



POLITECNICO DI MILANO



Politecnico di Milano
Applied Statistics
May 2023



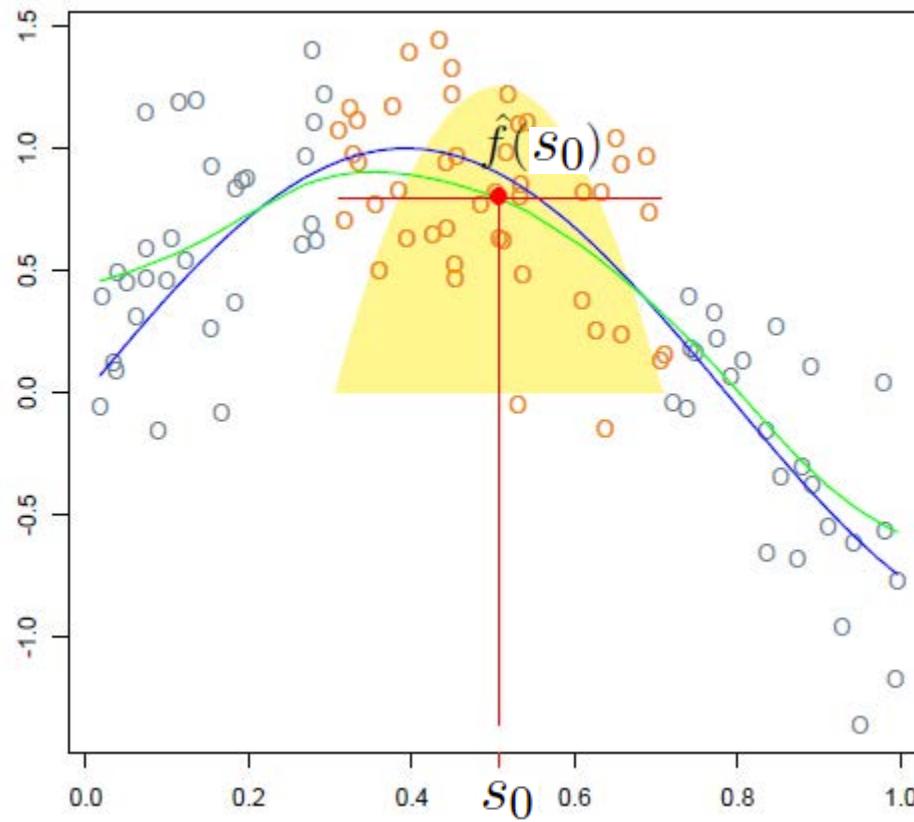
An introduction to functional data analysis

Laura M. SANGALLI

MOX - Dipartimento di Matematica, Politecnico di Milano

Part 3 - Smoothing

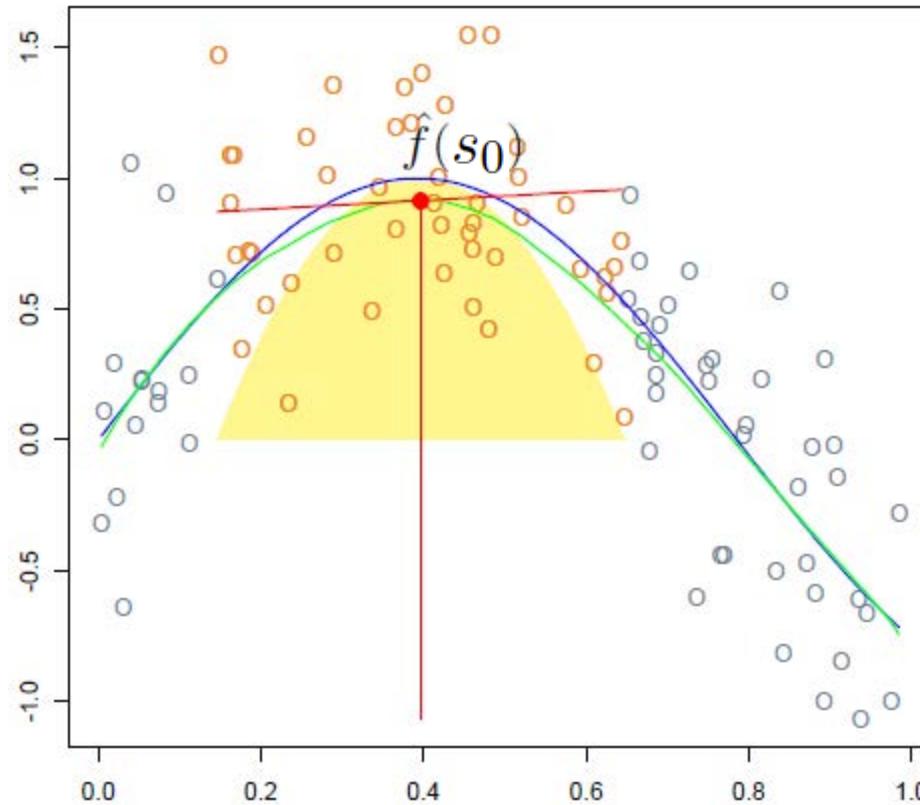
Local polynomial regression (kernel smoother)



Picture taken from
Friedman, Tibshirani and
Hastie (2013) The Elements of
Statistical Learning, Springer

At each abscissa s_0 , find (c_0, \dots, c_L) that minimize

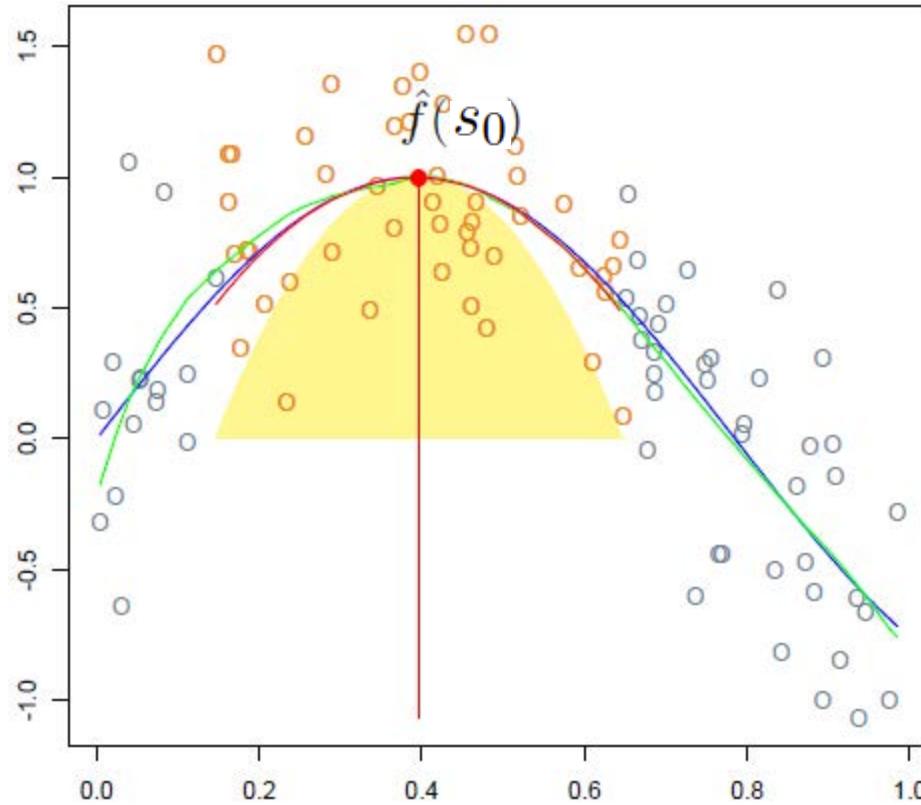
$$\sum_{i=1}^n \text{Kern}_h(s_0, s_i)[(z_i - \sum_{l=0}^L c_l(s_0 - s_i)^l)]^2 \quad \text{where } \text{Kern}_h(s_0, s_i) = D\left(\frac{|s_0 - s_i|}{h}\right)$$



Picture taken from
Friedman, Tibshirani and
Hastie (2013) The Elements of
Statistical Learning, Springer

At each abscissa s_0 , find (c_0, \dots, c_L) that minimize

$$\sum_{i=1}^n \text{Kern}_h(s_0, s_i)[(z_i - \sum_{l=0}^L c_l(s_0 - s_i)^l)]^2 \quad \text{where } \text{Kern}_h(s_0, s_i) = D\left(\frac{|s_0 - s_i|}{h}\right)$$



Picture taken from
Friedman, Tibshirani and
Hastie (2013) The Elements of
Statistical Learning, Springer

At each abscissa s_0 , find (c_0, \dots, c_L) that minimize

$$\sum_{i=1}^n \text{Kern}_h(s_0, s_i)[(z_i - \sum_{l=0}^L c_l(s_0 - s_i)^l)]^2 \quad \text{where } \text{Kern}_h(s_0, s_i) = D\left(\frac{|s_0 - s_i|}{h}\right)$$

Positive functions:

$$f(s) = e^{W(s)} \quad \text{where } W(s) = \sum_k c_k \psi_k(s)$$

Estimate f by minimizing

$$\sum_{i=1}^n (z_i - e^{W(s_i)})^2 + \lambda \int W''(s) ds$$

Increasing functions:

$$f(s) = C + \int_{s_0}^s \exp\{W(t)\} dt \quad \text{where } W(s) = \sum_k c_k \psi_k(s)$$

Densities (B^2 space):

Machalová, J., Hron, K., Monti, G., 2015. Preprocessing of centred logratio transformed density functions using smoothing splines. *J. Appl. Stat.* 43.



POLITECNICO DI MILANO



Politecnico di Milano
Applied Statistics
May 2023



An introduction to functional data analysis

Laura M. SANGALLI

MOX - Dipartimento di Matematica, Politecnico di Milano

Part 4 - Alignment

Decoupling and studying Phase and Amplitude variabilities

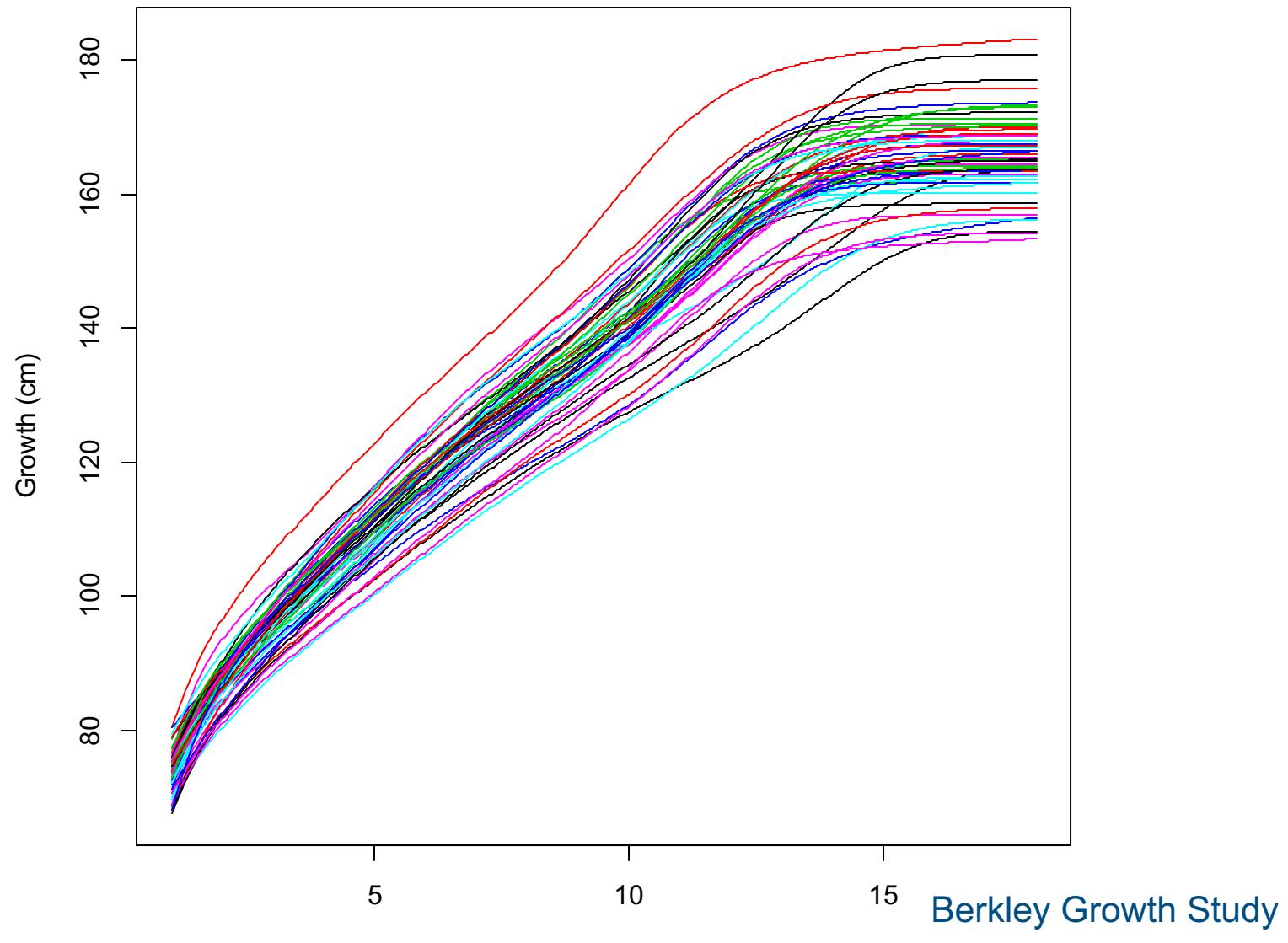
Registration, Alignment, Warping

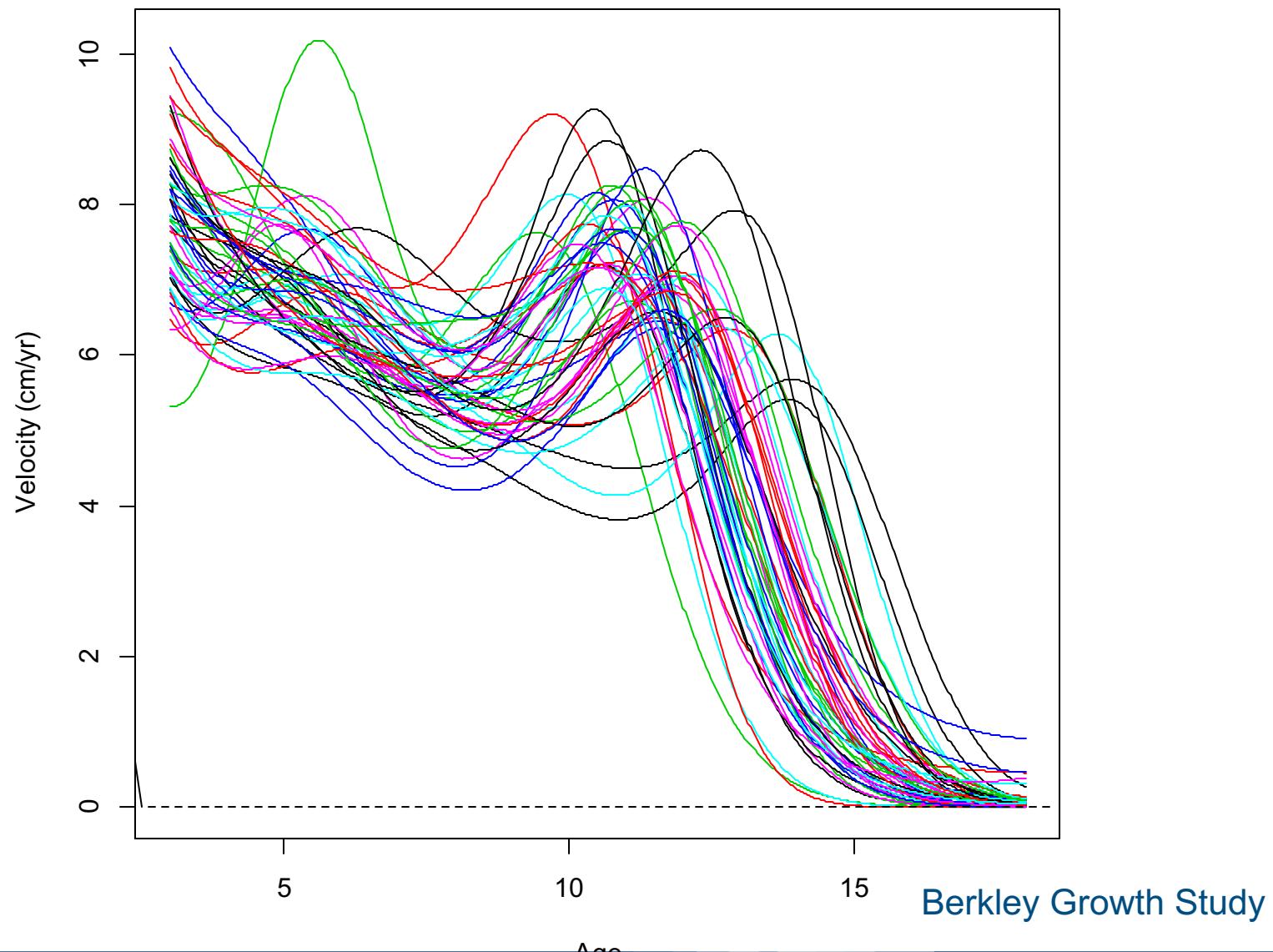
Chapter 7, Ramsay and Silverman 2005, Springer

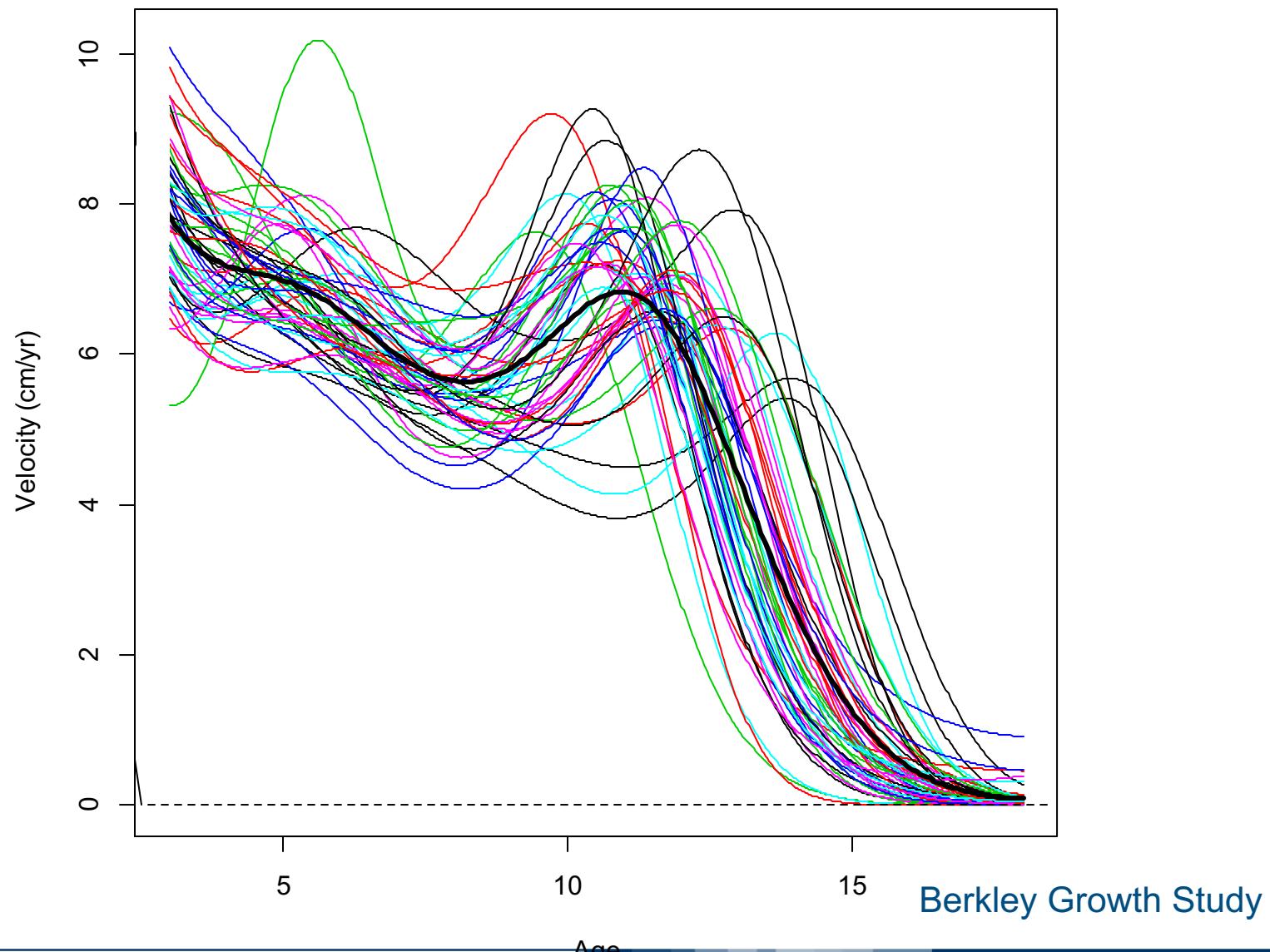
Review article:

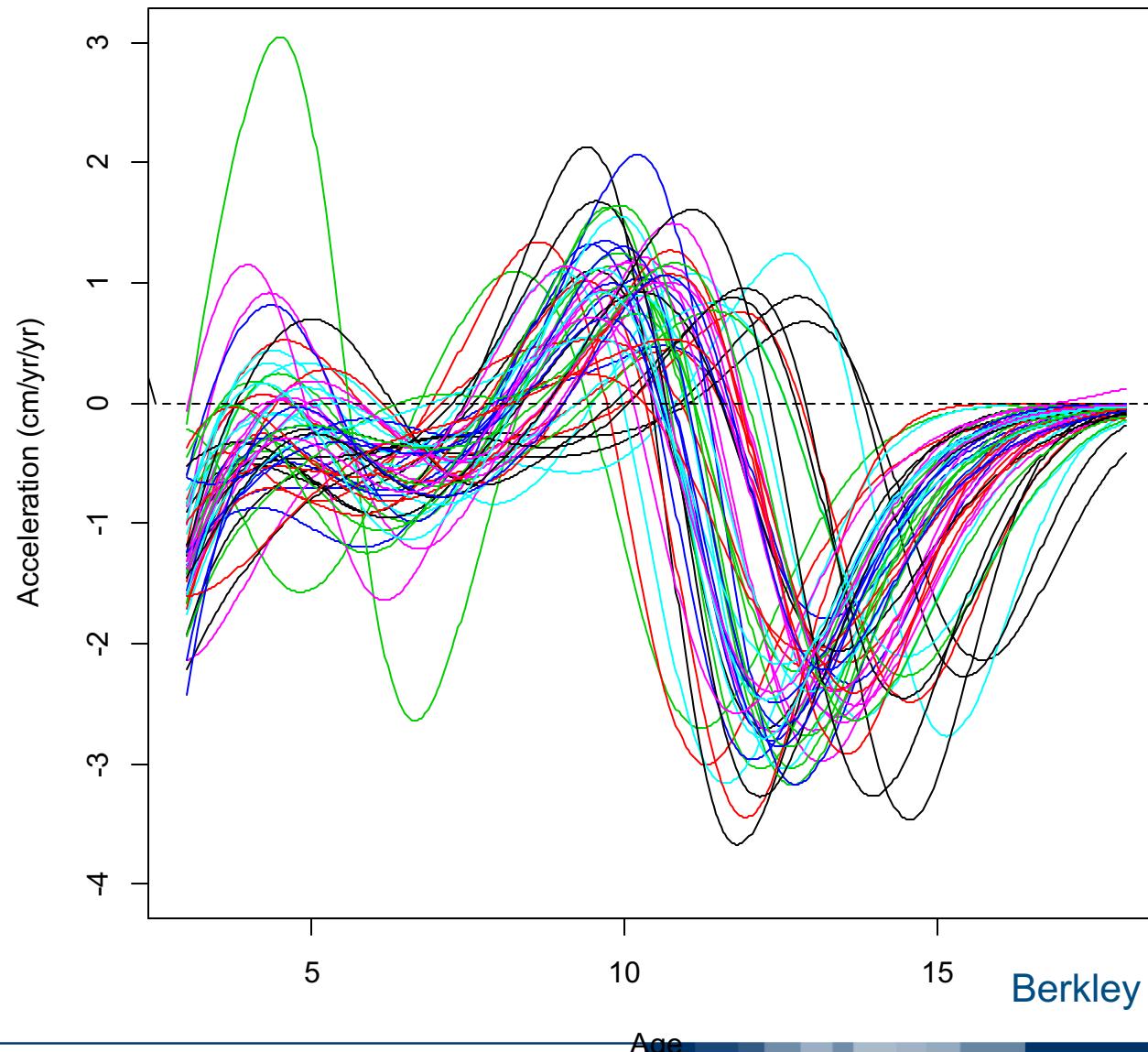
Marron, J.S., Ramsay, J.O., Sangalli, L.M., Srivastava, A. (2015), Functional Data Analysis of Amplitude and Phase Variation, *Statistical Science*, 30 (4), 468-484.

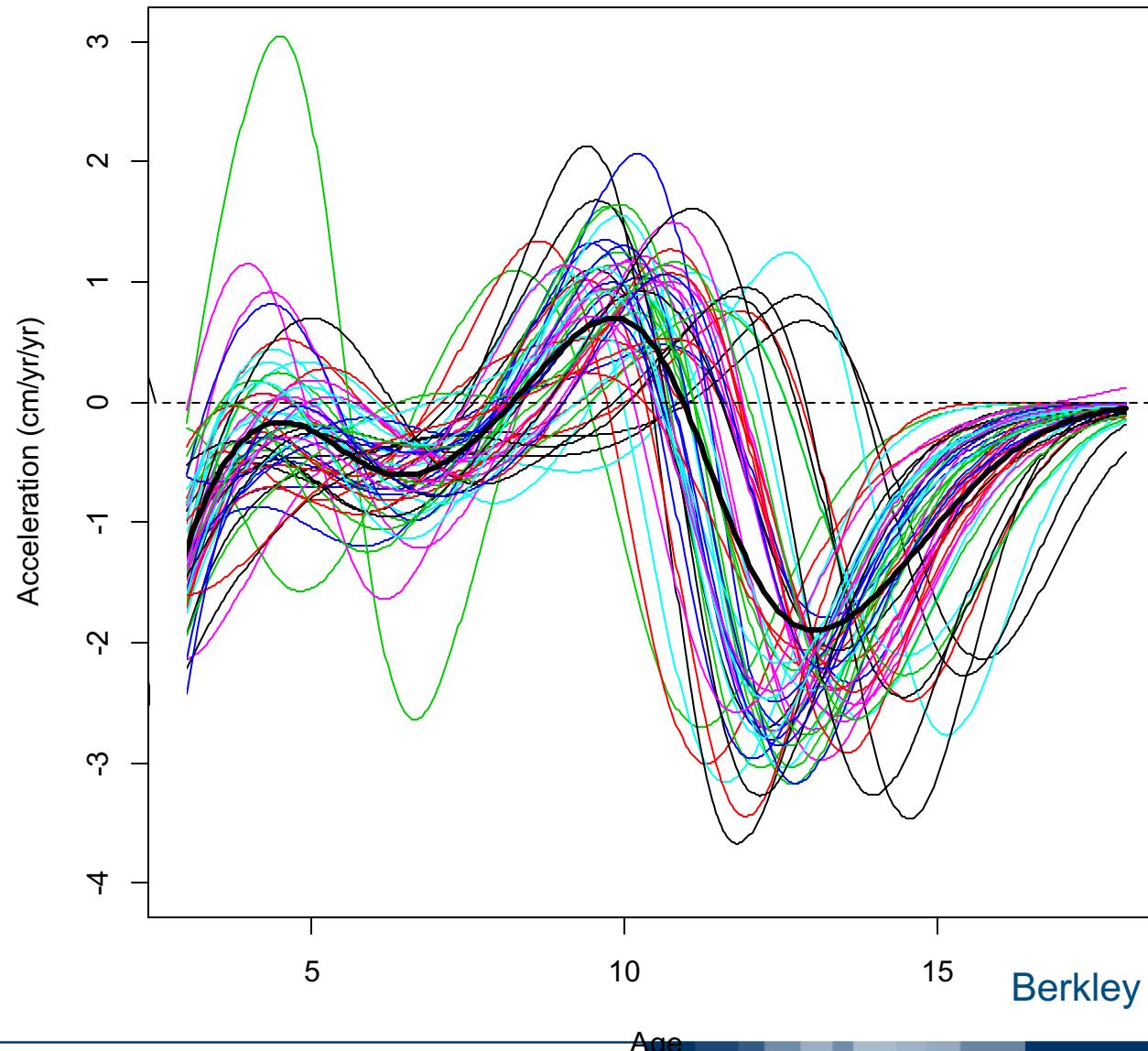




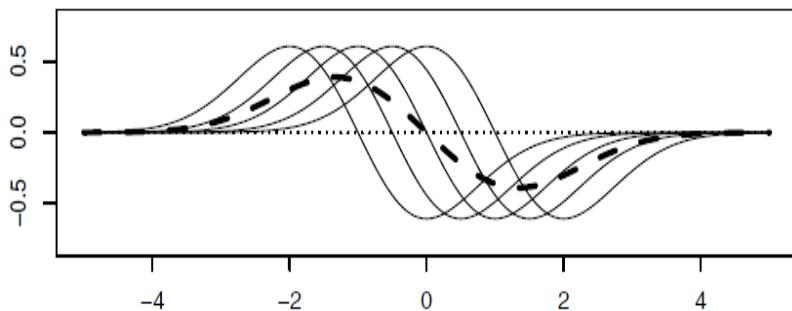




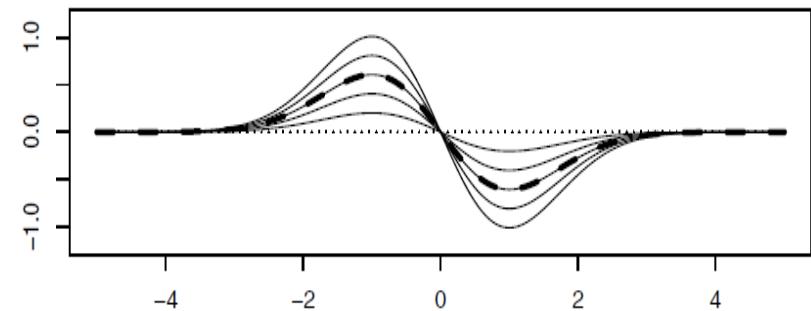




Phase variability



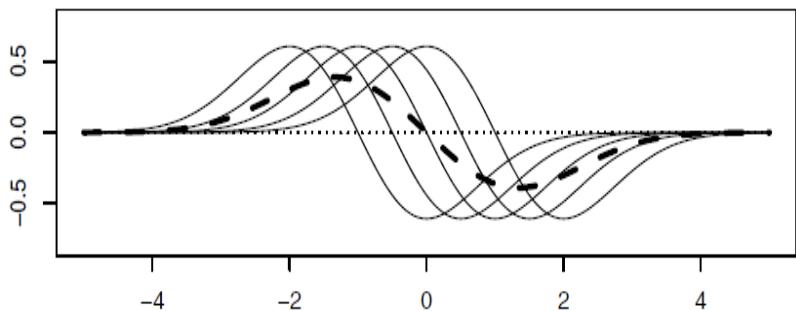
Amplitude variability



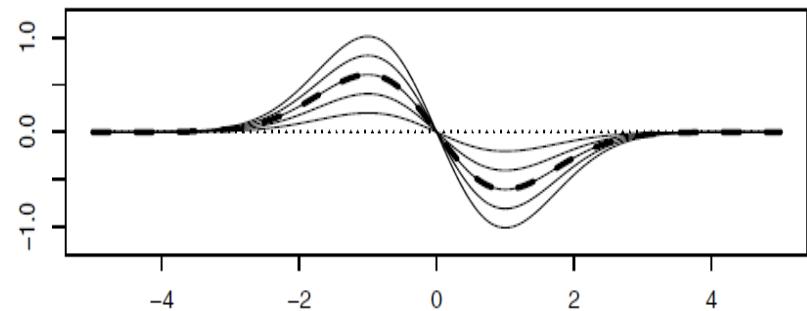
Phase variability: different curves exhibit more or less the same features but that these features occur at different times or space locations for different statistical units.

