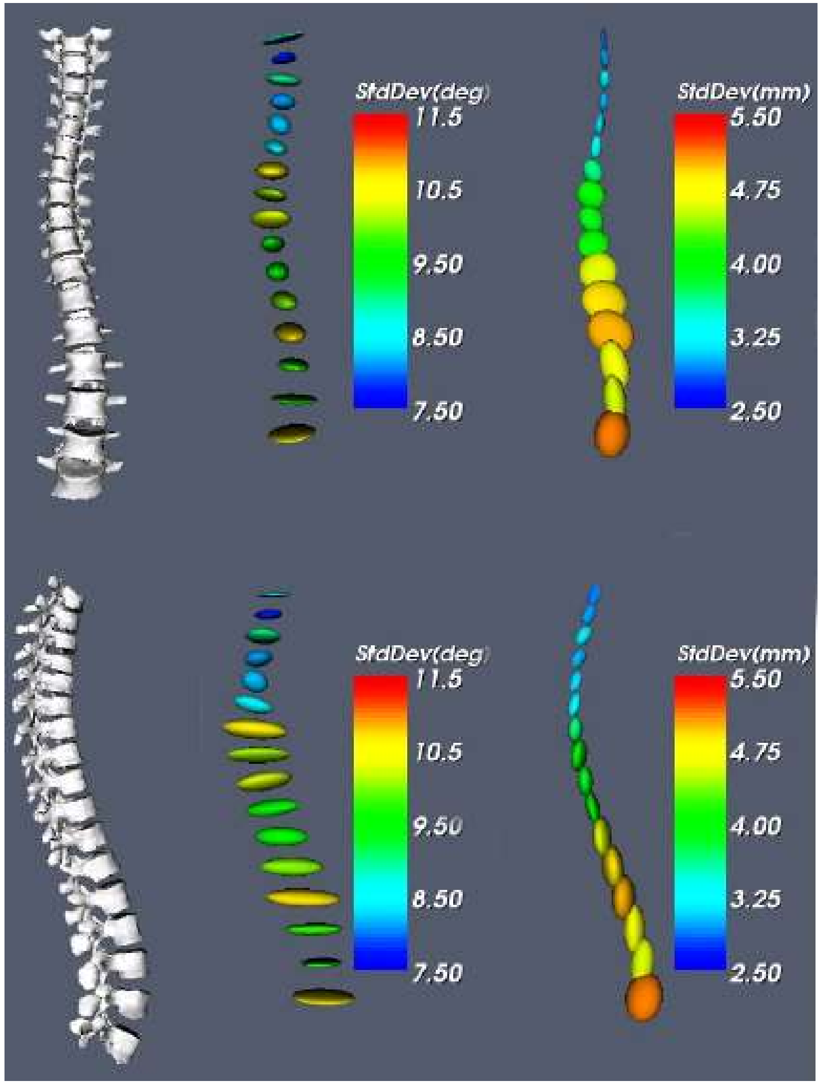
- AUC criterion on **flags** generalizes PCA [XP, AoS 2018]

[XP, Barycentric subspace analysis on Manifolds, Annals of Statistics, 2018]

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

Left invariant Mean on $(SO_3 \ltimes R^3)^{16}$

- 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)

Geometric Statistics for Computational Anatomy

Motivations

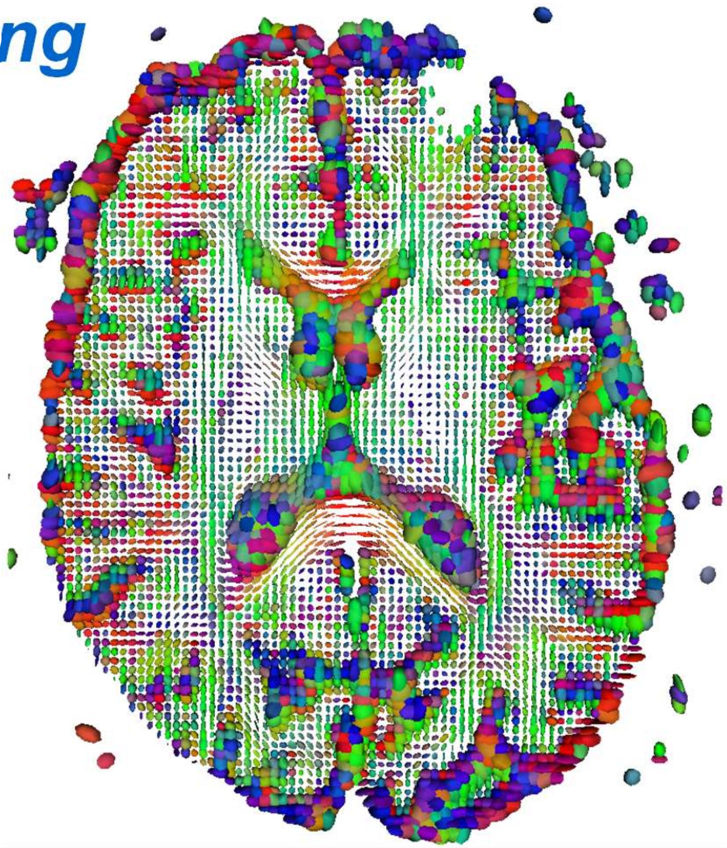
Simple statistics on Riemannian manifolds

- Bases for computing
- Extending statistics
- Manifold-valued image processing

Extension to transformation groups with affine spaces

Perspectives, open problems

Manifold-valued image processing: Diffusion Tensor Imaging



Covariance of the Brownian motion of water

- Filtering, regularization
- Interpolation / extrapolation
- Architecture of axonal fibers

Symmetric positive definite matrices

- Cone in Euclidean space (not complete)
- Convex operations are stable
  - mean, interpolation
- More complex operations are not
  - PDEs, gradient descent…

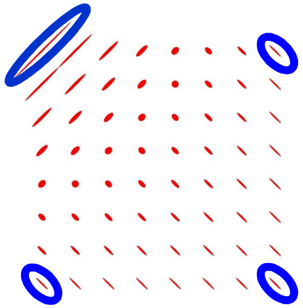
All invariant metrics under GL(n)

$$\langle W_1 \mid W_2 \rangle_{Id} = \mathrm{Tr}\left(W_1^T W_2\right) + \beta\, \mathrm{Tr}(W_1).\mathrm{Tr}(W_2) \quad (\beta > \text{-}1/n)$$

- Exponential map $\quad Exp_\Sigma(\overrightarrow{\Sigma\Psi}) = \Sigma^{1/2} \exp(\Sigma^{-1/2}.\overrightarrow{\Sigma\Psi}.\Sigma^{-1/2})\Sigma^{1/2}$
- Log map $\quad \overrightarrow{\Sigma\Psi} = Log_\Sigma(\Psi) = \Sigma^{1/2} \log(\Sigma^{-1/2}.\Psi.\Sigma^{-1/2})\Sigma^{1/2}$
- Distance $\quad dist(\Sigma,\Psi)^2 = \left\langle \overrightarrow{\Sigma\Psi} \mid \overrightarrow{\Sigma\Psi} \right\rangle_\Sigma = \left\| \log(\Sigma^{-1/2}.\Psi.\Sigma^{-1/2}) \right\|_{Id}^2$

Manifold-valued image algorithms

Integral or sum in M: weighted Fréchet mean



- Interpolation
  - Linear between 2 elements: interpolation geodesic
  - Bi- or tri-linear or spline in images: weighted means
- Gaussian filtering: **convolution = weighted mean**

  [Pennec, Fillard, Arsigny, IJCV 66(1), 2006]

$$\Sigma(x) = \min \sum_i G_\sigma(x - x_i)\ \mathrm{dist}^2(\Sigma, \Sigma_i)$$

PDEs for regularization and extrapolation: the exponential map (partially) accounts for curvature

- Gradient of Harmonic energy = Laplace-Beltrami

$$\Delta\Sigma(x) = \frac{1}{\varepsilon} \sum_{u \in S} \overrightarrow{\Sigma(x)\Sigma(x+\varepsilon u)} + O\left(\varepsilon^2\right)$$

