Michele Allegra [#]

A physicist and the brain

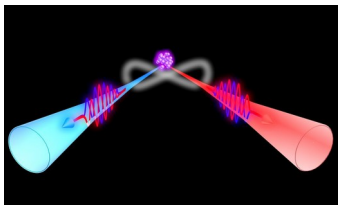
exploratory analysis tools from physics to neuroscience

Institut de Neurosciences de la Timone, Aix Marseille Université, CNRS, Marseille, France

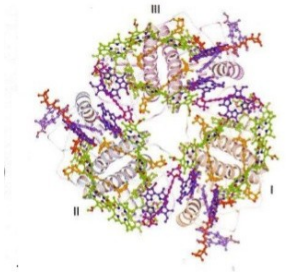
About myself

- **2011-2014 PhD in Physics, University of Turin** and ISI, Turin (with P Giorda)
  visiting MIT, Cambridge MA (with S Lloyd)

  information loss in quantum systems interacting with macroscopic environment



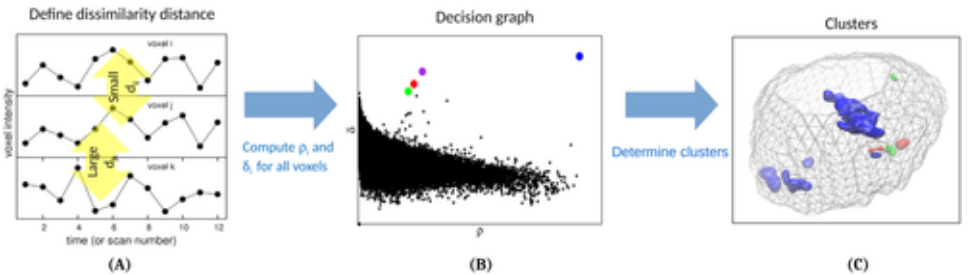
Photons
M Allegra, P Giorda, MGA Paris, PRL 105 (10), 100503



Photosynthetic Trimer
M Allegra, P Giorda, S Lloyd, PRA 93 (4), 042312

- **2015-2017 Postdoc, SISSA, Trieste** (with D Amati, A Laio)

  Advanced clustering techniques and applications to fMRI

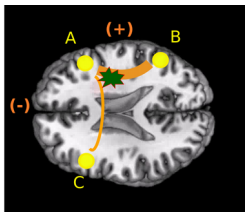


M Allegra, S Seyed-Allaei, …., C Reverberi, A Laio, D Amati, HMB 38 (3), 1421-1437

M Allegra, S Seyed-Allaei, NW Schuck, D Amati, A Laio, C Reverberi, NeuroImage, 116854

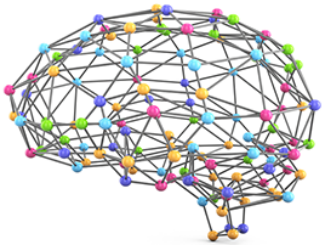
- **2018-2020 Postdoc, Intitute Neurosciences Timone, CNRS, Marseille** (with A Brovelli, M Corbetta)

  Information transfer and brain dynamics in stroke

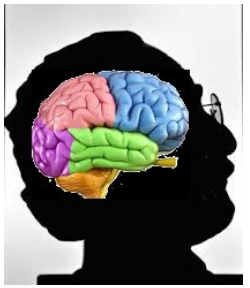


M Allegra, C Favaretto, .., M Corbetta, A Brovelli, in prep.

Physics of the brain?

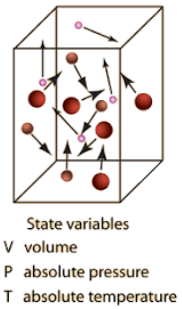


Physics of the brain

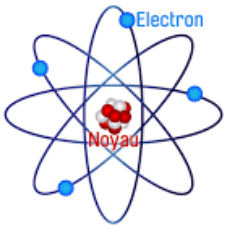


A physicist's brain

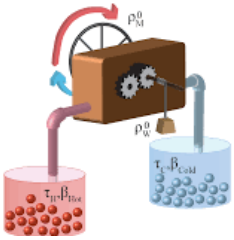
Abstract concepts are largely metaphorical George Lakoff, *Metaphors we live by*



Galaxies are a gas



Electrons are waves



Black holes are a heat engine

Formalism ⟵ Translation ⟶ Intuition

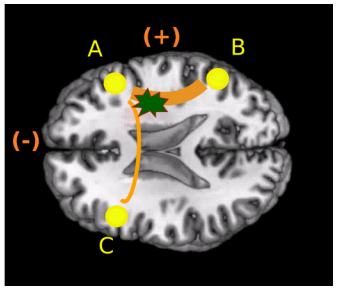
Exploring dynamics from physics to neuroscience

Characterize brain dynamics

Ideas and formalism familiar from study of chaotic sytems

- **Part 1: Information transfer and brain dynamics upon stroke**

  [M. Allegra, C. Favaretto, M. Corbetta, A. Brovelli, in prep. (2020)]

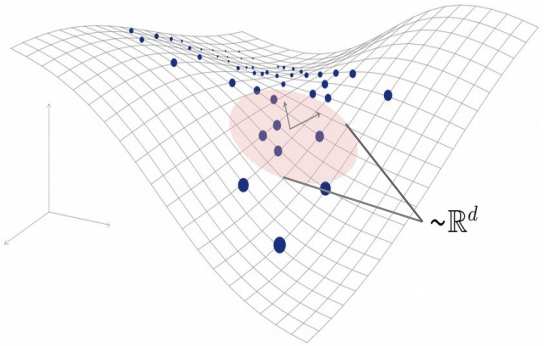


  - Measures of information transfer
  - Post-stroke anomalies in information transfer

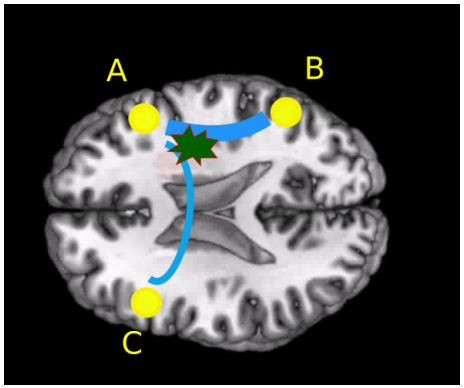
- **Part 2: the dimension of data and «attractors» of complex systems**

  [M. Allegra, E. Facco, A. Laio and A. Mira, sub. (2020); arXiv:1902.10459]

  - Estimates of the intrinsic dimension of data
  - Possible applications to brain dynamicss



Brain dynamics upon stroke



- Stroke induces focal lesions
- Lesions are subcortical and affects structural connections
- Several areas cannot communicate information (disconnection syndrome)

- **How is information transfer perturbed in the brain after stroke?**
  [M. Allegra, C. Favaretto, M. Corbetta, A.Brovelli, in prep. (2020)]

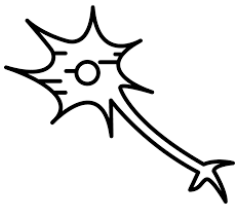
  From previous studies with functional connectivity [Siegel et al., PNAS 113.30 (2016)]

- The two hemispheres are less synchronized with one another
- The two hemispheres are more synchronized internally
- Does this correspond to shifts in information transfer?
- Goal : understand behavioral deficits and how to restore correct functionality

Information production

- The concept of **information production** was born in the study of chaotic systems

  [M. Falcioni, V. Loreto, A. Vulpiani. in *The Kolmogorov legacy in physics,* Springer].

