



UNIVERSITÀ DEGLI STUDI  
DI TRENTO

ICT  
Doctoral School

CiMeC  
Center for Mind/Brain Sciences

FONDAZIONE  
BRUNO KESSLER

# Machine Learning for Tract Segmentation in dMRI Data

Bao Nguyen

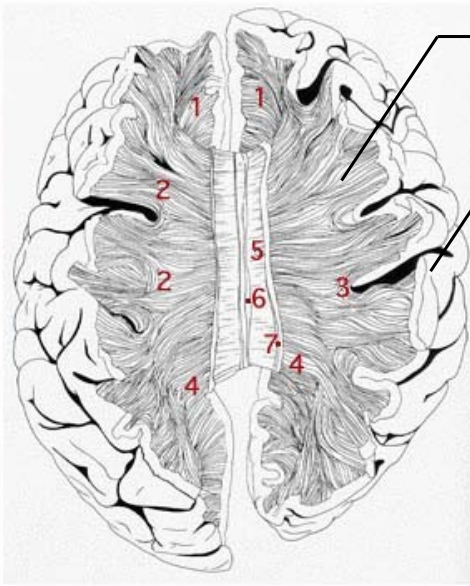
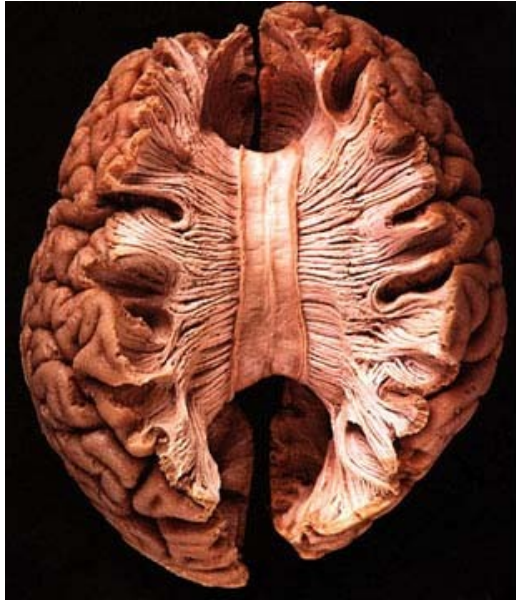
NeuroInformatics Laboratory (NILab)

*Trento, January 2013*

# Contents

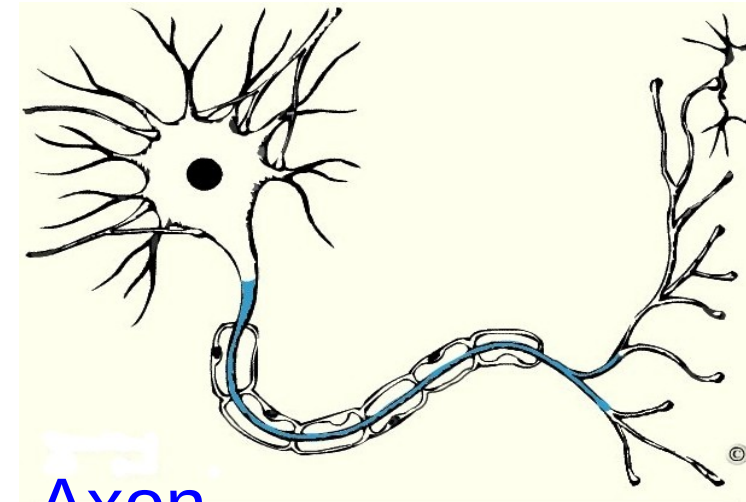
- Introduction
- State of the art (SoA)
- Problem statement
- Proposed solution
- Preliminary results
- Conclusion and Future works

# White matter



white matter

grey matter



Axon

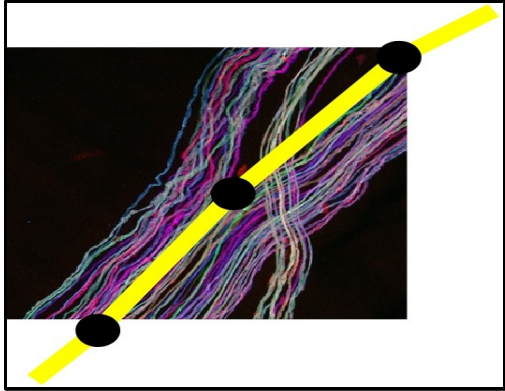
## dMRI technique

(diffusion Magnetic Resonance Imaging)  
in vivo (not invasive)

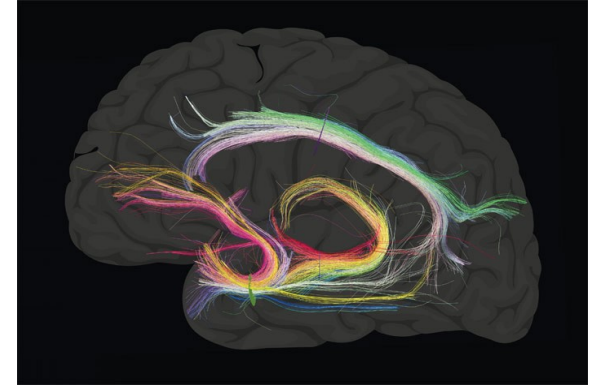
*Denis Le Bihan, 1984*

- Number:  $\sim 10^{12}$  axons
- Size:  $\sim 2\text{-}20\mu\text{m}$

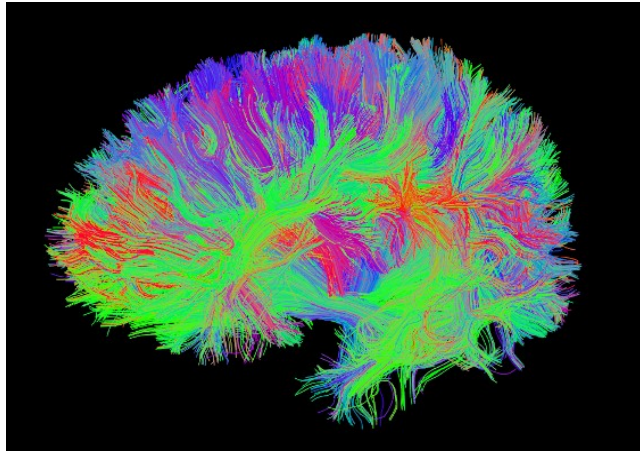
# Streamline & Tractography



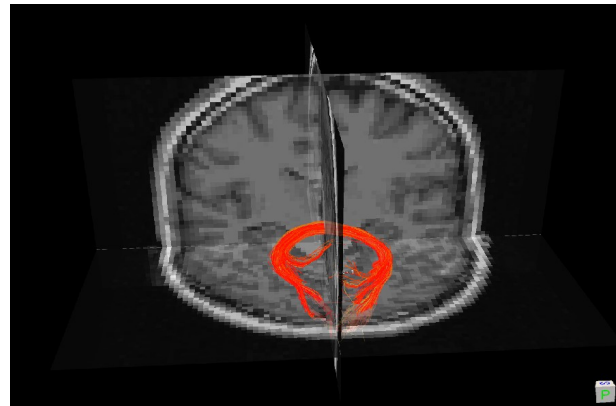
**Streamline**: a polyline representing thousands of axons.  
(fiber, track)