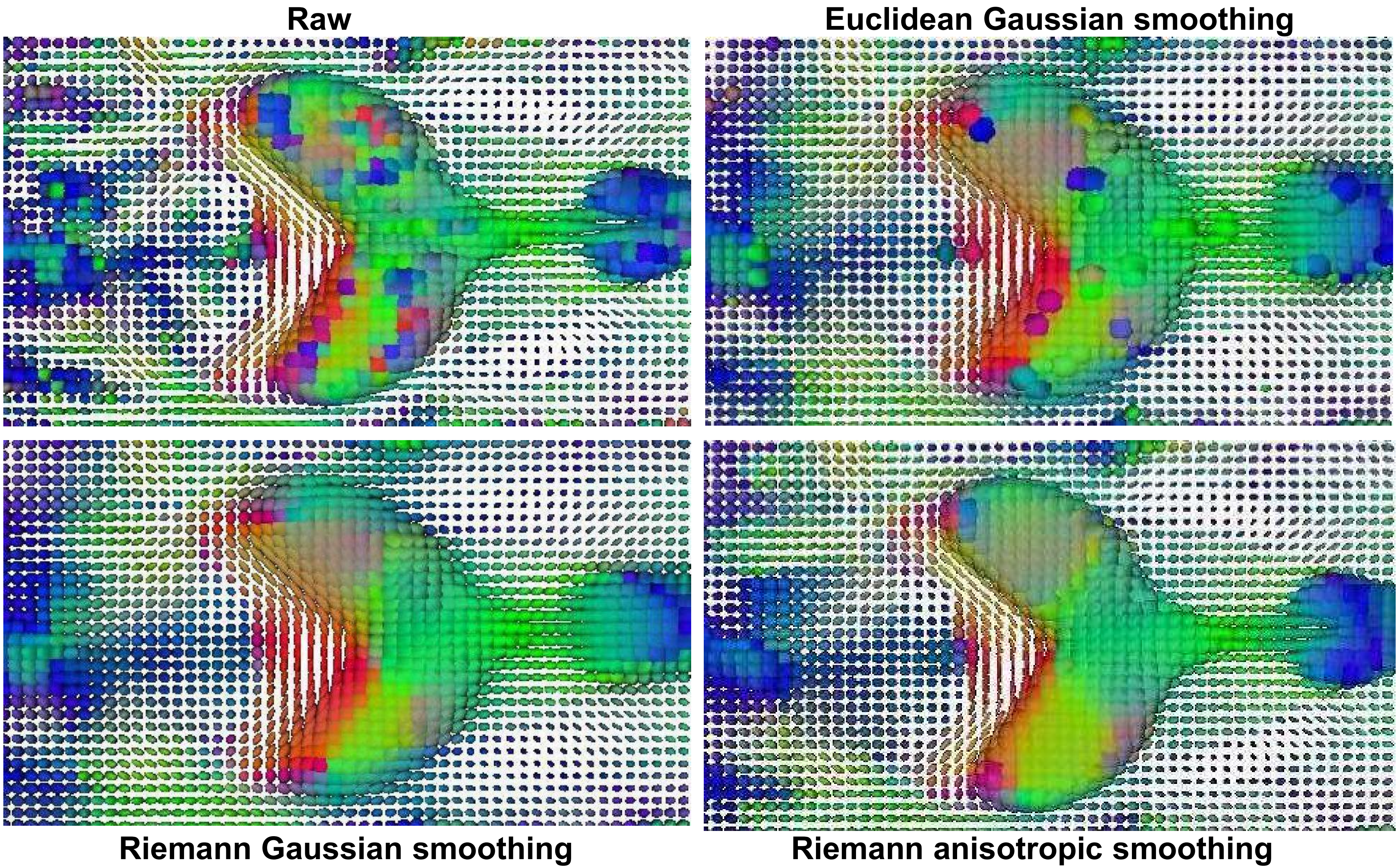
- Anisotropic regularization using robust functions

$$\mathrm{Reg}(\Sigma) = \int \Phi\left(\left\|\nabla\Sigma(x)\right\|^2_{\Sigma(x)}\right) dx$$

- Simple intrinsic numerical schemes thanks the exponential maps!

[Pennec, Fillard, Arsigny, IJCV 66(1), 2005, ISBI 2006]

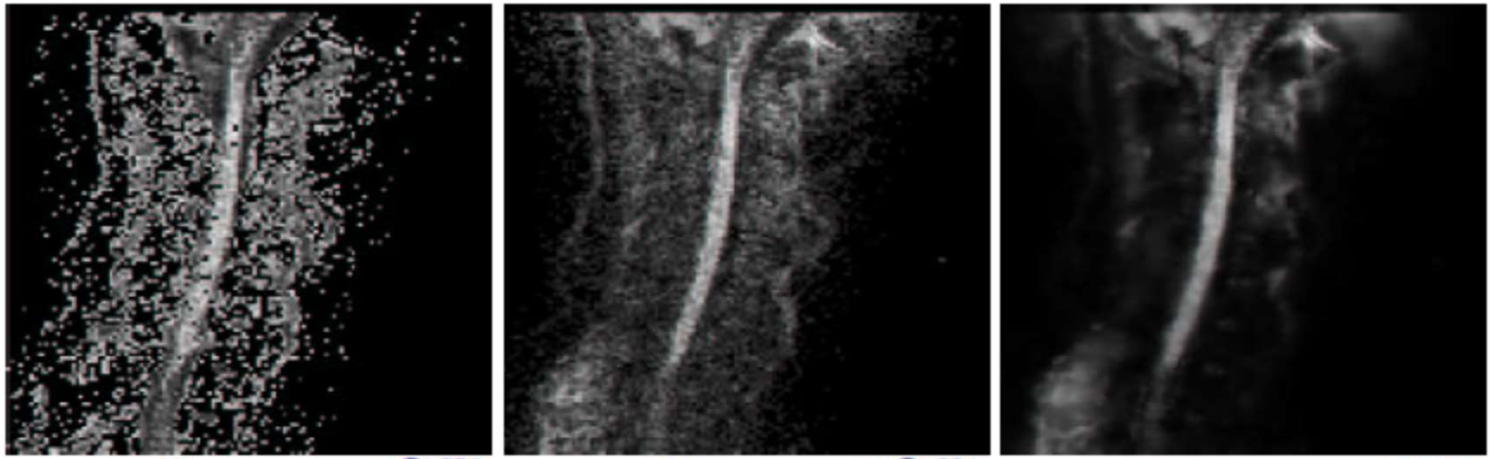
Filtering and anisotropic regularization of DTI



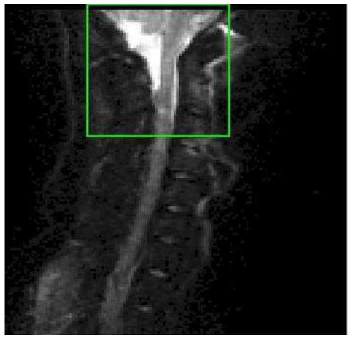
Rician MAP estimation with Riemannian spatial prior

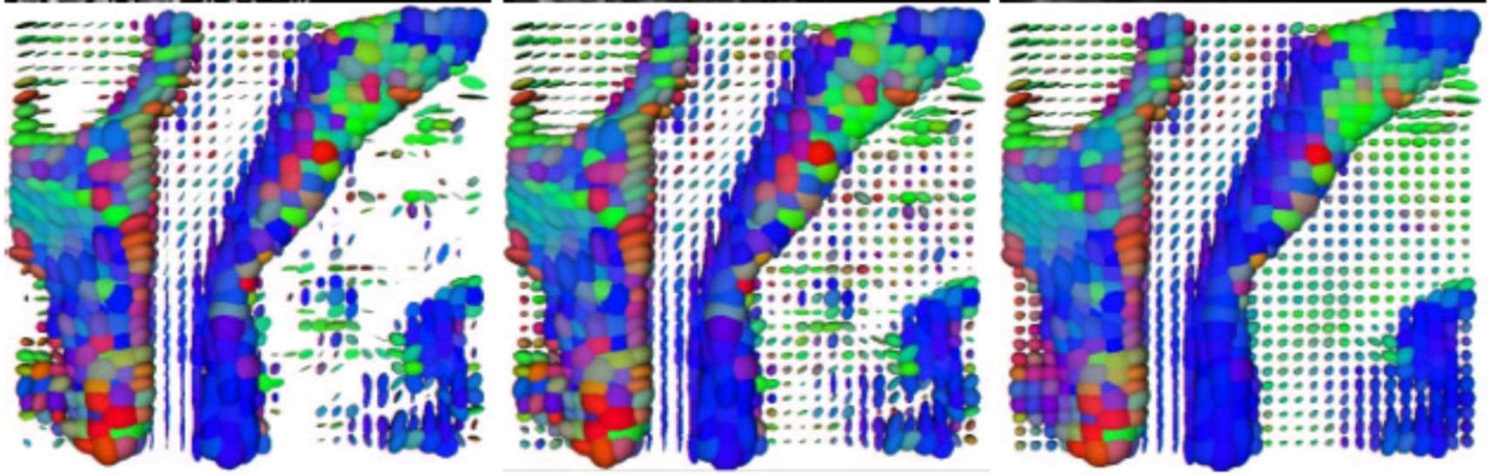
$$MAP(\Sigma) = -\sum_{i=1}^{N} \int \log\left(\frac{\hat{S}_i}{\sigma^2} \exp\left(-\frac{\hat{S}_i^{\,2} + S_i(\Sigma)^2}{2\sigma^2} \right) I_0\left(\frac{S_i(\Sigma)\hat{S}_i}{\sigma^2} \right) \right).dx + \int \Phi\left(\left\| \nabla\Sigma(x) \right\|^2_{\Sigma(x)} \right).dx$$