If not taken properly into account, the misalignment acts as a confounding factor and may blur subsequent analyses.

Phase variability



Amplitude variability



Registration of a set of functions

Find suitable warping functions $h_1(t), \dots, h_N(t)$ such that $c_1(h_1(t)), \dots, c_N(h_N(t))$ are the most similar.

The functions h_i should be increasing; they capture the phase variability. Amplitude variability is the remaining variability in vertical direction among the aligned curves.

In some cases, time or location is merely shifted from curve to curve, for example, because the measurements are started at random time points. For these situations, it is natural to use $h_i(t) = t + \delta t_i$. In other situations, phase variation is a matter of dilation, in which case $h_i(t) = \alpha_i t$ is a natural choice of warping function. In yet other situations, the time or space deformation is more complex.

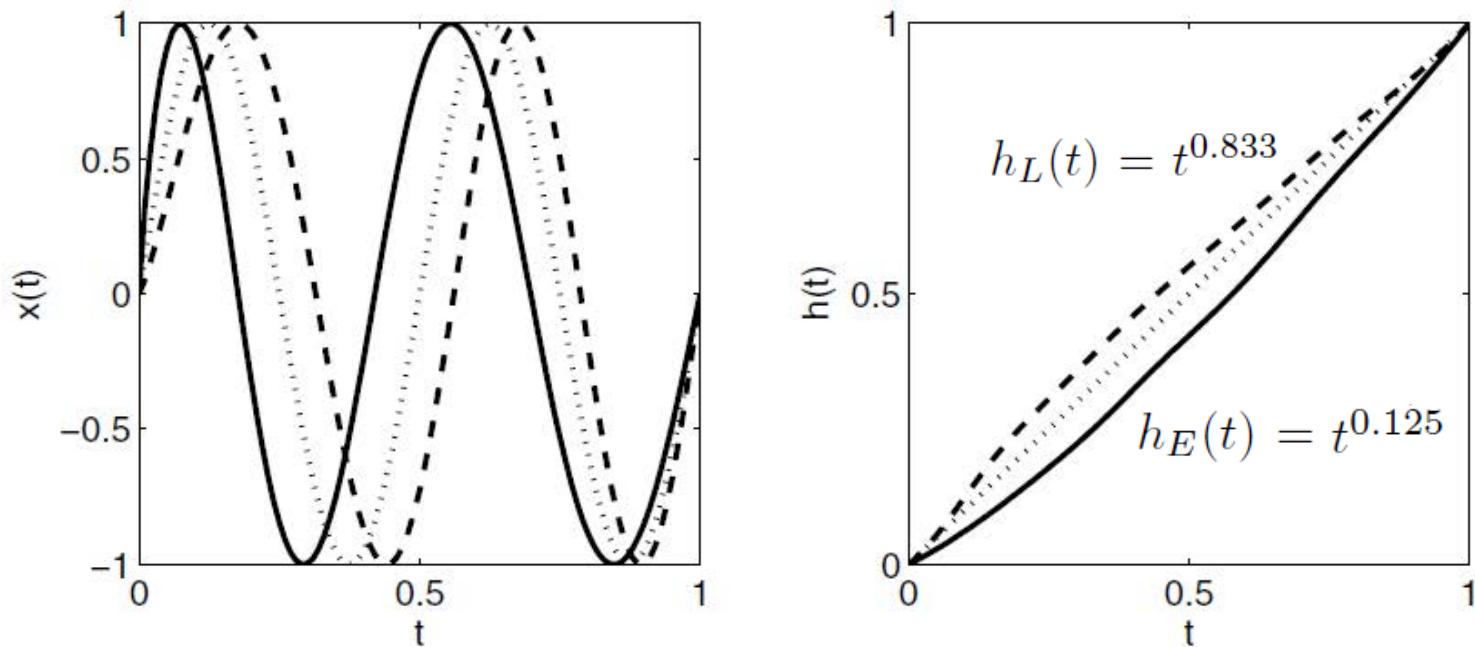
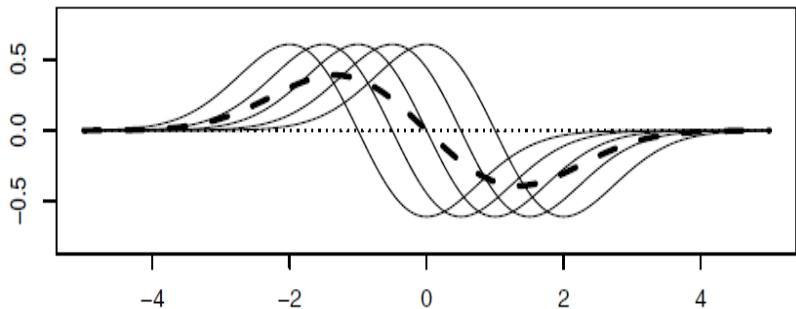


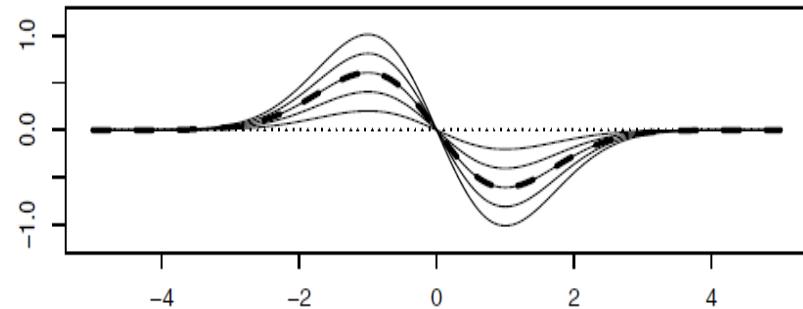
Figure 7.9. The left panel shows the target function, $x_0(t) = \sin(4\pi t)$, as a dotted line; an early function, $x_E(t) = \sin(4\pi t^{0.8})$, as a solid line; and a late function, $x_L(t) = \sin(4\pi t^{1.2})$, as a dashed line. The corresponding warping functions that register the early and late curves to the target are shown in the right panel.



Phase variability



Amplitude variability



Registration of a set of functions

Find suitable warping functions h_1, \dots, h_N such that $c_1 \circ h_1, \dots, c_n \circ h_N$ are the most similar.

→ **Landmark Approach:** known **landmarks** along the curves that are aligned so that landmarks occurs at the same abscissa points.

→ **Continuous Approach:** define a measure of similarity/dissimilarity between curves, that are aligned in order to maximize/minimize their similarity/dissimilarity.

Landmarks: significant (univocally identifiable) shape-events in a curve, e.g. crossings of zero, peaks, valleys, points of inflection.

c_1, \dots, c_N , where $c_i : [0, T] \rightarrow \mathbb{R}^d$

Suppose

- L landmarks; for the i -th curve, located at t_{i1}, \dots, t_{iL}
- a template curve c_0 is available with landmark locations t_{01}, \dots, t_{0L}
If not, we can define t_{0j} as the average of the t_{ij} 's

Warping function for the i -th curve: any strictly increasing function h_i s.t.

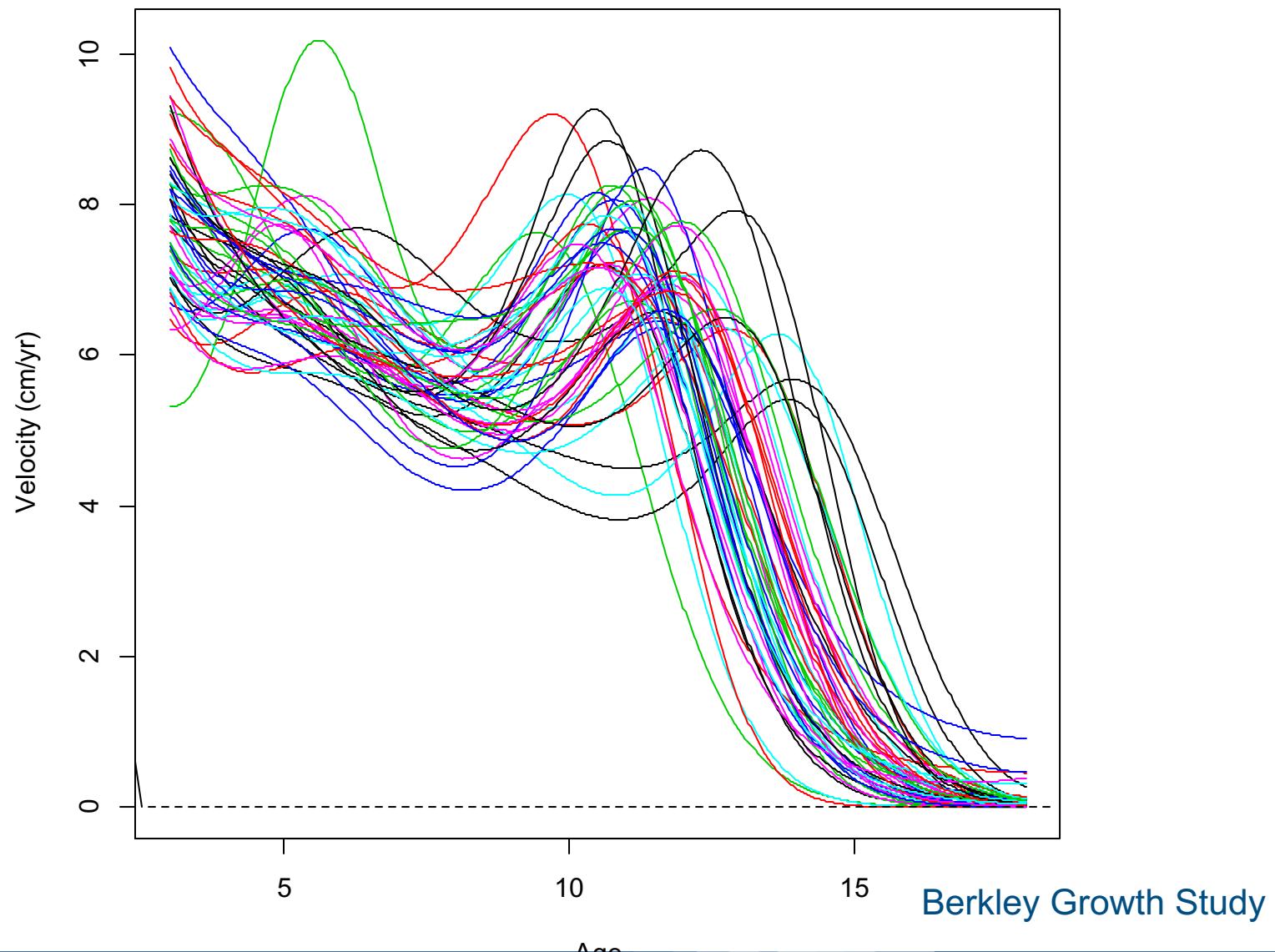
- $h_i(0) = 0$
 - $h_i(t_{0j}) = t_{ij}$, for $j = 1, \dots, L$
 - $h_i(T) = T$
- $(0, 0), (t_{01}, t_{i1}), \dots, (t_{0L}, t_{iL}), (T, T)$: interpolated by a piece-wise line, a polygon or higher order monotone splines (strictly increasing)

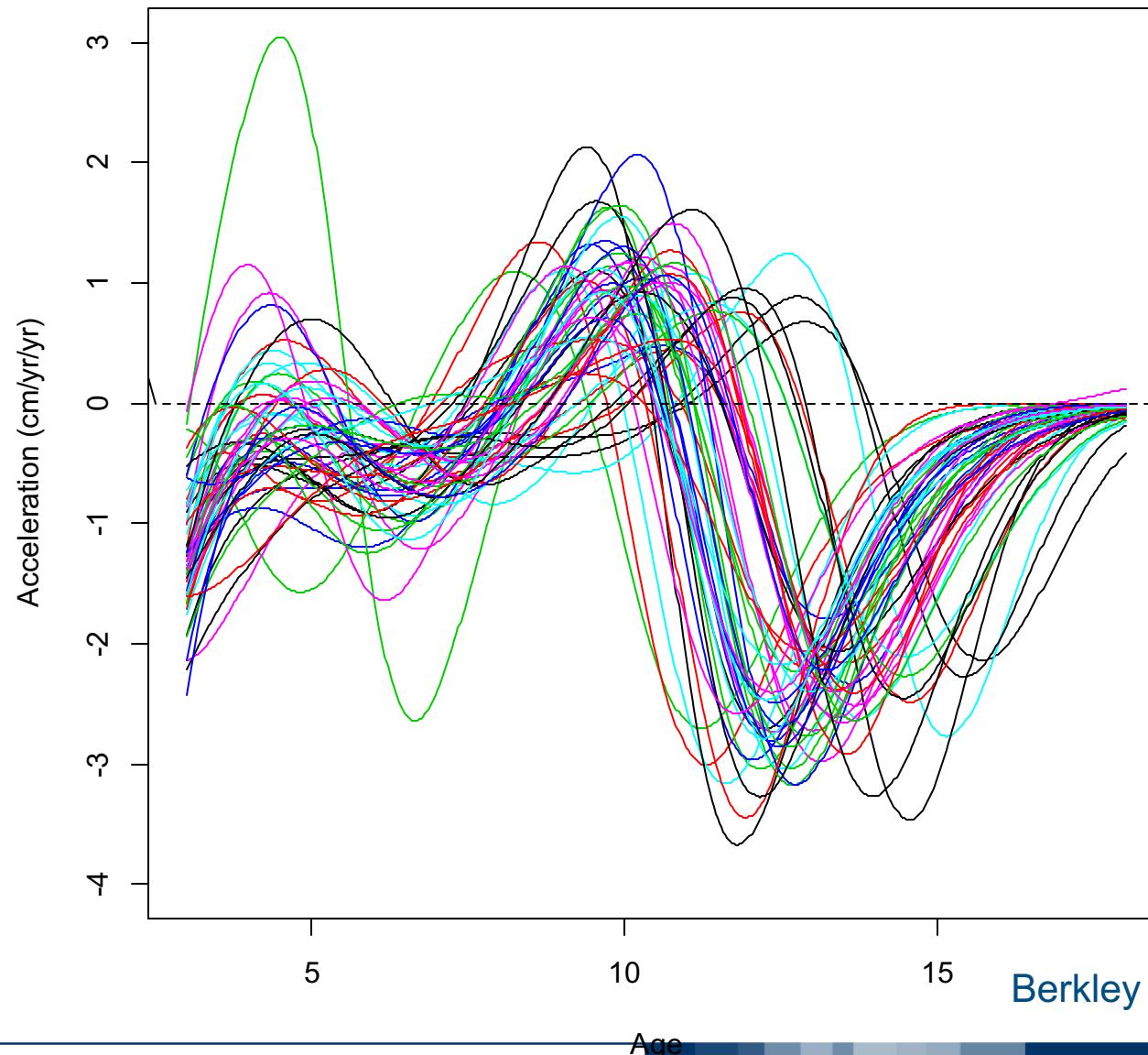
A model for warping functions

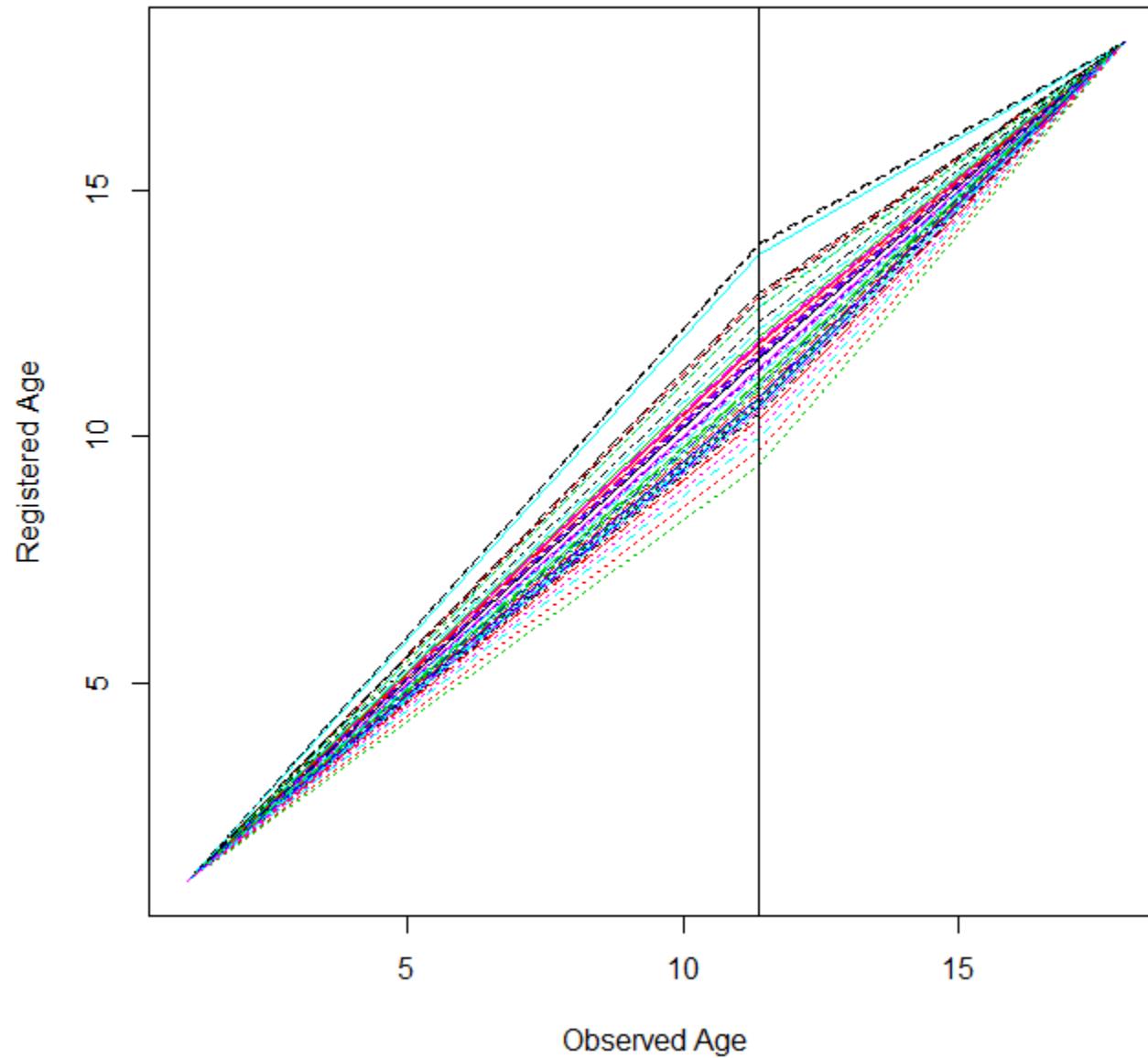
$$h_i(t) = C_{0i} + C_{1i} \int_0^t \exp\{W_i(u)\} du$$

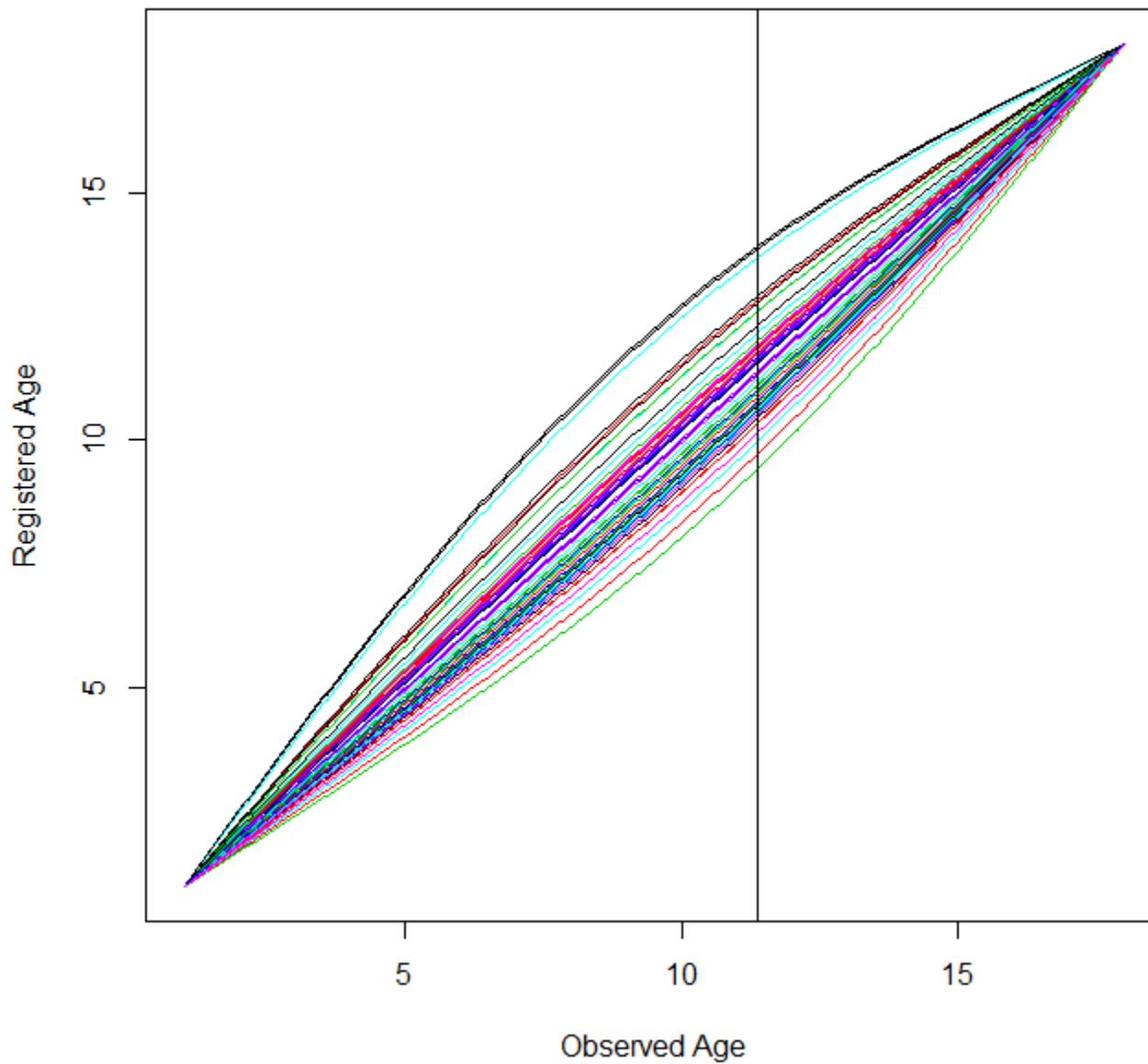
- C_{0i}, C_{1i} fixed by the requirement that $h_i(0) = 0$ and $h_i(T) = T$
- if shift registration is a possibility, C_{0i} can be allowed to pick phase shift
- Clock time corresponds to $W(u) = 0$.
- If $W_i(u)$ is positive, then $h_i(t) > t$, and warped time is growing faster than clock time, and this is what we want if our observed process is running late
- If $W_i(u)$ is negative, then $h_i(t) < t$, and clock time is being slowed down for a process that is running ahead of some target

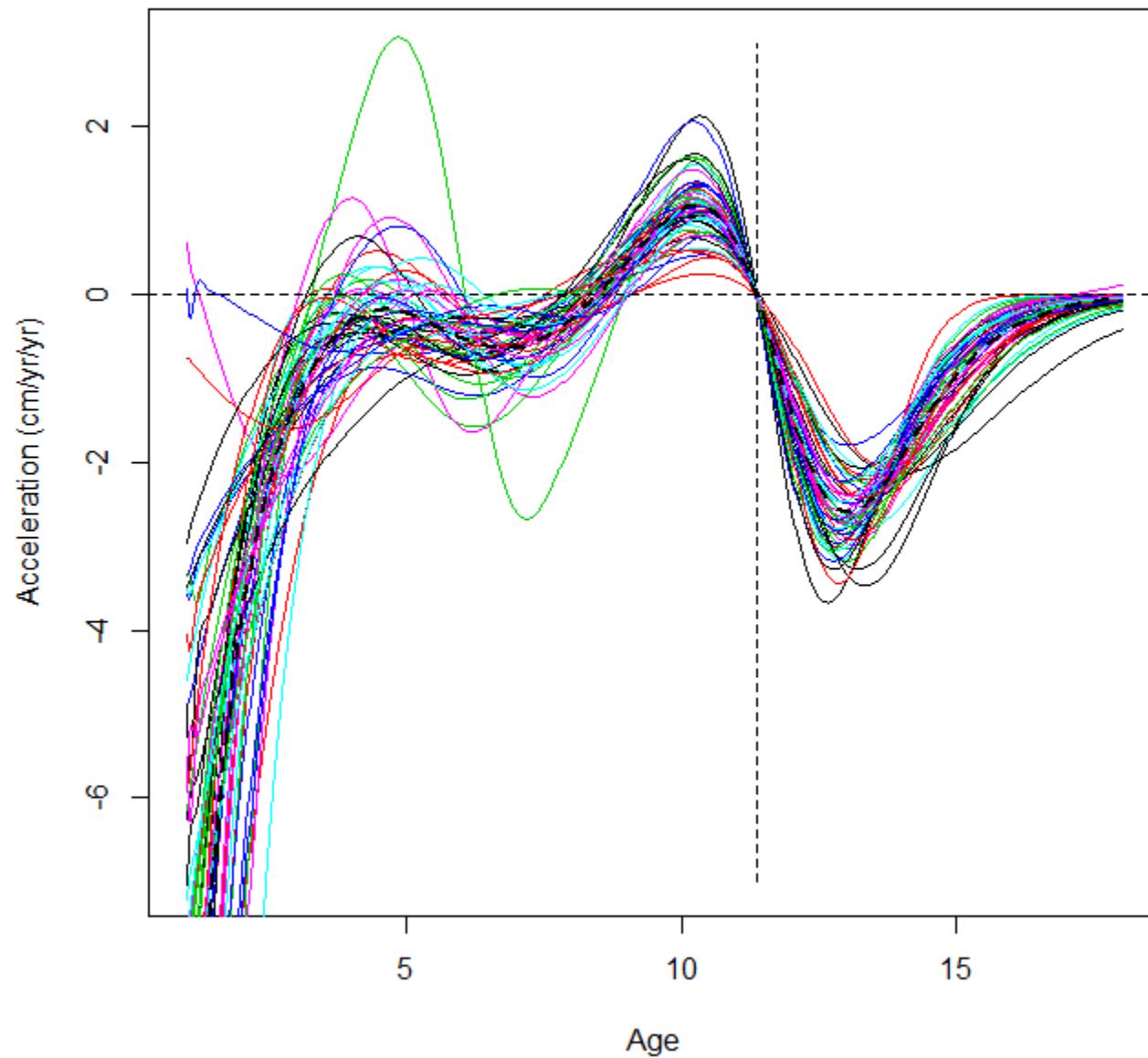


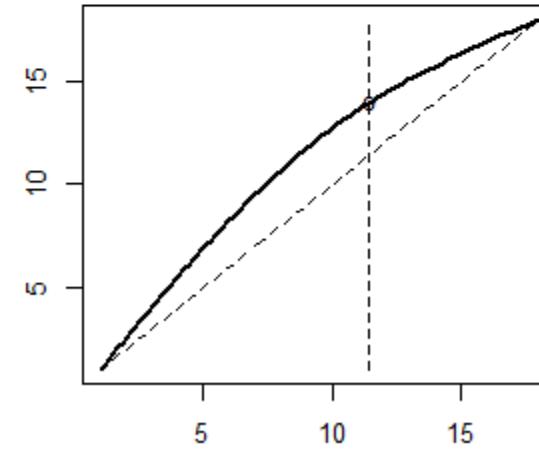
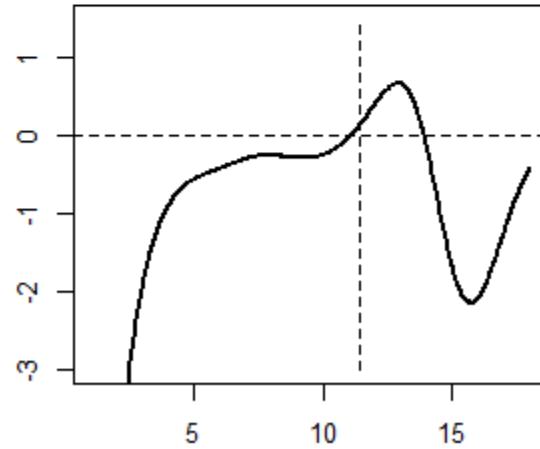
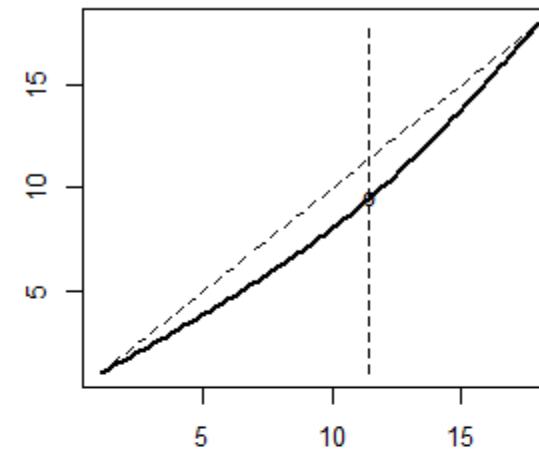
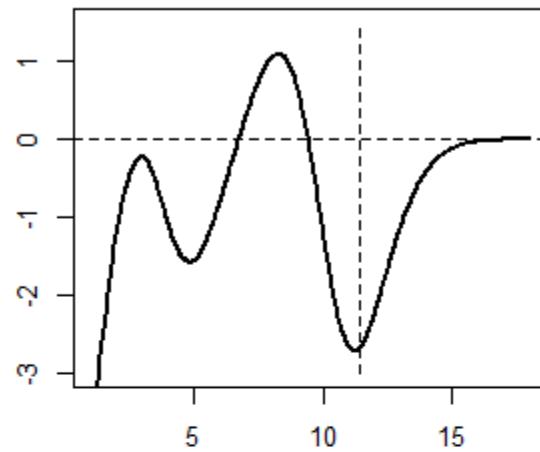