**Bundle**: a group of 'close' streamlines

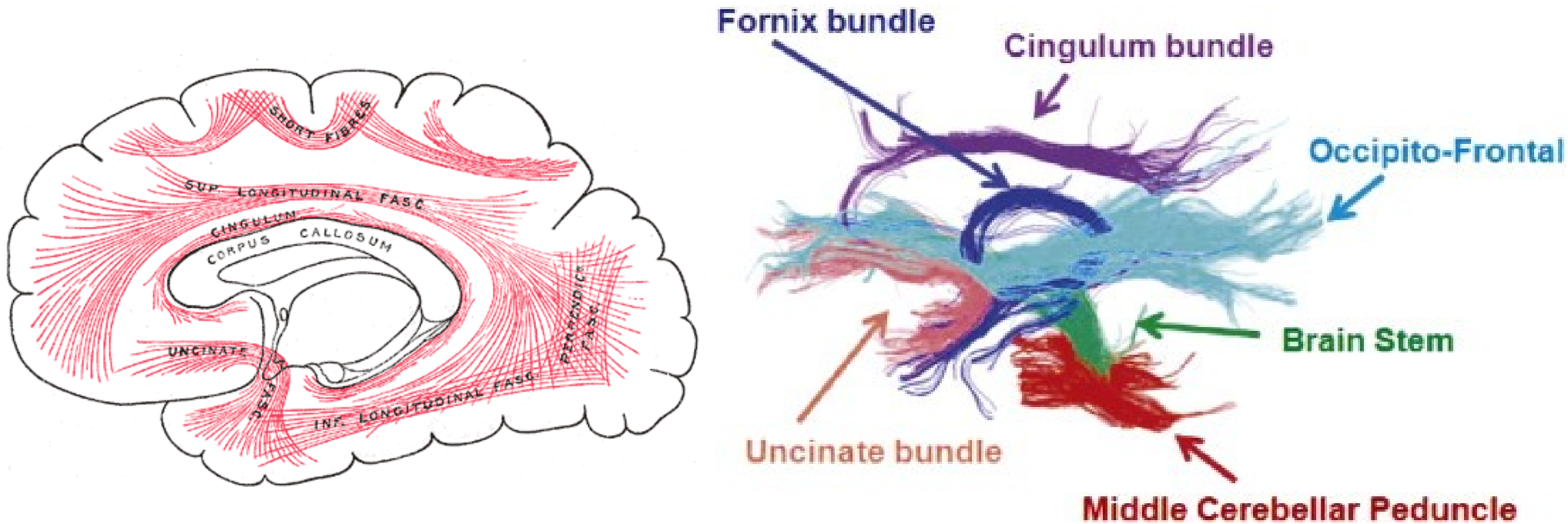


**Tractography**: presentation of whole brain by streamlines.



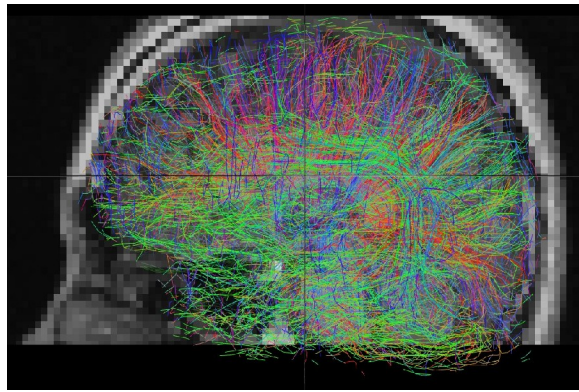
**Tract**: the real anatomical group of axons.

# Tractography Segmentation



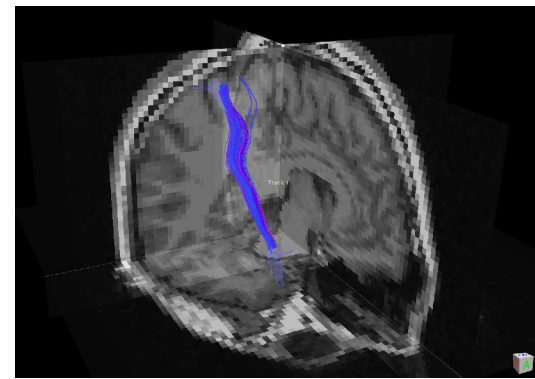
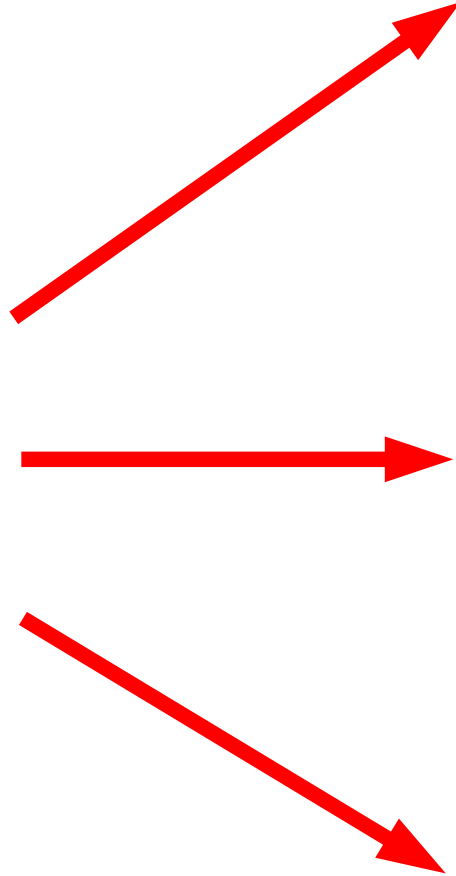
**To group** streamlines belonging to a **common anatomical area** into **one segmentation**



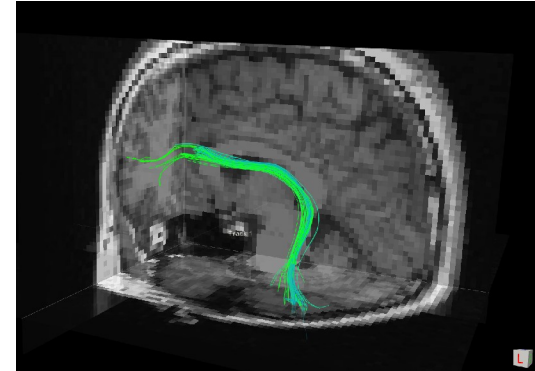


Tractography

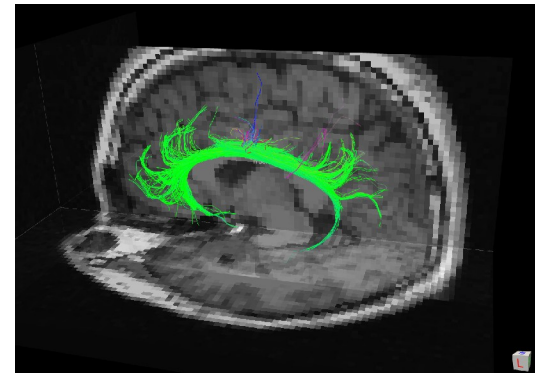
## Example of tractography segmentation



Corticospinal Tract (CST)