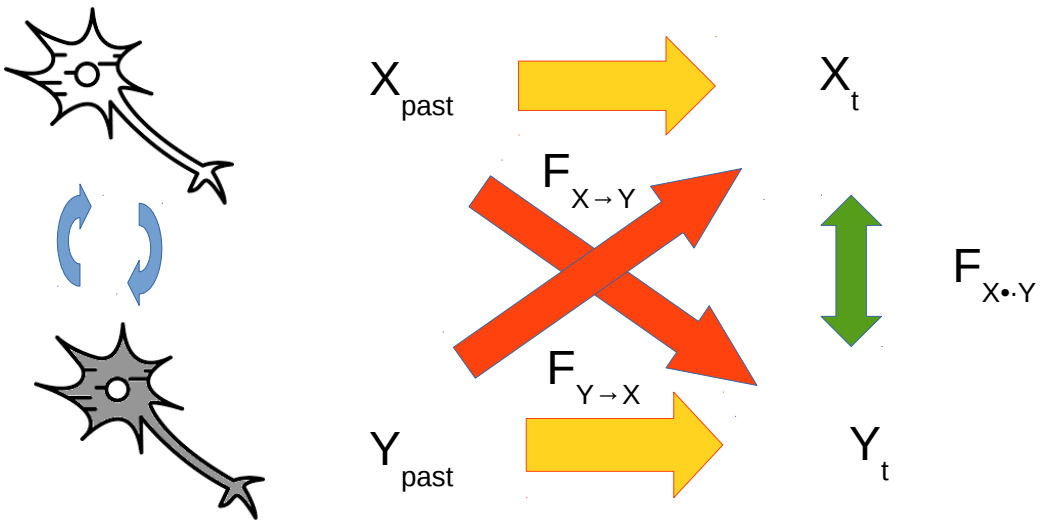


$X_{past} = X_{t-1}, X_{t-2}, \ldots, X_{t-L}$ $\xrightarrow{F_{X \to X}}$ X_t

- Uncertainty of $X_t = H(X_t)$ [H=Shannon entropy]

- **prediction:** X_{past} gives information $F_{X \to X}$ about X_t**:**
- new **information production**: $H(X_t|X_{past})$
- $F_{X \to X} = H(X_t) - H(X_t|X_{past})$

Information production in composite systems



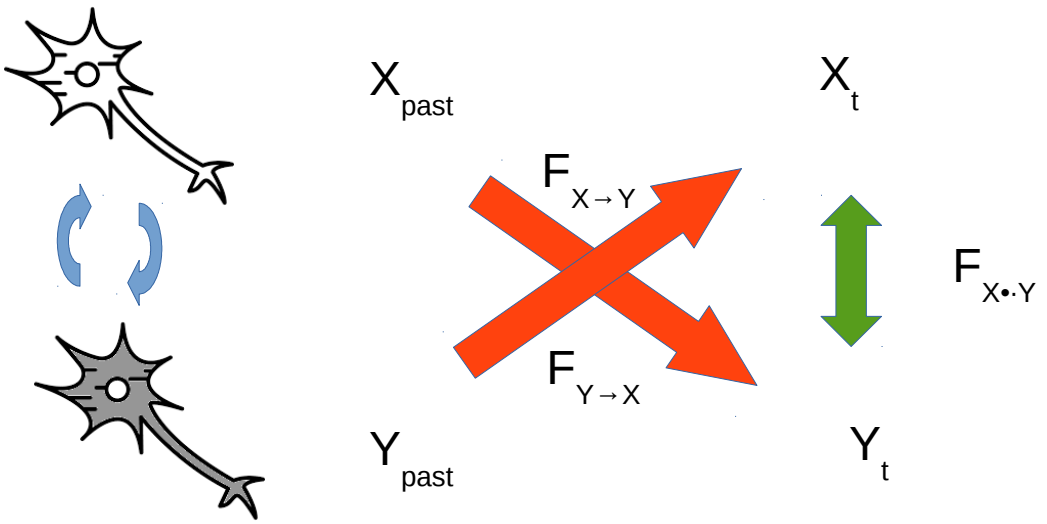
- information produced *jointly* $H(X_t Y_t | X_{past} Y_{past})$ vs *independently* $H(X_t|X_{past}) + H(Y_t|Y_{past})$
- difference since X predicts information about Y and viceversa [T. Schreiber, PRL 85, 461 (2001)]
- Three terms explain the difference:

- **$F_{Y\to X} = H(X_t|X_{past}) - H(X_t|X_{past} Y_{past})$**

  information about X_t predicted exclusively by Y_{past} (and not X_{past}) [reverse for $F_{X\to Y}$]

- **$F_{X\bullet\cdot Y} = H(X_t|X_{past} Y_{past}) - H(X_t|Y_t X_{past} Y_{past})$**

  information about X_t predicted exclusively by Y_t (and not X_{past}, Y_{past}) [symmetric]



- $F_{X \to Y}$ (or $F_{Y \to X}$) was called *transfer entropy* by Shreiber [T. Schreiber, Phys. Rev. Lett. 85, 461, 2001)]

  directed information flow from X to Y (or Y to X)

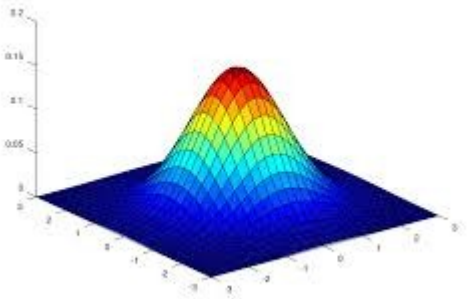
- $F_{X \cdot Y}$ was called *instantaneous feedback* by Geweke [J. Geweke, J. Am. Stat. Ass. 77.378 (1982)]