FA



Estimated tensors





Standard ML Rician MAP Rician

[Fillard, Arsigny, Pennec, Ayache ISBI'06, TMI 26(11) 2007]

Geometric Statistics for Computational Anatomy

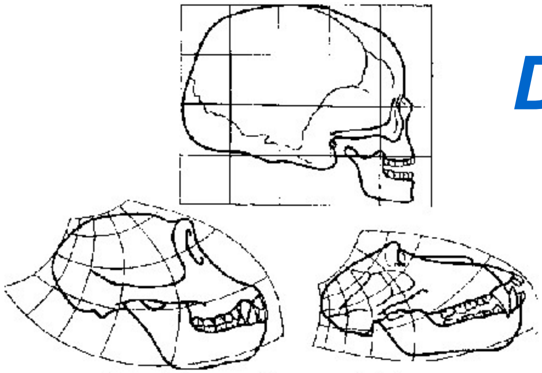
Motivations

Simple statistics on Riemannian manifolds

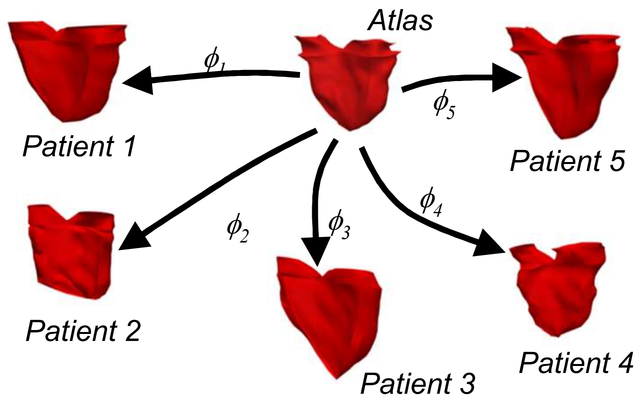
Extension to transformation groups with affine spaces

- The bi-invariant affine Cartan connection structure
- Longitudinal modeling with parallel transport on diffeomorphisms

Perspectives, open problems


Skulls of a human, a chimpanzee and a baboon and transformations between them

Diffeomorphometry



Lie group: Smooth manifold with group structure

- Composition $g \circ h$ and inversion g^{-1} are smooth
- Left and Right translation $L_g(f) = g \circ f$ $R_g(f) = f \circ g$
- Natural Riemannian metric choices : left or right-invariant metrics

Lift statistics to transformation groups

- [D'Arcy Thompson 1917, Grenander & Miller]
- LDDMM = right invariant kernel metric (Trouvé, Younes, Joshi, etc.)

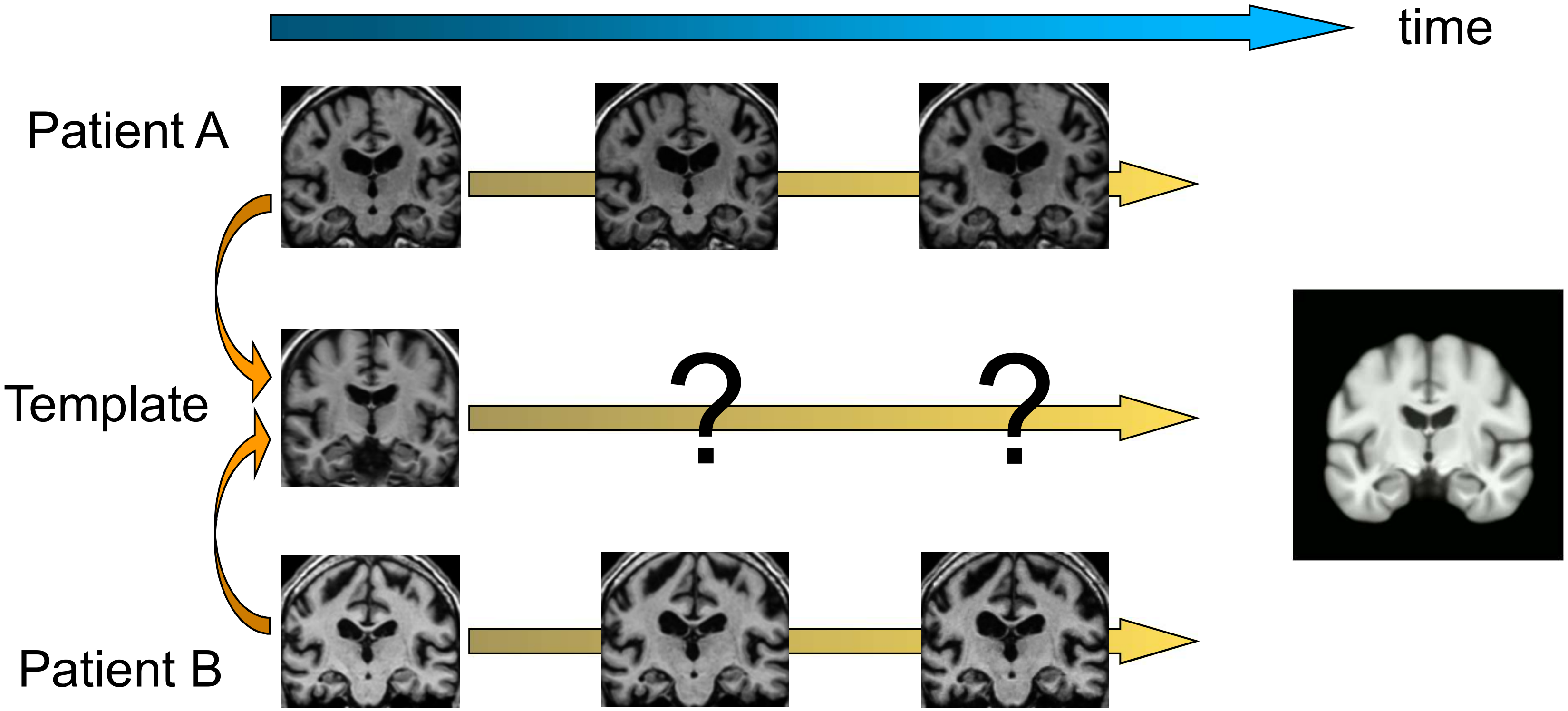
No bi-invariant metric in general for Lie groups

- **Incompatibility of the Fréchet mean with the group structure**
- Examples with simple 2D rigid transformations

Is there a more natural structure for statistics on Lie groups?

Longitudinal deformation analysis

Dynamic obervations



How to transport longitudinal deformation across subjects?

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

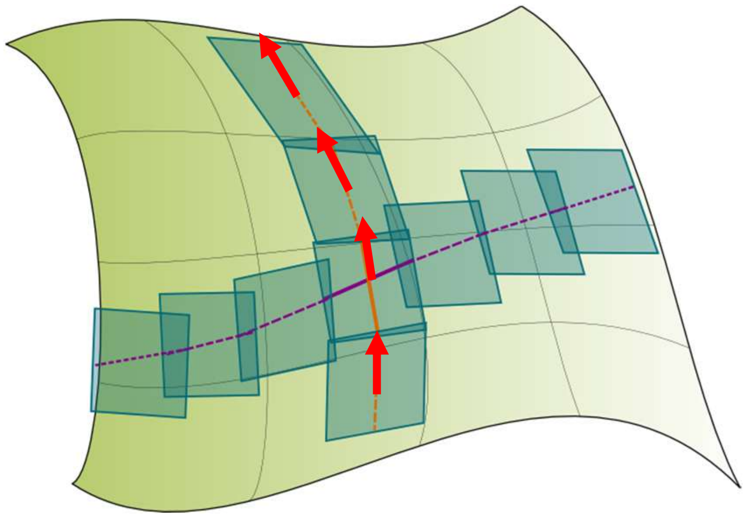
- Left / Right-invariant geodesics, one-parameter subgroups

Question: Can one-parameter subgroups be geodesics?