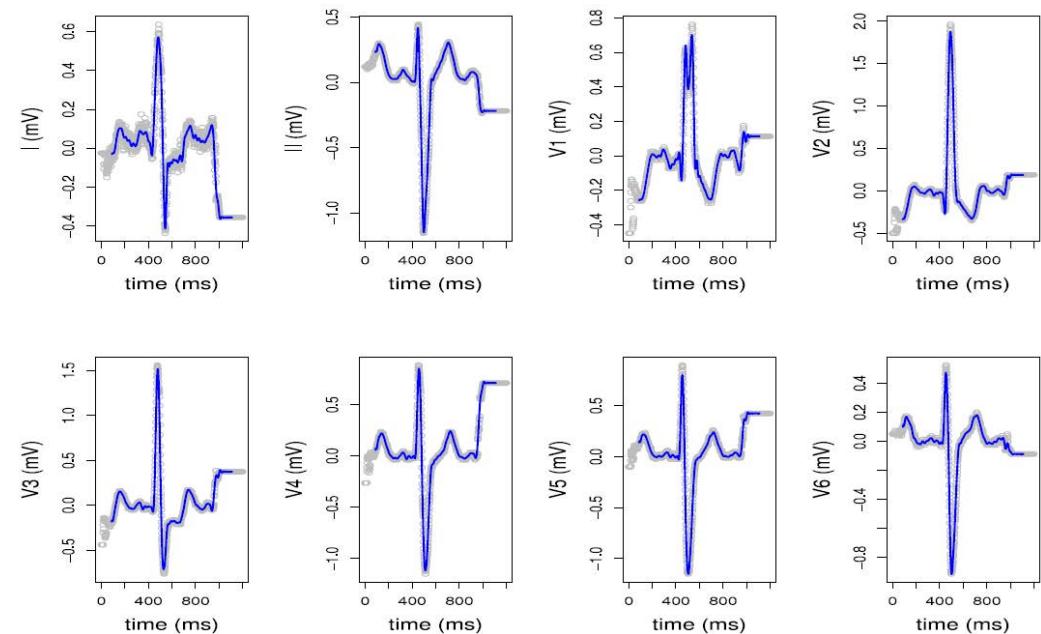
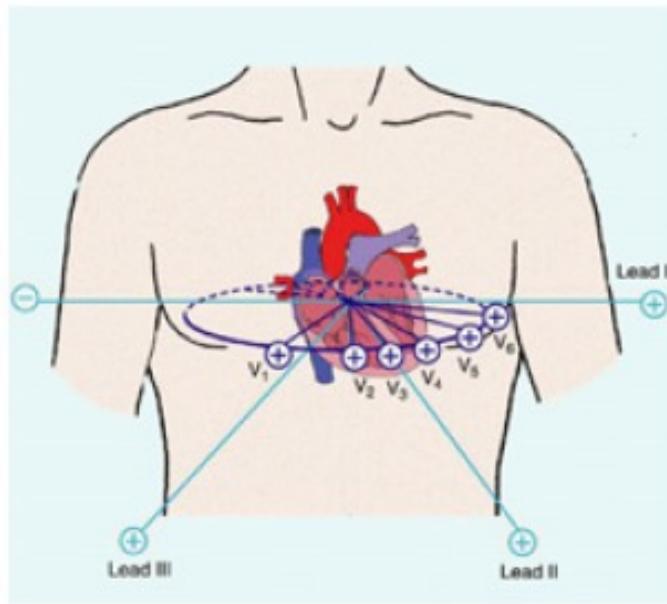






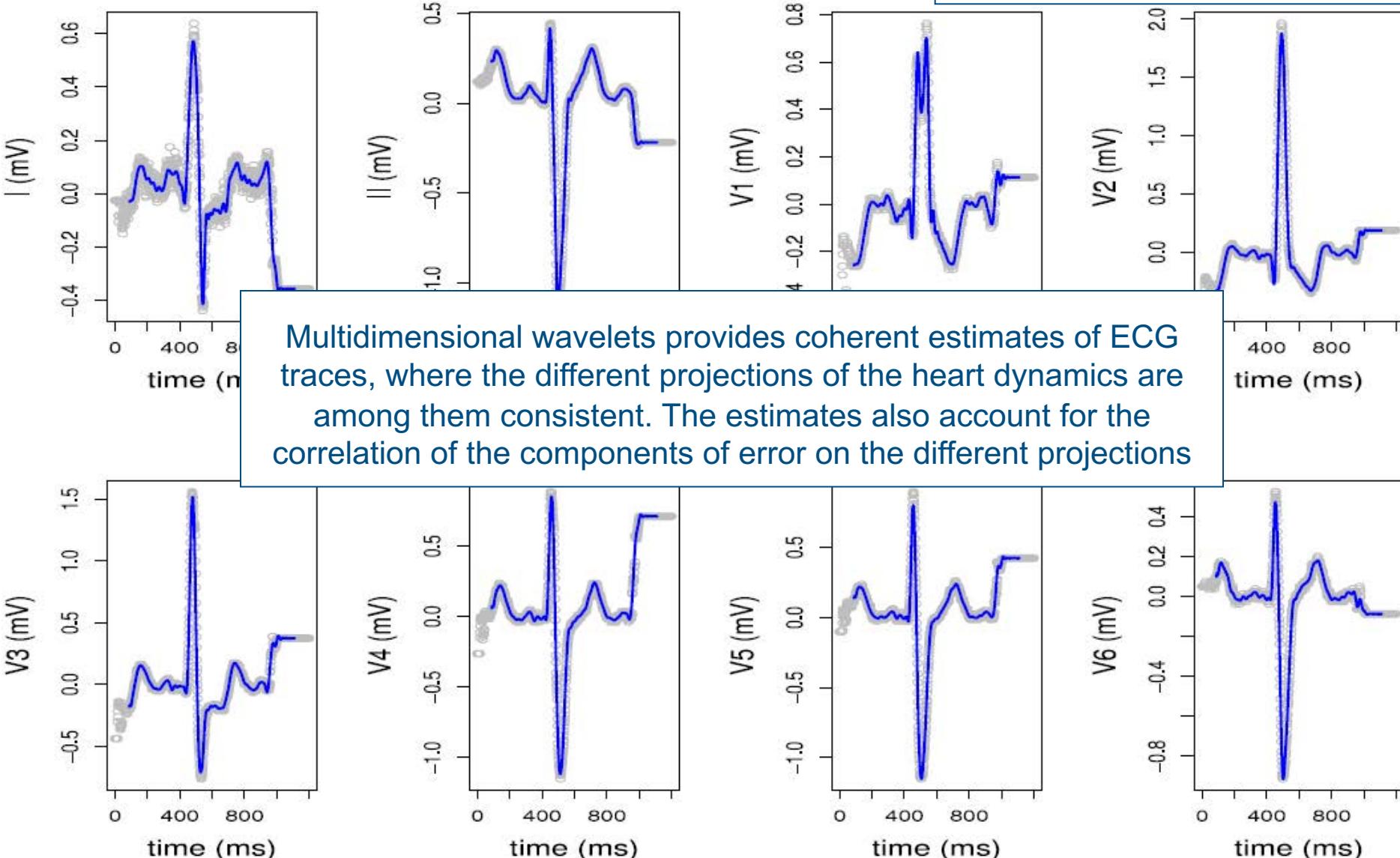
Case study:

Electro Cardio Gram (ECG) records alignment



Multi-leads ECG: eight-dimensional functional data, whose eight coordinates measure different projections of the heart dynamics in different directions

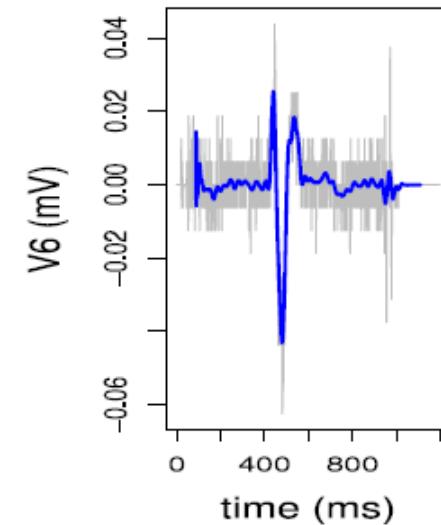
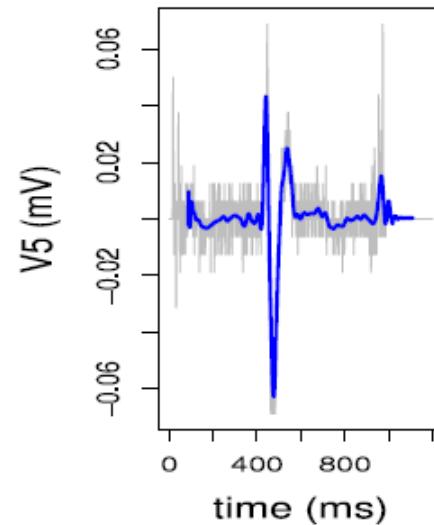
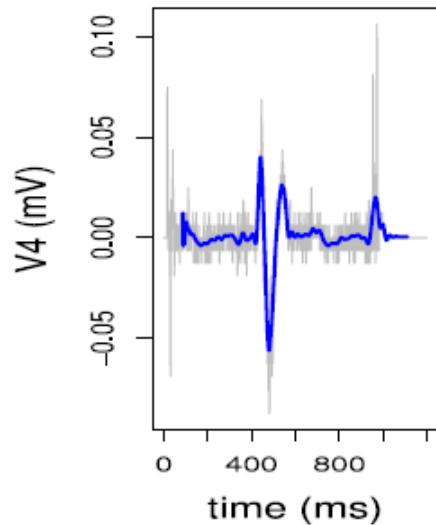
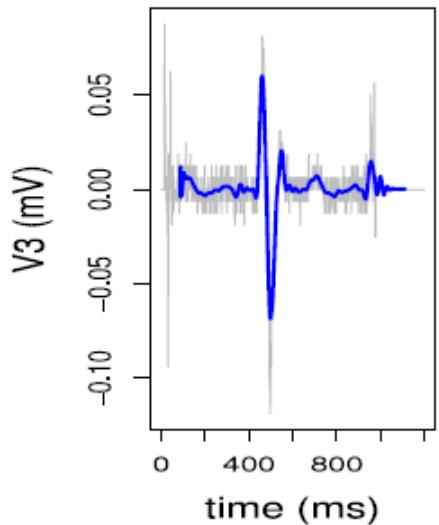
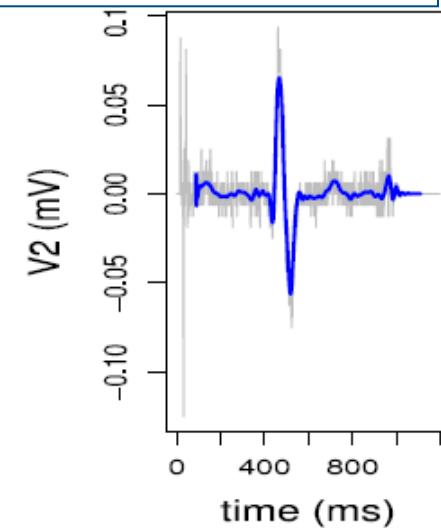
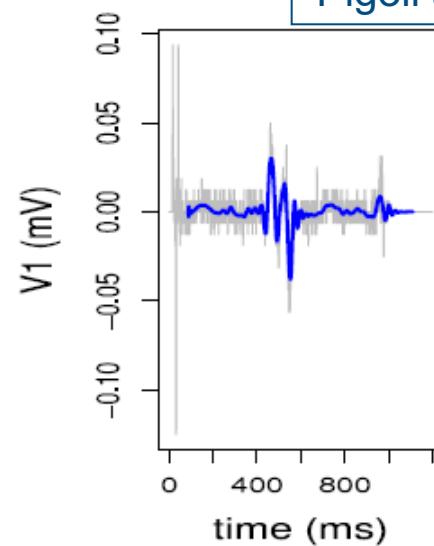
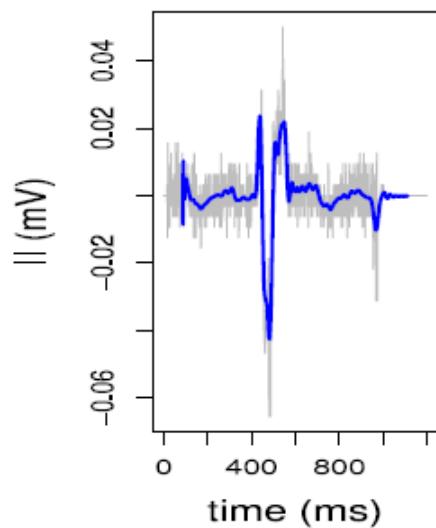
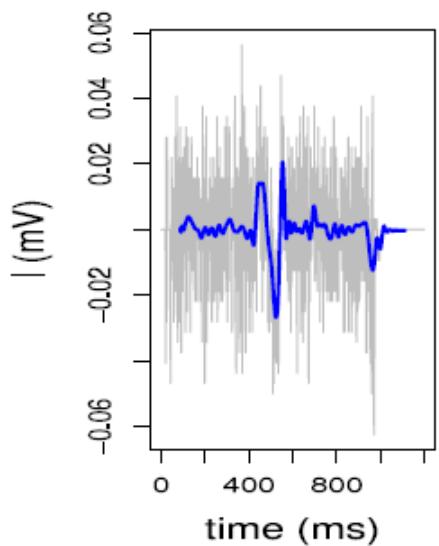
Pigoli and Sangalli 2012 CSDA



Derivatives of multidimensional wavelet estimates

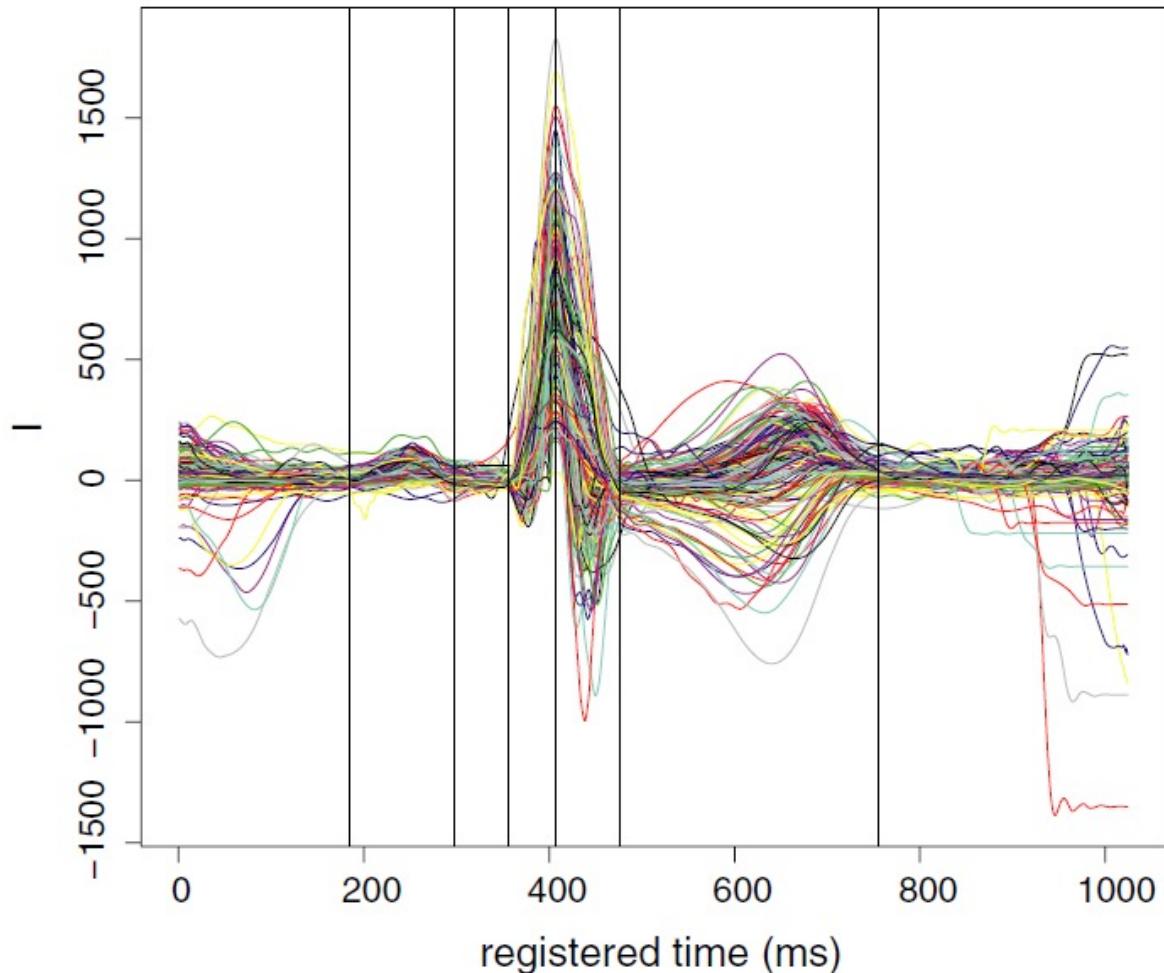
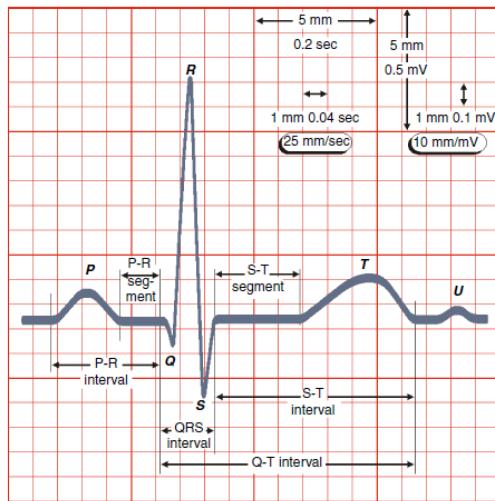
82

Pigoli and Sangalli 2012 CSDA





The 6 landmarks identify the *P*-wave (*P*, onset, *P*, offset), the *QRS*-complex (*QRS*, onset, *QRS*, offset), the *T*-wave (*T*, offset), the *R*-peak identified on lead I (*I*, peak).



Semi-automatic diagnostic procedure, based on the ECG morphology, that is able to classify physiological and pathological traces

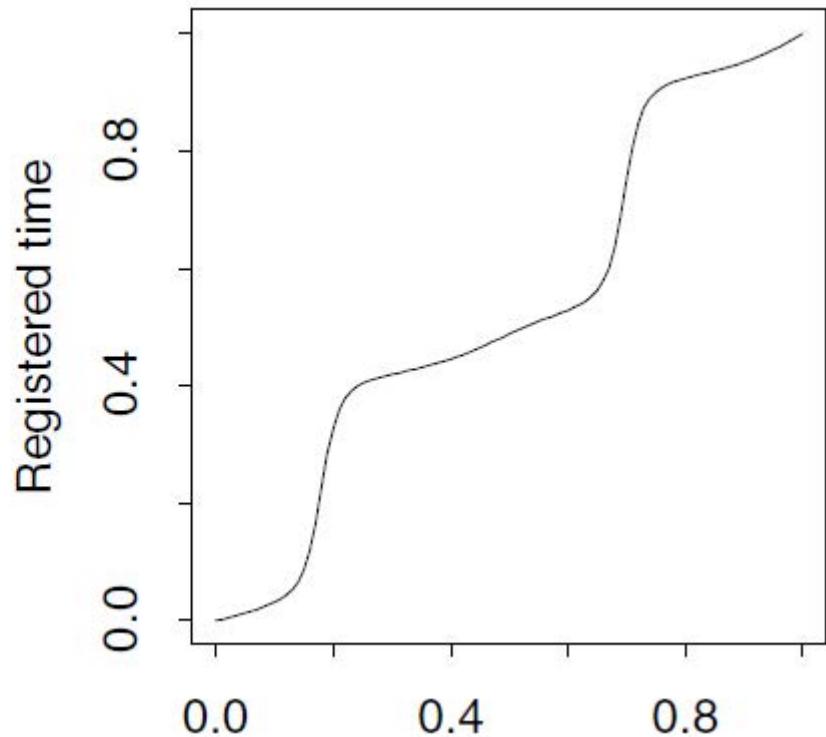
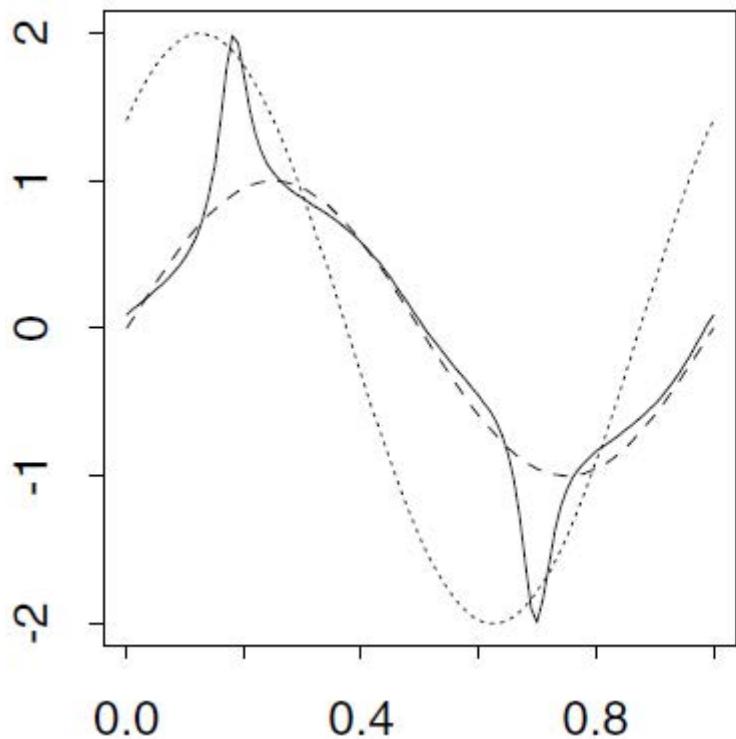
- Landmark-based registration may require significant user input and can be sensitive to the accuracy of the landmark identification.
- In some applications it is not possible to identify well-defined features that can be taken as landmarks

Alternative strategy: Continuous registration

Main idea:

- definition of a suitable distance (or closeness) measure between curves, which measures dissimilarity (or similarity) between curves.
- the curves are thus aligned by warping their time or space abscissa parameters choosing the optimal warping function in some class of admissible warping functions in order to minimize the final distance among the curves or, equivalently, maximize their final similarity.





The problem of decoupling amplitude and phase variability is not univocally defined as different measures of distance or similarity between curves can be considered, as well as different classes of admissible warping functions (e.g., simple translations or dilations, increasing linear transformations or more complex increasing transformations), leading to different registration results.

The choice of the couple formed by dissimilarity/similarity measure and admissible warping functions defines the distinction between phase variability and amplitude variability in the specific problem under analysis.

This choice must thus be problem specific.



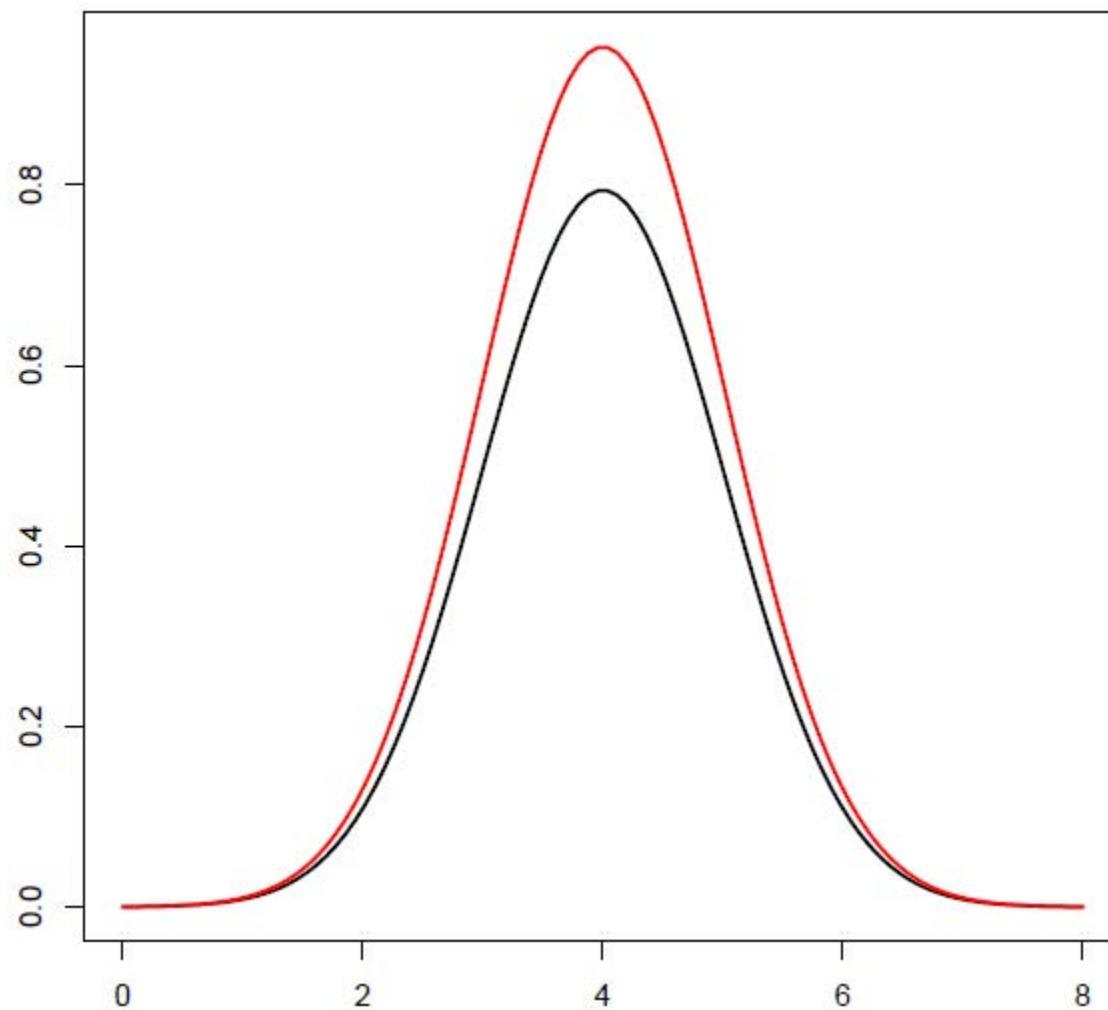
(ρ, W) must satisfy properties that ensure that the aligning problem is well-posed and the corresponding procedure is coherent

- ▶ ρ | Bounded
Reflexive
Symmetric
Transitive
- ▶ W Convex vector space
Group structure with respect to function composition
- ▶ (ρ, W) Properties of coherence
 - $\rho(\mathbf{c}_1, \mathbf{c}_2) = \rho(\mathbf{c}_1 \circ h, \mathbf{c}_2 \circ h), \quad \forall h \in W$ W -invariance of the index
(Isometry of the group, parallel orbits)
- $\rho(\mathbf{c}_1 \circ h_1, \mathbf{c}_2 \circ h_2) = \rho(\mathbf{c}_1 \circ h_1 \circ h_2^{-1}, \mathbf{c}_2) = \rho(\mathbf{c}_1, \mathbf{c}_2 \circ h_2 \circ h_1^{-1})$

(ρ, W) defines on the considered set of functions \mathcal{C} a partition in equivalence classes

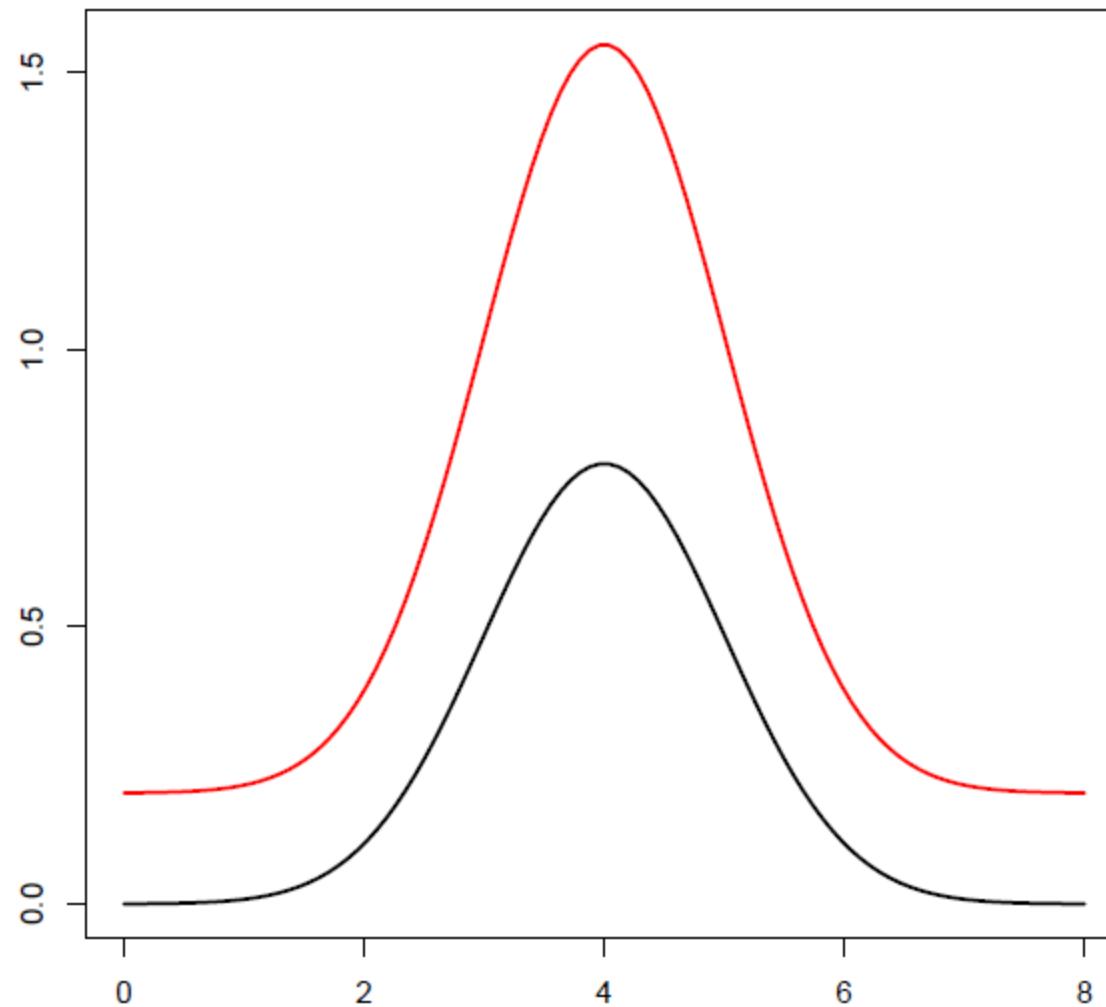
dissimilarity d	warpings W
$\ c_1 - c_2\ $	W_{shift}
$\ c'_1 - c'_2\ $	W_{shift}
$\ (c_1 - \bar{c}_1) - (c_2 - \bar{c}_2)\ $	W_{shift}
$\ (c'_1 - \bar{c}'_1) - (c'_2 - \bar{c}'_2)\ $	W_{shift}
$\left\ \frac{c_1}{\ c_1\ } - \frac{c_2}{\ c_2\ } \right\ $	$W_{affinity}$
$\left\ \frac{c'_1}{\ c'_1\ } - \frac{c'_2}{\ c'_2\ } \right\ $	$W_{affinity}$
$\left\ \text{sign}(c'_1)\sqrt{ c'_1 } - \text{sign}(c'_2)\sqrt{ c'_2 } \right\ $	$W_{diffeomorphism}$





$$\frac{\langle c_0, c_1 \rangle}{\|c_0\| \|c_1\|} = 1$$

$$\left\| \frac{c_0}{\|c_0\|} - \frac{c_1}{\|c_1\|} \right\| = 0$$



$$\frac{\langle c'_0, c'_1 \rangle}{\|c'_0\| \|c'_1\|} = 1$$

$$\left\| \frac{c'_0}{\|c'_0\|} - \frac{c'_1}{\|c'_1\|} \right\| = 0$$

If a template (prototype) curve φ is known, then it is enough to align each cuve to this template

If the template is unknown then it must be estimated from the data, leading to a complex optimization problem

find $\varphi \in \mathcal{C}$ and $\underline{h} = \{h_1, \dots, h_N\} \subset W$ such that

$$\frac{1}{N} \sum_{i=1}^N \rho(\varphi, \mathbf{c}_i \circ h_i) \geq \frac{1}{N} \sum_{i=1}^N \rho(\psi, \mathbf{c}_i \circ g_i)$$

for any other $\psi \in \mathcal{C}$ and $\underline{g} = \{g_1, \dots, g_N\} \subset W$

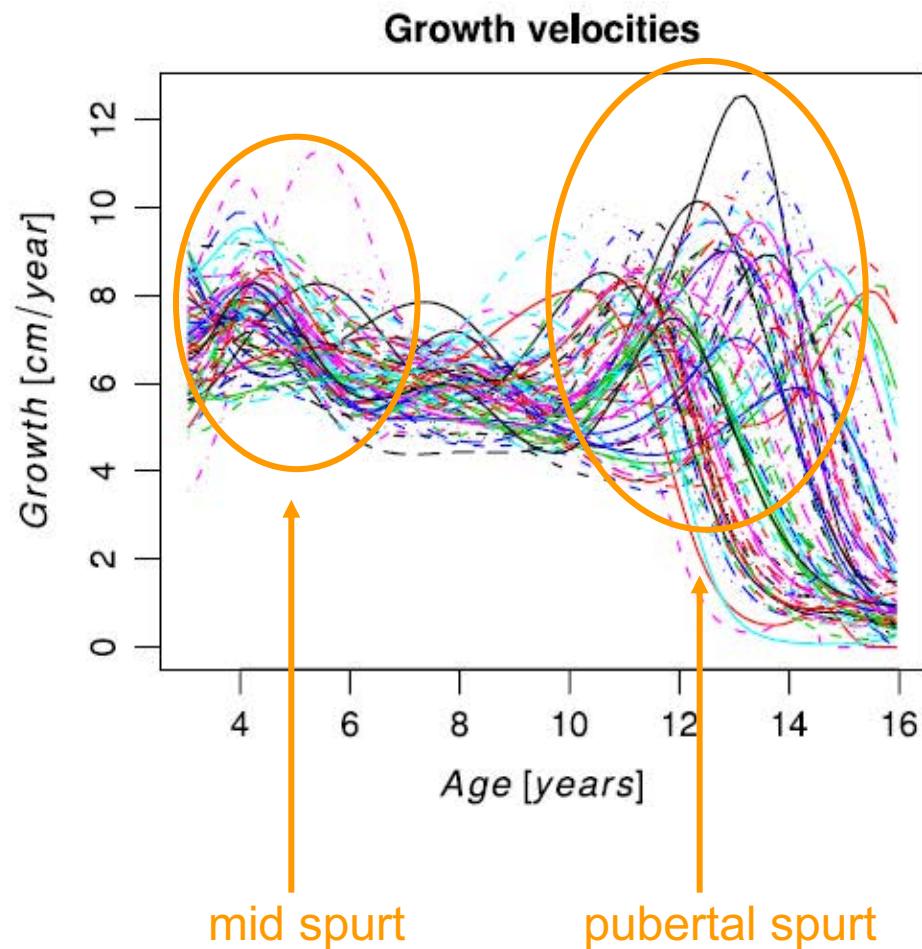
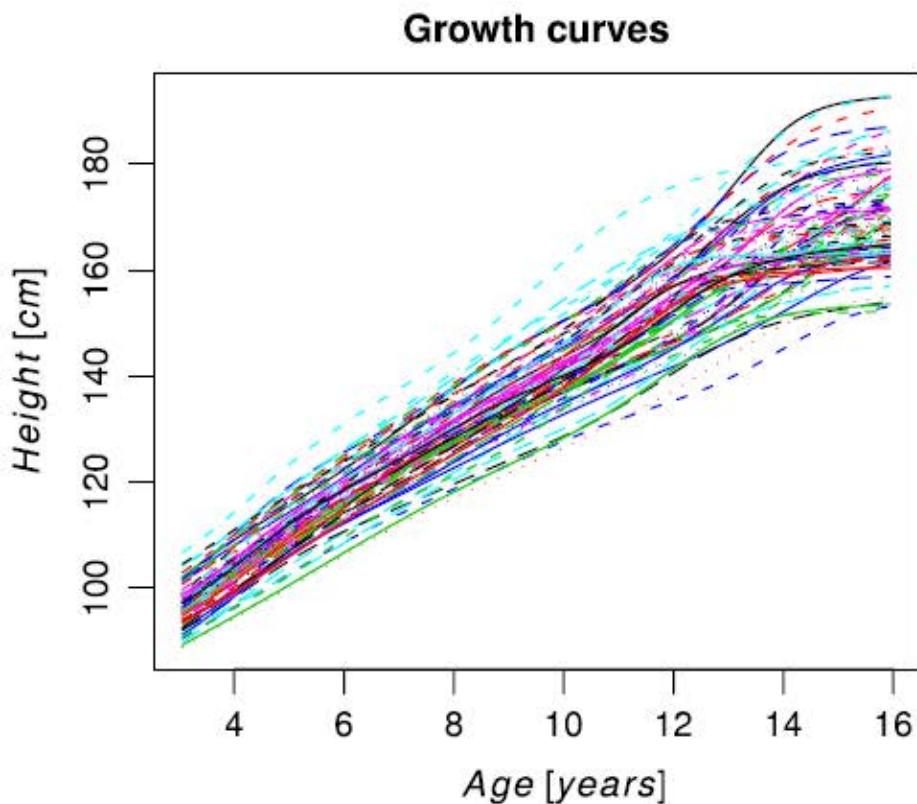
→ Iterative Procrustes procedure that alternates

- *template estimation step*: the template curve is estimated from the curves obtained in the previous alignment step
- *alignment step*: the curves are aligned to the template centerline estimated in the previous template estimation step

It should be noticed that both the phase variation and/or the amplitude variation may be associated to the phenomenon (for instance, the pathology) under study.