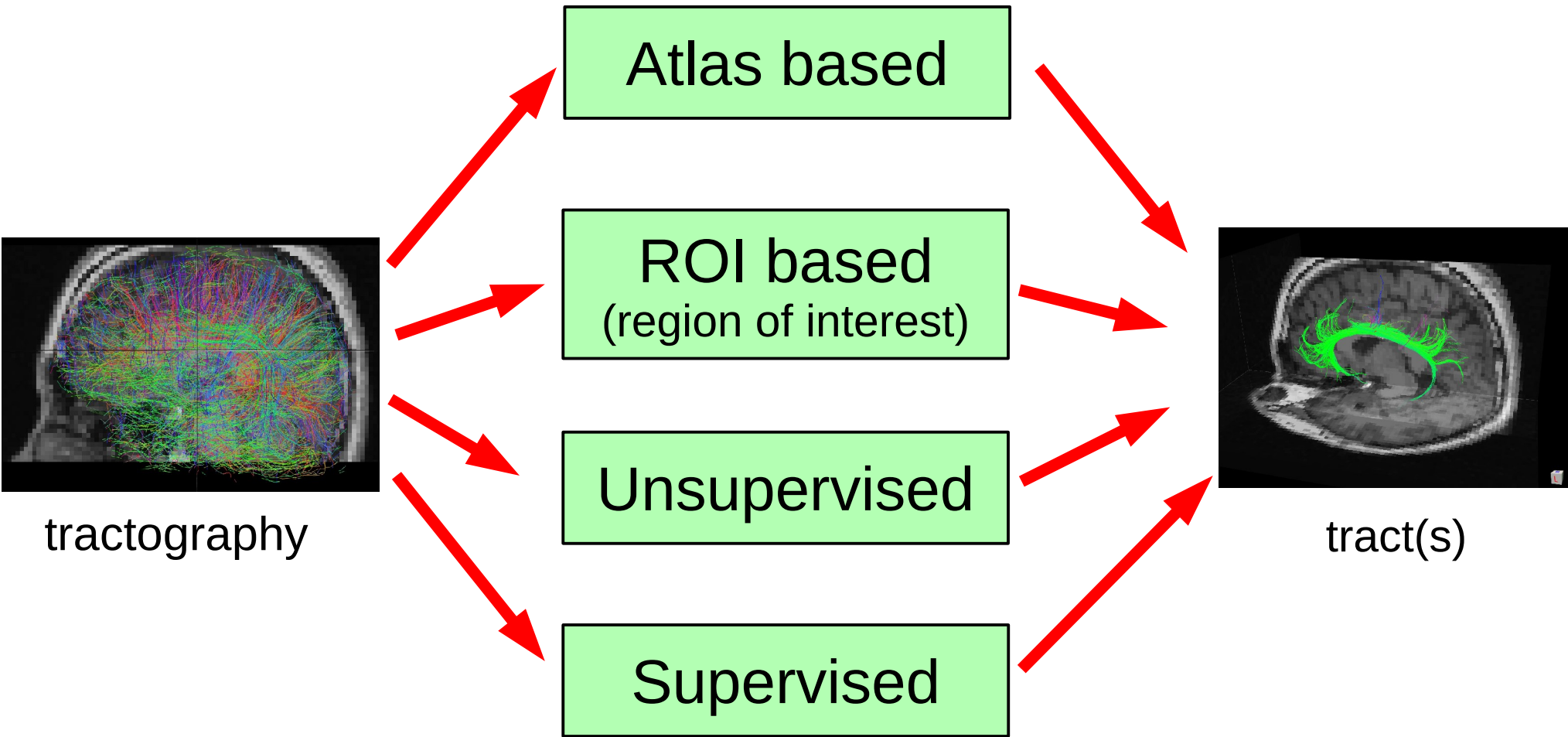


Arcuate Fasciculus Tract (AFT)

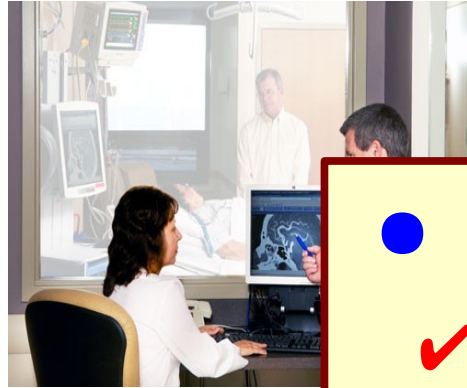


Cingulum Tract (CGT)

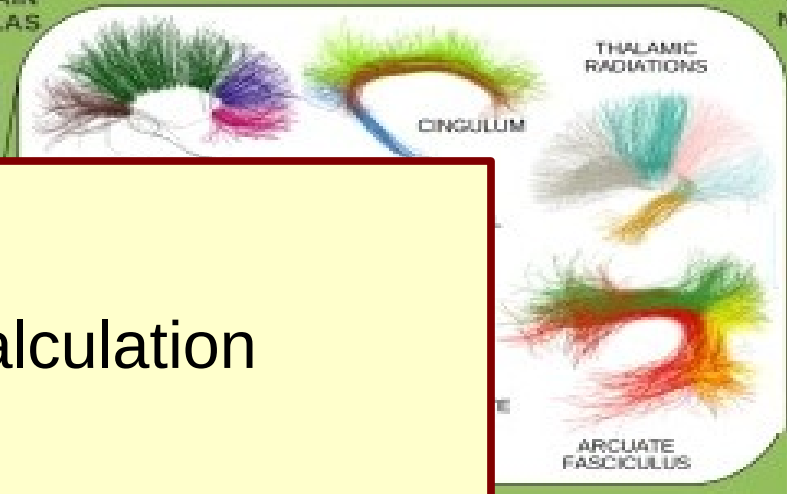
# Tract segmentation approaches



# Atlas based *Tract Segmentation*



MULTI-SUBJECT HUMAN BRAIN  
WHITE MATTER BUNDLE ATLAS

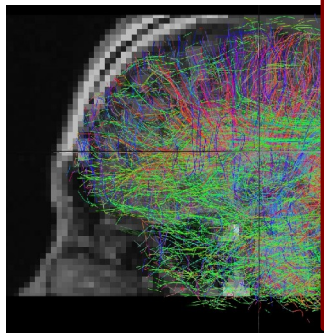


- Pros

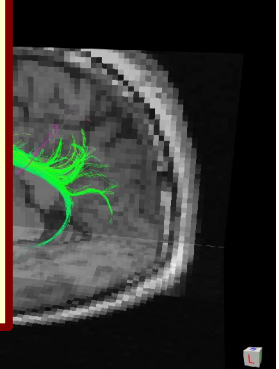
- ✓ No pairwise distance calculation
- ✓ Relate to anatomy