  «instantaneous» ($\Delta t<1$) information flow/sharing between X and Y

- $F_{X \to Y}$ $F_{Y \to X}$ coincide with notion of *Granger causality* for Gaussian systems
  [L. Barnett et al., Phys. Rev. Lett. 103, 238701 (2009)]

Covariance-based Granger causality

- To estimate $F_{X \to Y}$, $F_{Y \to X}$, and $F_{X \cdot Y}$ from data samples, we need to compute Shannon entropies from $P(X_t Y_t X_{past} Y_{past})$ and its marginal distributions



- assume multivariate Gaussian, P=N(μ,Σ) [Σ=covariance matrix]

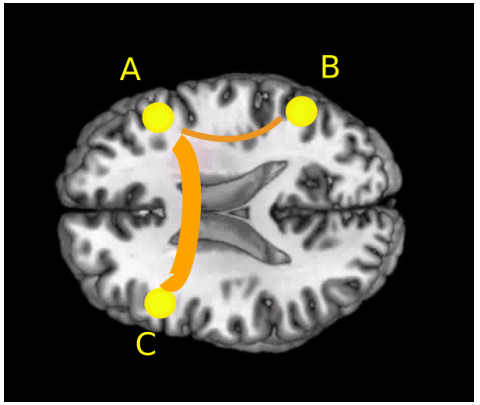
  H = 1/2 log(det(Σ)) + const.

- all entropies can be computed from Σ and its submatrices

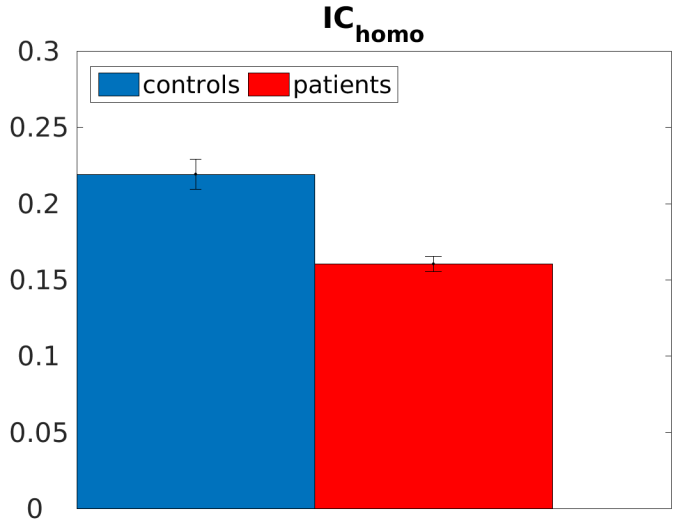
- *covariance-based* Granger causality measures

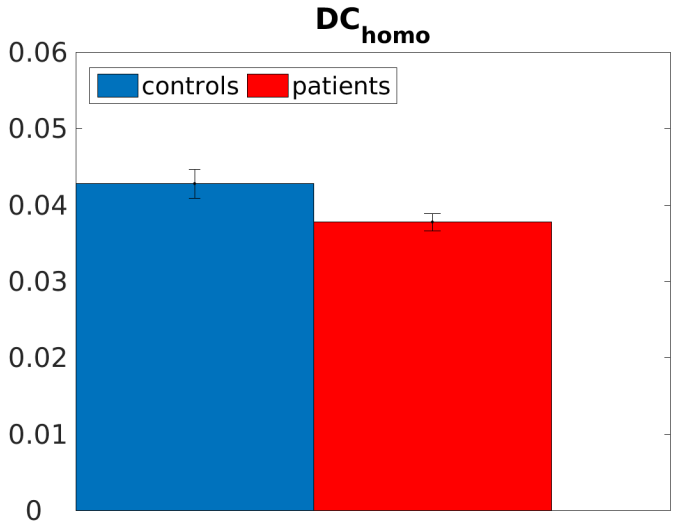
- $F_{X \cdot Y}$: Instantaneous Causality, **IC**
- $F_{X \to Y}$, $F_{Y \to X}$: Directed Causality, **DC**

Stroke and information transfer in the brain



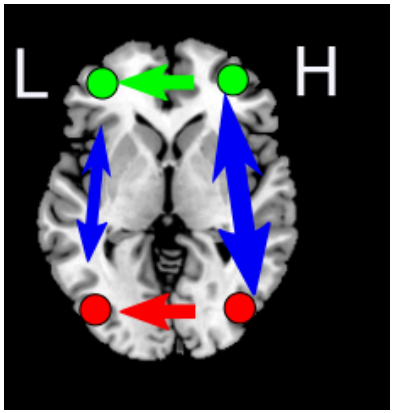
- **We used covariance-based GC combined with fMRI**
- We used a large (n>100) stroke database from St. Louis)
  [Siegel et al., PNAS 113.30 (2016); Corbetta et al. Neuron 85.5 (2015)]