It is thus necessary to study both types of variations to see how they relate to the problem being investigated.

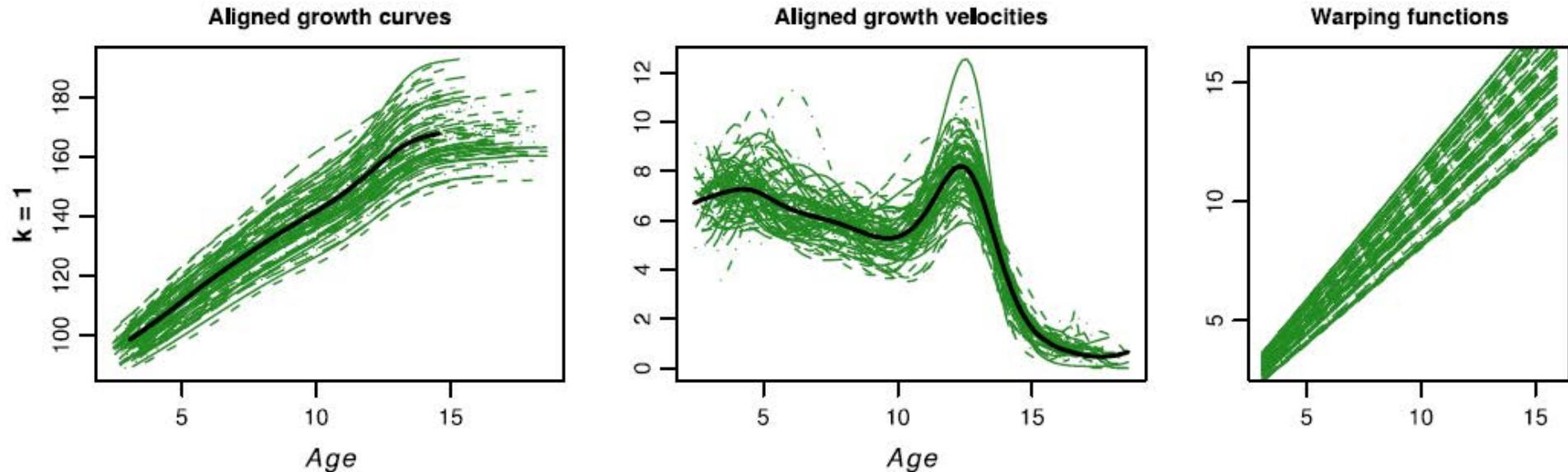




93 children, 39 boys and 54 girls

Curves estimated by monotonic cubic regression splines, implemented using the R package *fda*

Does the analysis point out some differences in the growth of boys and girls?

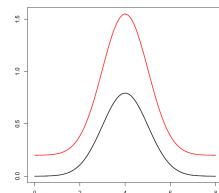


Results of continuous alignment using the following similarity index and class of warping functions (ρ, W) :

$$\rho(c_i, c_j) = \frac{\int_{S_{ij}} c'_i(s)c'_j(s)ds}{\sqrt{\int_{S_{ij}} c'_i(s)^2 ds} \sqrt{\int_{S_{ij}} c'_j(s)^2 ds}}$$

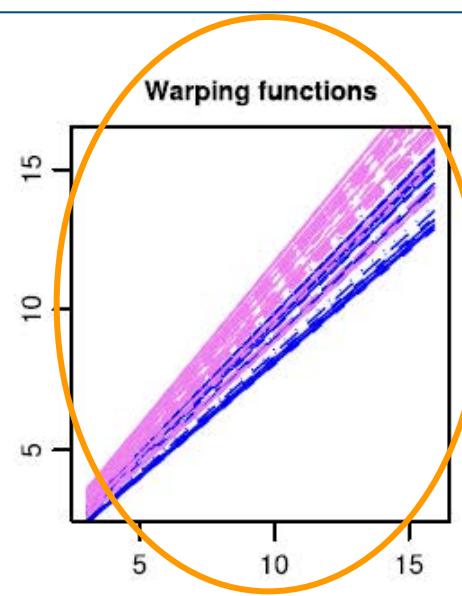
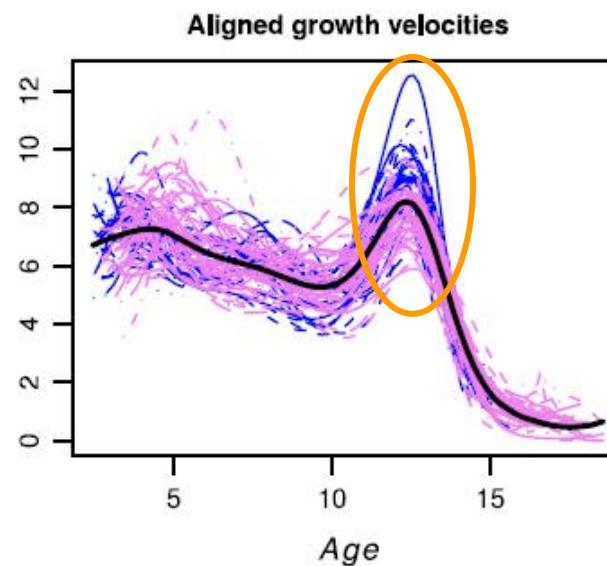
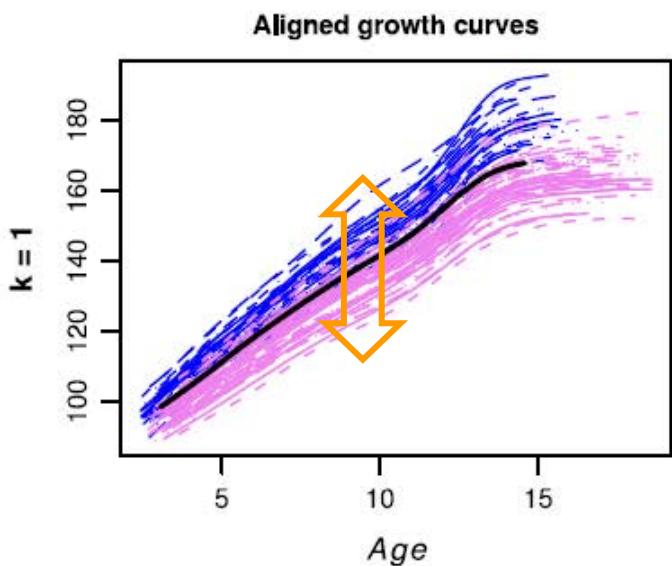
$$\rho(c_i, c_j) = 1 \Leftrightarrow \exists a \in \mathbb{R}, b \in \mathbb{R}^+ : c_i(t) = a + b c_j(t)$$

the focus is on growth patterns, rather than on the absolute heights of the children or on their more or less pronounced growths



$$W = \{h : h(t) = mt + q \text{ with } m \in \mathbb{R}^+, q \in \mathbb{R}\}$$

constant modifications of the running speeds of the children biological clocks



Once the biological clocks are aligned
the height of boys
stochastically dominates the
one of girls for any registered
biological age

boys have a more
pronounced growth
during puberty (more
prominent growth
velocity peak)

Neat separation
of boys and girls
in the phase.
The biological
clocks of boys
and girls run at
different speeds



POLITECNICO DI MILANO



Politecnico di Milano
Applied Statistics
May 2023



An introduction to functional data analysis

Laura M. SANGALLI

MOX - Dipartimento di Matematica, Politecnico di Milano

Part 5 – Alignment & Clustering

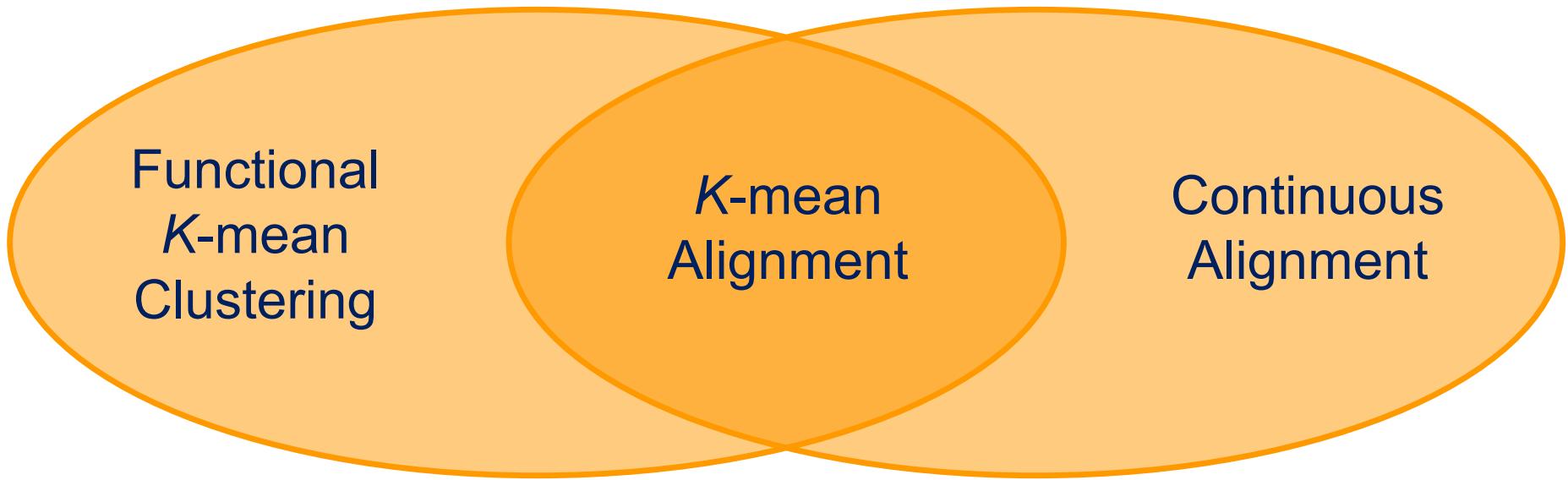


Curve clustering (functional k-mean clustering of curves)

- Heckman, N.E., Zamar, R.H. (2000), "Comparing the shapes of regression functions", Biometrika 87, 135-144.
- Tarpey, T., Kinateder, K.K.J. (2003), "Clustering functional data", J. Classification 20, 93-114.
- Shimizu, N., Mizuta, M. (2007), "Functional clustering and functional principal points", In: Lecture Notes in Artificial Intelligence, vol. 4693. Springer-Verlag, Berlin Heidelberg, 501-508.
- Cuesta-Albertos, J.A., Fraiman, R. (2007), "Impartial trimmed k-means for functional data". Comput. Statist. Data Anal. 51, 4864-4877.

Simultaneous clustering and alignment of curves

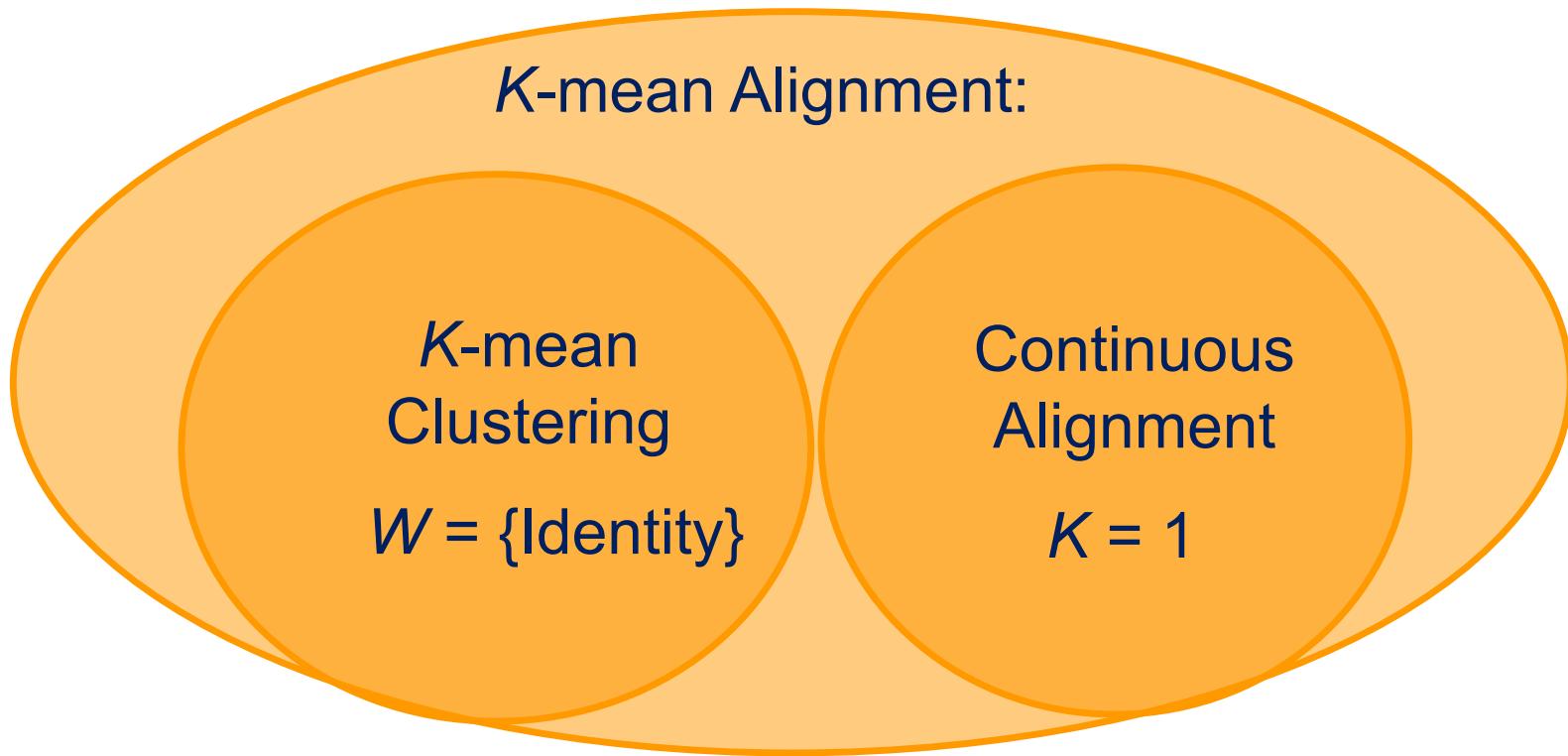
- Tang, R., Müller, H.-G. (2009), "Time-synchronized clustering of gene expression trajectories". Biostatistics 10, 32-45.
- Boudaoud, S., Rix, H., Meste, O. (2010), "Core Shape modelling of a set of curves". Comput. Statist. Data Anal. 54, 308-325.
- Liu, X., Yang, M.C.K. (2009), "Simultaneous curve registration and clustering for functional data". Comput. Statist. Data Anal. 53, 1361-1376.
- Sangalli, L.M., Secchi, P., Vantini, S., Vitelli, V. (2010), "K-means alignment for curve clustering", Comput. Statist. Data Anal., 54, 1219-1233.
- Sangalli, L.M., Secchi, P., Vantini, S. (2014), "Analysis of AneuRisk65 data: K-mean Alignment", Electronic Journal of Statistics, 8 (2), 1891-1904.



→ K-mean Clustering
with warping allowed

→ Continuous Alignment
with K templates

Code for K-mean alignment: R package fdakma, available from CRAN



Code for K-mean alignment: R package fdakma, available from CRAN

Goal of **Alignment**:
Decoupling Phase and Amplitude Variability



Goal of **K-mean** Clustering:
Decoupling Within and Between-cluster (Amplitude) Variability



Goal of **K-mean Alignment**:
**Identifying Phase Variability, Within-cluster Amplitude Variability
and Between-cluster Amplitude Variability**

(disclosing clustering in the phase)

Aligning and clustering a set of N curves

$$\{\mathbf{c}_1, \dots, \mathbf{c}_N\}$$

with respect to k template curves

$$\underline{\varphi} = \{\varphi_1, \dots, \varphi_k\}$$

Domain of attraction of φ_j

$$\Delta_j(\underline{\varphi}) = \{\mathbf{c} \in \mathcal{C} : \sup_{h \in W} \rho(\varphi_j, \mathbf{c} \circ h) \geq \sup_{h \in W} \rho(\varphi_r, \mathbf{c} \circ h), \forall r \neq j\}, \quad j = 1, \dots, k$$

Labelling function

$\lambda(\underline{\varphi}, \mathbf{c})$: indicates a cluster the curve \mathbf{c} should be assigned to

$\lambda(\underline{\varphi}, \mathbf{c}) = j$: the similarity index obtained by aligning \mathbf{c} to φ_j is at least as large as the similarity index obtained by aligning \mathbf{c} to any other template φ_r , with $r \neq j$

$\varphi_{\lambda(\underline{\varphi}, \mathbf{c})}$: indicates a template the curve \mathbf{c} can be best aligned to

Curve clustering when curves are misaligned

Trivial case: $\underline{\varphi} = \{\varphi_1, \dots, \varphi_k\}$ known

In order to cluster and align the set of N curves $\{\mathbf{c}_1, \dots, \mathbf{c}_N\}$ with respect to $\underline{\varphi}$:

for $i = 1, \dots, N$

- assign \mathbf{c}_i to the cluster $\lambda(\underline{\varphi}, \mathbf{c}_i)$
- align it to the corresponding template $\varphi_{\lambda(\underline{\varphi}, \mathbf{c}_i)}$

Non-trivial case: $\underline{\varphi} = \{\varphi_1, \dots, \varphi_k\}$ unknown

need to be themselves estimated from the data, leading to a complex optimization problem

Given $\{\mathbf{c}_1, \dots, \mathbf{c}_N\} \subset \mathcal{C}$, find $\underline{\varphi} = \{\varphi_1, \dots, \varphi_k\} \subset \mathcal{C}$,
 $\{\lambda_1, \dots, \lambda_N\} \subset \{1, \dots, k\}$ and $\underline{h} = \{h_1, \dots, h_N\} \subset W$
 that maximise $\frac{1}{N} \sum_{i=1}^N \rho(\varphi_{\lambda_i}, \mathbf{c}_i \circ h_i)$

An approximate solution to
 this optimization problem is
 given by the following
 iterative procedure

$\underline{\varphi}^{[q-1]} = \{\varphi_1^{[q-1]}, \dots, \varphi_k^{[q-1]}\}$: set of templates after iteration $q-1$

$\{\mathbf{c}_1^{[q-1]}, \dots, \mathbf{c}_N^{[q-1]}\}$: N curves aligned and clustered to $\underline{\varphi}^{[q-1]}$

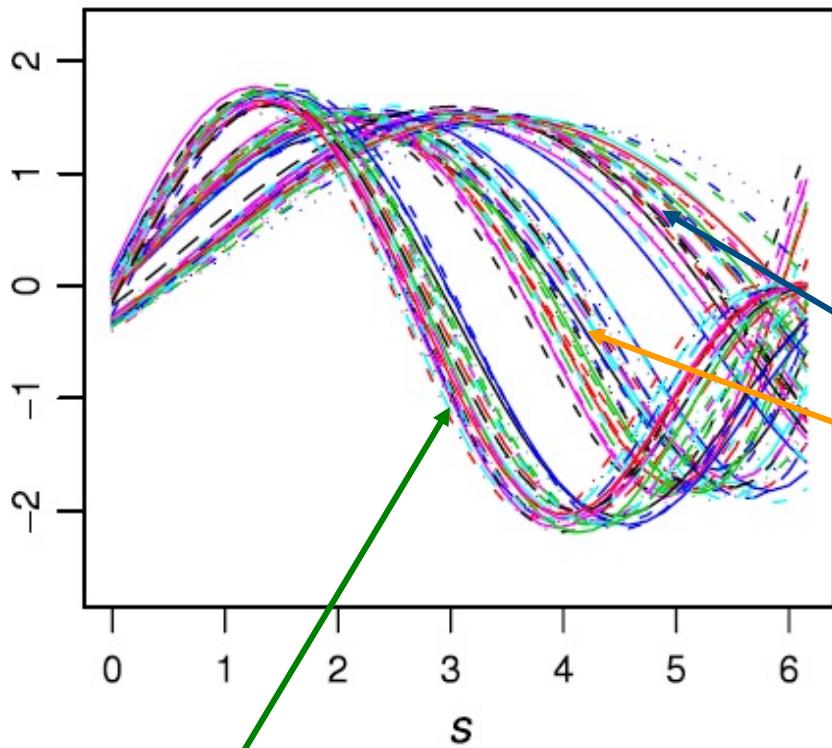
Template identification step. For $j = 1, \dots, k$, the template of the j th cluster $\varphi_j^{[q]}$ is estimated using all curves assigned to cluster j at iteration $q-1$.

$$\varphi_j^{[q]} = \arg \max_{\varphi \in \mathcal{C}} \sum_{i: \lambda_i=j} \rho(\varphi, \mathbf{c}_i^{[q-1]})$$

Assignment and alignment step. The set of curves $\{\mathbf{c}_1^{[q-1]}, \dots, \mathbf{c}_N^{[q-1]}\}$ is clustered and aligned to the set of templates $\underline{\varphi}^{[q]} = \{\varphi_1^{[q]}, \dots, \varphi_k^{[q]}\}$.

Normalization step. For $j = 1, \dots, k$, all curves assigned to cluster j are warped along a common warping function, so that the average warping undergone by curves assigned to the same cluster is the identity transformation (thus avoiding the drifting apart of clusters or the global drifting of the overall set of curves).

The algorithm is stopped when, in the assignment and alignment step, the increments of the similarity indexes are all lower than a fixed threshold.



2 AMONG THEM ARE CLUSTERS
probably have generated these data?

ONE has associated a
further CLUSTERING IN
THE PHASE

$$1 * \sin(s) + 1 * \sin\left(\frac{s^2}{2\pi}\right)$$

$$2 * \sin(s) - 1 * \sin\left(\frac{s^2}{2\pi}\right) + (1 + \varepsilon_{4i})s + (1 + \varepsilon_{2i}) * \sin\left(\frac{(\varepsilon_{3i} + (1 + \varepsilon_{4i})s)^2}{2\pi}\right)$$

$$\rho(c_i, c_j) = \frac{\int_{S_{ij}} c'_i(s)c'_j(s)ds}{\sqrt{\int_{S_{ij}} c'_i(s)^2 ds} \sqrt{\int_{S_{ij}} c'_j(s)^2 ds}}$$

$$\rho(c_i, c_j) = 1 \Leftrightarrow \exists a \in \mathbb{R}, b \in \mathbb{R}^+ : c_i(t) = a + b c_j(t)$$

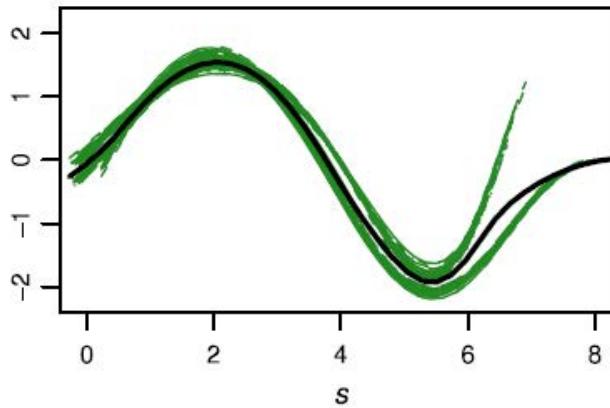
$$W = \{h : h(t) = mt + q \text{ with } m \in \mathbb{R}^+, q \in \mathbb{R}\}$$



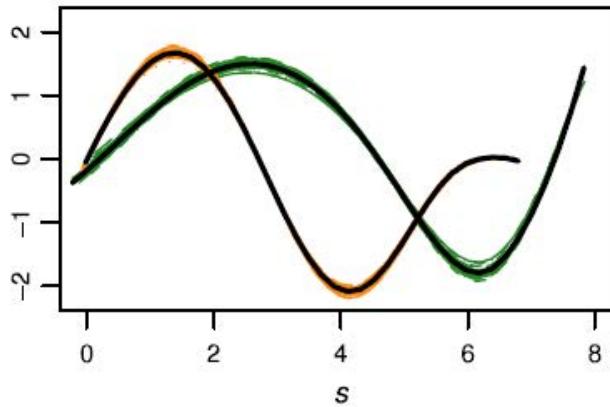


Aligned and clustered curves

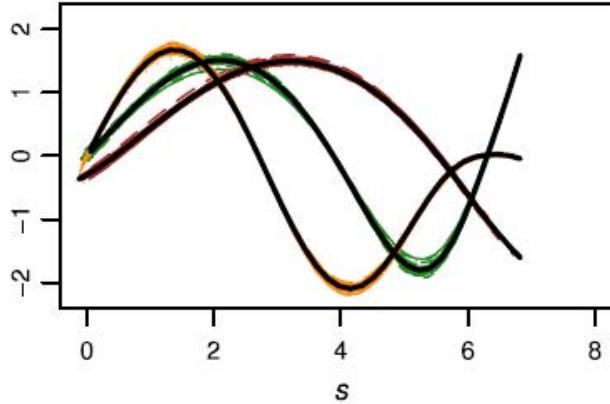
$K = 1$



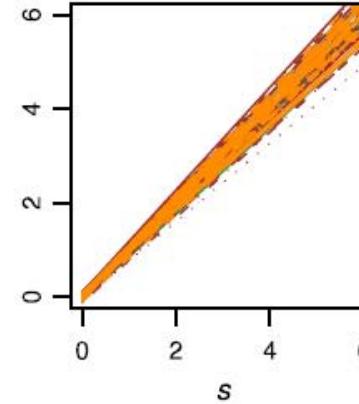
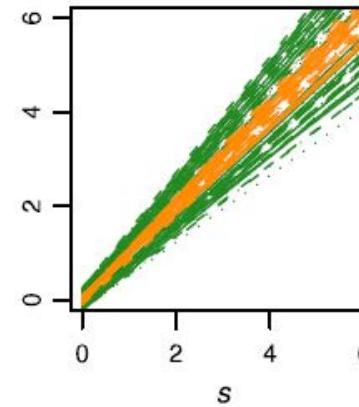
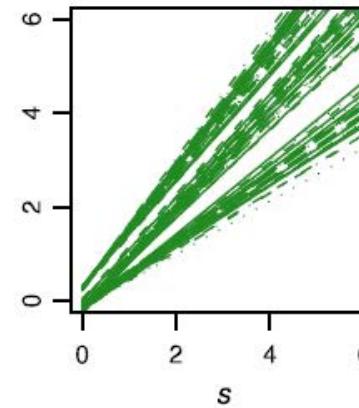
$K = 2$