- Cons

- ✓ Whole brain, not target tract
- ✓ Brain variety → co-registration problem



tractography



tract(s)



# ROI based *Tract Segmentation*

(ROI:Region of Interest)  
Wakana et al., 2007

- Pros

- ✓ No pairwise distance calculate
- ✓ Relate to anatomy (indirect)
- ✓ Target tract

- Cons

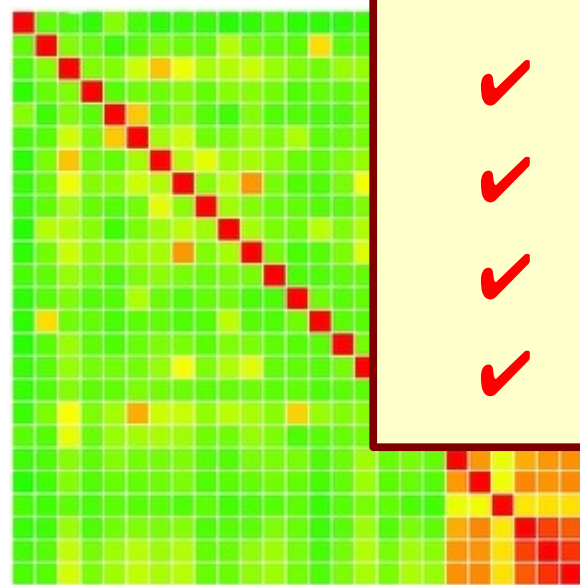
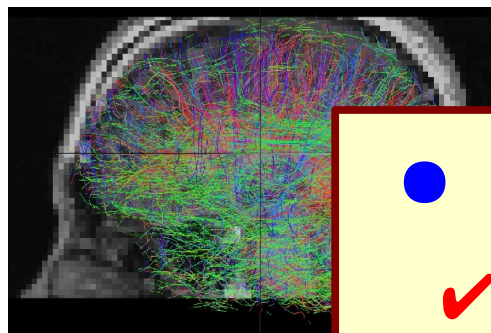
- ✓ Prior knowledge of trajectory
- ✓ Work on well characterized tracts
- ✓ Co-registration

ROIs  
are  
drawn

tract

tractography

# Unsupervised *Tract Segmentation*



similarity matrix

- Pros

- ✓ No co-registration

- Cons

- ✓ Calculate pairwise distance
- ✓ Whole brain not target tract
- ✓ Not relate to anatomy
- ✓ Number of cluster

k-means, etc.)

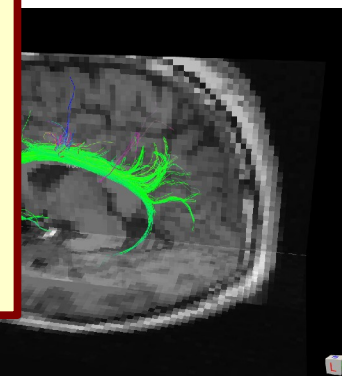
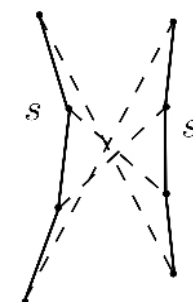
## Streamlines distance

$$d_{\text{mean}}(s, s')$$

$$d_{\text{mean}}(s', s)$$

$$d_{\text{mean}}(s, s')$$

$$d_{\text{flipped}}(s, s')$$



tract(s)

# *Supervised* Tract Segmentation

- Pros

- ✓ Target tract
- ✓ Relate to anatomy

- Con

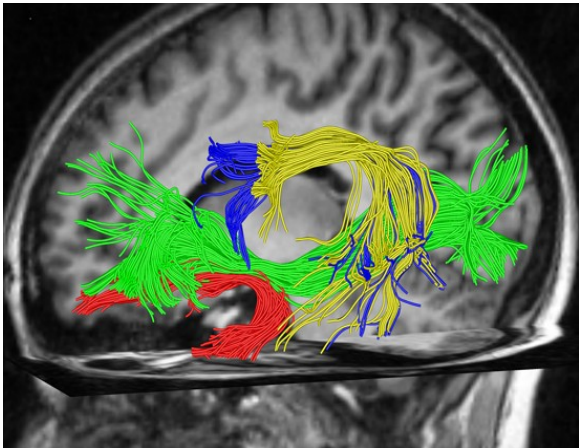
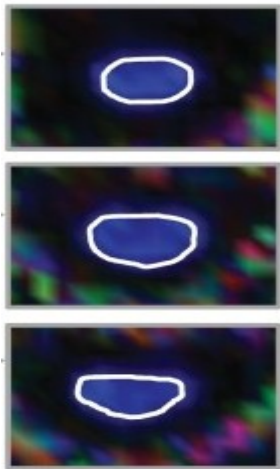
- ✓ Calculate pairwise distance
- ✓ Brain variety → co-registration problem
- ✓ Ground truth data

tractography

tract(s)

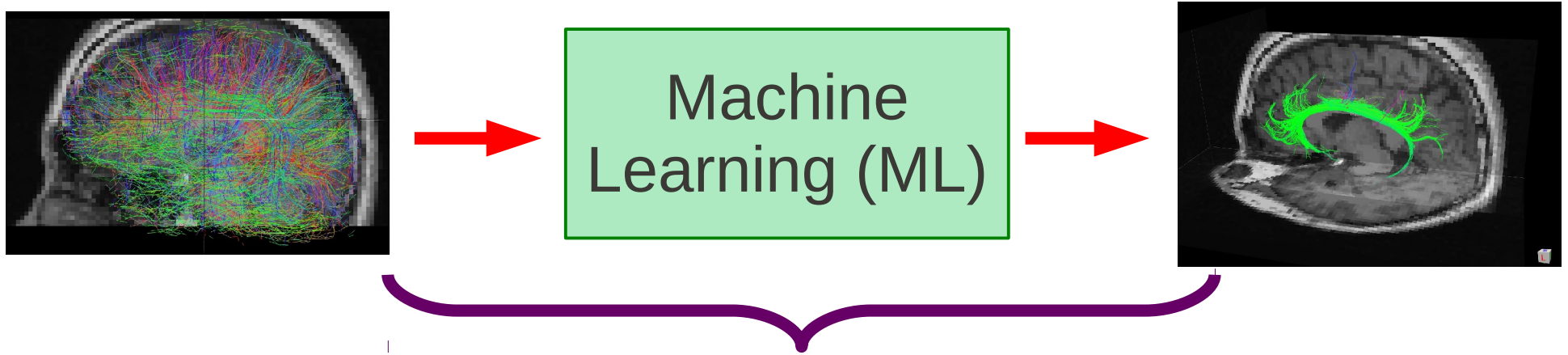
# New strategy: BOI - Bundle of Interest

Focus **directly** on which **bundle** (cluster of streamlines) that user wants to **work on**