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



Adapted from Lê Nguyên Hoang, science4all.org

Geodesics = straight lines

- Null acceleration: $\nabla_{\dot{\gamma}}\dot{\gamma} = 0$
- 2nd order differential equation: Normal coordinate system
- Local exp and log maps

[XP & Arsigny, 2012, XP & Lorenzi, Beyond Riemannian Geometry, 2019]
[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: $S_A\big(M(t)\big) = A\exp(-tV)A^{-1} = M(-t)$

[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]

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

- Locus of points *x* such that $\sum Log\left(x^{-1}.y_i\right) = 0$
- 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**
  - If m is a mean of $\{g_i\}$ and h is any group element, then

    $h \circ m$ = mean$\{h \circ g_i\}$, $m \circ h$ = mean$\{g_i \circ h\}$ and $m^{(-1)}$ = mean$\left\{g_i^{(-1)}\right\}$

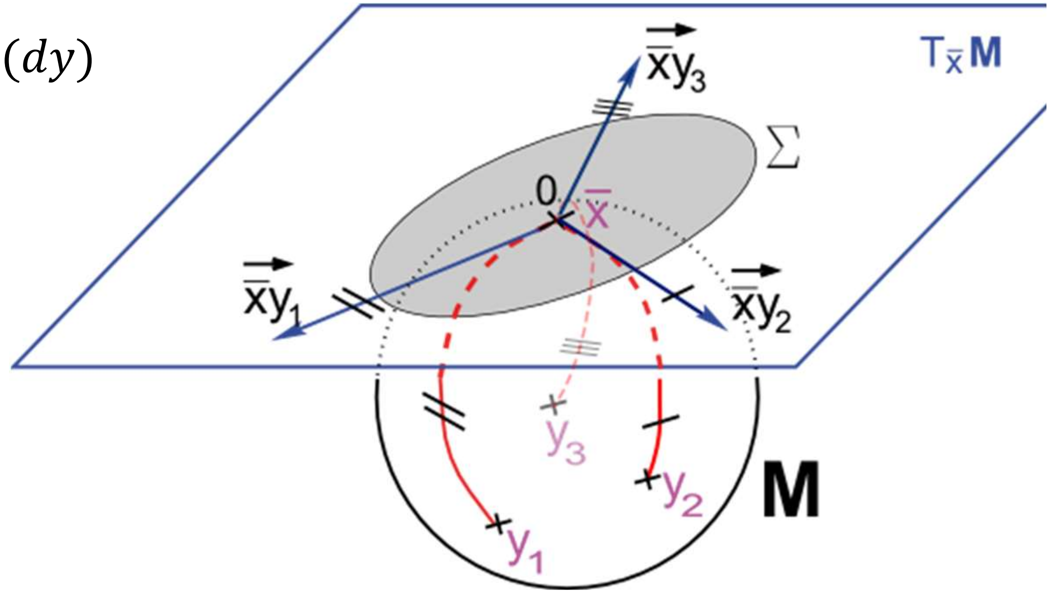
Generalization of the Statistical Framework

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 according to the basis



Other statistical tools

- Previous therorem on CLT holds
- Mahalanobis distance well defined and bi-invariant
- ~~Tangent Principal Component Analysis (t-PCA)~~
- PGA & BSA, provided a data likelihood

Geometric Statistics for Computational Anatomy

Motivations

Simple statistics on Riemannian manifolds

Extension to transformation groups with affine spaces

- The bi-invariant affine Cartan connection structure
- **Longitudinal modeling with parallel transport on diffeomorphisms**

Perspectives, open problems

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

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

- Lagrangian formalism
- Sobolev Norm H_k (or RKHS in LDDMM) → diffeomorphisms
  [Miller, Trouve, Younes, Holm, Dupuis, Beg… 1998 – 2009]
- Geometric Mechanics [Arnold, Smale, Souriau, Marsden, Ratiu, Holmes, Michor…]

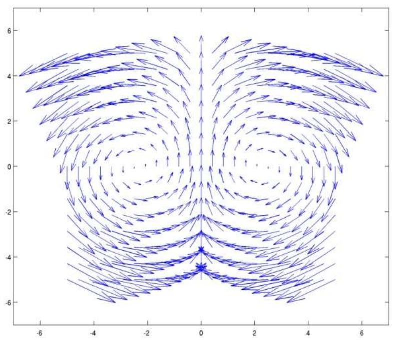
Geodesics determined by optimization of a time-varying vector field

- 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 by shooting/adjoint or path-straightening methods

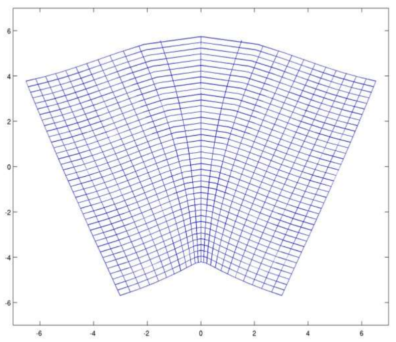
The SVF framework for Diffeomorphisms

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

- Exponential of a smooth vector field is a diffeomorphism
- Use ~~time-varying~~ Stationary Velocity Fields to parameterize deformation



Stationary velocity field



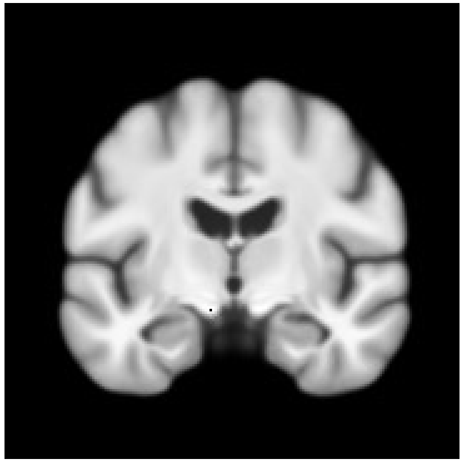
Diffeomorphism

Efficient numerical algorithms

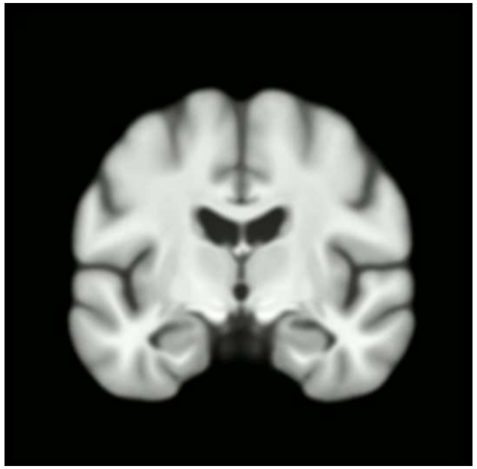
- Recursive **Scaling and squaring algorithm** **[Arsigny MICCAI 2006]**
  - Deformation: exp($\boldsymbol{v}$)=exp($\boldsymbol{v}$/2) o exp($\boldsymbol{v}$/2)
  - Jacobian: Dexp(v) = Dexp(v/2) o exp(v/2) . Dexp(v/2)
- Optimize deformation parameters: **BCH formula** **[Bossa MICCAI 2007]**
  - $\exp(\boldsymbol{v}) \circ \exp(\varepsilon \boldsymbol{u}) = \exp(\, \boldsymbol{v} + \varepsilon \boldsymbol{u} + [\boldsymbol{v},\varepsilon \boldsymbol{u}]/2 + [\boldsymbol{v},[\boldsymbol{v},\varepsilon \boldsymbol{u}]]/12 + \ldots\,)$ where $[\boldsymbol{v},\boldsymbol{u}](p) = \mathrm{Jac}(\boldsymbol{v})(p).\boldsymbol{u}(p) - \mathrm{Jac}(\boldsymbol{u})(p).\boldsymbol{v}(p)$

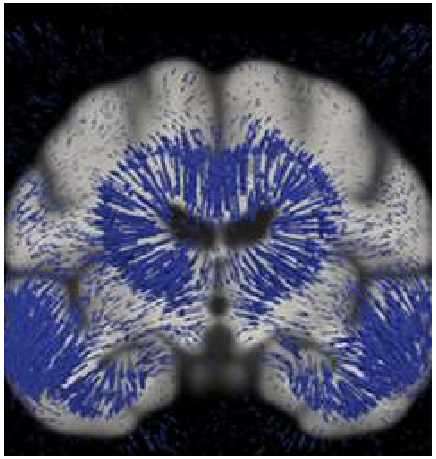
Measuring Temporal Evolution with deformations: Deformation-based morphometry