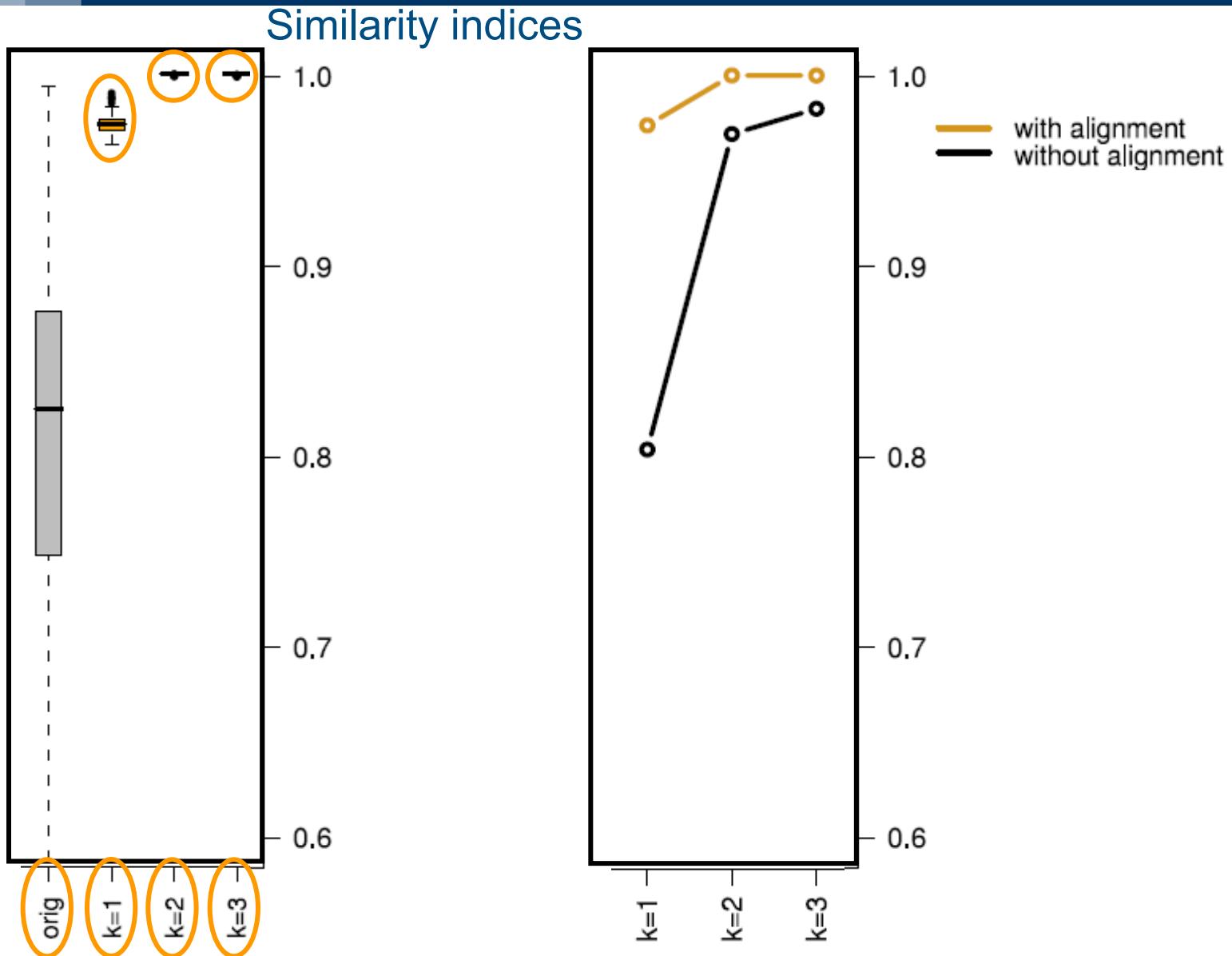


$K = 3$

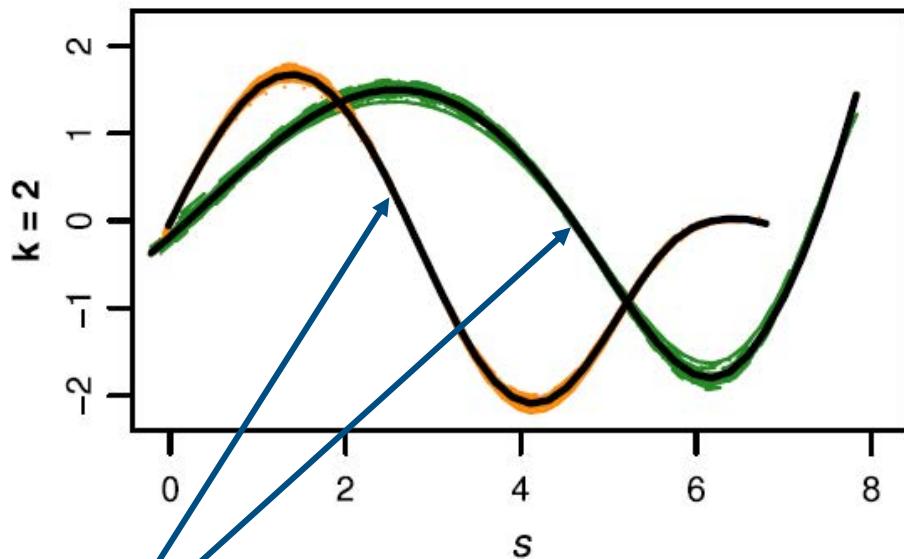


Warping functions¹⁰⁶



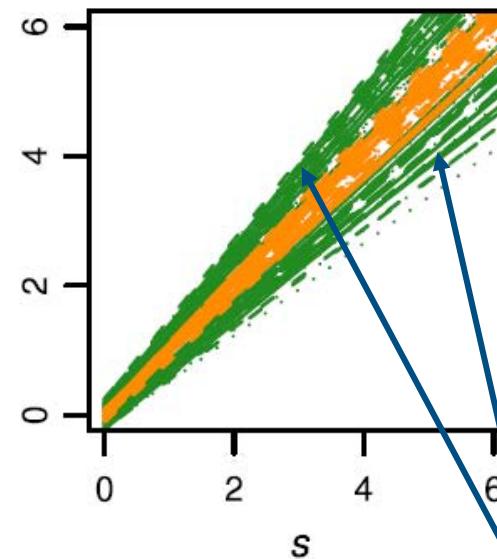


Aligned and clustered curves



k-mean alignment is able to efficiently detect true amplitude clusters and also to disclose clustering structures in the phase

Warping functions





POLITECNICO DI MILANO



Politecnico di Milano
Applied Statistics
May 2023



An introduction to functional data analysis

Laura M. SANGALLI

MOX - Dipartimento di Matematica, Politecnico di Milano

Part 6 – Case Study: AneuRisk



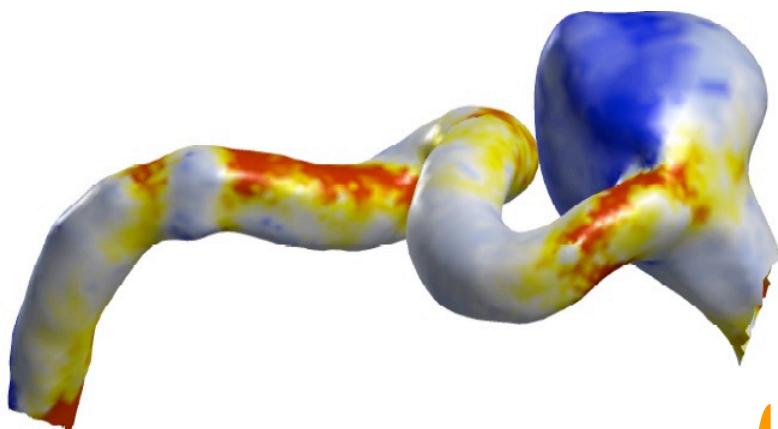
AneuRisk data

Available from: <https://statistics.mox.polimi.it/aneurisk/>

Main references:

- Sangalli, L.M., Secchi, P., Vantini, S. (2014a), "AneuRisk65: three-dimensional cerebral vascular geometries", *Electronic Journal of Statistics*, 8, 2, 1879–1890.
- Sangalli, L.M., Secchi, P., Vantini, S. (2014b), "Analysis of AneuRisk65 data: Kmean Alignment", *Electronic Journal of Statistics*, 8, 2, 1891–1904.
- Sangalli, L.M., Secchi, P., Vantini, S., Vitelli, V. (2010), "K-mean alignment for curve clustering", *Computational Statistics and Data Analysis*, 54, pp. 1219-1233.
- Sangalli, L.M., Secchi, P., Vantini, S., Veneziani, A. (2009), "Efficient estimation of 3-dimensional curves and their derivatives by free-knot regression splines, applied to the analysis of inner carotid artery centerlines", *Journal of the Royal Statistical Society Ser C*, 58, 285-306.
- Sangalli, L.M., Secchi, P., Vantini, S., Veneziani, A. (2009), "A Case Study in Exploratory Functional Data Analysis: Geometrical Features of the Internal Carotid Artery", *Journal of the American Statistical Association*, 104, 37-48.

Code for K-mean alignment: R package fdakma, available from CRAN



A CONJECTURE

The pathogenesis of cerebral aneurysms is conditioned by the geometry of the cerebral vessels through its effects on blood fluid dynamics



Laboratory of Biological Structure Mechanics

Department of Structural Engineering



Statistics

Numerical Analysis

Bio-Engineering

Computer Sciences

Neurosurgery

Neuroradiology

Numerical Analysis



Alessandro Veneziani

Principal Investigator



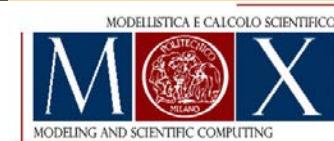
Now at



Tiziano Passerini



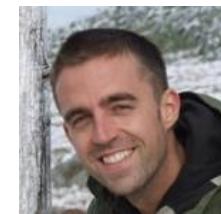
Marina Piccinelli



Statistics



Piercesare Secchi



Simone Vantini



Laura Sangalli



Valeria Vitelli

(now at University of Oslo)



Edoardo Boccardi

Medicine



<https://s...>

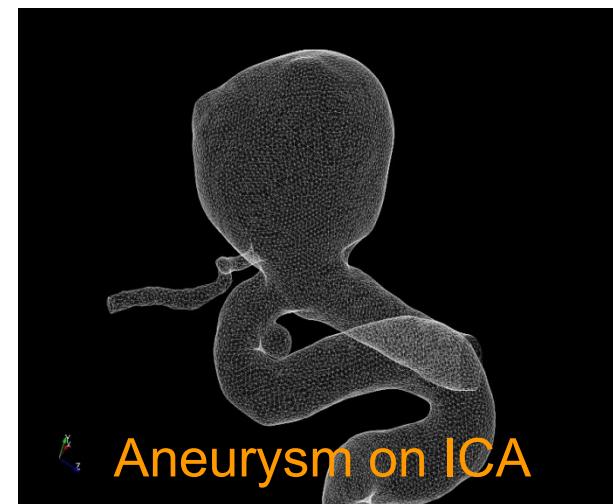
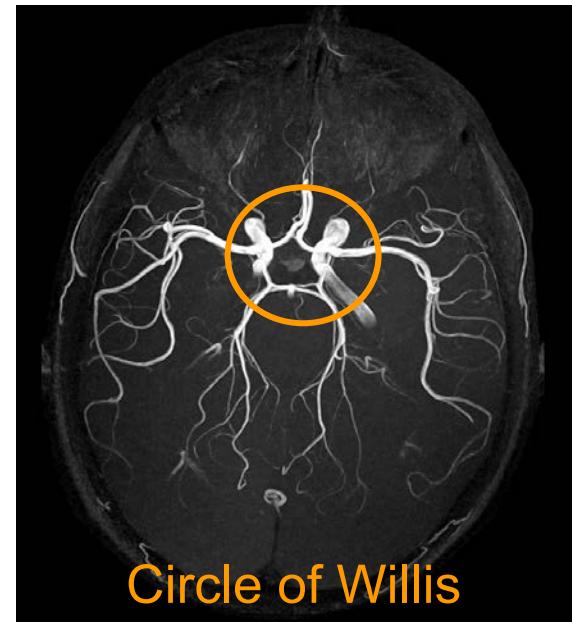
polimi.it

POLITECNICO DI MILANO

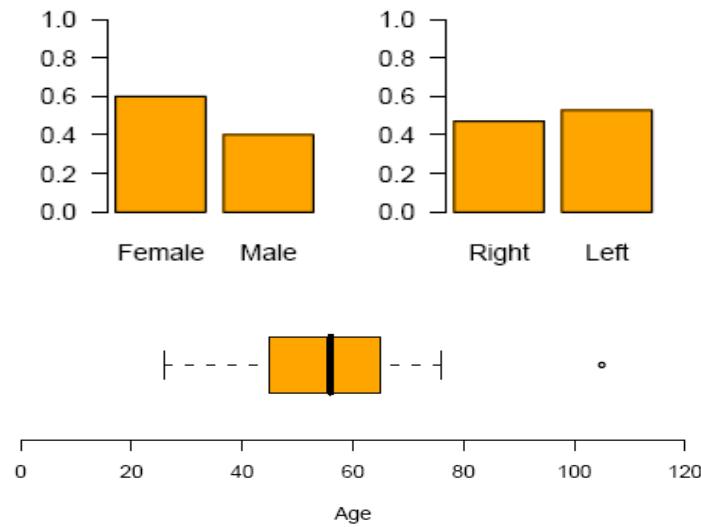
- Cerebral aneurysms: deformations of cerebral arteries, mostly placed on vessels belonging to or connected to the Circle of Willis

Aneurysms EPIDEMIOLOGY

- Incidence rate of cerebral aneurysms:
1/20 people
- Incidence rate of ruptured cerebral aneurysms per year:
1/10000 people per year
- Mortality due to a ruptured aneurysm:
> 50%: Out of 9 patients with a ruptured aneurysm:
 - 3 are expected to die before arriving at the hospital
 - 2 to die after having arrived at the hospital
 - 2 to survive with permanent cerebral damages
 - 2 to survive without permanent cerebral damages



Observational Study conducted at Ospedale Ca' Granda Niguarda – Milano relative to 65 patients hospitalized from September 2002 to October 2005.



Upper group	Lower group	
Aneurym at or after ICA biforc	Aneurysm before ICA biforc	No aneurysms
33	25	7



Data: 3D-angiographies

Observational Study conducted at Ospedale Ca' Granda Niguarda – Milano relative to 65 patients hospitalized from September 2002 to October 2005.

Sequence of X-Rays



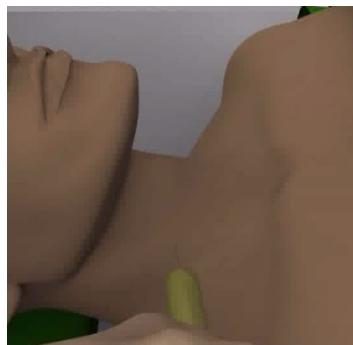
3D-array of
gray scaled pixels



From X-rays to Centerlines and Local Maximal Inscribed Sphere Radius

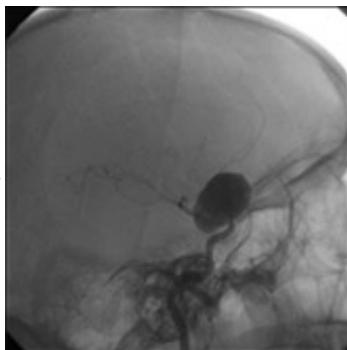
116

Contrast Fluid Injections



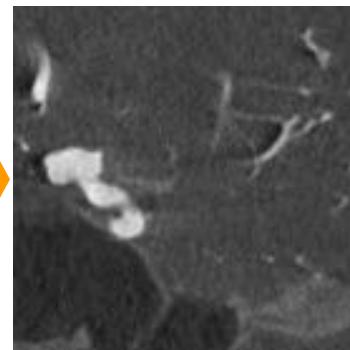
1

X-rays (one direction)



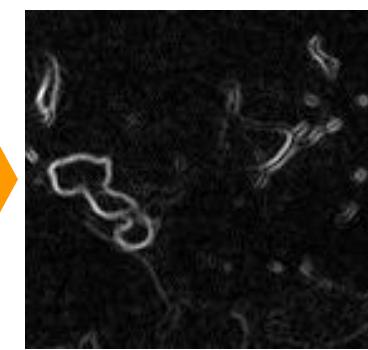
2

3d-array (one slice)



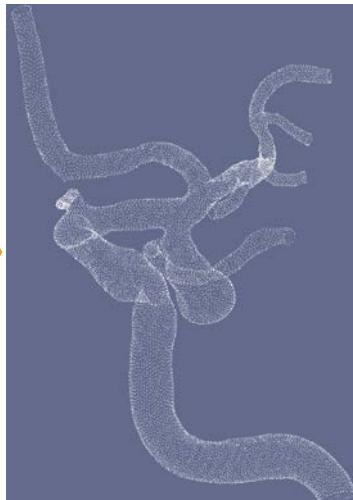
3

Gradient 3d-array (one slice)



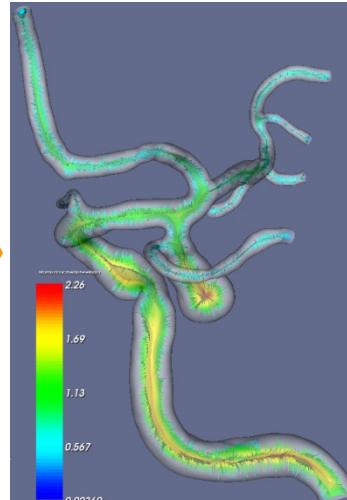
4

Surface Points



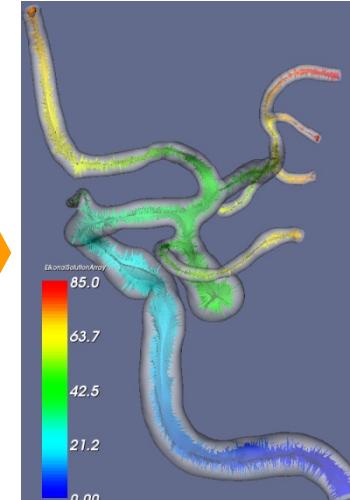
4

Voronoi Diagram



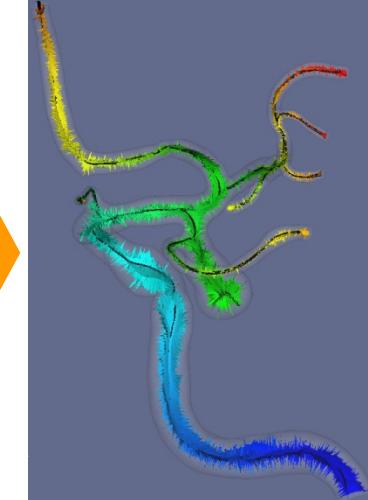
5

Eikonal Equation

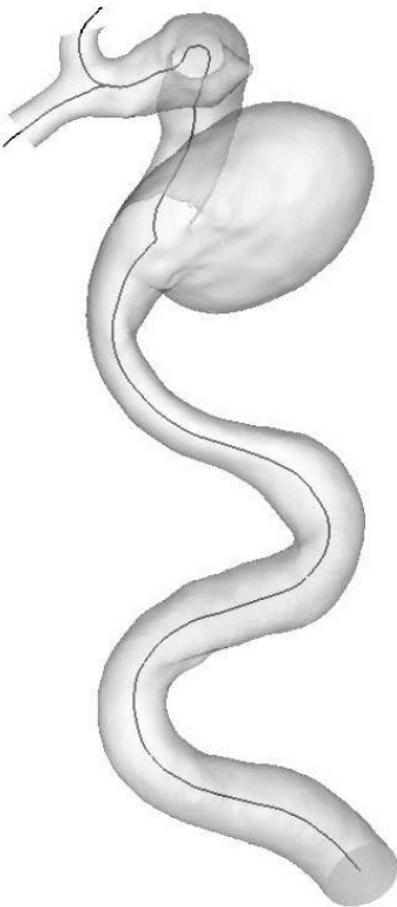


6

Centerline+MISR



7



Focus on Internal Carotid Artery (ICA)

For each patient i elicitation of 3-spatial coordinates of ICA centerline

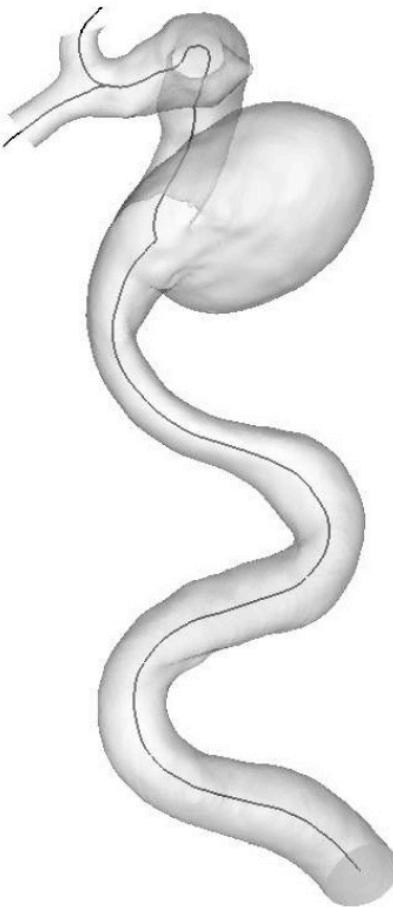
$$\{(x_{ij}, y_{ij}, z_{ij}) : j = 1, \dots, n_i\}$$

and vessel radius

$$\{R_{ij} : j = 1, \dots, n_i\}$$

alone a fine grid of points $(350 \leq n_i \leq 1380)$

Two geometric quantities that greatly influence the haemodynamics: vessel **radius** and **curvature** (curvature of its centerline)



Focus on Internal Carotid Artery (ICA)

For each patient i elicitation of 3-spatial coordinates of ICA centerline

$$\{(x_{ij}, y_{ij}, z_{ij}) : j = 1, \dots, n_i\}$$

and vessel radius

$$\{R_{ij} : j = 1, \dots, n_i\}$$

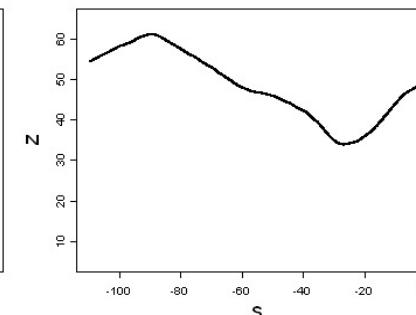
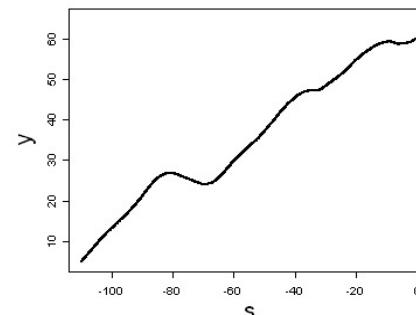
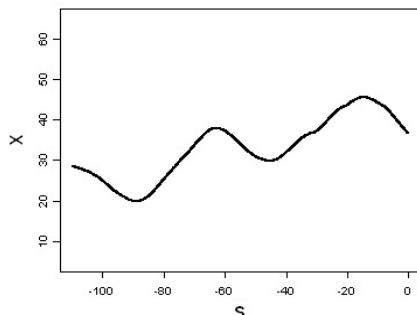
alone a fine grid of points $(350 \leq n_i \leq 1380)$

Approximate curvilinear abscissa: $\{s_{ij} : j = 1, \dots, n_i\}$ $s_{i1} = 0$

$$s_{ij} - s_{ij-1} = -\sqrt{(x_{ij} - x_{ij-1})^2 + (y_{ij} - y_{ij-1})^2 + (z_{ij} - z_{ij-1})^2}, \quad j = 2, \dots, n_i$$

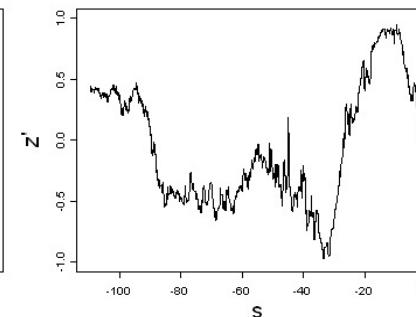
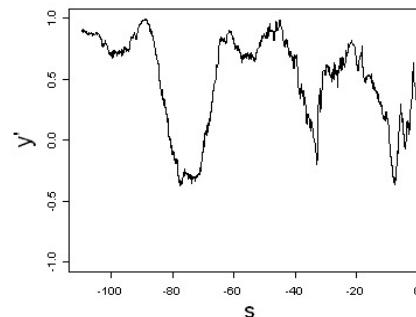
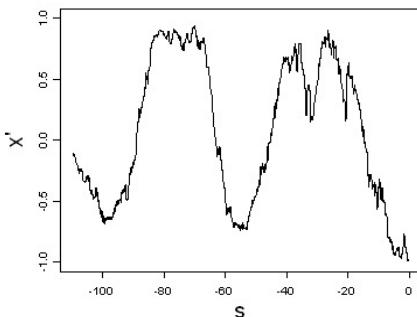


COORDINATES
PATIENT 1



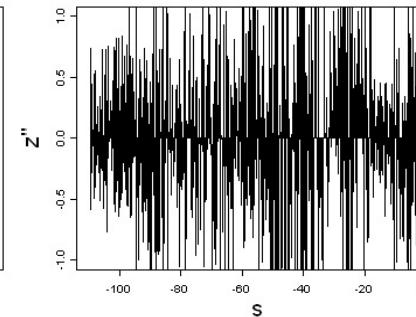
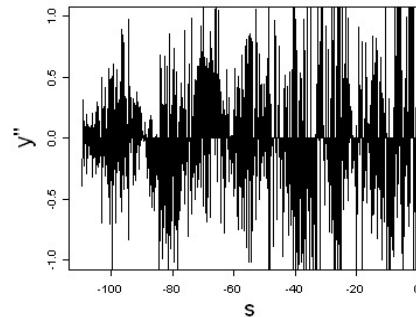
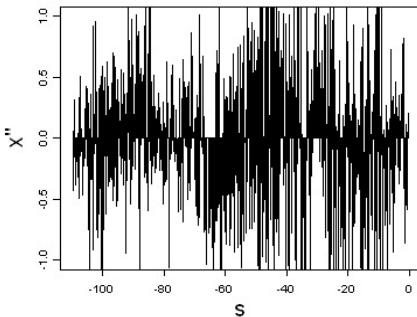
Very high signal-to-noise ratio
Fine grid of observed points

FIRST
DIFFERENCES



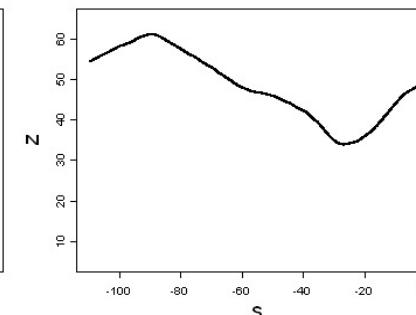
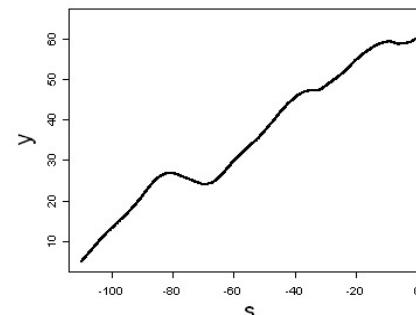
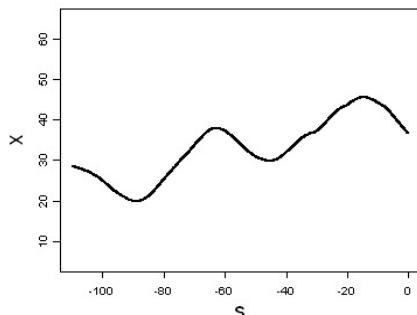
Rough estimates of first and second derivatives by means of central differences

SECOND
DIFFERENCES



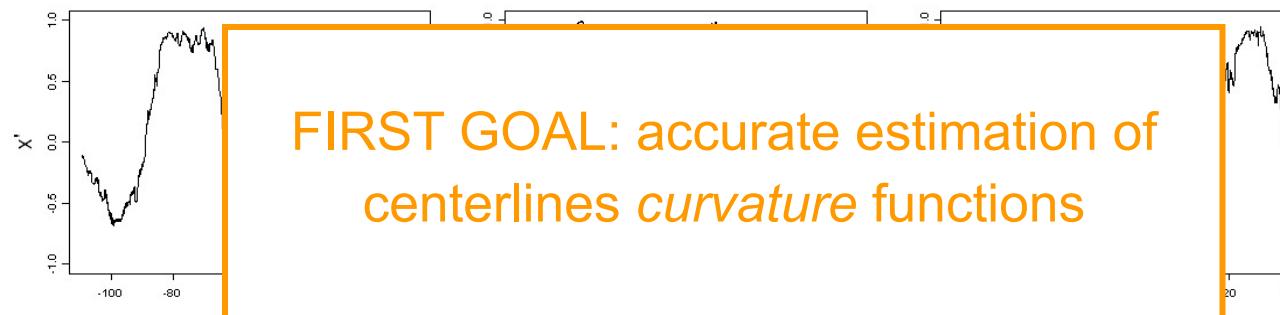
COORDINATES

PATIENT 1



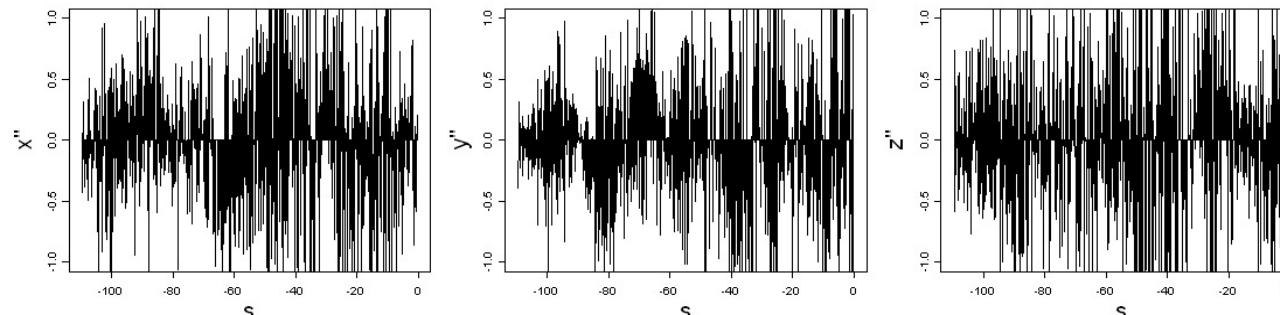
Very high signal-to-noise ratio
Fine grid of observed points

FIRST DIFFERENCES



Rough estimates of first and second derivatives by means of central differences

SECOND DIFFERENCES



b-spline basis system for the vector space

$\{b_{r,m}^{[k]}(s) : r = 1, \dots, m + n_k\}$ of splines of order m

with knot vector $\mathbf{k} = (k_1, \dots, k_{n_k})$

Functional estimates of the 3-spatial coordinates
by 3D free-knot regression splines

$$(\hat{x}(s), \hat{y}(s), \hat{z}(s))$$

$$\hat{x}(s) = \sum_{r=1}^{m+n_k} \hat{\lambda}_r^{[x]} b_{r,m}^{[\hat{k}]}(s) \quad \hat{y}(s) = \sum_{r=1}^{m+n_k} \hat{\lambda}_r^{[y]} b_{r,m}^{[\hat{k}]}(s) \quad \hat{z}(s) = \sum_{r=1}^{m+n_k} \hat{\lambda}_r^{[z]} b_{r,m}^{[\hat{k}]}(s)$$

FIND

$$\hat{n}_k, \hat{\mathbf{k}} = (\hat{k}_1(s), \dots, \hat{k}_{n_k}(s)), \hat{\lambda}^{[x]}, \hat{\lambda}^{[y]}, \hat{\lambda}^{[z]}$$

by minimizing

$$\sum_{j=1}^n \left(x_j - \sum_{r=1}^{m+n_k} \lambda_r^{[x]} b_{r,m}^{[\mathbf{k}]}(s_j) \right)^2 + \sum_{j=1}^n \left(y_j - \sum_{r=1}^{m+n_k} \lambda_r^{[y]} b_{r,m}^{[\mathbf{k}]}(s_j) \right)^2 + \sum_{j=1}^n \left(z_j - \sum_{r=1}^{m+n_k} \lambda_r^{[z]} b_{r,m}^{[\mathbf{k}]}(s_j) \right)^2 + \mathcal{C}(m+n_k)$$

FIX

$m=5$ to obtain smooth estimates of the curvature (function of second derivative)