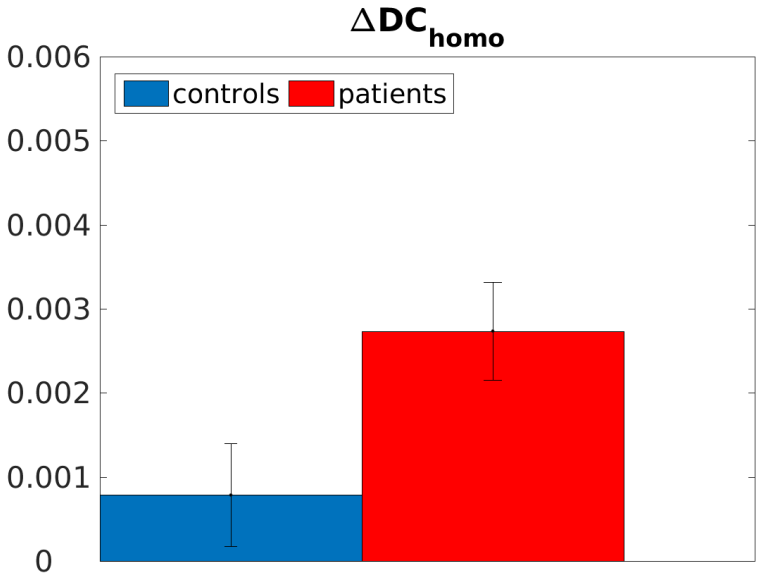


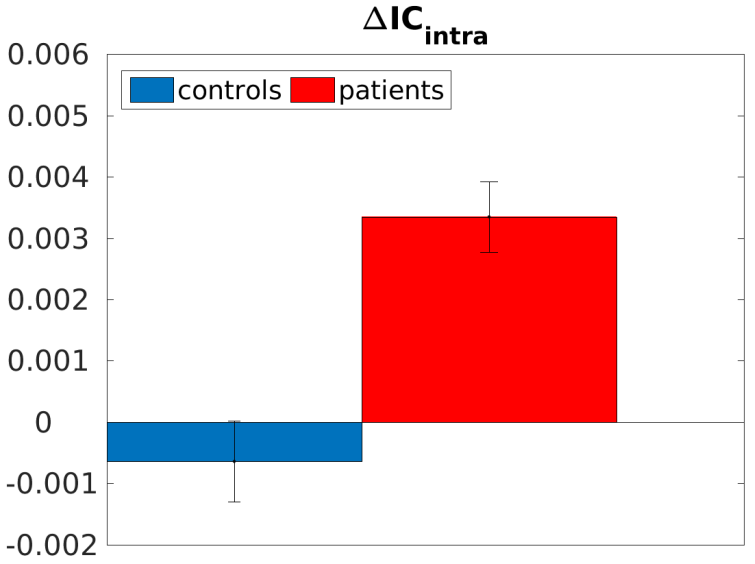
- Instantaneous Causality for homologous regions (IC_{homo}): reduced in patients
- Directed Causality for homologous regions (DC_{homo}): reduced in patients
- **Interhemispheric information flow (IC and DC) is reduced in patients**

Brain dynamics upon stroke



- **Unbalances between healthy and lesioned hemisphere**





- Net direction of homotopic DC (healthy to lesioned - lesioned to healthy) : ΔDC_{homo}
  DC from healthy to lesioned hemisphere higher than reverse in patients
- IC for regions of same hemisphere, then difference healthy – lesioned (ΔIC_{intra}):
  intra-hemispheric IC higher in the healthy hemisphere for patients [...similar effect for DC]
- **Information flow within and from the lesioned hemisphere is reduced**

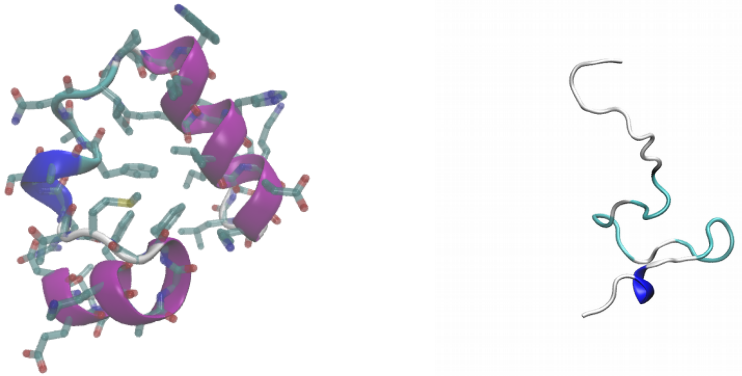
Main findings and future developments

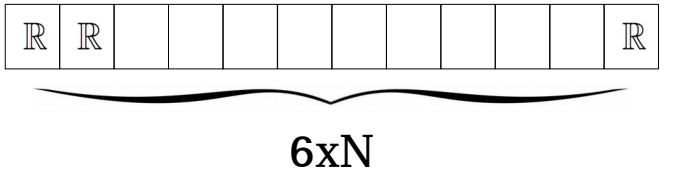
- Information production in composite systems yields measures of information transfer between subsystems, equivalent to instantaneous and directed «Granger causality»
- These measures can be used to characterize brain dynamics upon stroke
- After stroke, the interhemispheric information flow is reduced
- After stroke, the information flow within and from the lesioned hemisphere is reduced

- Q1) Effects due to neural activity or spurious influence of hemodynamics?
- Q2) Reduced information transfer: reduced communication or reduced activity in lesioned hemisphere?
- We need a *model* to discriminate contributions of hemodynamics, effective connectivity (communication strength), activity

Intrinsic dimension of dynamics

The state of a molecule (e.g. villin headpiece) is described by 6N variables





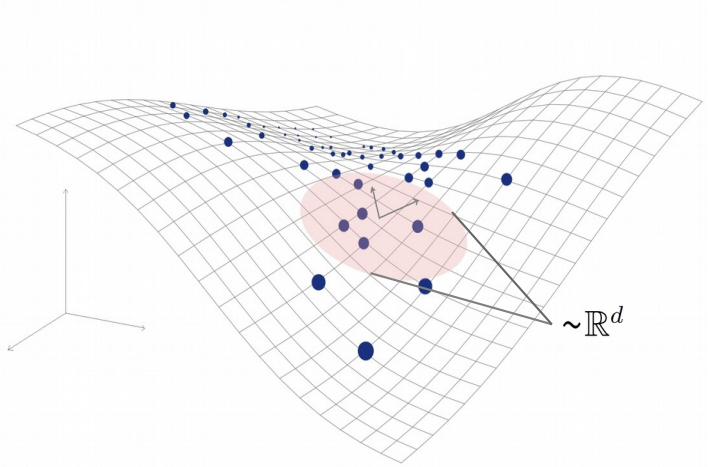
Due to soft and hard constraints, the system “moves” on some directions (hypersurface of dimension d << 6N)

d is called **intrinsic dimension**

What is *d* for this system?

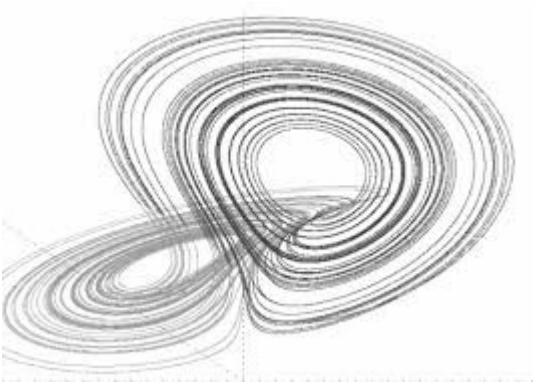
Is *d* the same for different «attractors» of the system (e.g. folded and unfolded state)?