Approach	ROI	BOI
Anatomy related	Yes (indirect)	Yes
Visualization	No	Yes
Interaction	No	Yes
No prior knowledge of trajectory	No	Yes

# Goals



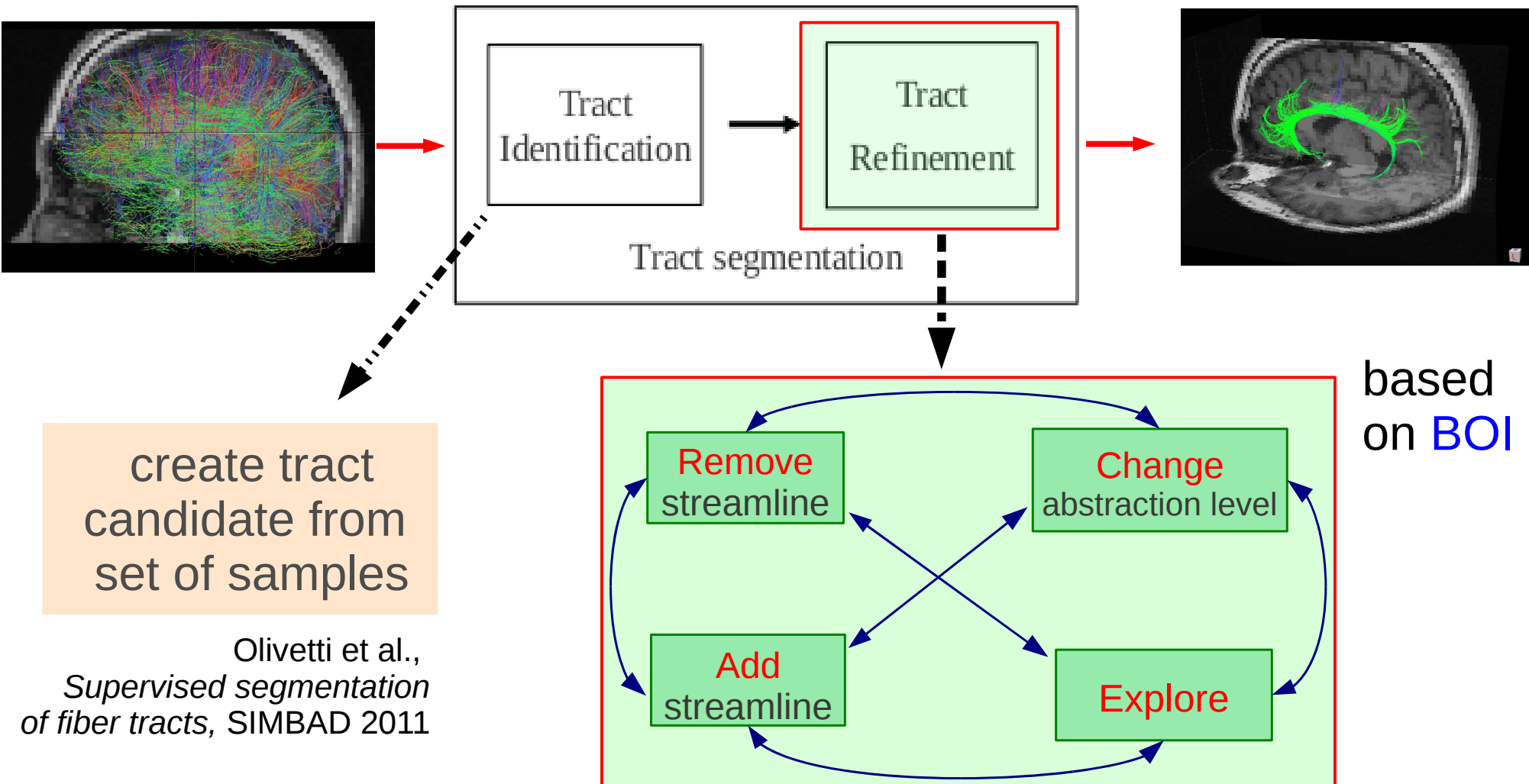
Improve the **support** of **ML** for tract segmentation

## Challenges

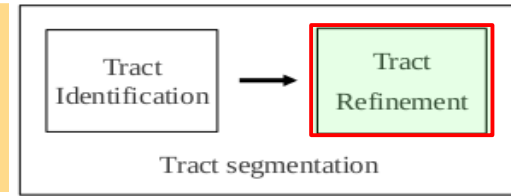
- **overcome** disadvantages of Atlas, ROI
- **combine** both un-supervised and supervised
- design an **effective method** for tract segmentation



# Process design: *interactive segmentation*



# Interactive tract refinement



## Demo of Spaghetti

# Problem statement

- Given a set of  $N$  objects  $\mathcal{X} = \{x_1, \dots, x_N\}$
- Traditional clustering: find **one partition** of  $\mathcal{X}$

$$C = \{C_1, \dots, C_K\} \text{ with } K \leq N$$

where  $C_i$  is a cluster of  $\mathcal{X}$ :  $C_i = \{x_1^i, \dots, x_j^i\}$ ,  $j \leq N$

i  $C_i \neq \emptyset, i \in [1, \dots, K]$

ii  $\bigcup_{i=1}^K C_i = \mathcal{X}$

iii  $C_i \cap C_j = \emptyset, i, j \in [1, \dots, K], i \neq j$

# Interactive clustering

- Our approach: find **a set m partitions** of  $\mathcal{X}$

$$\mathcal{P} = \{ P_1, \dots, P_m \}$$

where  $P_i$  is one partition of  $\mathcal{X}$  :  $P_i = \{C_1^i, \dots, C_{d_i}^i\}$