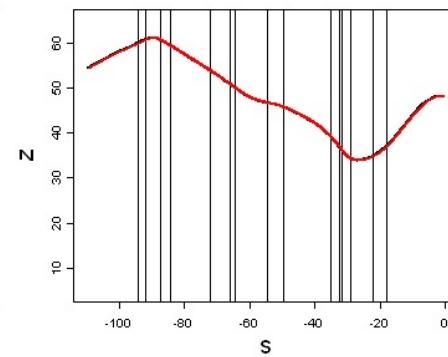
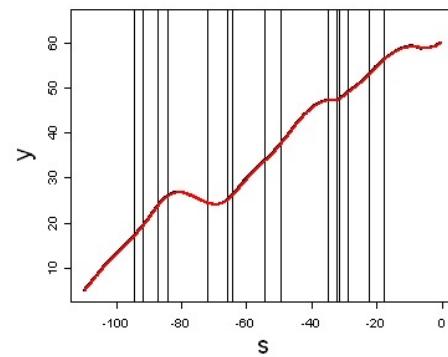
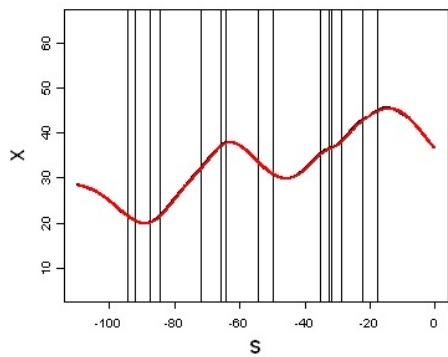


Accurate curve estimation by 3D free-knot splines

123

Sangalli Secchi Vantini Veneziani, 2009, JRSS-C

Curve estimate



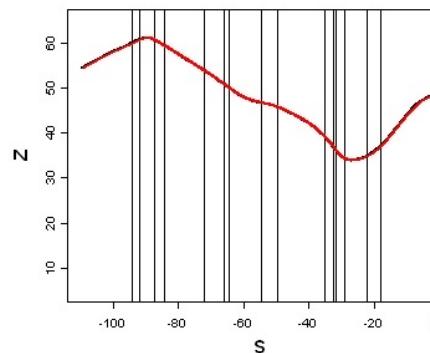
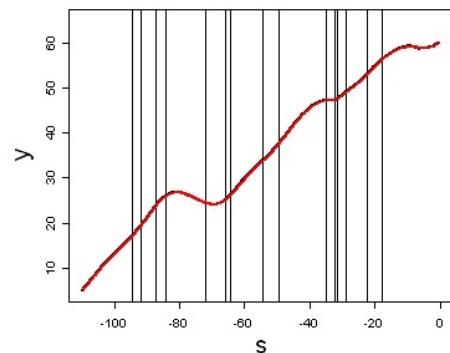
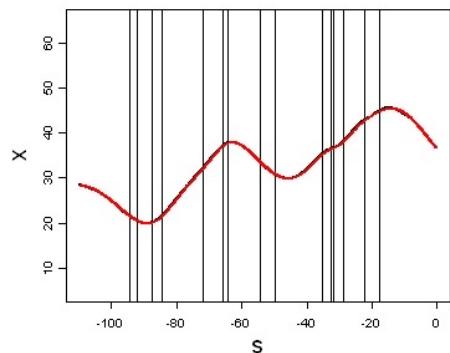
Derivatives of splines are still splines with the same knot vector and coefficients directly computed from the coefficients of the original spline

Accurate curve estimation by 3D free-knot splines

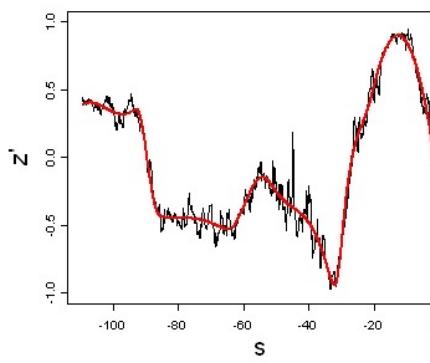
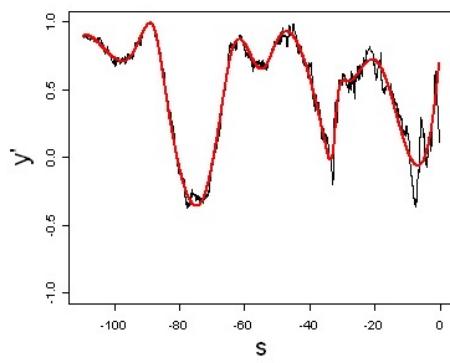
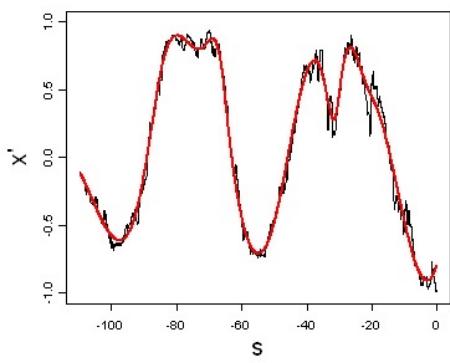
124

Sangalli Secchi Vantini Veneziani, 2009, JRSS-C

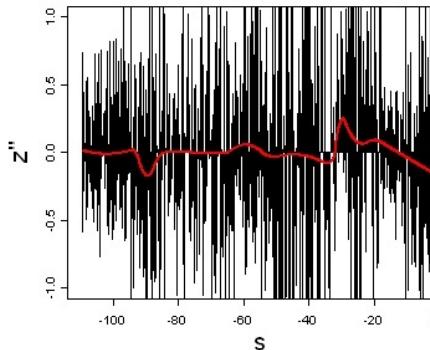
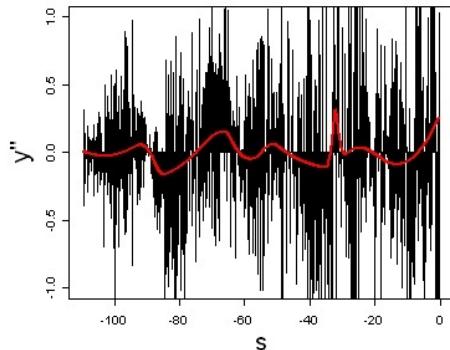
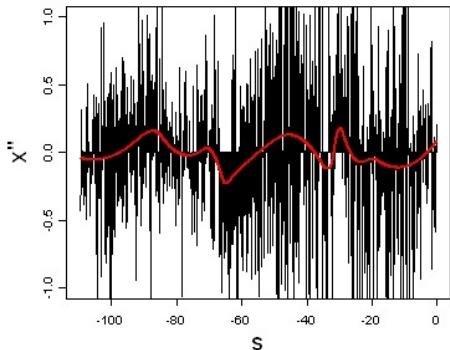
Curve estimate

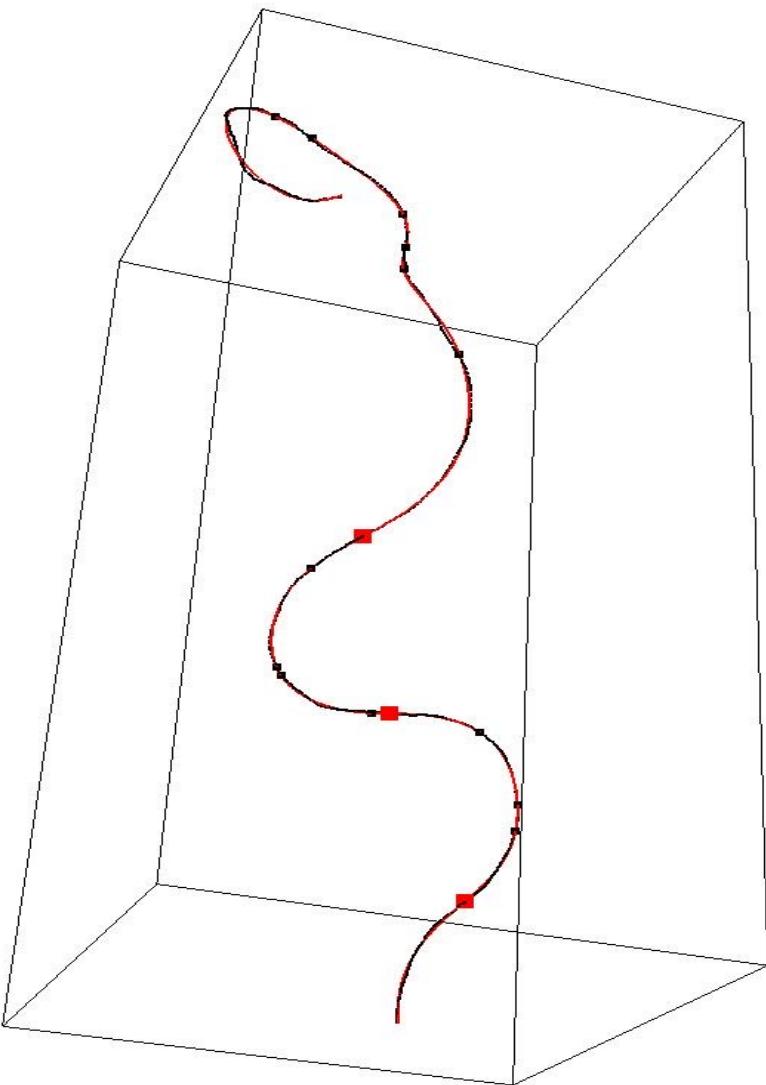


First deriv

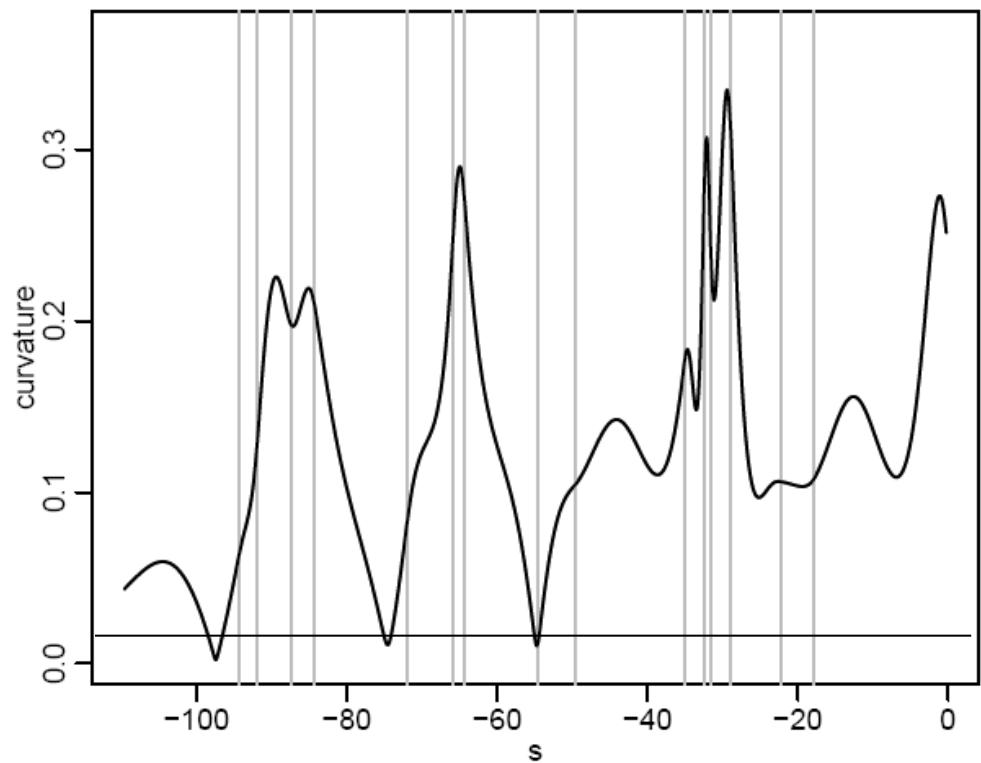


Second deriv.



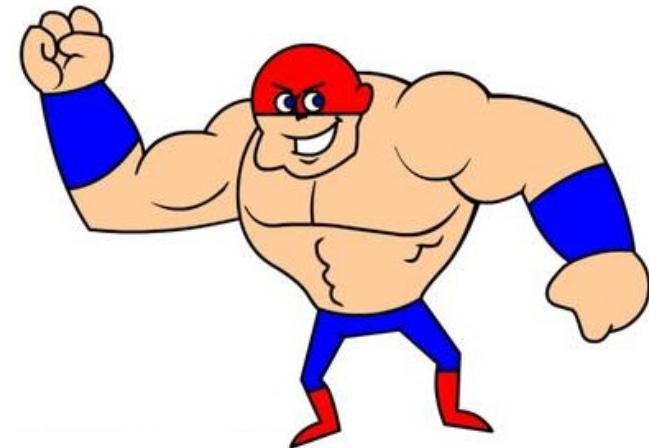
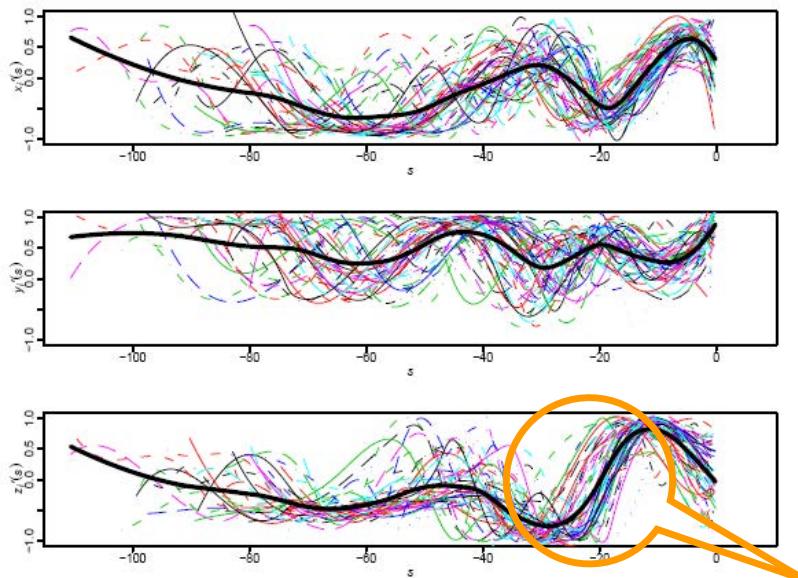


Curvature function



$$C_i(s) = \frac{\|(x'_i(s) \ y'_i(s) \ z'_i(s)) \times (x''_i(s) \ y''_i(s) \ z''_i(s))\|}{\|(x'_i(s) \ y'_i(s) \ z'_i(s))\|^3}$$

Centerline first derivatives



Visualise data

Phase Variability
(strongly dependent on dimensions of body structure and arteries)

- To enable meaningful comparisons across patients we need to decouple between-patients *phase variability* and between-patients *amplitude variability*

due to *differences in the dimensions* of patients carotids

due to *differences in the morphological shapes* of patients carotids

\mathcal{C} : set of curves $\mathbf{c}(s) : \mathbb{R} \rightarrow \mathbb{R}^d$

Similarity index

$$\rho(\mathbf{c}_1, \mathbf{c}_2) = \frac{1}{d} \sum_{p=1}^d \frac{\int c'_{1p}(s)c'_{2p}(s)ds}{\sqrt{\int c'_{1p}(s)^2 ds} \sqrt{\int c'_{2p}(s)^2 ds}}$$

c_{ip} : pth component of $\mathbf{c}_i = (c_{i1}, \dots, c_{id})$

Class W of warping functions

$$W = \{h : h(s) = ms + q \text{ with } m \in \mathbb{R}^+, q \in \mathbb{R}\}$$

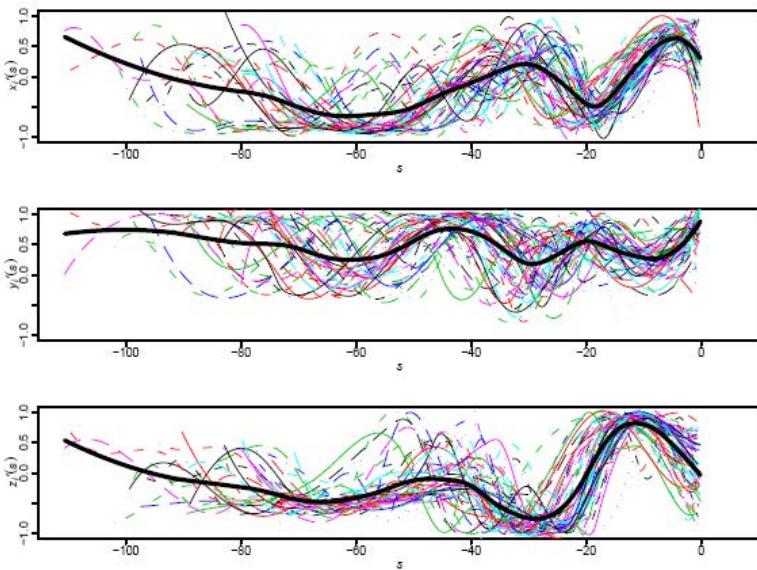
$$\rho(\mathbf{c}_1, \mathbf{c}_2) = 1 \Leftrightarrow \text{for } p = 1, \dots, d, \exists \theta_{0p} \in \mathbb{R}, \theta_{1p} \in \mathbb{R}^+ : \\ c_{1p}(s) = \theta_{0p} + \theta_{1p} c_{2p}(s).$$

Aneurysm location on aligned ICA radius and curvature profiles

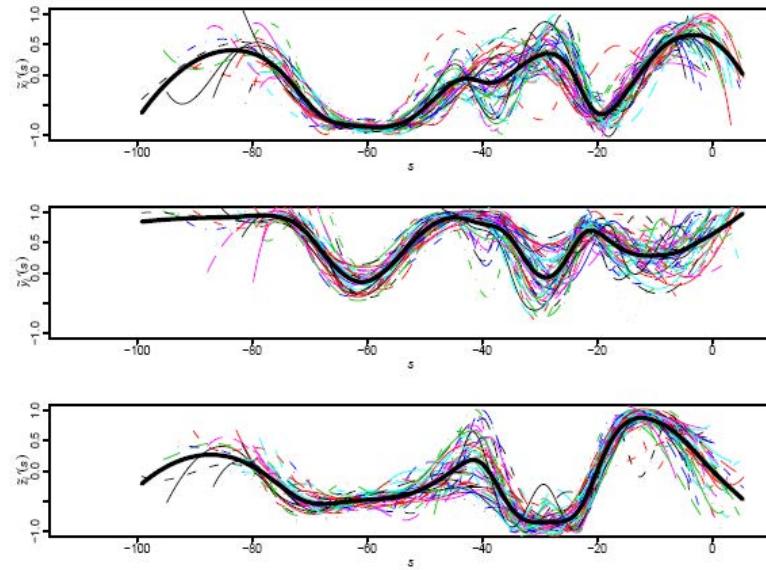
128

Sangalli Secchi Vantini Veneziani, 2009, JASA

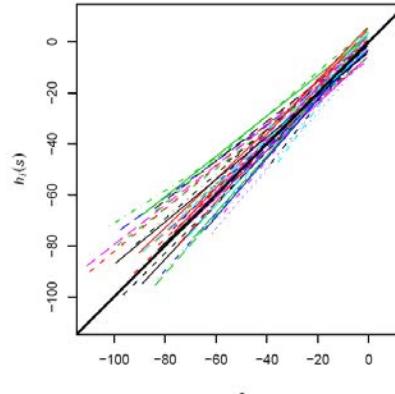
Original centerlines



Aligned centerlines



Warping functions (phase variab)

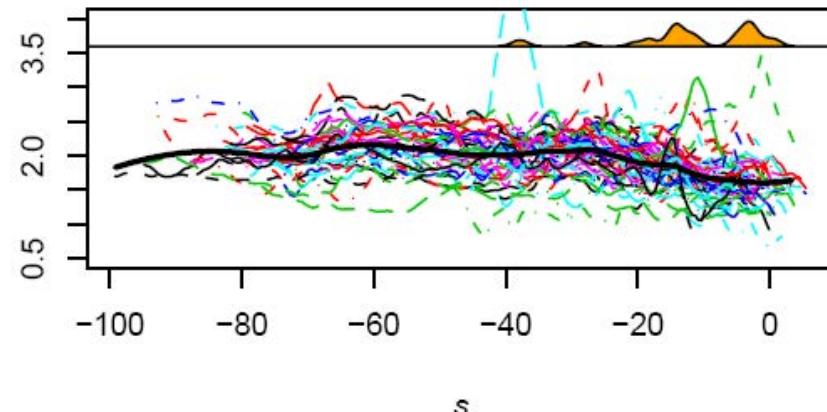
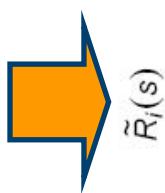
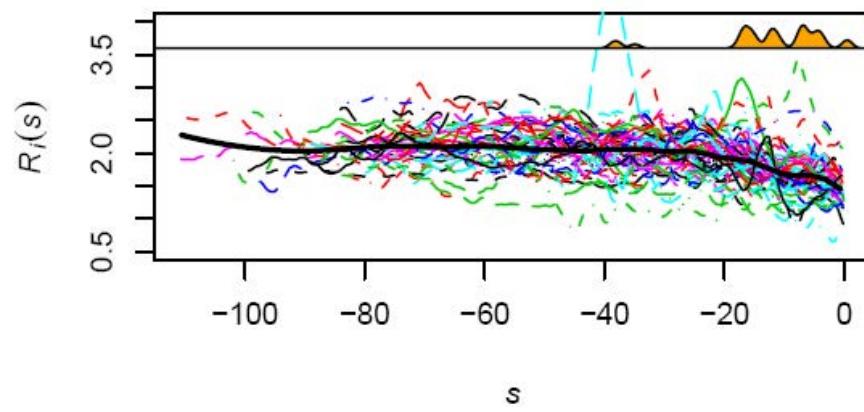


Aneurysm location on aligned ICA radius and curvature profiles

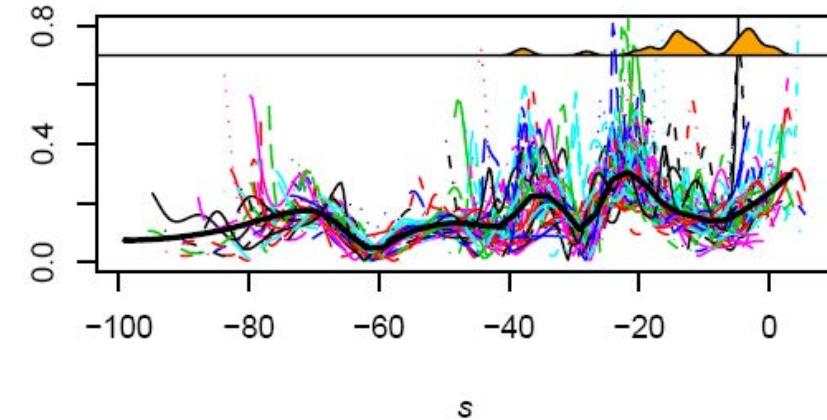
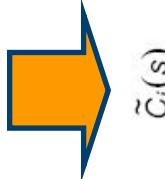
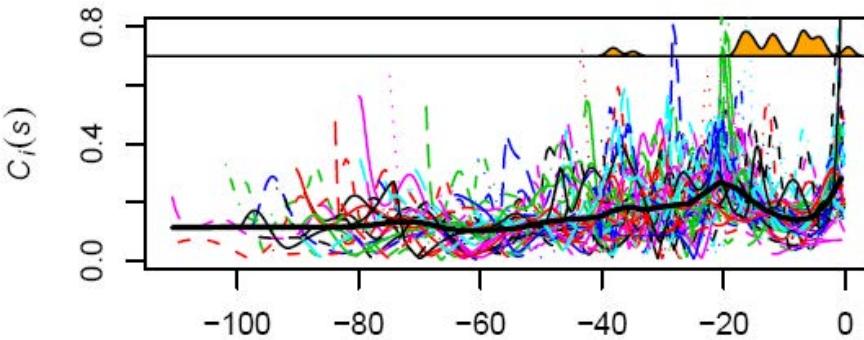
129

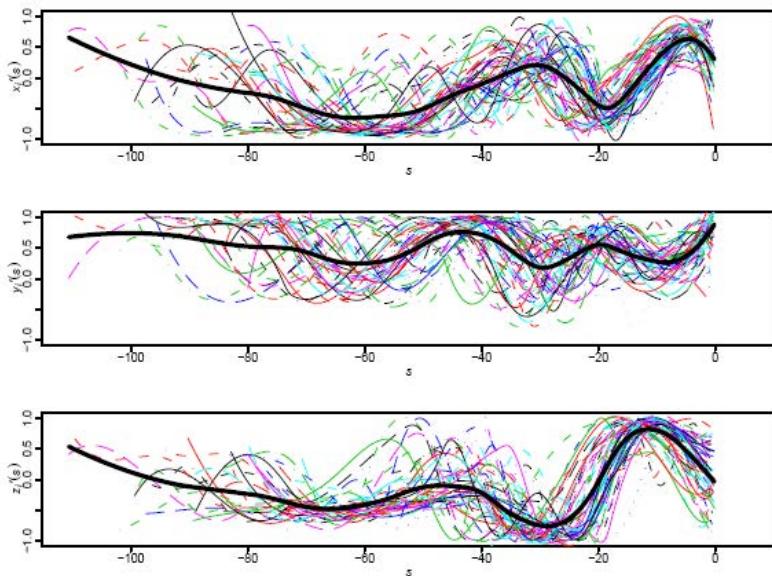
Sangalli Secchi Vantini Veneziani, 2009, JASA

Radius



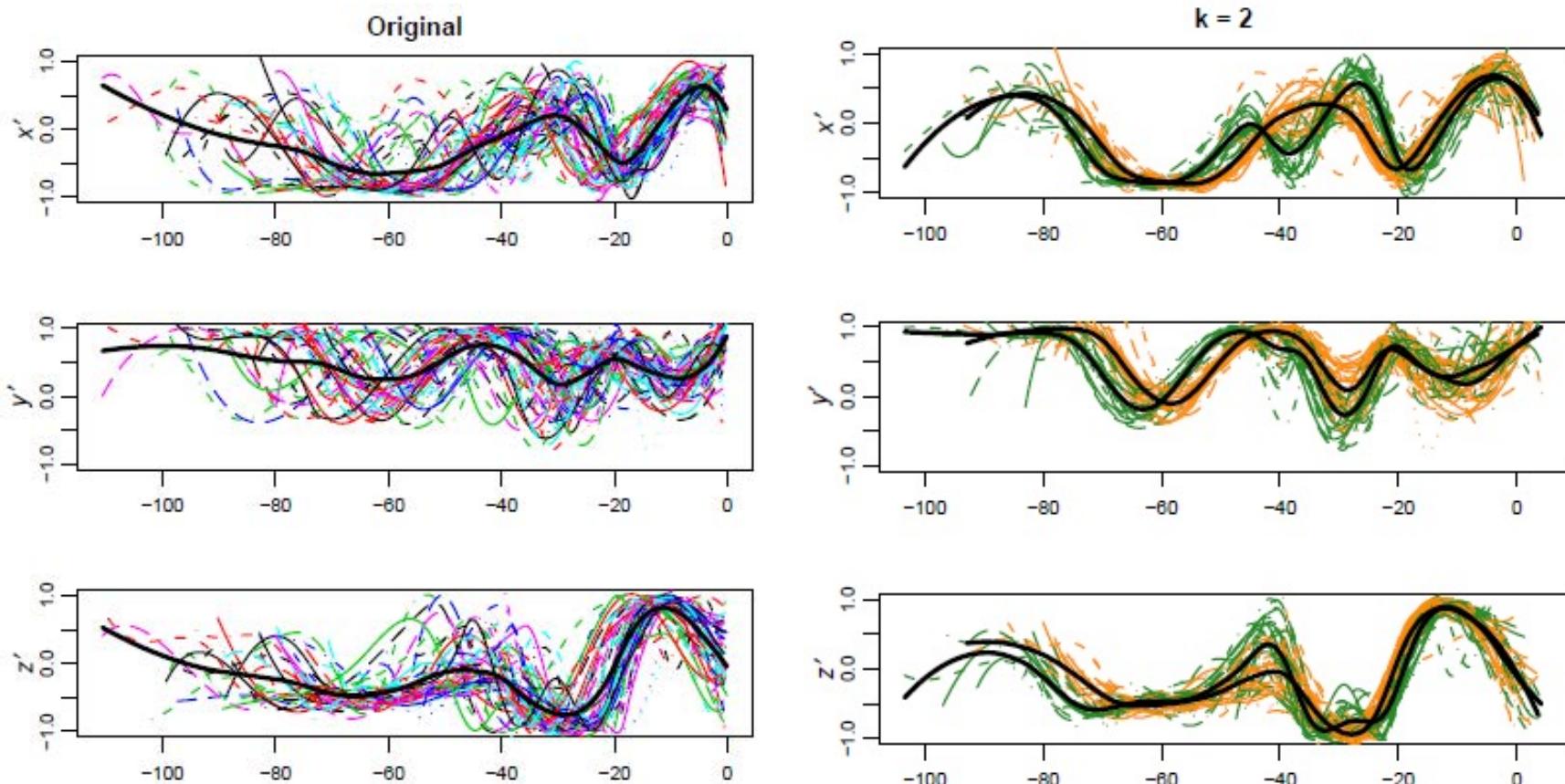
Curvature

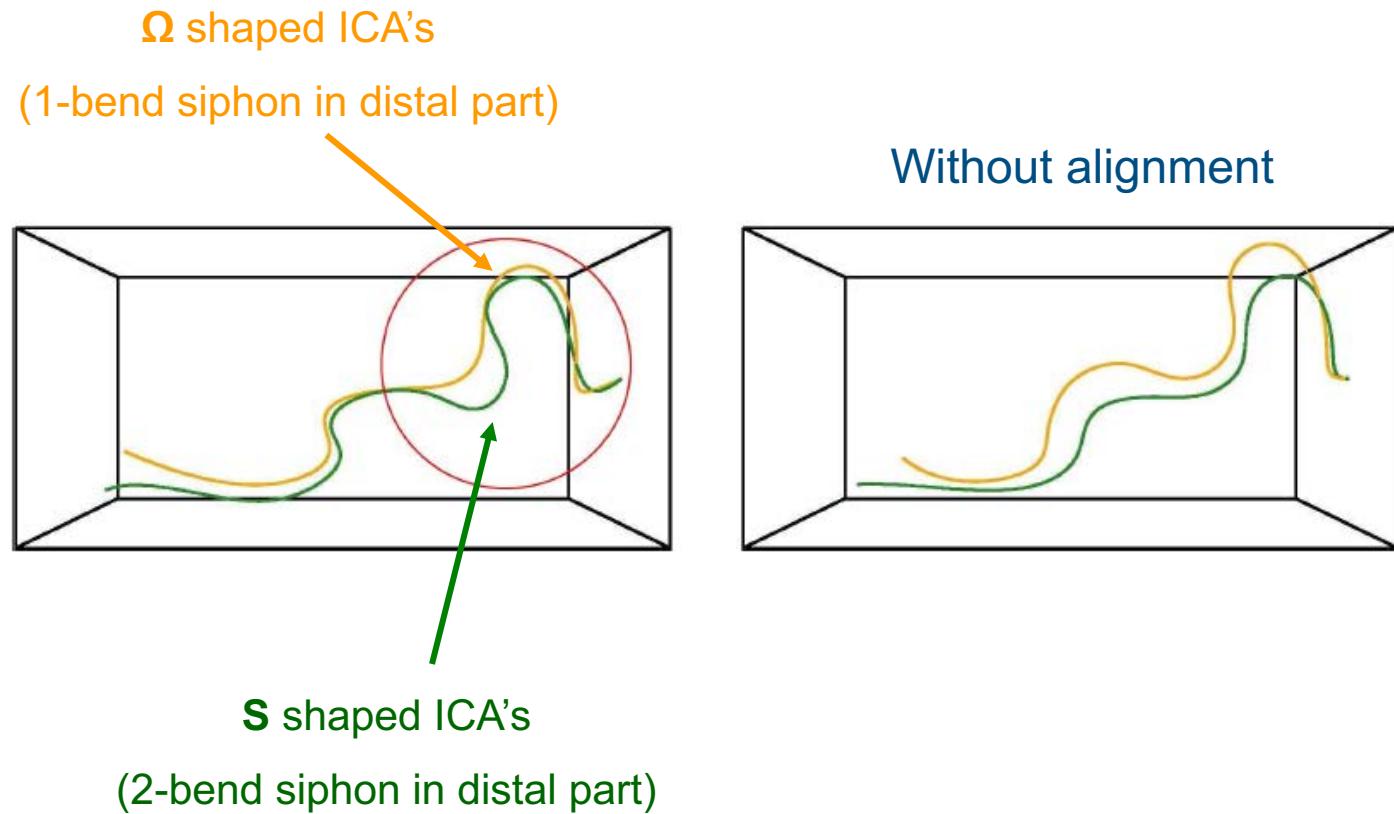




GOAL: Identify ICA's with different morphological shapes

Need to be able to:
jointly align and cluster
the N centerlines
in multiple groups (k groups)
having unknown templates





The procedure identify two prototype shapes of ICA's that are described in the medical literature
Krayenbuehl et. Al. (1982)

NO interesting insights
Simple clustering without alignment is driven by phase variability and fails to identify different morphological shapes

	Aneurysm at or after ICA biforc. (33)	Aneurysm before ICA biforc. (25)	No aneurysms (7)
S shaped ICA's	30%	52%	100%
Ω shaped ICA's	70%	48%	0%

► The ICA siphon acts as a flow energy dissipator to steady blood flow in the brain

→ S shaped ICA's seems to be more effective in making the blood-flow steadier with respect to Ω shaped ICA's



Mathematical Biosciences Institute

[Home](#) [About](#) [News](#) [Events](#) [People](#) [Visitors](#) [Postdoctoral](#) [Committees](#) [Institute Partners](#) [Education](#) [Publications](#)

- [Calendar](#)
- [Apply for Workshop](#)
- [Workshops](#)
- [Visiting Lecturer Program](#)
- [Colloquia/Seminars](#)
- [Summer Programs](#)
- [Public Lectures](#)
- [Visitor Info](#)
- [Annual Programs](#)
- [Propose MBI Programs](#)

CTW: Statistics of Time Warpings and Phase Variations (November 13-16, 2012)

Organizers: J. S. Marron (UNC), J. O. Ramsay (McGill), L. Sangalli (Politecnico di Milano), A. Srivastava (Florida State)

[Description](#) [Schedule](#) [Participants](#) [Titles & Abstracts](#) [Resources](#) [Apply for Event](#)
[Flyer](#)

Background: A common feature of functional measurements of data over time, space and other continua, is that salient features in the resulting curves and surfaces vary in position from one recording to the next. For example, the growth patterns of children vary in the timing of puberty, human movements in activities like handwriting and golf swings speed up and slow down from one instance to another, seasonal events like hurricanes arrive early some years and late in others, and traffic jams vary in location over city streets from one day to another. At the same time, each of the events can also vary in intensity. We refer to positional variation as phase variation, and intensity variation as amplitude variation. It is now evident that many processes unfold over a system time that not only does not unroll at the same rate as physical clock time, but also tends to vary in a significant way from one realization of a functional event to another.