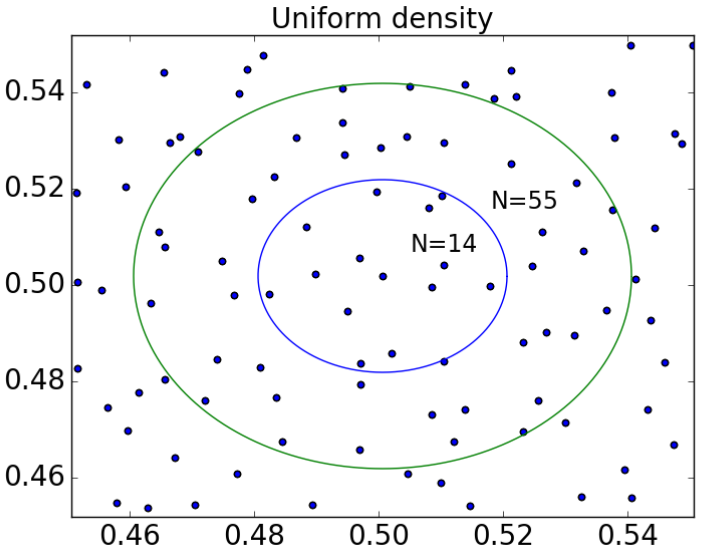
...characterize folding free energy landscape

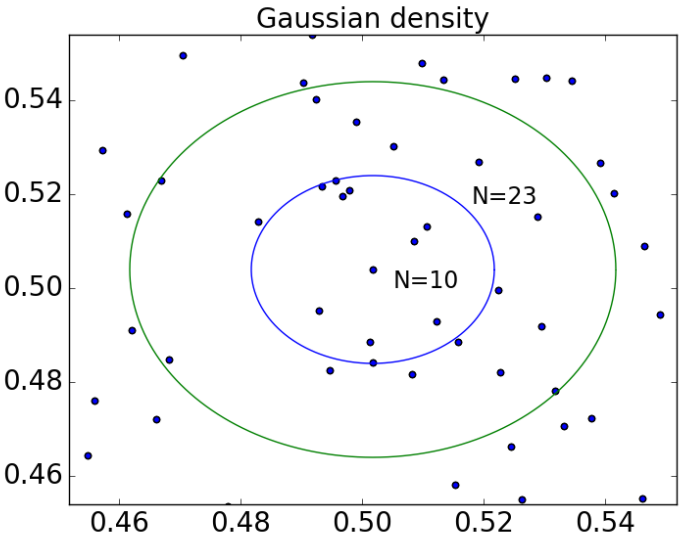


ID estimation: scaling approach

- What is the ID of attractors?
- **distances between points in the dataset follow a *scaling law* that depends on *d***
- Example: correlation dimension
  (Grassberger & Procaccia, PRL 50, 1983)
- The number of nearest neighbors at distance $<\varepsilon$ from point i scales as $N_i(\varepsilon) \sim \varepsilon^d$
- However, the density of points ρ should not have large variations, or the estimation can fail







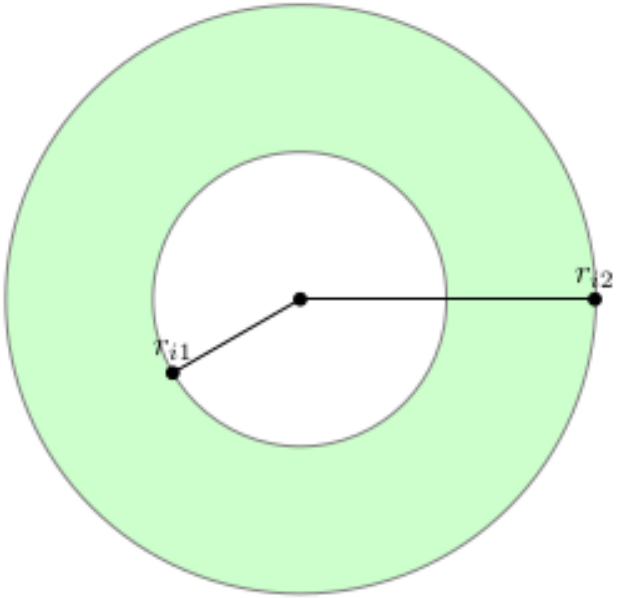
ID estimation: TWO-NN

- TWO-NN: estimating the ID in case of (strongly) variable density
  [E Facco, M D'Errico, A Rodriguez, A Laio, Sci. Rep. 7, 12140 (2017)]

  Make two weak assumptions:

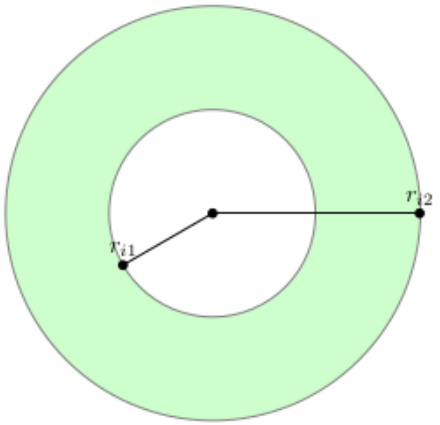
- **H1)** the data points x_i are independent samples from a probability density ρ(x).
- **H2)** local uniformity**:** ρ(x) ~ const. in the region containing the first 2 neighbors of x_i

- r_{i1}, r_{i2} distances of 1st and 2nd neighbor of point i
- **$\mu_i = r_{i2}/r_{i1}$ follows a Pareto distribution: $P(\mu) = d\mu^{-d-1}$**
- **The distribution of μ depends only on *d***