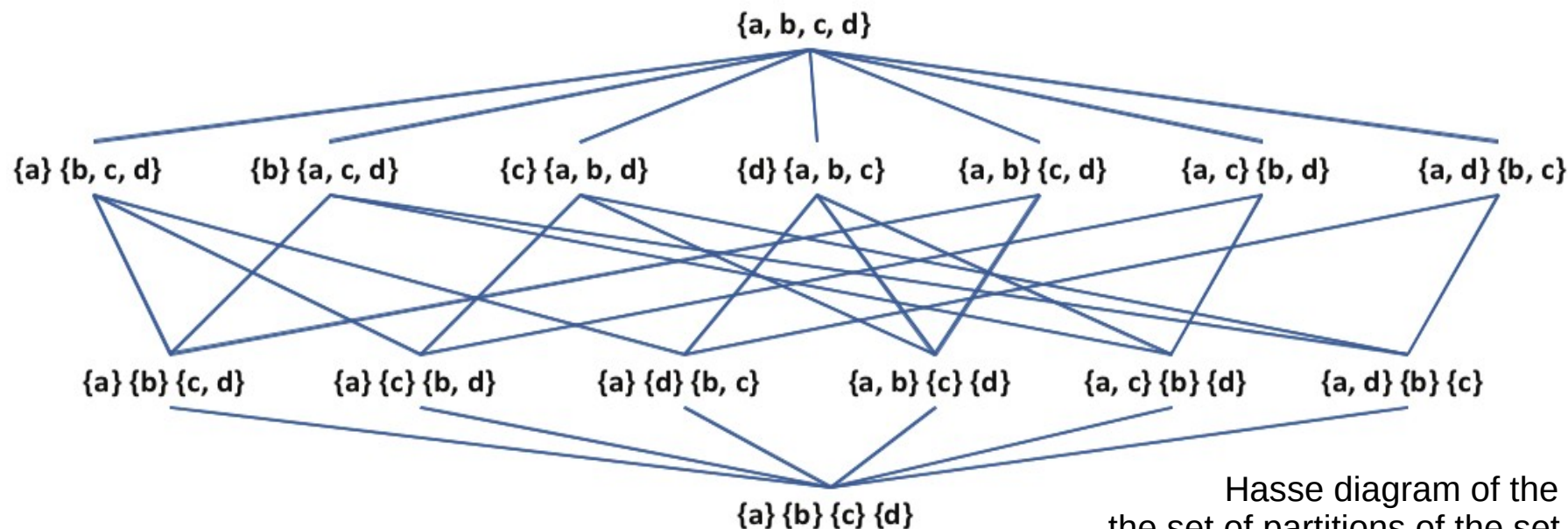
- i  $P_i$  represents the *ith abstraction level* of  $\mathcal{X}$
- ii constraint  $\gamma$ :  $\forall i \in [1, m-1], P_i \preceq P_{i+1}$  ("*nested in*")

- Denoted as a triple  $\langle \mathcal{P}, \mathcal{X}, \gamma \rangle$

# Interactive clustering: partial order relation

- $\mathcal{P}_{\mathcal{X}}$ : set of all possible partitions of  $\mathcal{X}$
- Over  $\mathcal{P}_{\mathcal{X}}$ , a partial order relation  $\preceq$  ("nested in")

$$\forall P_a, P_b \in \mathcal{P}_{\mathcal{X}}, P_a \preceq P_b \Leftrightarrow \forall C_i^b \in P_b, \exists C_{i_1}^a, \dots, C_{i_k}^a \in P_a: \mathbf{C}_i^b = \bigcup_{t=1}^k \mathbf{C}_{i_t}^a$$



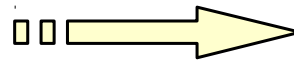
Hasse diagram of the lattice of the set of partitions of the set  $\{a, b, c, d\}$



# Interactive clustering: **update** partitions

- **Remove** an old object  $x_{r.m} \in \mathcal{X}$

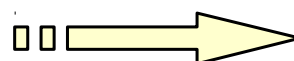
$$\mathcal{X} = \{x_1 \dots, x_N\}$$



$$\mathcal{X}' = \mathcal{X} \setminus \{x_{r.m}\}$$

- **Add** a new object  $x_{add}$

$$\mathcal{X} = \{x_1 \dots, x_N\}$$



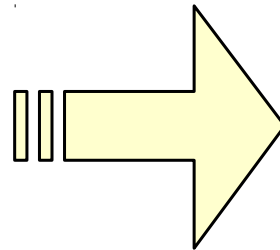
$$\mathcal{X}' = \mathcal{X} \cup \{x_{add}\}$$

$$\mathcal{P} = \{P_1, \dots, P_m\}$$

$$\langle \mathcal{P}, \mathcal{X}, \gamma \rangle, \gamma: P_i \preceq P_{i+1}$$

$i \in [1, m-1]$

current partitions



$$\mathcal{P}' = \{P'_1, \dots, P'_m\}$$

$$\langle \mathcal{P}', \mathcal{X}', \gamma' \rangle, \gamma': P'_i \preceq P'_{i+1}$$

$i \in [1, m-1]$

updating partitions

# Interactive clustering: add new object

- Current viewing abstraction level is *ith*
- $C(x, P_i)$ : cluster in partition  $P_i$  having object  $x$

- All upper partitions

$\gamma'_1: \forall j \in [i, \dots, m], \forall x \in \mathcal{X}:$

$$C(x, P'_i) = C(x_{add}, P'_i) \rightarrow C(x, P'_j) = C(x_{add}, P'_j)$$

- All lower partitions

$\gamma'_2: \forall k \in [1, \dots, i-1], \forall x \in \mathcal{X}:$

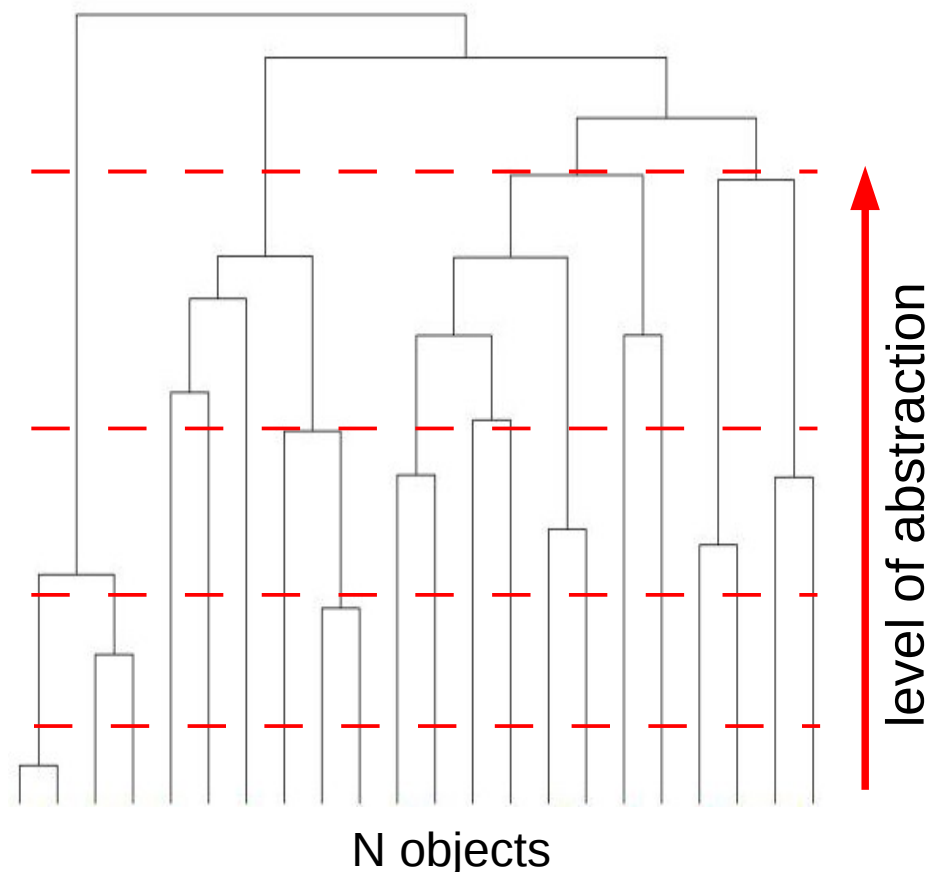
$$C(x, P'_i) \neq C(x_{add}, P'_i) \rightarrow C(x, P'_k) \neq C(x_{add}, P'_k)$$