Fast registration with deformation parameterized by SVF



$$\varphi_t(x) = exp(t.v(x))$$

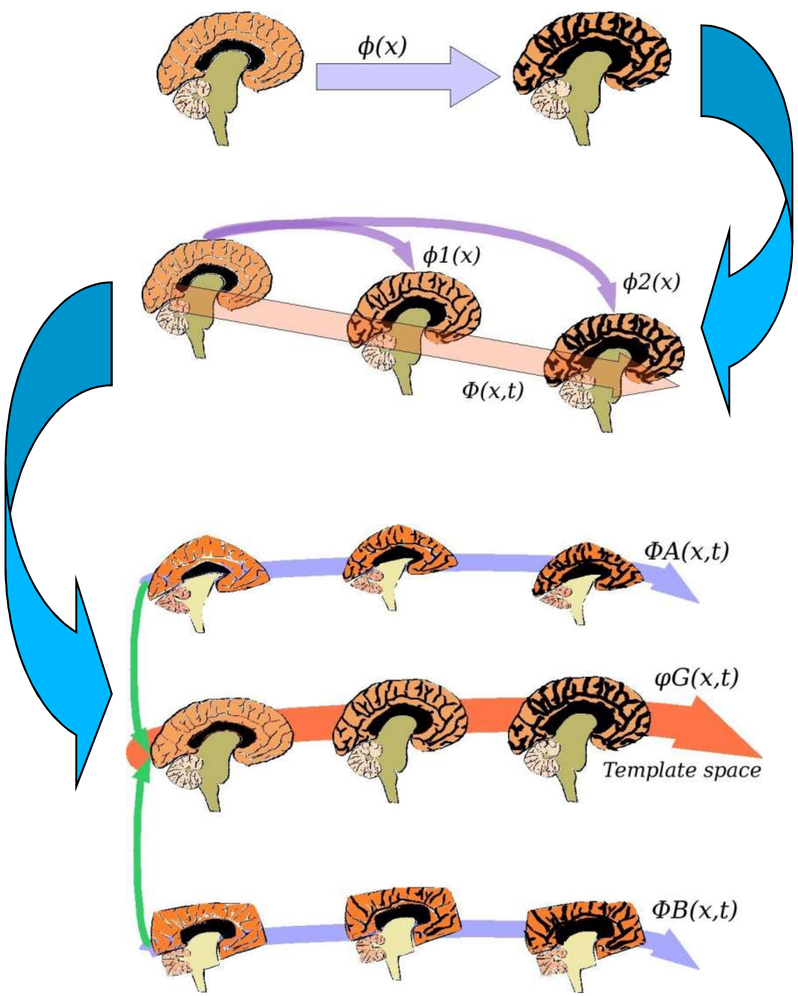




https://team.inria.fr/asclepios/software/lcclogdemons/

[LCC log-demons for longitudinal brain imaging.
Lorenzi, Ayache, Frisoni, Pennec, Neuroimage 81, 1 (2013) 470-483]

Analysis of longitudinal datasets



Single-subject, two time points

Log-Demons (LCC criteria)

Single-subject, multiple time points

4D registration of time series within the Log-Demons registration: geodesic regression

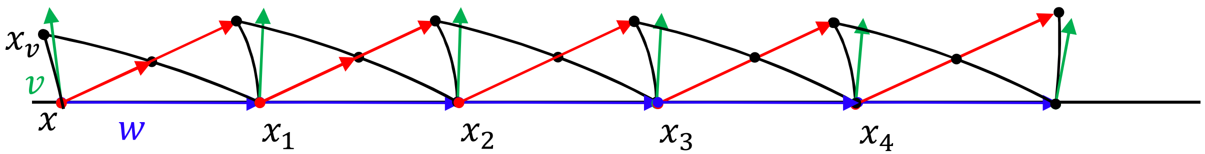
Multiple subjects, multiple time points

Population trend with parallel transport of SVF along inter-subject trajectories

[Lorenzi et al, IPMI 2011, JMIV 2013]

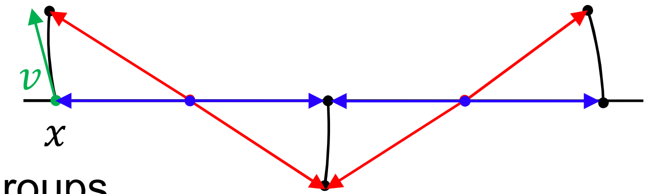
Discrete approximations of Parallel transport

Schild's Ladder [Lecture at Princeton 60ies, Elhers et al 1972]



- Build geodesic parallelogram
- Iterate along the curve
- One step is a 1st order approximation [Kheyfets et al 2000]

Pole ladder: [Lorenzi, XP, JMIV 50 (1-2), 2013]



- Simpler method with piecewise geodesics
  - Closed form expression for Cartan connection on Lie groups
- One step is of order 4 in general affine manifolds [XP, Arxiv 1805.11436, 2018]

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

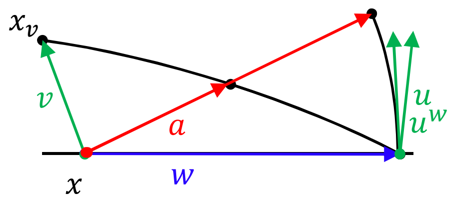
  - **Exact in symmetric spaces (transvection)!**

→ No approximation formula beyond 1st order for SL
→ No results for the iterated SL and PL schemes
→ No results for approximate geodesics

Convergence of Schild's Ladder

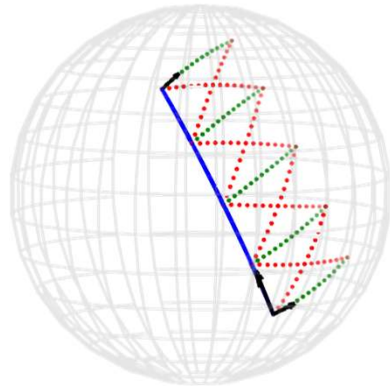
Gavrilov's Taylor expansion of one Schild's ladder step

- Taylor series for mid-point rule

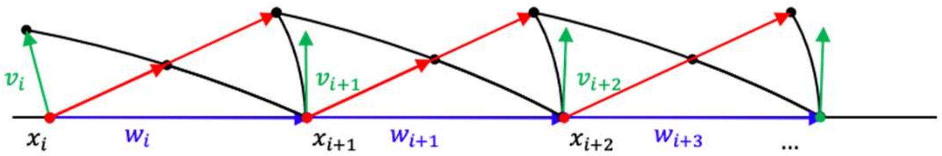


$$2a = w + v + \frac{1}{6} R(v,w)(w-v) + O(4)$$

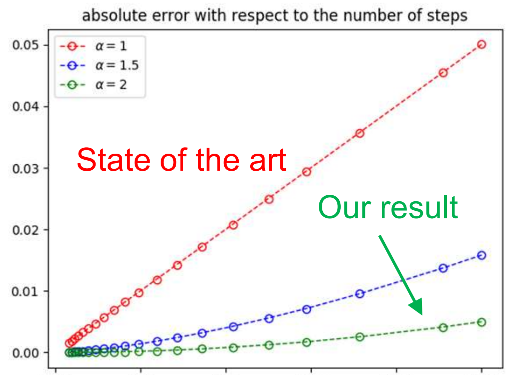
$$u - u^w = \frac{1}{2} R(w,v)v + O(4)$$



Convergence of the iterated Schild's ladder



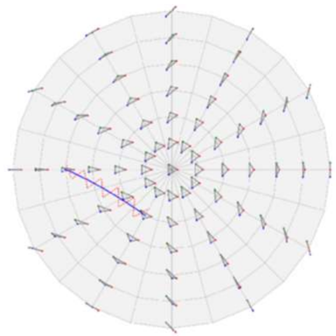
$$v_{i+1} = n^\alpha \cdot \operatorname{schild}\left(x_i, \frac{w_i}{n}, \frac{v_i}{n^\alpha}\right),$$



Theorem: the scheme converge at speed $\|v_n - \Pi_x^{x_n} v\| \leq \frac{\tau}{n^\alpha} + \frac{\beta}{n^2}.$

[N. Guigui, XP, Numerical Accuracy of Ladder Schemes for Parallel Transport on Manifolds. Foundations of Computational Mathematics, 2021. Arxiv 2007.07585.]