The registration or alignment of features in curves and images by smooth, one-to-one transformations of time or space, respectively, is an emerging hot topic that presents many challenges. From its beginnings with dynamic time warping in the late 50's, followed by the landmark registration methods of Fred Bookstein, the registration of brain images to a fixed atlas, and its widespread application in functional data analysis, statisticians have realized that nonlinear phase

The Electronic Journal of Statistics

**Special Section on Statistics of Time
Warpings and Phase Variation,
Volume 8 , Number 2, 2014**



POLITECNICO DI MILANO



Politecnico di Milano
Applied Statistics
May 2023



An introduction to functional data analysis

Laura M. SANGALLI

MOX - Dipartimento di Matematica, Politecnico di Milano

Part 7 – Functional Principal Component Analysis

Functional Principal Component Analysis

Chapter 8 of Ramsay and Silverman (2005), *Functional Data Analysis*, Springer



Problem: Given a dataset of N zero-mean multivariate observations in \mathbb{R}^p , X_1, \dots, X_N find the orthonormal directions a_1, \dots, a_p of maximum variability (for the dataset).

Equivalently, for $k=1, \dots, p$, find:

$$\begin{aligned} a_k &= \operatorname{argmax}_{\mathbf{a} \in \mathbb{R}^p} \operatorname{Var}(\mathbf{a}' \mathbf{X}) \\ \text{subject to: } &\mathbf{a}' \mathbf{a} = 1, \mathbf{a}'_j \mathbf{a} = 0 \text{ for } j < k \end{aligned}$$

- We can re-write the problem as

$$\begin{aligned} a_k &= \operatorname{argmax}_{\mathbf{a} \in \mathbb{R}^p} \frac{1}{N} \sum_{i=1}^N (\mathbf{a}' \mathbf{X}_i)^2 \\ \text{subject to: } &\mathbf{a}' \mathbf{a} = 1, \mathbf{a}'_j \mathbf{a} = 0 \text{ for } j < k \end{aligned}$$

or, equivalently

$$\begin{aligned} a_k &= \operatorname{argmax}_{\mathbf{a} \in \mathbb{R}^p} \frac{1}{N} \sum_{i=1}^N \langle \mathbf{a}, \mathbf{X}_i \rangle^2 \\ \text{subject to: } &\|\mathbf{a}\| = 1, \langle \mathbf{a}_j, \mathbf{a} \rangle \geq 0 \text{ for } j < k \end{aligned}$$

Note 1. We assume $N > p$ and absence of collinearity, i.e. the data matrix is full rank.

Note 2 . If X_1, \dots, X_N are not zero-mean, they can be centered by subtracting the (sample) mean. For unbiasedness, divide by $N-1$ instead of N .



Problem: Given a dataset of N zero-mean multivariate observations in \mathbb{R}^p , X_1, \dots, X_N find the orthonormal directions a_1, \dots, a_p of maximum variability, i.e., those solving for $k=1, \dots, p$,

$$a_k = \operatorname{argmax}_{\mathbf{a} \in \mathbb{R}^p} \frac{1}{N} \sum_{i=1}^N \langle \mathbf{a}, X_i \rangle^2$$

subject to: $\|\mathbf{a}\| = 1, \langle a_j, \mathbf{a} \rangle = 0 \text{ for } j < k$

Solution: Call S the sample covariance matrix of X_1, \dots, X_N . Then, the **principal components** are found as the eigenvectors of the matrix S ; for $k=1, \dots, p$, they solve the eigen-equation

$$Se_k = \lambda_k e_k$$

The eigenvalue λ_k associated with the eigenvector e_k represents the variability along the direction e_k .

Note. We call score u_{ik} the projection of the observation X_i along the direction e_k , i.e.,

$$u_{ik} = \langle X_i, e_k \rangle = X'_i e_k$$

Problem: Given a dataset of N zero-mean functional observations in H , X_1, \dots, X_N , find the directions of maximum variability (in H) of the dataset, i.e., for $k=1, \dots, N$, find ξ_k maximizing

$$\frac{1}{N} \sum_{i=1}^N \langle \xi, X_i \rangle_H^2$$

subject to: $\|\xi\| = 1, \langle \xi_j, \xi \rangle_H = 0$ for $j < k$

- We look for an orthonormal system in H maximizing the *variability* of the corresponding projections
- Indeed, $\langle \xi, X_i \rangle_H$ is the projection of X_i «along the direction» ξ (i.e., a «direction» in H). Note that $\langle \xi, X_i \rangle_H$ is a scalar, hence $\frac{1}{N} \sum_{i=1}^N \langle \xi, X_i \rangle_H^2$ is a sample variance in the usual sense.

Note 1. If the data are not zero-mean, they can be centered by subtracting the (sample) mean. N should then be replaced by $N-1$.

Note 2. If data are linearly independent and centered on the sample mean, we can find at most $N-1$ principal components.

Problem: Given a dataset of N zero-mean functional observations in H , X_1, \dots, X_N , find the directions of maximum variability (in H) of the dataset, i.e., for $k=1, \dots, N$, find ξ_k maximizing

$$\frac{1}{N} \sum_{i=1}^N \langle \xi, X_i \rangle_H^2$$

subject to: $\|\xi\| = 1, \langle \xi_j, \xi \rangle_H = 0$ for $j < k$

- As in multivariate principal component analysis, **functional principal components** are related with the eigen-decomposition of the functional counterpart of the (sample) covariance matrix
- Recall that the **sample covariance operator** is defined as

$$Sx = \frac{1}{N} \sum_{i=1}^N \langle X_i, x \rangle X_i, \quad x \in H$$

In L^2 it is equivalently defined as

$$[Sx](t) = \int_T \widehat{c}(s, t)x(s)d(\varepsilon)], \quad x \in L^2 \quad \text{with} \quad \widehat{c}(s, t) = \frac{1}{N} \sum_{i=1}^N X(s)X(t)$$

Note. If data are centered on the sample mean, divide by $N-1$ for unbiasedness.

Problem: Given a dataset of N zero-mean functional observations in H, X_1, \dots, X_N , find the directions of maximum variability (in H) of the dataset, i.e., for $k=1, \dots, N$, find ξ_k maximizing

$$\frac{1}{N} \sum_{i=1}^N \langle \xi, X_i \rangle_H^2$$

subject to: $\|\xi\| = 1, \langle \xi_j, \xi \rangle_H = 0$ for $j < k$

Solution: Let S be the sample covariance operator of X_1, \dots, X_N . Then, the **functional principal components** ξ_1, \dots, ξ_N are found as the eigenfunctions of the operator S , i.e., they solve the eigen-equations

$$S\xi_k = \lambda_k \xi_k$$

The eigenvalue λ_k associated with the eigenvector ξ_k represents the variability along the direction ξ_k

We call functional score u_{ik} the projection of the observation X_i along the direction ξ_k , i.e.,

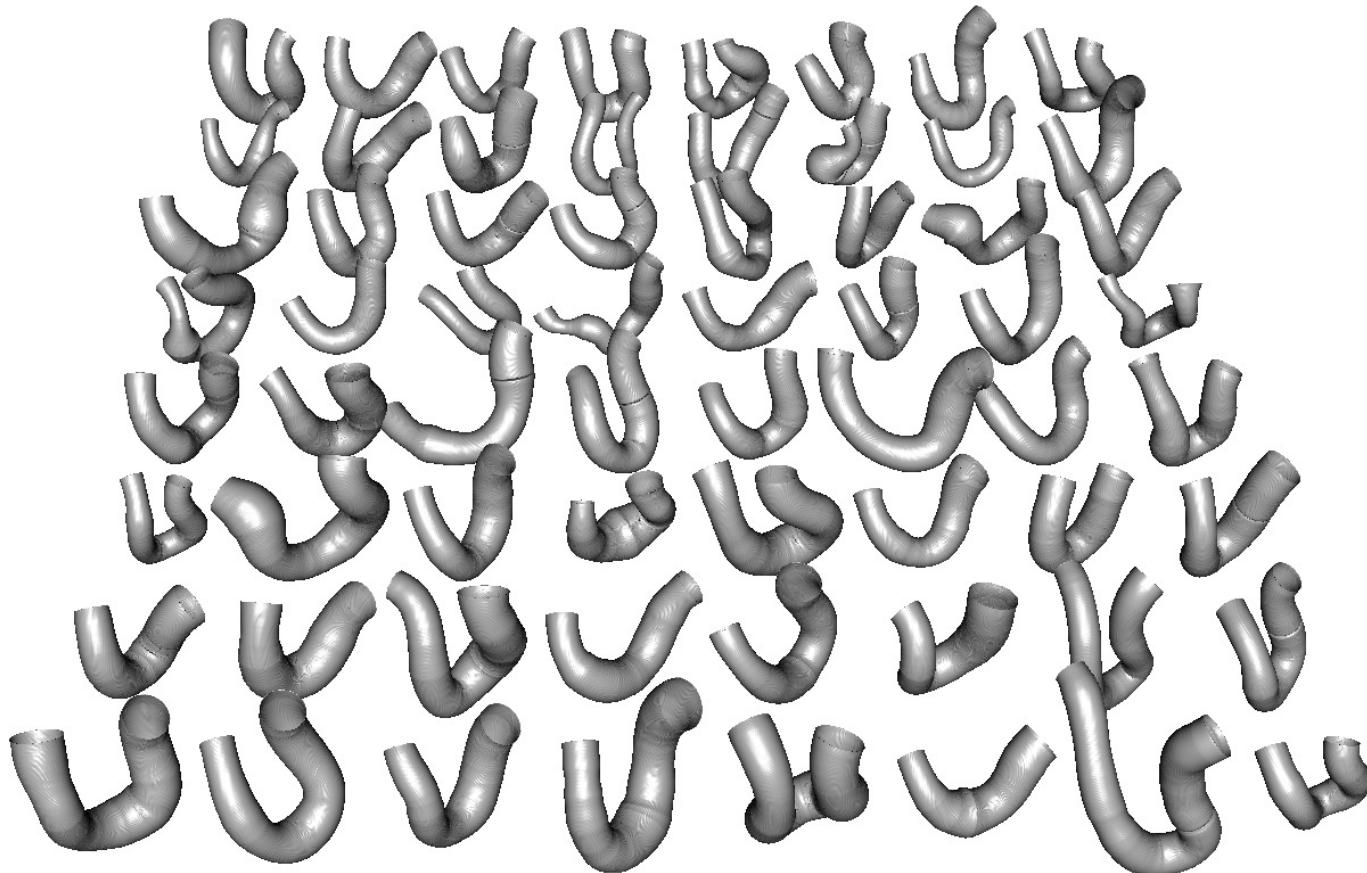
$$u_{ik} = \langle X_i, \xi_k \rangle$$

Note. If data are centered on the sample mean, we can find at most $N-1$ principal components

- !!! We only have discrete and noisy realizations of X_1, \dots, X_N !!!
for the i -th statistical unit:
 $\{x_{i1}, \dots, x_{in}\}$: x_{ij} is the value observed for the i -th statistical unit at s_j
→ for each statistical unit we obtain a functional representation by smoothing
PRE-SMOOTHING APPROACH
- Dimensional reduction: look for an elbow in the cumulative percentage of total variance explained by the first p functional principal components
- Useful plots: boxplots of scores along the first p directions, to investigate the possible presence (and influence) of outliers, clustering structures, etc
- Interpretation of the loadings:
 - Plot directly loadings (*only for expert users*)
 - Plot mean +/- eigenfunction multiplied by a proper constant, e.g., std. along the component, which corresponds to sqrt of the eigenvalue: $\bar{X} \pm \sqrt{\lambda_k} \xi_k$
 - Plotting the projection of the dataset along each component: $\bar{X} + u_{ik} \xi_k$
or along the first p components: $\bar{X} + \sum_{k=1}^p u_{ik} \xi_k$

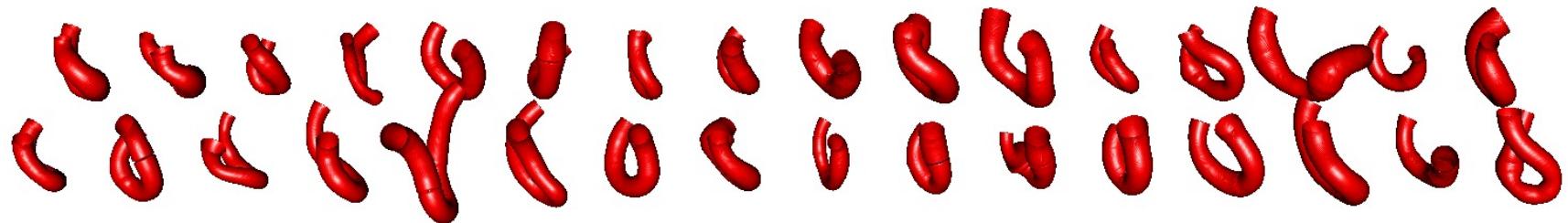
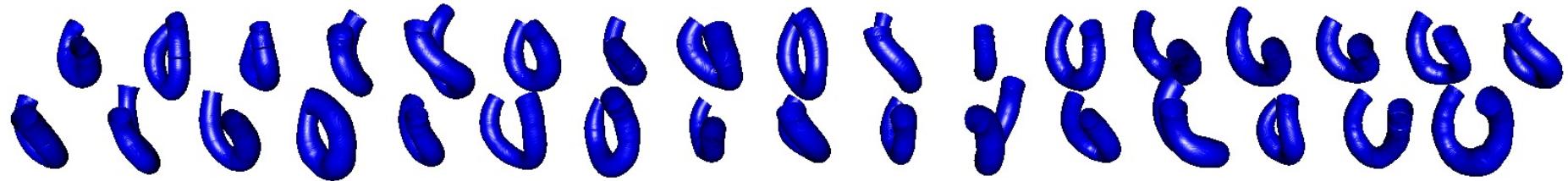
$$CPV(p) = \frac{\sum_{k=1}^p \hat{\lambda}_k}{\sum_{k=1}^N \hat{\lambda}_k}.$$

The sample of 65 ICA: each patient is represented by the centerline and radius of ICA



Aneurysm at or after ICA bifurcation

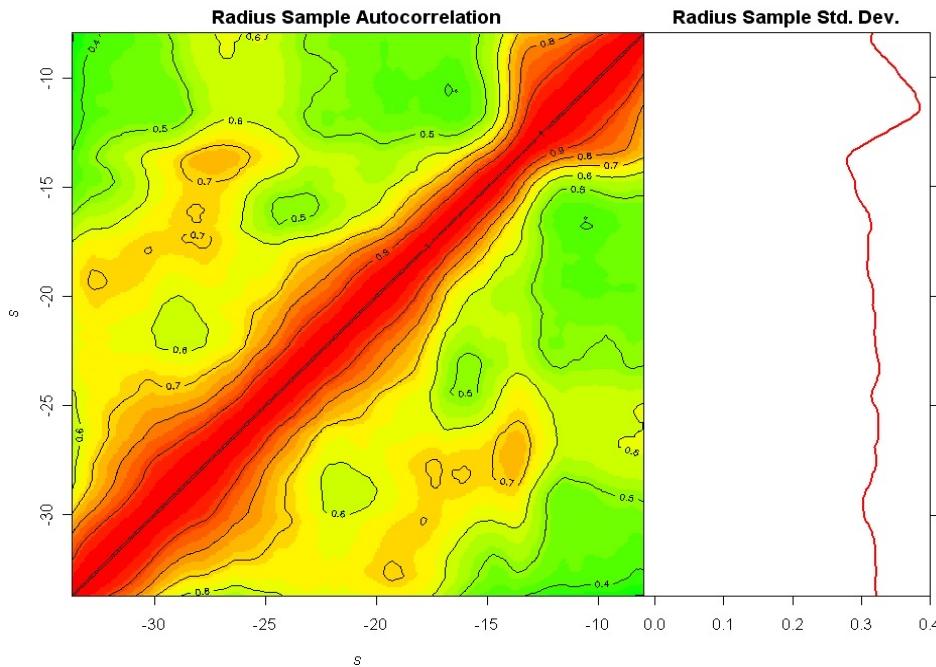
Upper Group: 33



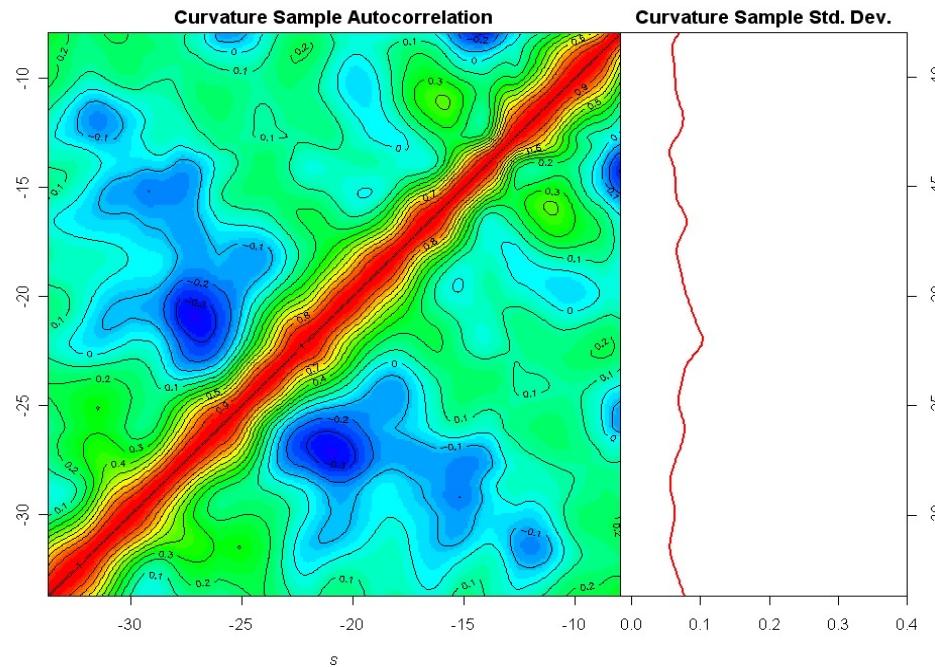
Lower Group: 32

Aneurysm before ICA bifurcation or no aneurysm

Sample Autocorrelation Function and Std. Dev. for Radius Profiles



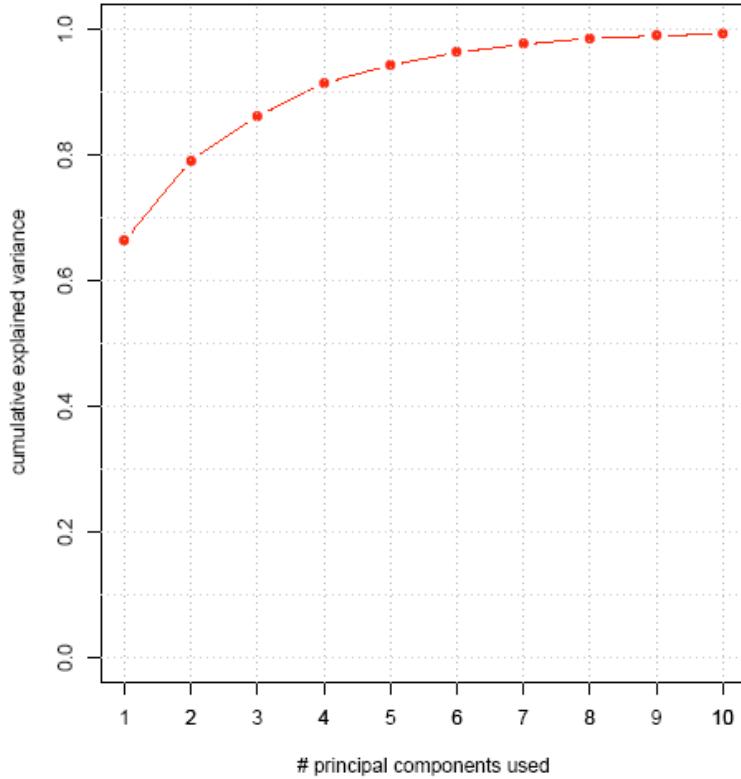
Sample Autocorrelation Function and Std. Dev. for Curvature Profiles



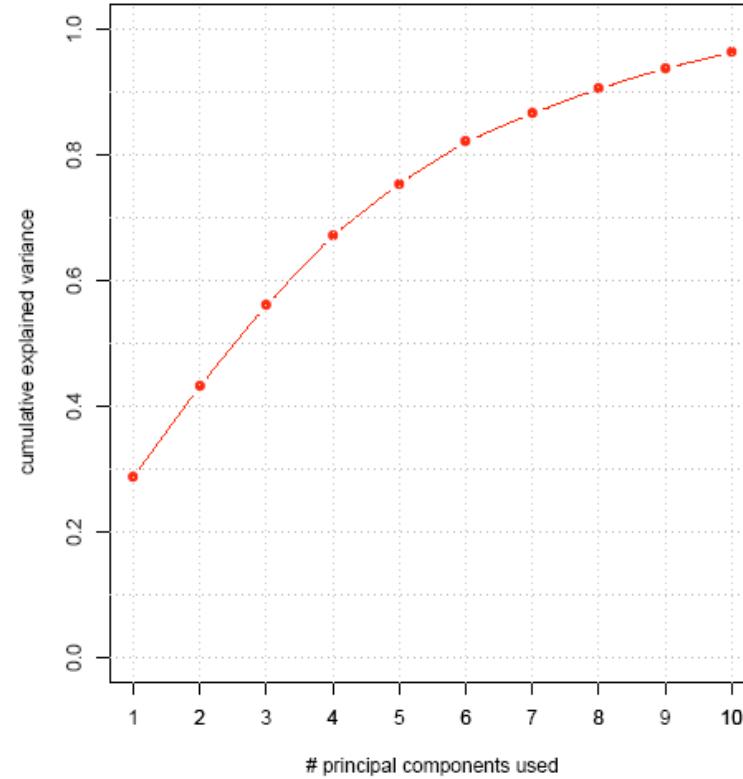
Aligned data

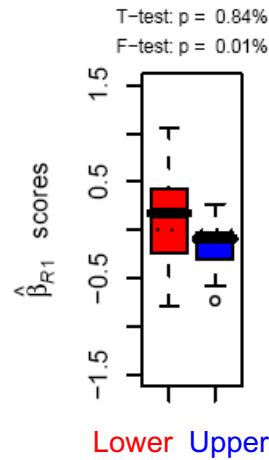
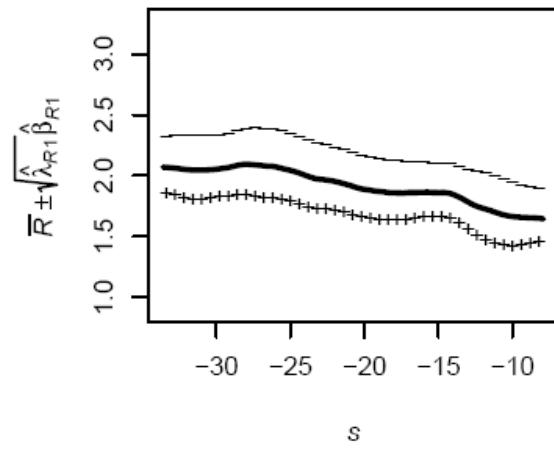
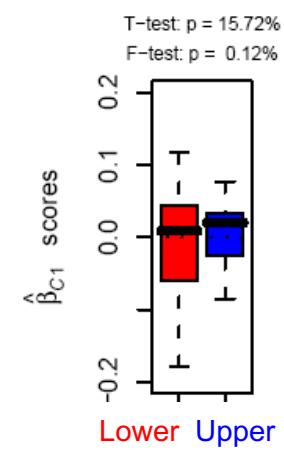
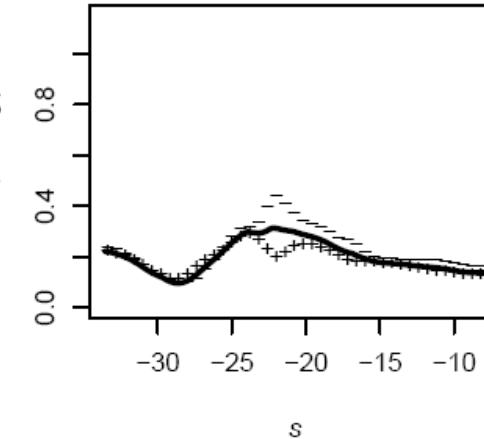
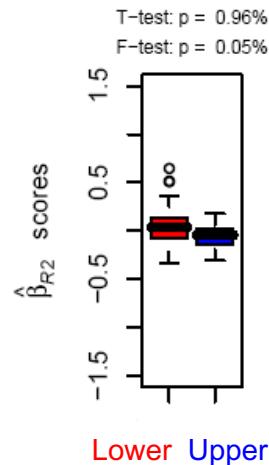
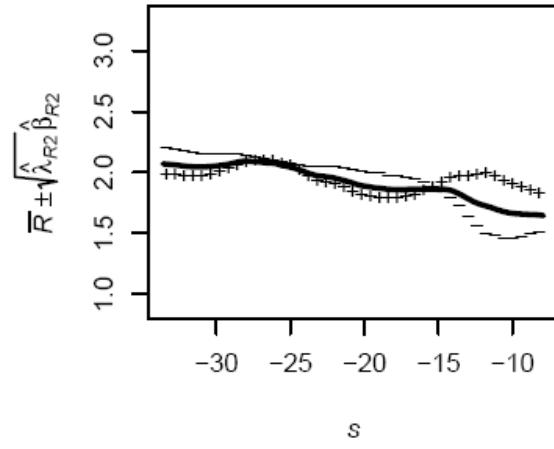
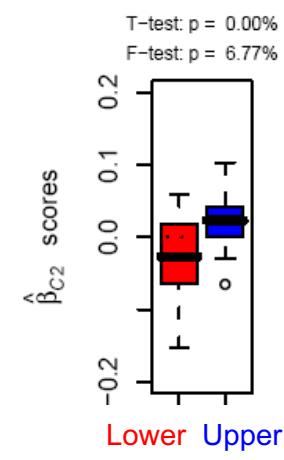
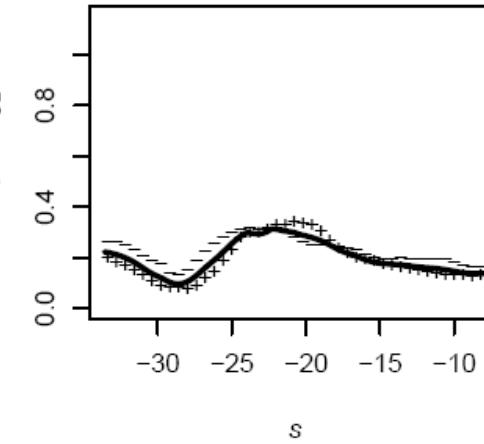
Courtesy of S. Vantini

Radius

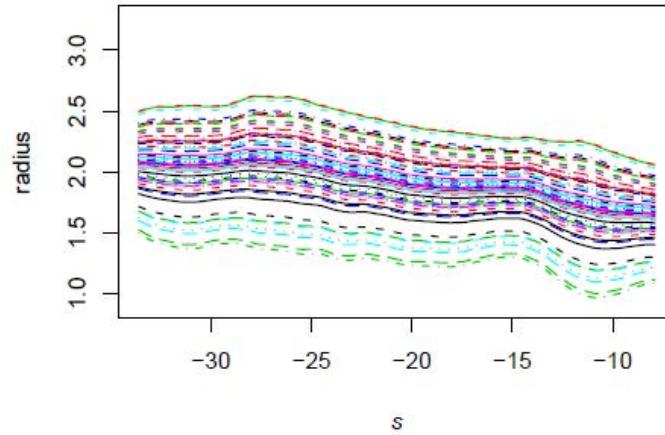


Curvature

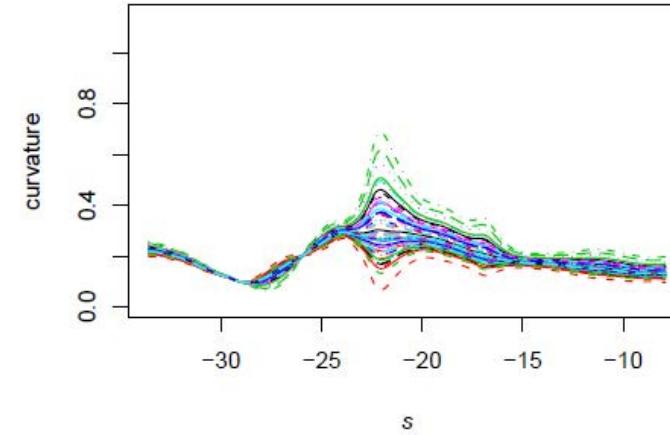


1st PC for Radius (65.9%)**1st PC for Curvature (21.0%)****2nd PC for Radius (13.0%)****2nd PC for Curvature (14.9%)**