# Hierarchical clustering

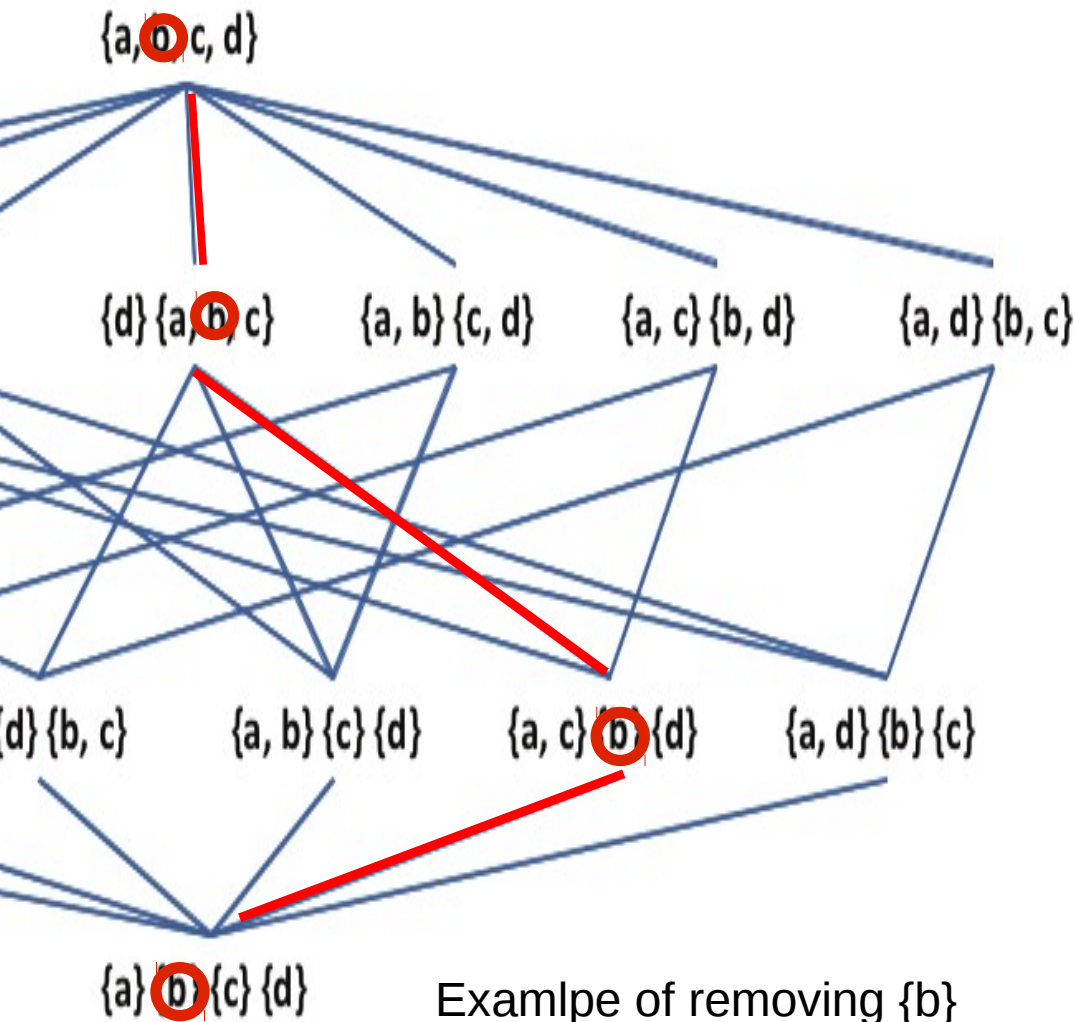
Find **m partitions** of  $\mathcal{X}$ :  $\mathcal{P} = \{P_1, \dots, P_m\}$ ,  $P_i \preceq P_{i+1}$ ,  $i \in [1, m-1]$

## Algorithm

1. Assign each  $s_i$  to one cluster
2. Merge two closest clusters
3. Compute distances
4. Repeat until all in one cluster



# Interactive clustering: *remove object*



Example of removing  $\{b\}$

Removing object  $x_{r.m}$   
from all partitions

$\forall P_i, i \in [1, \dots, m],$

$\forall j \in [1, \dots, d_i]:$

if  $x_{r.m} \in C_j^i$  then

$C_j^i = C_j^i \setminus \{x_{r.m}\},$

if  $C_j^i = \emptyset, P_i = P_i \setminus \{C_j^i\}$

# Interactive clustering: *add object*

Denote  $\delta_{\max}$  as the maximum distance in cluster  $C_i$

$$\delta_{\max}(C_i) = \max \{d(x, x')\}, \forall x, x' \in C_i, x \neq x'$$

1. Find the closet cluster of  $x_{\text{add}}$  in  $P_1$ :  $C_{\text{close}}(x_{\text{add}}, P_1)$

2. Start from the direct parent of  $C_{\text{close}}(x_{\text{add}}, P_1)$ :

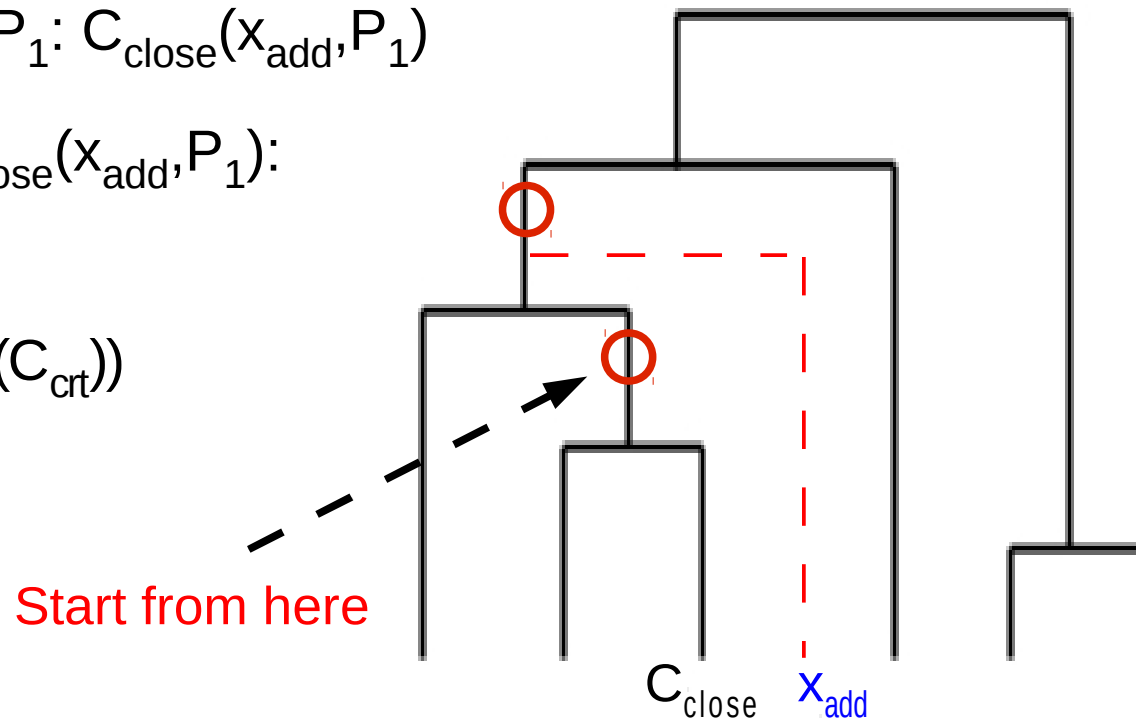
$$C_{\text{crt}} = C_{\text{close}}(x_{\text{add}}, P_1).parent$$