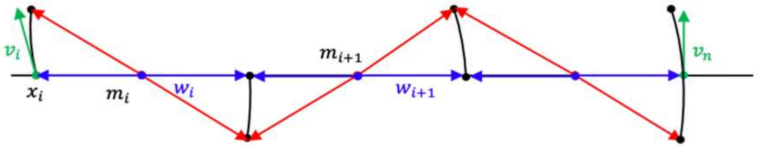
Convergence of pole Ladder

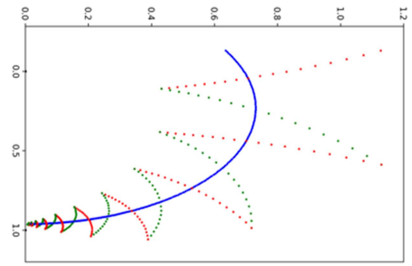


Taylor expansion of one pole ladder step

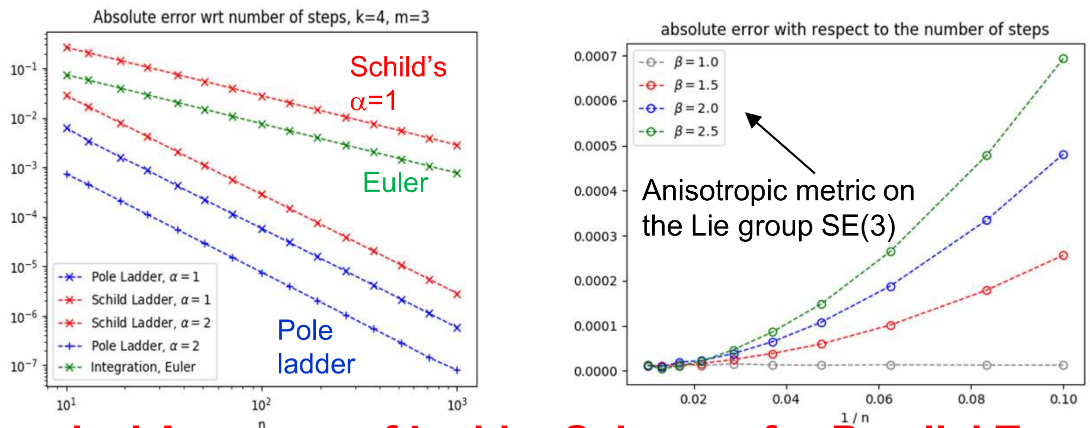
$$\text{pole}(u) = \Pi(u) + \frac{1}{12}\nabla_v R(u,v)(5u-2v) + \frac{1}{12}\nabla_u R(u,v)(v-2u) + O(5)$$

Convergence of the iterated pole ladder





Theorem: the scheme converge at speed $\|v_n - \Pi_m^{m_n} v\| \leq \frac{\gamma}{n^2}$

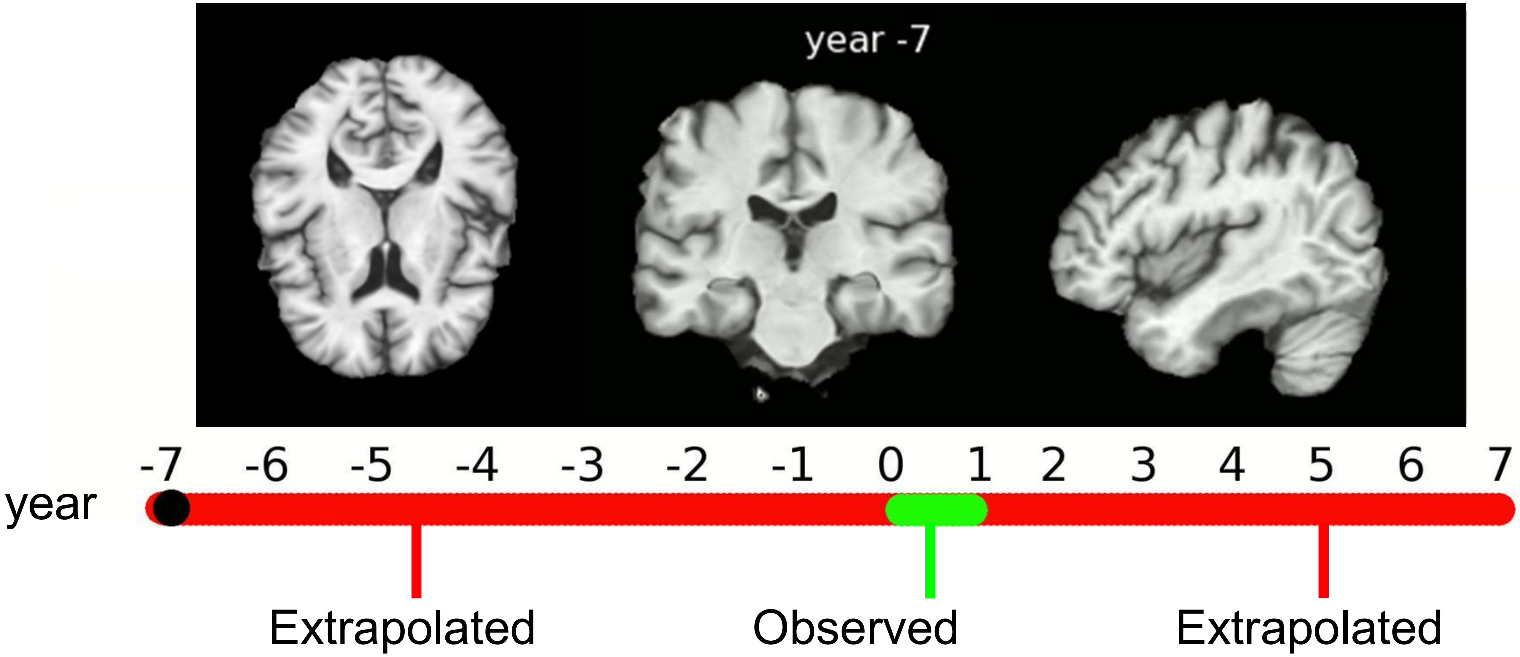


[N. Guigui, XP, Numerical Accuracy of Ladder Schemes for Parallel Transport on Manifolds. Foundations of Computational Mathematics, 2021. Arxiv 2007.07585.]

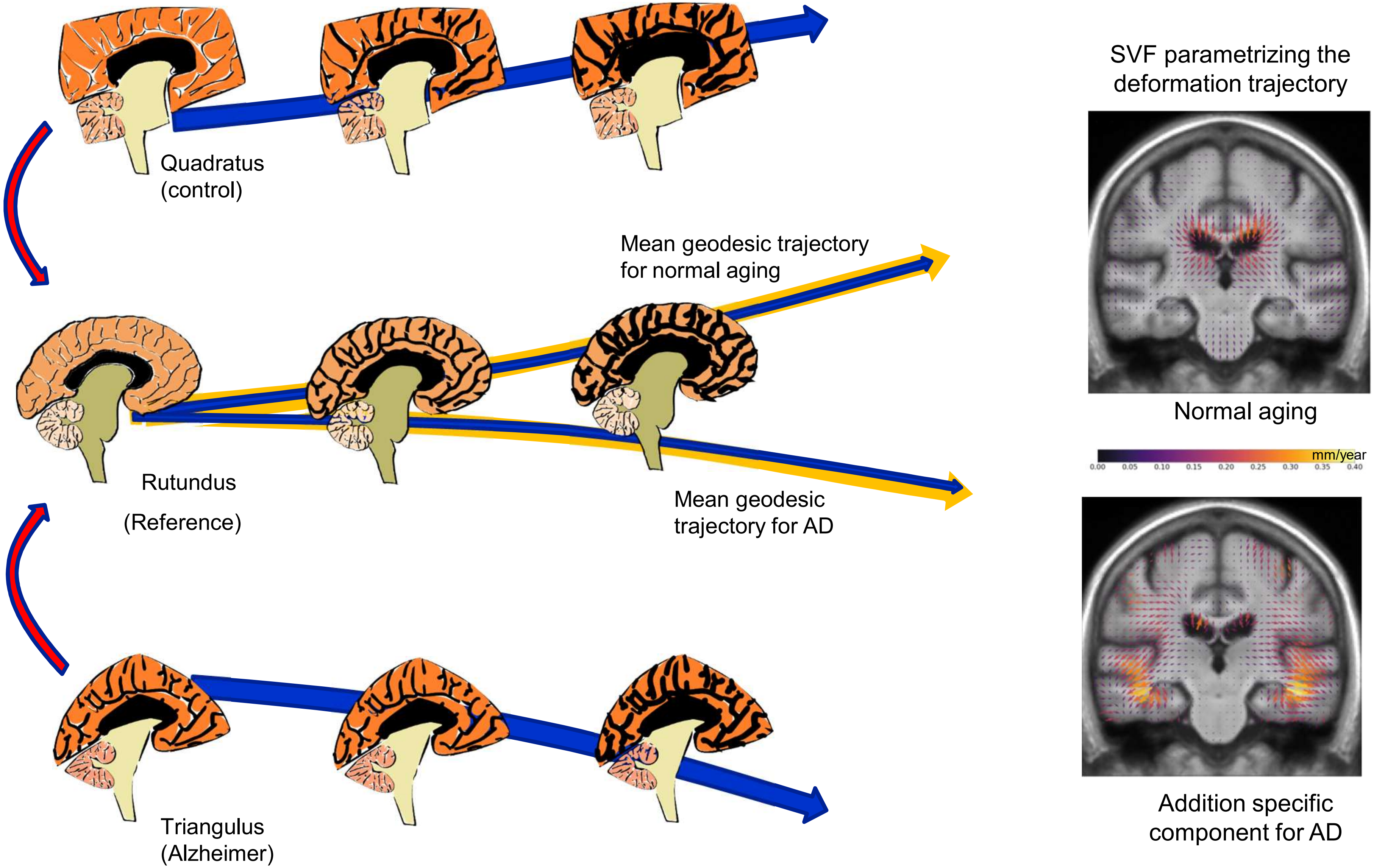
The Stationnary Velocity Fields (SVF) framework for diffeomorphisms