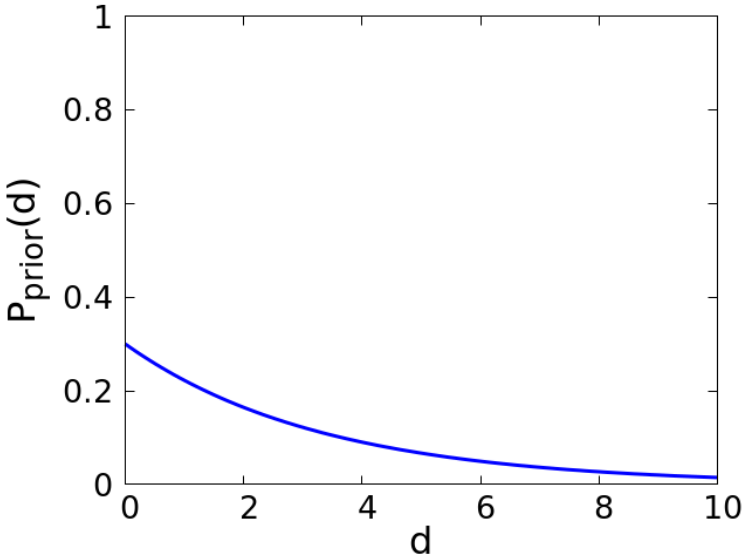


ID estimation: TWO-NN

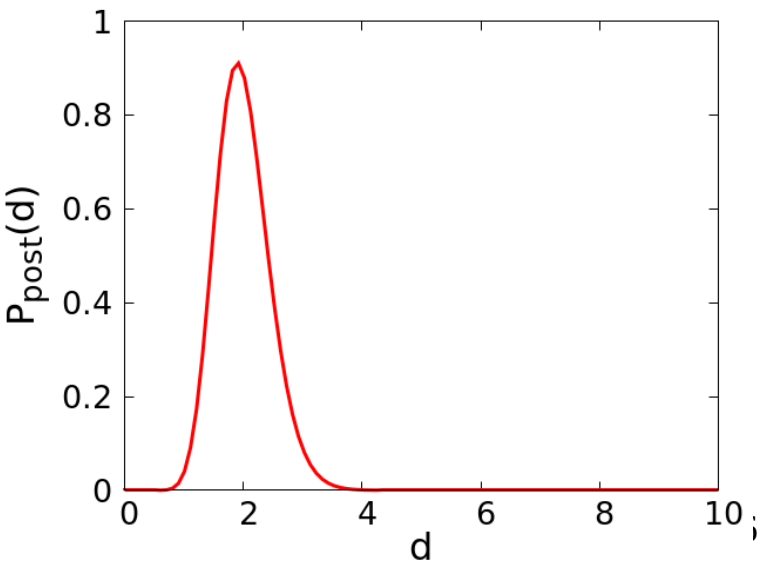
- $\mu_i = r_{i2}/r_{i1}$ follows a Pareto distribution: $P(\mu|d) = d\mu^{-d-1}$
- Bayesian estimations of d : assume $P_{prior}(d) \sim Gamma(a,b)$
- Given the $\{\mu_i\}$, $P_{post}(d) \sim Gamma(a+N, b+\sum \log \mu_i)$



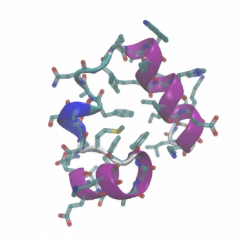
- **d estimate**: posterior average $\langle d \rangle_{post} = (a+N)/(b+\sum \log \mu_i)$

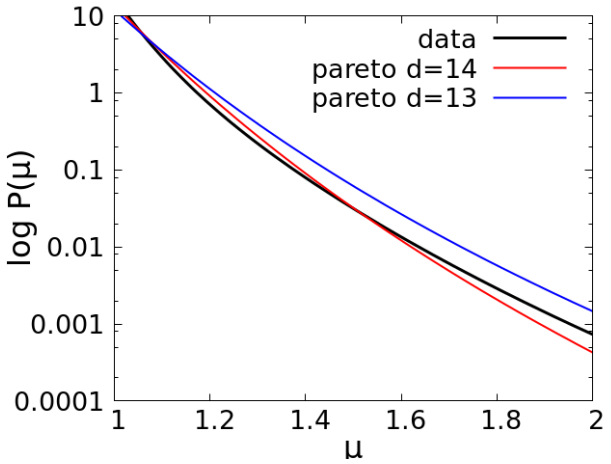


data $\{\mu_i\}$



The problem of multiple IDs



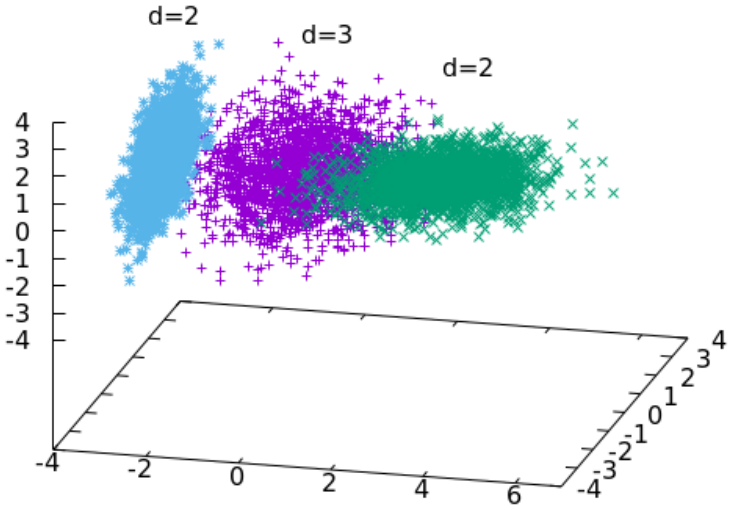


d ~ 14, but the data do not fit well a Pareto

What if *d* is not uniform?

the data may lie on several manifolds, each with different intrinsic dimension

[M. Allegra, E. Facco, A. Laio and A. Mira, sub. (2020); arXiv:1902.10459]



Extending TWO-NN to multiple IDs

- assume the points lie on K «manifolds» with different IDs **d**=$d_1,\ldots,d_K$
- the distribution of $\{\mu_i\}$ is simply a **mixture of Pareto distributions**

$$\mathcal{L}(\boldsymbol{\mu}|\boldsymbol{d},\boldsymbol{Z}) = \prod_{i=1}^{N} d_{Z_i}\mu_i^{-(d_{Z_i}+1)}$$

- d_k dimension of manifold k
- Z_i assignment of each point to a maniold [Z_i = 1,…,K]

$$P_{post}(\boldsymbol{d},\boldsymbol{Z}|\boldsymbol{\mu}) = \mathcal{L}(\boldsymbol{\mu}|\boldsymbol{d},\boldsymbol{Z})P_{prior}(\boldsymbol{d},\boldsymbol{Z})$$