3. If  $(C_{\text{crt}} = \emptyset) \vee (d(x_{\text{add}}, C_{\text{crt}}) < \delta_{\max}(C_{\text{crt}}))$

3.1. Merge:  $C_{\text{crt}} = C_{\text{crt}} \cup \{s_a\}$

3.2. Stop

4.  $C_{\text{crt}} = C_{\text{crt}}.parent$





# Preliminary Results

- (method) Dissimilarity representation

E. Olivetti, **T. B. Nguyen**, E. Garyfallidis,  
*The Approximation of the Dissimilarity Projection*,  
Pattern Recognition in NeuroImaging, PRNI 2012.



### The Approximation of the Dissimilarity Projection

\*Neurodynamics Laboratory (NE-Lab), Bruno Kessler Foundation, Trento, Italy  
 †Centro Interdipartimentale Memorie e Cervello (CIMeC), University of Trento, Italy  
 [MRF] Equitation and Brain Sciences Unit, University of Cambridge, UK

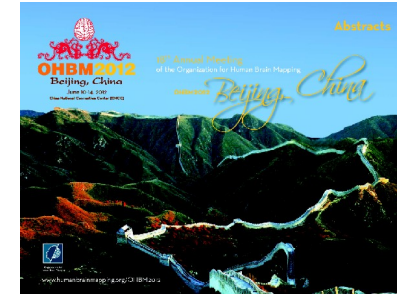
Abstract: *Staphylinidae* possess a common feature, known as the 'staphylinid gait', in which the 30 setae of the male aedeagus within the male endite of the beak act as a template. The setal pattern of testis epigrams has a close relation from the common ancestor and pattern recognition constitutes a powerful means of comparison as far as supergeneric representation goes for the family. Many of the current knowledge regarding the setal pattern is based on a few critical cases. This presentation contrasts with the extensive nature of the morphology between its members, the genera, called decemline or triads, have different lengths and different number of points and for this reason they are not commonly represented in a common ventral aspect. In this work we present a new method for the study of the setal pattern, based on an automatic unfolding technique based on achieving a

[illegible]

In the following we present a concise informal description of the dimension reduction aspects, with a detail of the algorithmic aspects in Section III. The dimensionality approximation is an Euclidean embedding technique designed for inferring a set of objects by a set of observed objects. The objects to be inferred are called object  $z$ , and are new (unknown) to the vector of dimensions of the observed objects. This representation (to be used) is usually presented as a cluster of dimensions and clustering parameters. In the next section, we will discuss the algorithmic aspects of this representation. In the next section, we will discuss the algorithmic aspects of this representation. In the next section, we will discuss the algorithmic aspects of this representation.

- (software/tool) Spaghetti: an interactive visualization tool for segmentation tractography

E. Garyfallidis, S. Gerhard, P. Avesani,  
**T. B. Nguyen**, V. Tsiaras, I. N. Smith, E. Olivetti,  
*A software application for real-time, clustering  
based exploration of tractographies*, OHBM 2012.



A software application for real-time, clustering-based exploration of fractographies

<sup>1</sup>University of Cambridge, Cambridge, United Kingdom; <sup>2</sup>University of Zurich and CTI Zurich, Zurich; <sup>3</sup>Bruno Kessler Foundation (FBK) and University of Trento (CIMO), Trento, Italy; <sup>4</sup>University of Cambridge, Cambridge, United Kingdom

Interacting with tectonographies is a difficult procedure for various reasons: (a) tectonographies are usually represented by hundreds of thousands of instantiated tectonemes (packets) (see fig. 2A), (b) Sika sets are often cluttered with noisy tracks, i.e. tracks which have no relevance in anatomy, (c) the size of the entire tectonography is often too large to load in the memory of the graphics card, (d) analysing specific regions of a tectonography in 3D space is cumbersome because of the unique shape (displacement) of the tectonemes.

We present a scientific visualization tool that solves the problem of interacting with hierarchies by creating relative replications in terms of the underlying bundle visualizations. The process that we propose works on a journey starting from a small number of clusters of streamlined the user decided which clusters to explore. Exploring a cluster means that the next set of clusters is chosen, its content is a finer grained level.

The application starts visualizing the basin as a set of a few representative tracks, i.e. the cluster representatives. These representative tracks are provided by a very fast trackography clustering algorithm called JuckiIndies [1,9].

see Fig. 10) which can be measured in tens of milliseconds. Each representative track acts as the access point in the observation column. The column width should be chosen to achieve the goal of the investigation, e.g. in a range only, which we call "bandwidth of interest" (BWI). After visually inspecting the associated topography (Fig. 11), the user interactively selects one or more representative tracks (Fig. 12). The selected representative track is in white. When one or more representative tracks are selected, the user chooses the content of the related column (Fig. 13). In order to inspect the detailed structure or content of the selected one or more tracks, the user can re-select the selected tracks and/or sub-tracks. This is done prior to re-selecting the content of the related column. Furthermore, the user can expand the small columns to a single track representation (Fig. 14). By adding the columns (except the visual repeat step Fig. 15), and the re-selecting step (Fig. 16) as required in order to unveil the local structures (Fig. 17).

- (case study) ALS (amyotrophic lateral sclerosis) disease (in preparation)

# Conclusion

- An effective method for **tract segmentation**:  
    **tract candidate** (supervised) and **tract refinement** (clustering)
- An **interactive visualization tool** for tract segmentation
- ALS case study

## Future works

- Implement the **modified HAC** for tractography
- Revise the solution for '**adding object**' to partitions
- Integrate tract candidate (**supervised**) into Spaghetti

# Credits

- **Nivedita Agarwal**, *S.Chiera Trento Hospital, Italy;*  
*University of Utah, USA*
- **Eleftherios Garyfallidis**, *University of Cambridge, UK;*  
*University of Sherbrooke, Canada*
- **Emanuele Olivetti**, *Fondazione Bruno Kessler, Italy*
- **Paolo Avesani**, *Fondazione Bruno Kessler, Italy*
- **Luigi Cattaneo**, *CiMeC, University of Trento, Italy*
- **Francesca Maule**, *CiMeC, University of Trento, Italy*

Thank you!