- 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

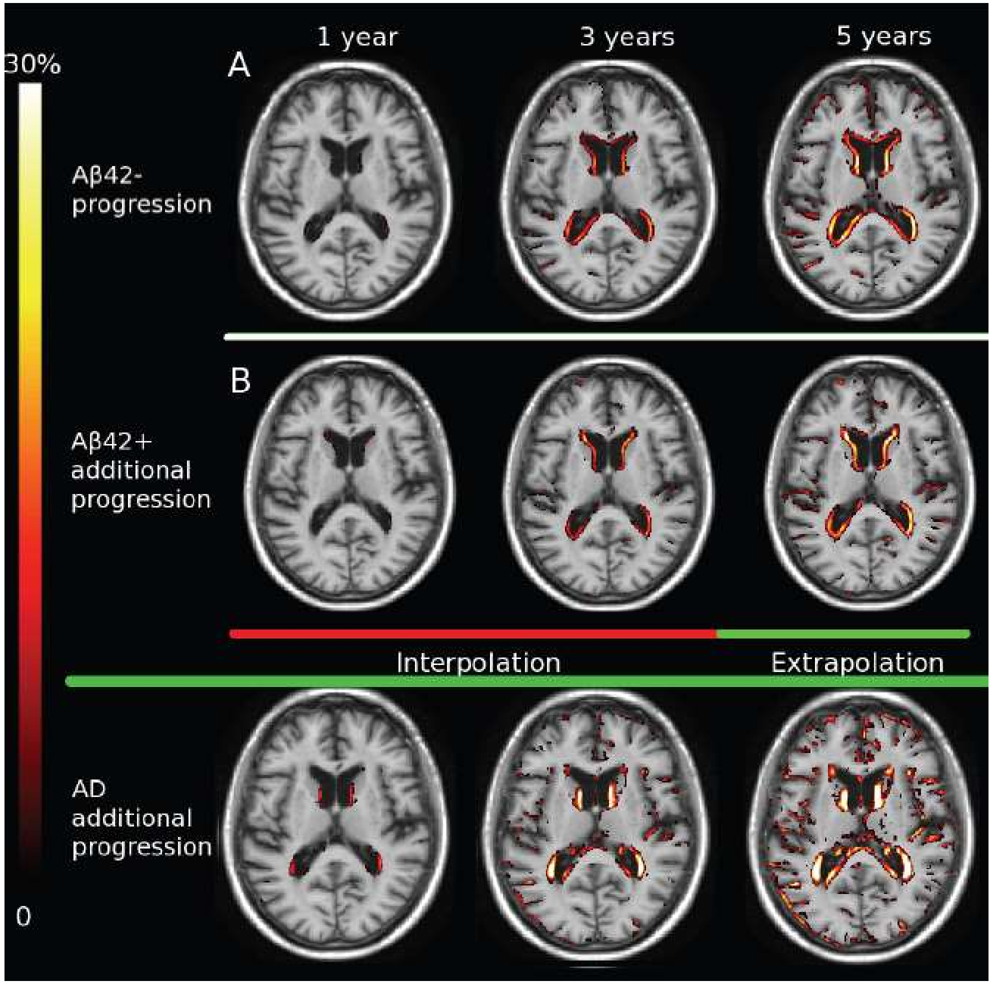


Modeling Normal and AD progression

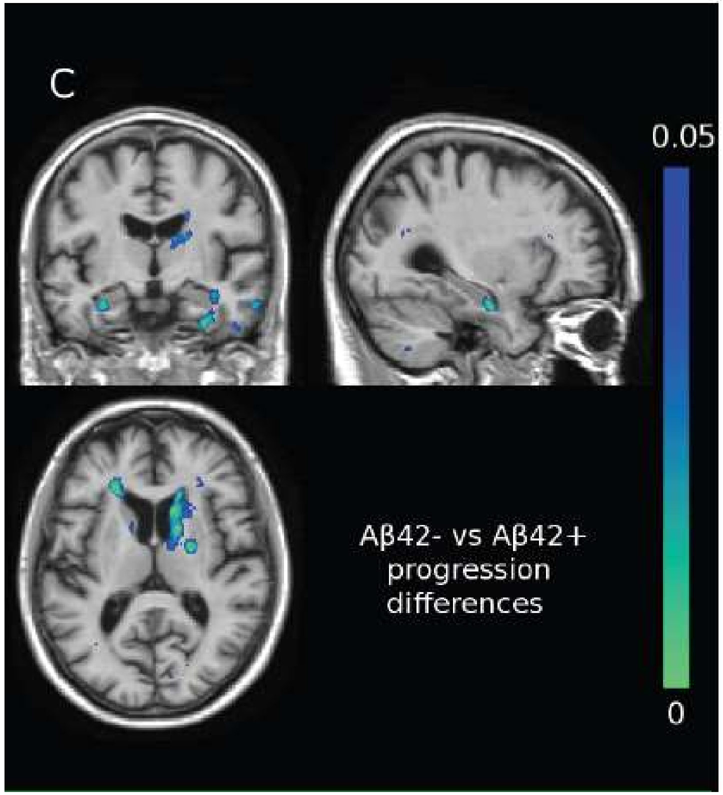


Study of prodromal Alzheimer's disease

Linear regression of the SVF over time: interpolation + prediction



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



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]

Advances in Geometric Statistics

Motivations

Simple statistics on Riemannian manifolds

Extension to transformation groups with affine spaces

Conclusion

Exp_x / Log_x and Fréchet mean are the basis of algorithms to compute on Riemannian/affine manifolds

Simple statistics

- Mean through an exponential barycenter iteration
- Covariance matrices and higher order moments
- Tangent PCA or more complex PGA / BSA

Efficient Discrete parallel transport using Schild / Pole ladder

- Quadratic convergence in #step for general (non-closed form) geodesics

Manifold-valued image processing [XP, IJCV 2006]

- Interpolation / filtering / convolution: weighted means
- Diffusion, extrapolation:
  Discrete Laplacian in tangent space = Laplace-Beltrami

Affine vs Riemannian connection for Lie groups

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 not always 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?

Pushing the frontiers of Geometric Statistics

Beyond the Riemannian / metric structure

- Riemannian manifolds, Non-Positively Curved (NPC) metric spaces
- **Affine connection, Quotient, Stratified spaces (trees, graphs)**

Beyond the mean and unimodal concentrated laws

- **Nested sequences (flags) of subspace in manifolds**
- A continuum from PCA to Principal Cluster Analysis?

Geometrization of statistics

- **Geometry of sample spaces** **[Harms, Michor, XP, Sommer, arXiv:2010.08039]**
  - Stratified boundary of the smooth manifold of probability densities
- Explore **influence of curvature, singularities** (borders, corners, stratifications) on non-asymptotic estimation theory

Make G-Statistics an effective discipline applications

http://geomstats.ai : a python library to implement generic algorithms on many Riemannian manifolds

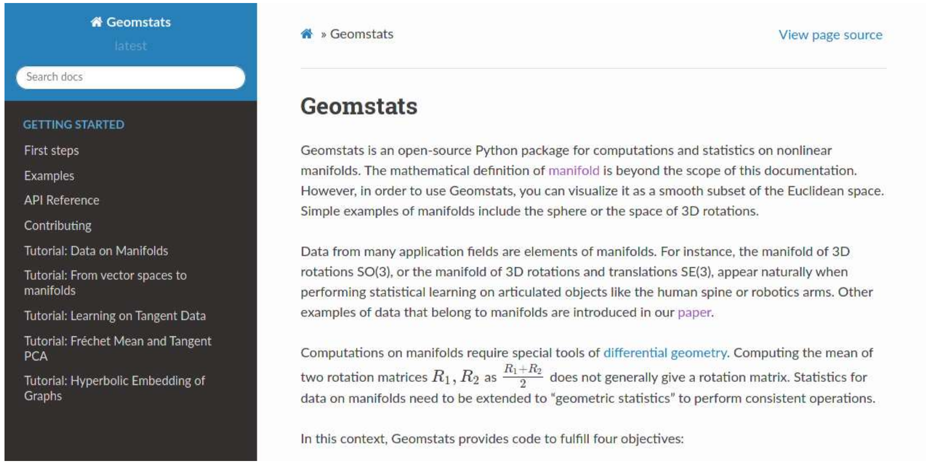
GitHub pypi package 2.2.0 build passing codecov 92%

- Mean, PCA, clustering, **parallel transport**…

- 15 manifolds / Lie groups already implemented (SPD, H(n), SE(n), etc)
- **Generic manifolds with geodesics by integration / optimization**

- scikit-learn API (hide geometry, compatible with GPU & learning tools).