- Estimate jointly **d,Z** by Gibbs Sampling of the posterior distribution

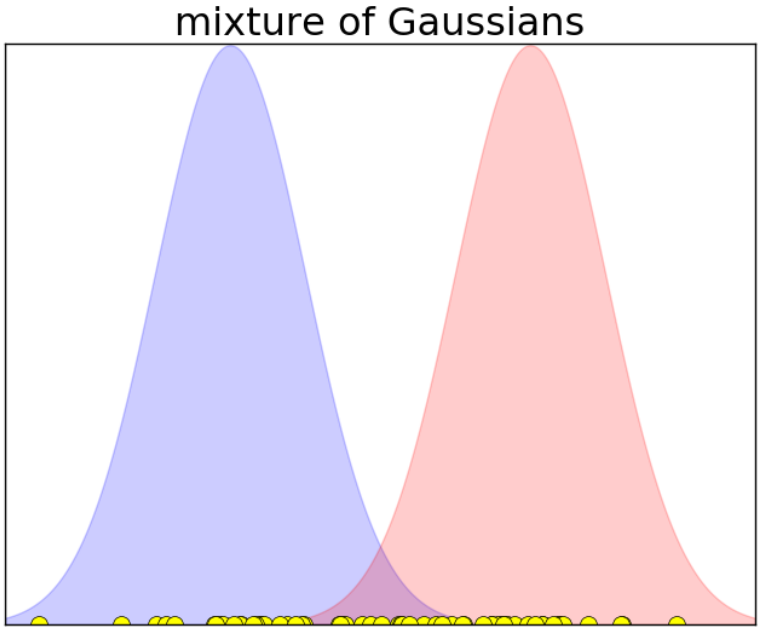
Extending TWO-NN to multiple IDs

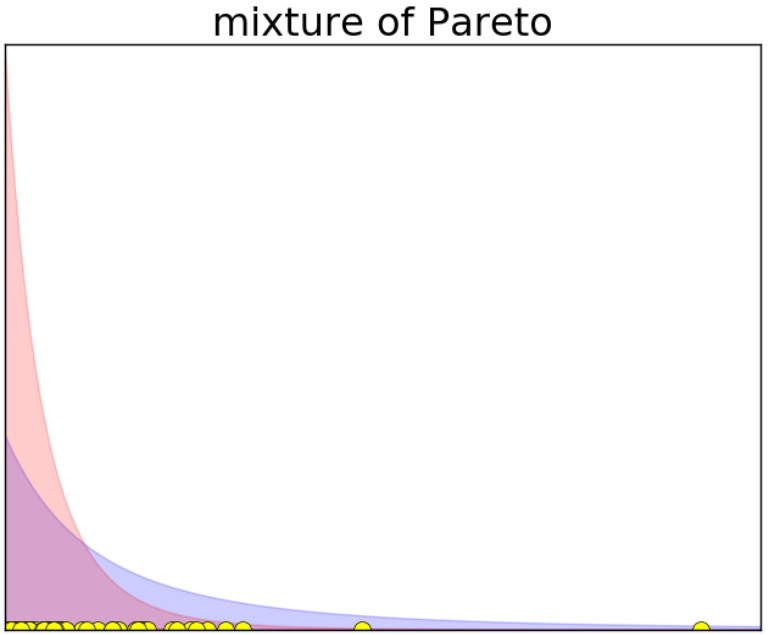
What is the problem?

Z are easy to assign only if mixture components are largely non-overlapping

But Pareto distributions with different *d* are highly overlapping!

The Z assigment is not reliable





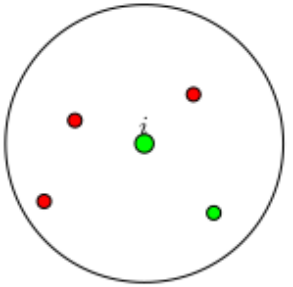
Extending TWO-NN to multiple IDs

Let the neighborhood of point *i* be defined by its first *q* neighbors

n_i^{in} # neighbors with same Z as *i*

n_i^{out} # neighbors with diffferent Z

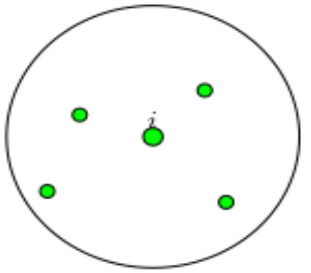
We get non-uniform neighborhoods: $n_i^{out} > n_i^{in}$



One more assumption: **neighborhoods must be approximately uniform**

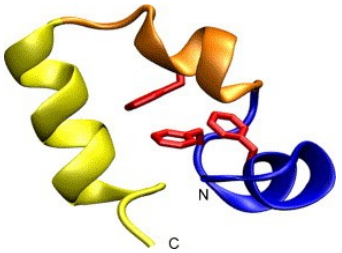
Enforce with **additional term in the likelihood («Potts interaction)**

$$\mathcal{L}(n^{in}|\boldsymbol{Z}) \propto \prod_i \xi^{n_i^{in}} (1-\xi)^{n_i^{out}} \qquad 0.5 < \xi < 1$$

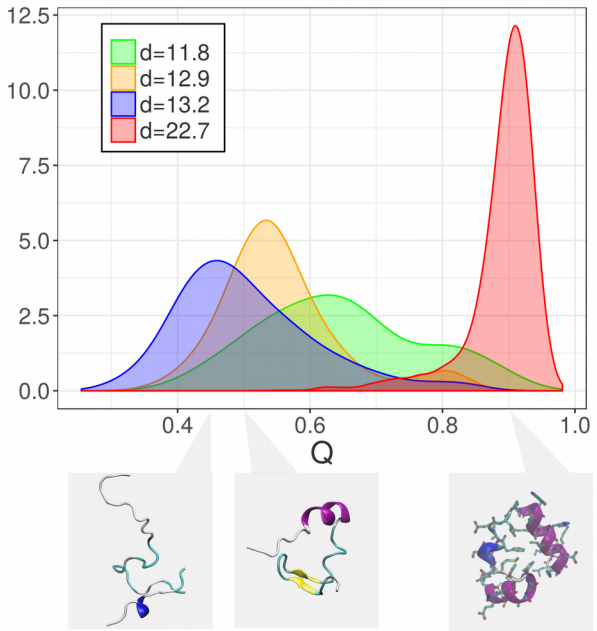


ξ : propensity of neighbors to be in the same manifold

ID of folding protein



- consider a simulation of unfolding/refolding protein (villin headpiece)
- for each configuration, M=32 dihedral *angles*.

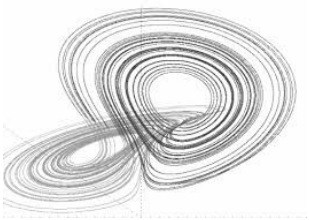


- three manifolds with low dimensions d~13
- one manifold with high dimension d~23

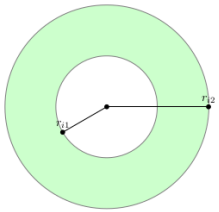
- q = fraction of native contacts (=degree of folding)
- **Folded configurations in high-dimensional manifold**
- local ID discriminates folded and unfolded configurations
- effective # of phase space directions the system can explore varies folded/unfolded «state»

Recap

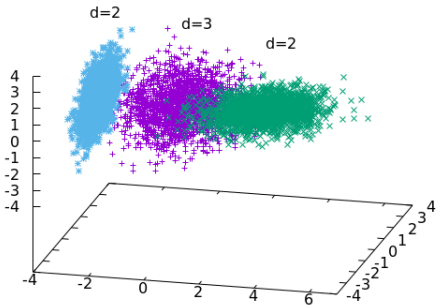
- Physics of chaotic systmes gave us the concept of ID estimators traditional ID estimators require constant density



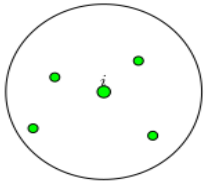
- TWO-NN estimates the ID based on the statistics of first neighbor distances



- We extended TWO-NN to the case where the data contains regions with different dimensionality



- Inferring the IDs is a non-trivial inference problem, requires modification of the basic model



- In real data, we find large variations of the ID, highlighting relevant structure in the data

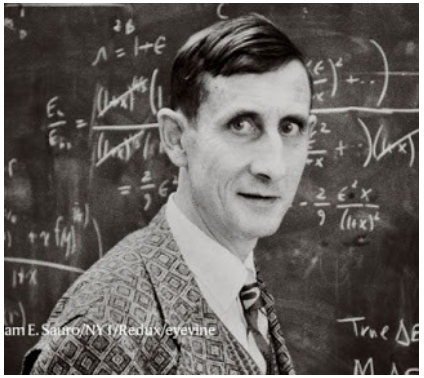
ID and brain dynamics?