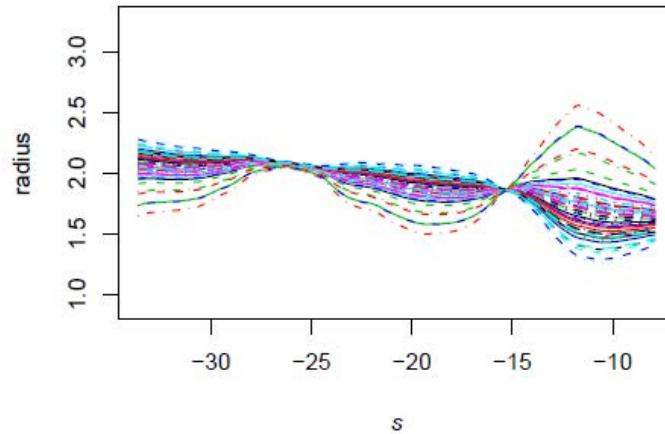
1st PC for Radius (65.9%)



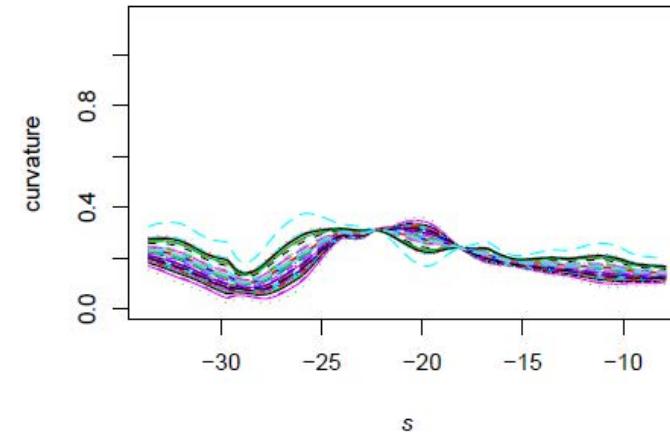
1st PC for Curvature (21.0%)



2nd PC for Radius (13.0%)



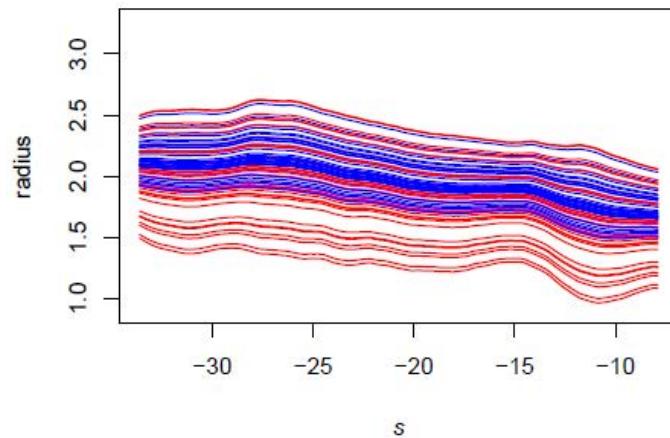
2nd PC for Curvature (14.9%)



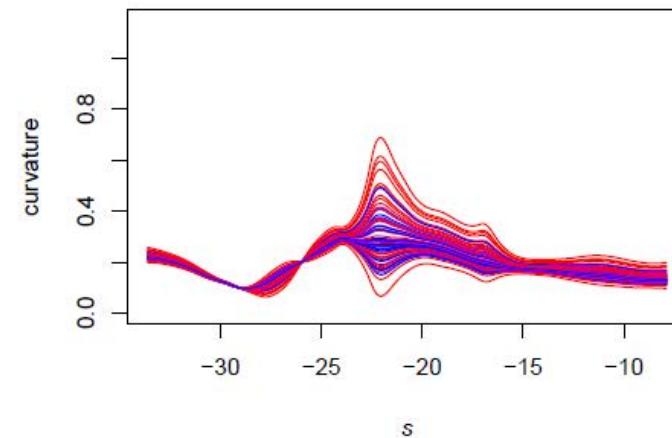
Courtesy of S. Vantini



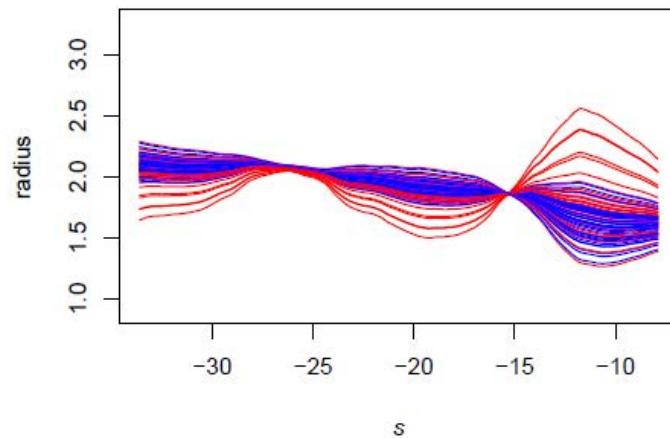
1st PC for Radius (65.9%)



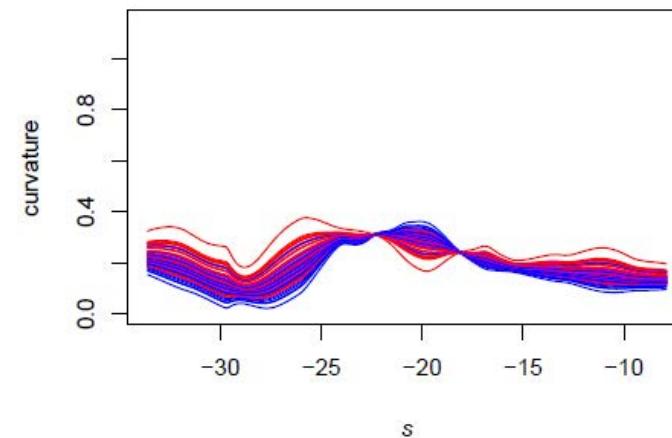
1st PC for Curvature (21.0%)



2nd PC for Radius (13.0%)



2nd PC for Curvature (14.9%)



Courtesy of S. Vantini



Discriminant analysis of functional principal component scores

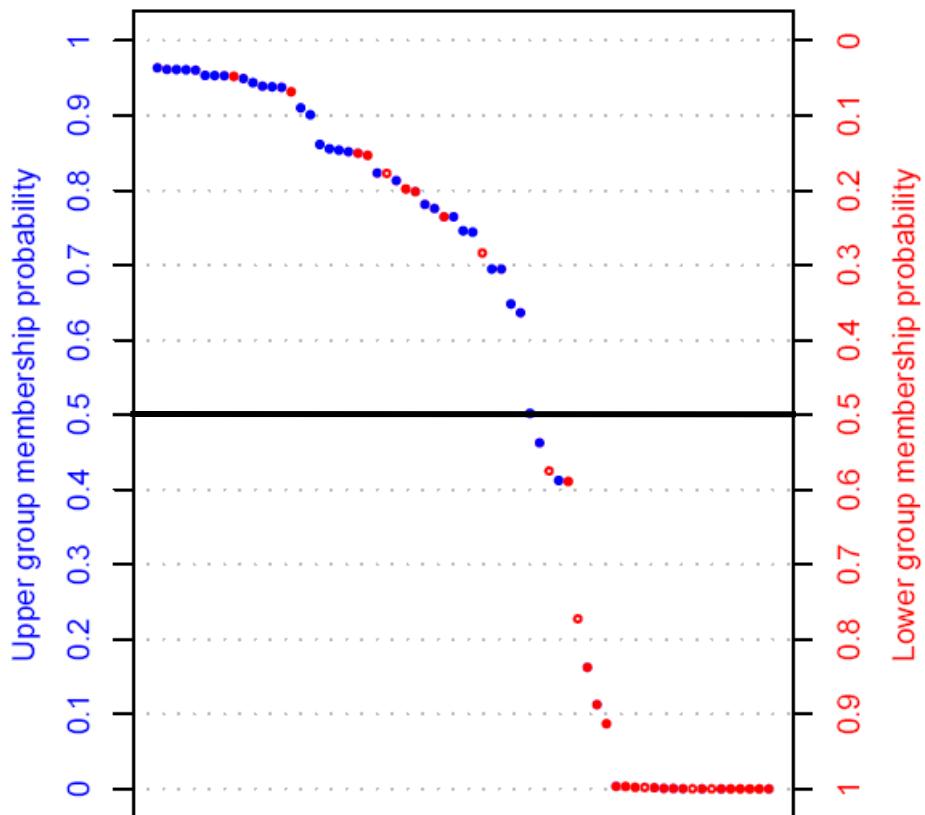
153

Sangalli, Secchi, Vantini, Veneziani 2009 JASA

→ **Upper group** patients are very well characterize by this two geometric features

A quadratic discriminant analysis of scores correctly identifies 31 out of the 33 patients in this group

- Large vessels
- Strong tapering
- High within-patient curvature variability
- Lower between-patient variability



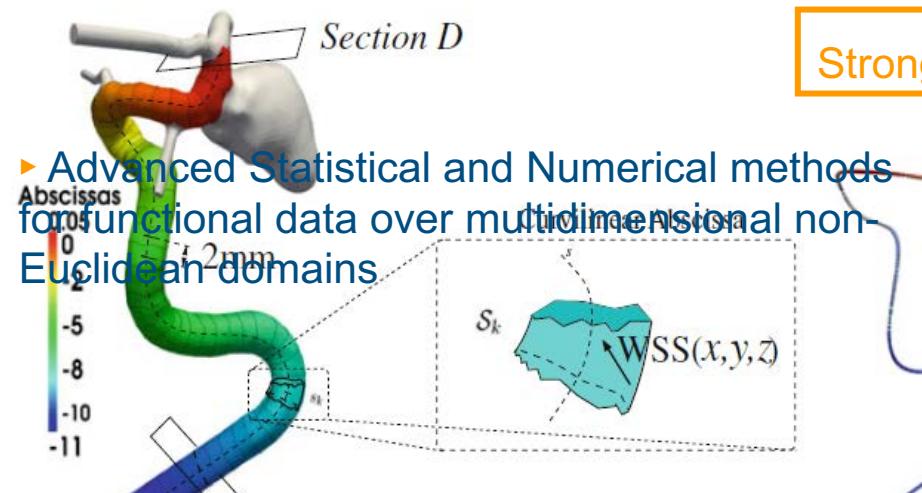
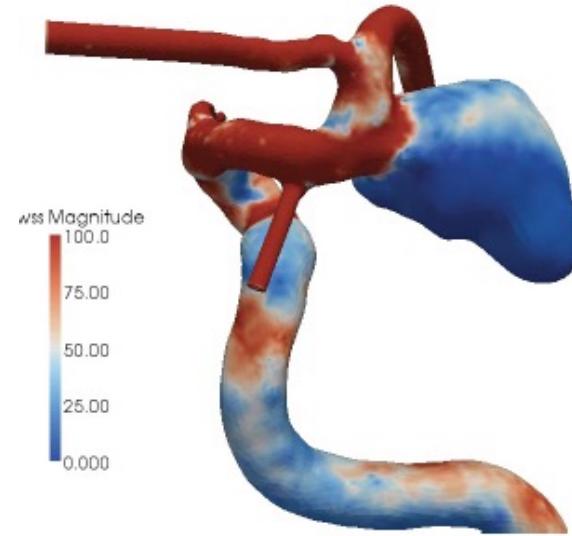
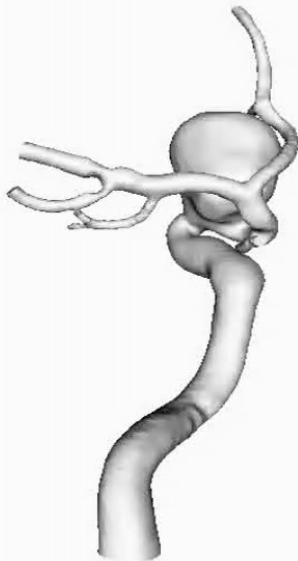
Blood Flow Numerical Simulations

Ettinger et al. 2016, Biometrika

Lila et al. 2016, AoAS

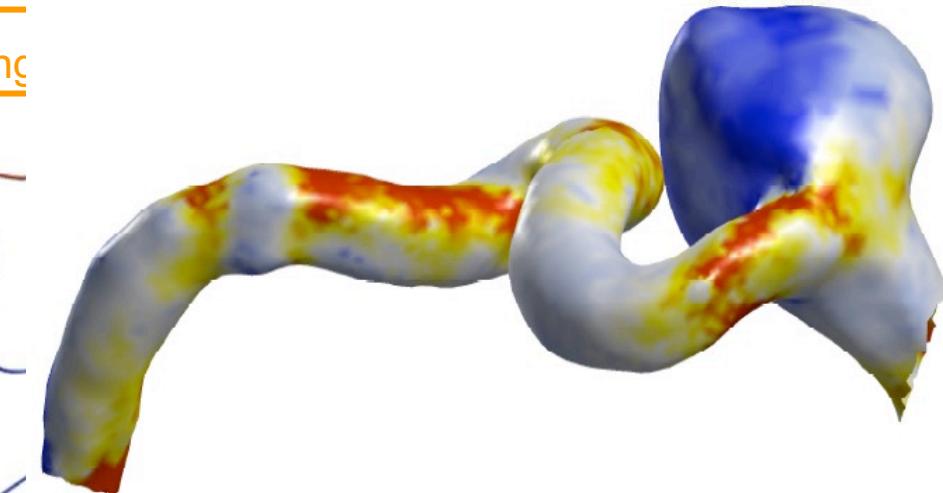
Wilhelm et al 2016, CMAME

Hemodynamic data obtained by Computational Fluid Dynamics



► Advanced Statistical and Numerical methods
for functional data over multidimensional non-
Euclidean domains

Strong



An alternative way to compute functional Principal Components

- Best K basis approximation property (zero-centered functional data)

$$(\xi_k)_{j=1}^K = \underset{(\{f_k\}_{k=1}^K : \langle f_k, f_l \rangle_H = \delta_{kl})}{\operatorname{argmin}} \mathbb{E} \left[\left\| X - \sum_{k=1}^K \langle X, f_k \rangle_H f_k \right\|_H \right]$$

▷ $H \equiv L^2$

$$(\xi_k)_{k=1}^K = \underset{(\{f_k\}_{k=1}^K : \langle f_k, f_l \rangle_{L^2} = \delta_{kl})}{\operatorname{argmin}} \mathbb{E} \int_{\mathcal{D}} \left\{ X - \sum_{k=1}^K \langle X, f_k \rangle_{L^2} f_k \right\}^2$$

- Empirical version for $K = 1$

$$\frac{1}{n} \frac{1}{N} \sum_{i=1}^N \sum_{j=1}^n (x_{ij} - u_i f(s_j))^2 + \text{smoothness}$$

An alternative way to compute functional Principal Components

- Best K basis approximation property (zero-centered functional data)

$$(\xi_k)_{j=1}^K = \underset{(\{f_k\}_{k=1}^K : \langle f_k, f_l \rangle_H = \delta_{kl})}{\operatorname{argmin}} \mathbb{E} \left[\left\| X - \sum_{k=1}^K \langle X, f_k \rangle_H f_k \right\|_H \right]$$

- ▷ $H \equiv L^2$

$$(\xi_k)_{k=1}^K = \underset{(\{f_k\}_{k=1}^K : \langle f_k, f_l \rangle_{L^2} = \delta_{kl})}{\operatorname{argmin}} \mathbb{E} \int_{\mathcal{D}} \left\{ X - \sum_{k=1}^K \langle X, f_k \rangle_{L^2} f_k \right\}^2$$

- Empirical version for $K = 1$

$$\frac{1}{n} \frac{1}{N} \sum_{i=1}^N \sum_{j=1}^n (x_{ij} - u_i f(s_j))^2 + \lambda \int (f'')^2$$

- ▶ Estimate first PC function ξ_1 and associated PC score vector $\mathbf{u} \in \mathbb{R}^N$

$$(\hat{\mathbf{u}}, \hat{\xi}_1) = \underset{\mathbf{u}, f}{\operatorname{argmin}} \left\{ \sum_{i=1}^N \sum_{j=1}^{n_i} (x_{ij} - u_i f(s_j))^2 + \lambda \mathbf{u}^\top \mathbf{u} \int (f'')^2 \right\}$$

Invariance properties:

$$\mathbf{u} \rightarrow c\mathbf{u} \text{ and } \xi \rightarrow \frac{1}{c}\xi$$

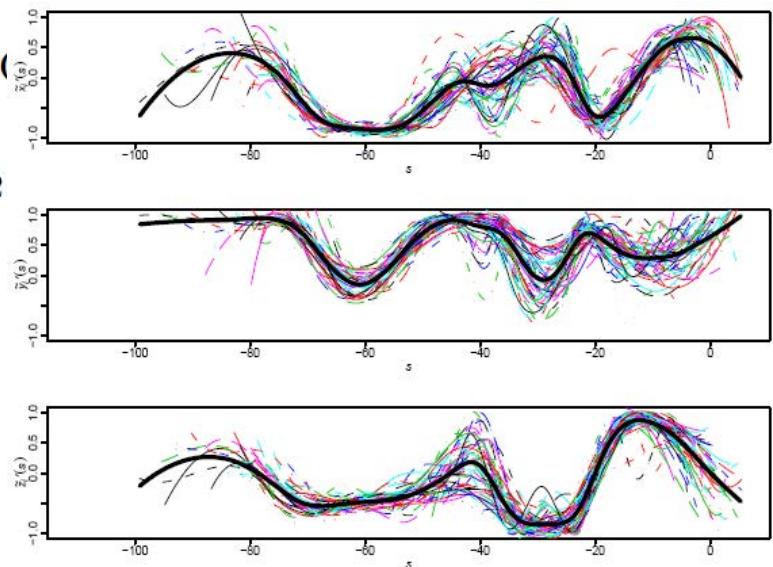
$$\mathbf{X} \rightarrow c\mathbf{X} \text{ and } \mathbf{u} \rightarrow c\mathbf{u}$$

- ▶ Subsequent PC functions are extracted sequentially by removing previous PC functions from data matrix \mathbf{X}
- ▶ Missing data / sparsely sampled data / incomplete functional data:
 Functional datum for i -statistical unit is observed at points $s_{i1}, \dots, s_{in_i} : (x_{i1}, \dots, x_{in_i})$

$$(\hat{\mathbf{u}}, \hat{\xi}) = \underset{\mathbf{u}, f}{\operatorname{argmin}} \sum_{i=1}^N \sum_{j=1}^{n_i} (x_{ij} - u_i f(s_{ij}))^2 + \lambda \mathbf{u}^\top \mathbf{u} \int (f'')^2$$

- ▶ Estimate first PC function ξ_1 and associated PC scores

$$(\hat{\mathbf{u}}, \hat{\xi}_1) = \operatorname{argmin}_{\mathbf{u}, f} \left\{ \sum_{i=1}^N \sum_{j=1}^{n_i} (x_{ij} - u_i f(s_j))^2 \right\}$$



- ▶ Subsequent PC functions are extracted sequentially by removing previous PC functions from data matrix \mathbf{X}
- ▶ Missing data / sparsely sampled data / incomplete functional data:
 Functional datum for i -statistical unit is observed at points $s_{i1}, \dots, s_{in_i} : (x_{i1}, \dots, x_{in_i})$

$$(\hat{\mathbf{u}}, \hat{\xi}) = \operatorname{argmin}_{\mathbf{u}, f} \sum_{i=1}^N \sum_{j=1}^{n_i} (x_{ij} - u_i f(s_{ij}))^2 + \lambda \mathbf{u}^\top \mathbf{u} \int (f'')^2$$

AneuRisk data: significantly decreases the cross-validation missclassification error



POLITECNICO DI MILANO



Politecnico di Milano
Applied Statistics
May 2023



An introduction to functional data analysis

Laura M. SANGALLI

MOX - Dipartimento di Matematica, Politecnico di Milano

Part 8 – Case Study: Neuronal connectivity
Functional data over complex multidimensional domains

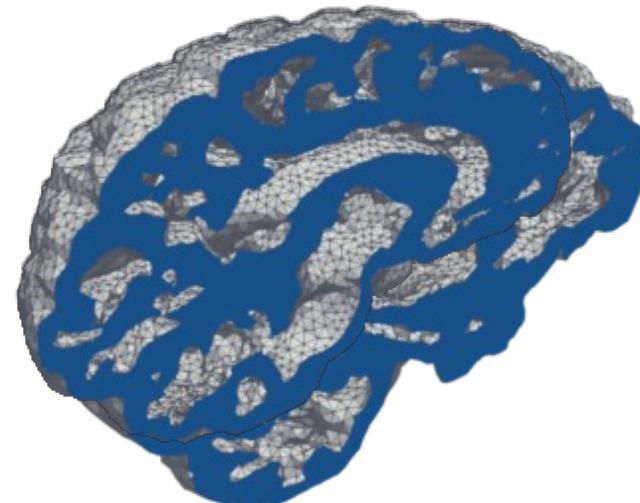
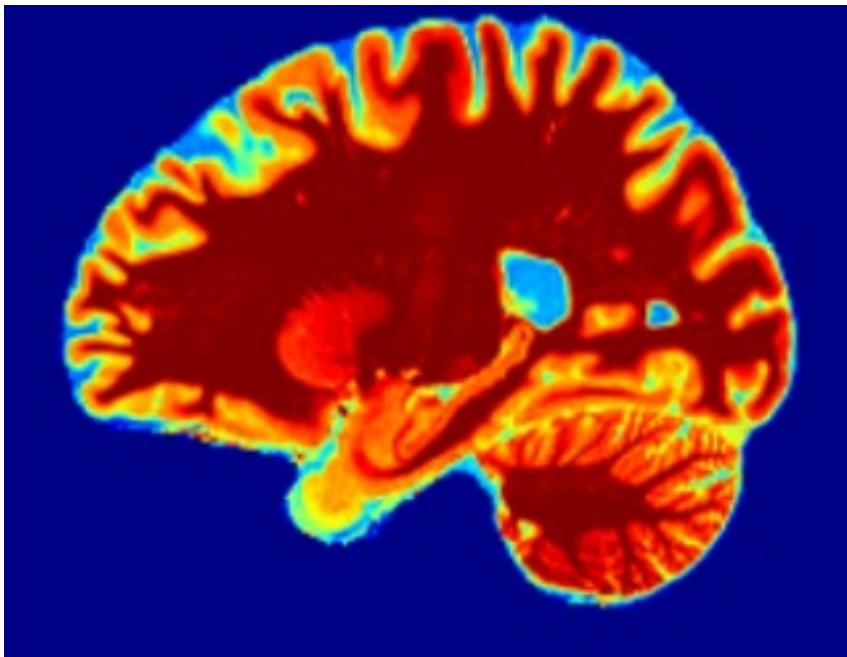
fMRI data measuring neuronal activity: functional connectivity *schizophrenic vs healthy*

164

Functional Magnetic Resonance Imaging (fMRI)

- ▶ 49 schizophrenic patients and 122 healthy subjects
- ▶ Blood Oxygen Level Dependent (BOLD) signal
- ▶ proxy measure of neuronal activity
- ▶ can be used to compute maps of connectivity with respect to a cerebral region of interest
- ▶ signal in *grey matter* (*non-convex volume*)

Arnone et al. 2023, Biometrics
Clementi et al. 2023+



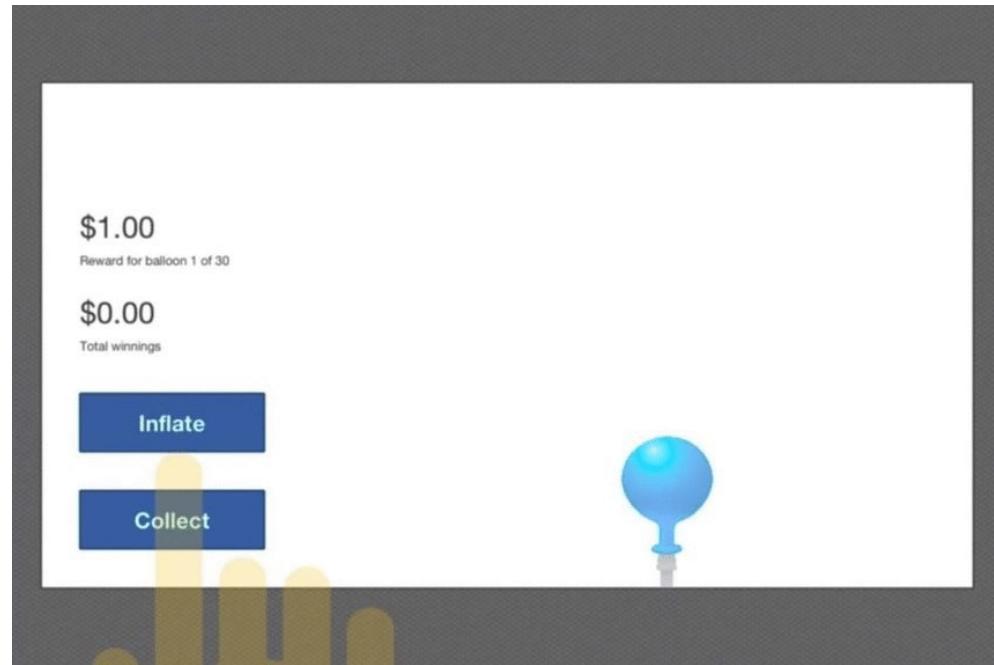
fMRI data measuring neuronal activity: functional connectivity *schizophrenic vs healthy*

165

Functional Magnetic Resonance Imaging (fMRI)

- ▶ 49 schizophrenic patients and 122 healthy subjects
- ▶ Task-based fMRI: Balloon Analog Risk Task (BART)
- ▶ schizophrenic patients have been reported to be more risk adverse than healthy subjects
- ▶ functional connectivity w.r.t. *left anterior cingulate gyrus*, key region involved in risk appraisal during decision making and in loss-aversion (Fukunaga, 2012, Cogn. Affect. Behav. Neurosci.)

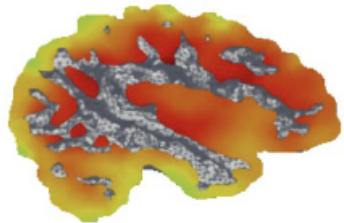
AIM: exploring neuronal activity/connectivity
of the various subjects while they perform
the task and evaluating possible differences
between schizophrenic and healthy subjects



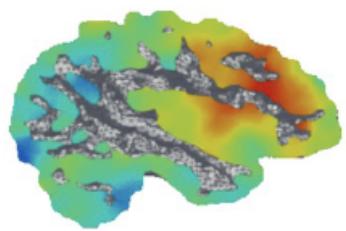
**Semel Institute for
Neuroscience and Human Behavior**

regularized fPCA

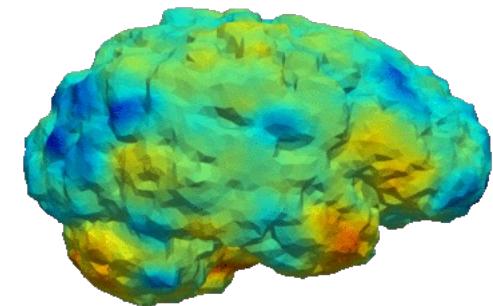
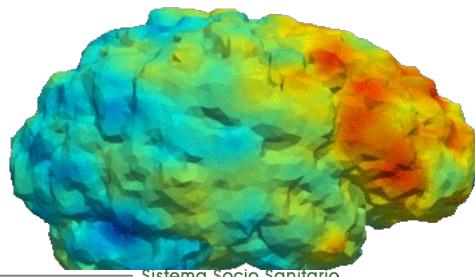
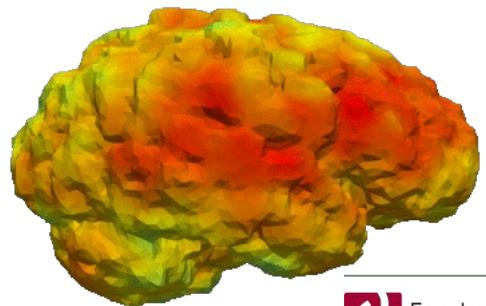
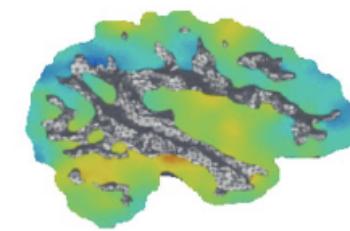
1st PC



2nd PC



3rd PC



Fondazione I.R.C.C.S.
Istituto Neurologico Carlo Besta



Sistema Socio Sanitario
Regione
Lombardia

The 1st PC displays greater variation in the inferior frontal gyrus (region implicated in go/no go tasks and in risk aversion)

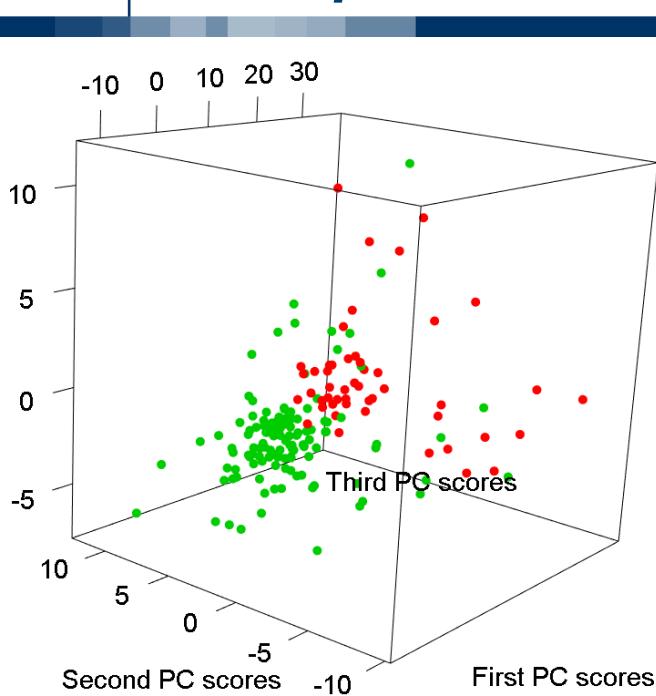
The 2nd PC contrasts Lingual gyrus & cunes (visual preprocessing) / insular cortex (emotional awareness)

F. Panzica

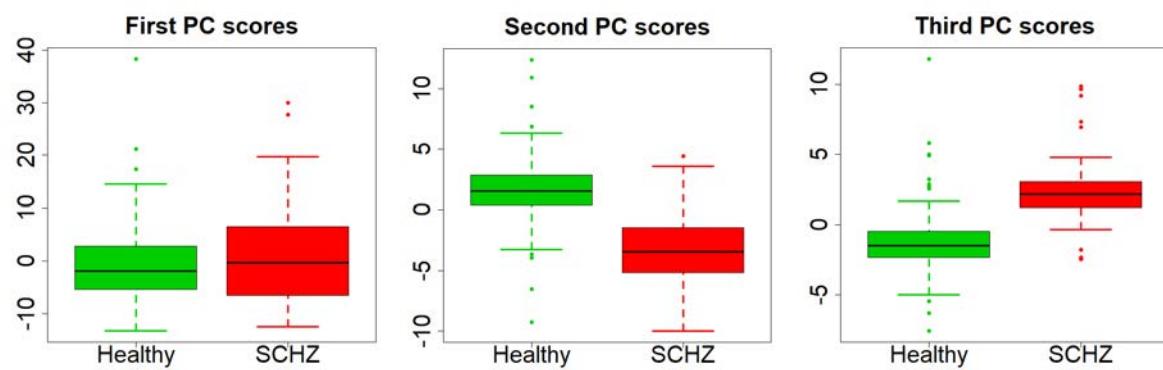
The 3rd PC contrasts the rear and the frontal part of the cerebellum

fMRI data: neuronal connectivity schizophrenic vs healthy

168



PC scores with *regularized fPCA*



Linear Discriminant analysis

regularized fPCA

accuracy: 93%; precision: 89%

True

Predicted	Healthy	SCHZ
Healthy	117	7
SCHZ	5	42

presmoothing PCA

accuracy: 82%; precision 76%

True

Predicted	Healthy	SCHZ
Healthy	113	21
SCHZ	9	28

These techniques may support advances in the study and analysis of the brain functionality

Credits:

Statistics Learning @ POLIMI

<https://mox.polimi.it/research-areas/statistics/people-statistics/>