- 10 introductory tutorials
- ~ 35000 lines of code
- ~20 academic developers
- 2 hackathons organized in 2020


Geomstats
latest
Search docs
GETTING STARTED
First steps
Examples
API Reference
Contributing
Tutorial: Data on Manifolds
Tutorial: From vector spaces to manifolds
Tutorial: Learning on Tangent Data
Tutorial: Fréchet Mean and Tangent PCA
Tutorial: Hyperbolic Embedding of Graphs

» Geomstats
View page source

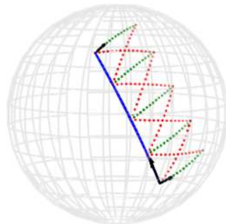
Geomstats

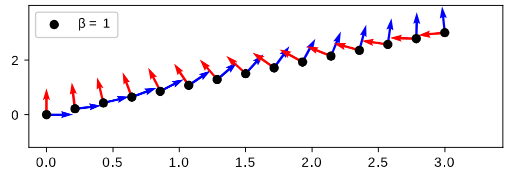
Geomstats is an open-source Python package for computations and statistics on nonlinear manifolds. The mathematical definition of manifold is beyond the scope of this documentation. However, in order to use Geomstats, you can visualize it as a smooth subset of the Euclidean space. Simple examples of manifolds include the sphere or the space of 3D rotations.

Data from many application fields are elements of manifolds. For instance, the manifold of 3D rotations SO(3), or the manifold of 3D rotations and translations SE(3), appear naturally when performing statistical learning on articulated objects like the human spine or robotics arms. Other examples of data that belong to manifolds are introduced in our paper.

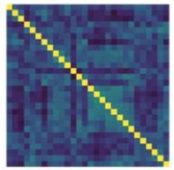
Computations on manifolds require special tools of differential geometry. Computing the mean of two rotation matrices R_1, R_2 as $\frac{R_1+R_2}{2}$ does not generally give a rotation matrix. Statistics for data on manifolds need to be extended to "geometric statistics" to perform consistent operations.

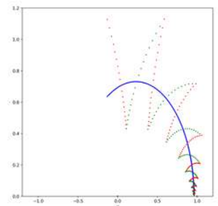
In this context, Geomstats provides code to fulfill four objectives:


Schild's/pole Ladders



Rotations-Translations


SPD



[Miolane et al, JMLR 2020, Scipy 2020]

The G-Statistics groups

Yann Thanwerdas

Geometric Statistics

Nicolas Guigui

Morten Pedersen

Dimbihery Rabenoro

Anna Calissano

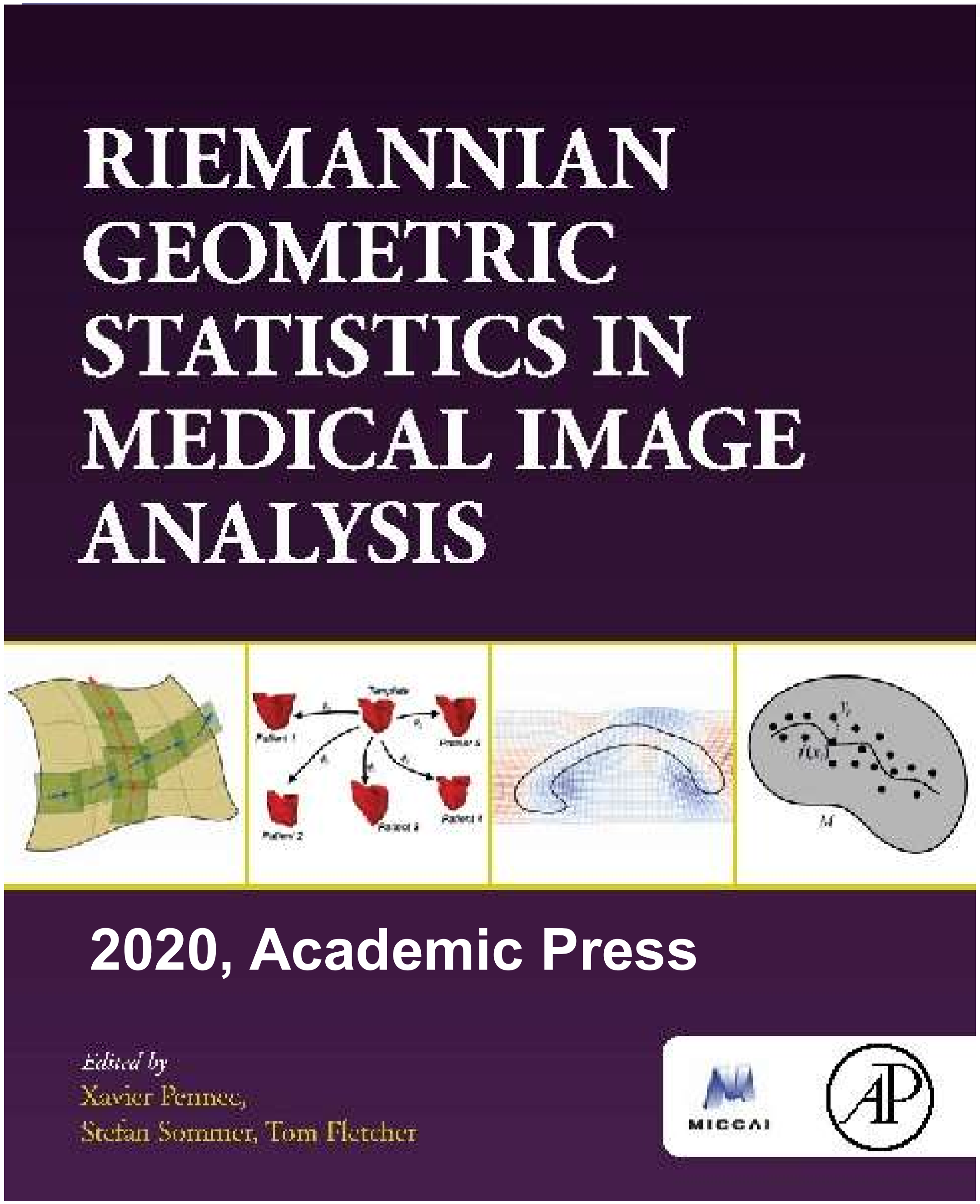
James Benn

Elodie Maignant

Luis G. Pereira

PhD positions available

Tom Szwagier



Thank you for your attention

Thanks to many collaborators

- Marco Lorenzi
- Jonathan Boisvert
- Pierre Fillard
- Vincent Arsigny
- Maxime Sermesant
- Nicholas Ayache
- Kristin McLeod
- Nina Miolane
- Loic Devillier
- Marc-Michel Rohé
- Yann Thanwerdas
- Nicolas Guigui
- …….

Selected References

Statistics on Riemannnian manifolds

- XP. Intrinsic Statistics on Riemannian Manifolds: Basic Tools for Geometric Measurements. JMIV, 25(1):127-154, July 2006.

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

- XP, Pierre Fillard, and Nicholas Ayache. A Riemannian Framework for Tensor Computing. IJCV, 66(1):41-66, Jan. 2006.

Bi-invariant means with Cartan connections on Lie groups

- XP 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.

Cartan connexion for diffeomorphisms:

- Marco Lorenzi and XP. Geodesics, Parallel Transport & One-parameter Subgroups for Diffeomorphic Image Registration. IJCV, 105(2), November 2013

Manifold dimension reduction (extension of PCA)

- XP. Barycentric Subspace Analysis on Manifolds. Annals of Statistics. 46(6A):2711-2746, 2018. [arXiv:1607.02833]