- **We could analyze the dimension of dynamical « attractors» in brain dynamics**

- Several models of cortical dynamic posit that the brain has several attractors (multistability) [Deco and Jirsa, J Neuro, 32(10):3366 –3375]
- Not easy to verify this in data
- Consider regional time series from different imaging modalities (fMRI, MeG,eeG)

- Can we find regions with different ID with our method? This could direcly identify «attractors»
- Is the ID of models and data the same? (hint about accuracy/non-accuracy of models, already used for protein evolution models)

Physics of the brain?



Freeman Dyson
15/12/1923-28/2/2020

Philip W. Anderson
13/12/1923-29/3/2020

« the surest way to save physics from some rather catastrophic stagnation or decline during the next 30 years is to keep young physicists working on the frontiers where physics overlaps other sciences, such as astronomy and biology.

Freeman Dyson, *The future of physics*, Physics Today 23, 9, 23 (1970)

« A movement is under way toward joining together into a general subject all the various ideas about ways new properties emerge. We call this subject the science of complexity. Within this topic, ideas equal in depth and interest to those in physics come from some of the other sciences. This movement is overdue and healthy»

Philp W. Anderson, *Is Complexity Physics?* Physics Today 44, 7, 9 (1991)

Acknowledgments

Andrea Brovelli
INT, CNRS, Marseille

Maurizio Corbetta
Università di Padova

Chiara Favaretto
Università di Padova

Alessandro Laio
SISSA, Trieste

Antonietta Mira
USI, Lugano

Elena Facco
ML Reply, Torino

Alex Rodriguez
ICTP, Trieste

Thank you for your invitation and attention!

Samir Suweis

Flavio Seno

Maurizio Corbetta

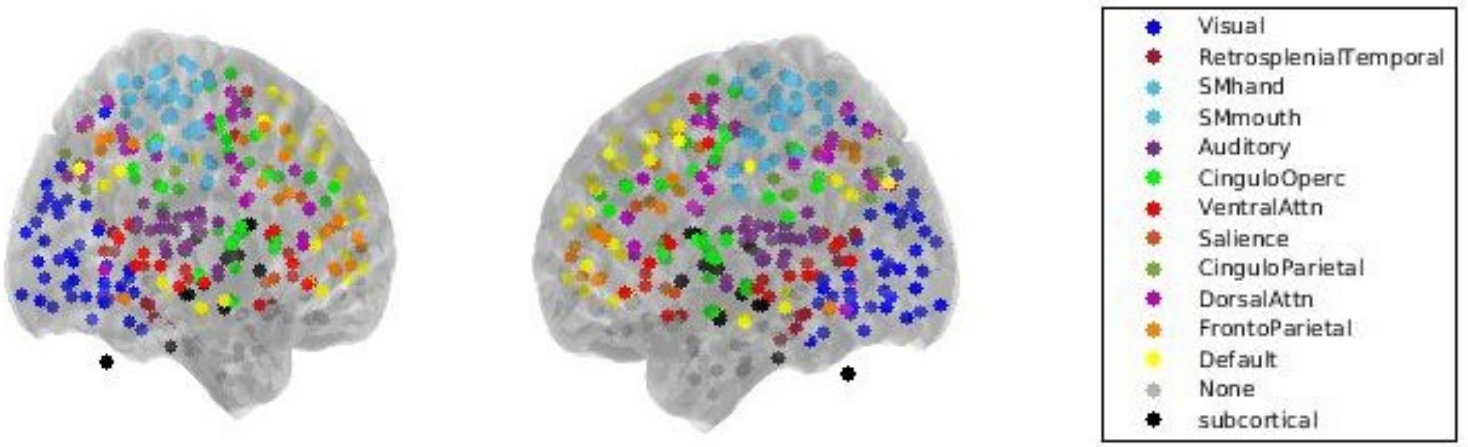
PADOVA neuroscience CENTER

UNIVERSITÀ DEGLI STUDI DI PADOVA

BACK-UP SLIDES

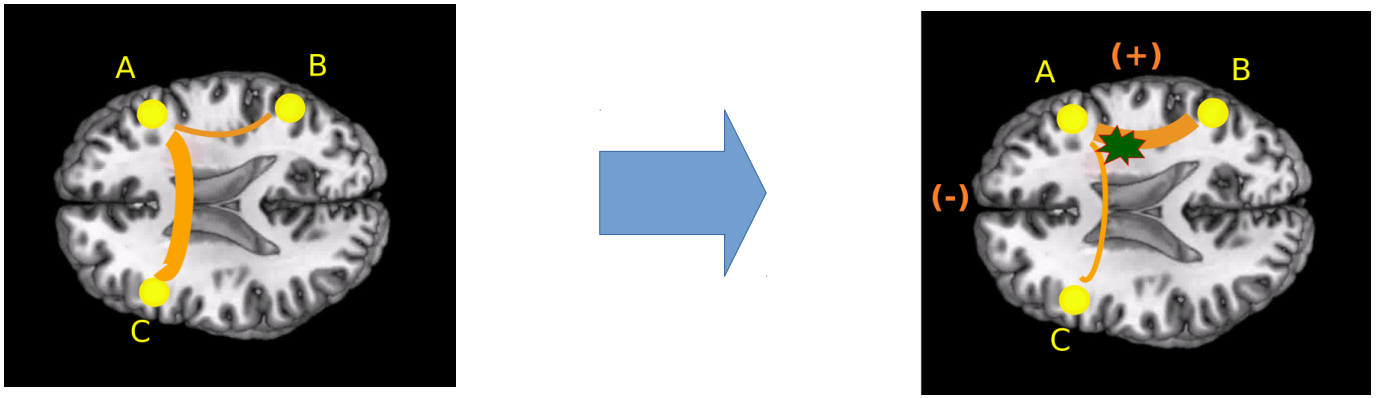
Stroke and information transfer: regions used

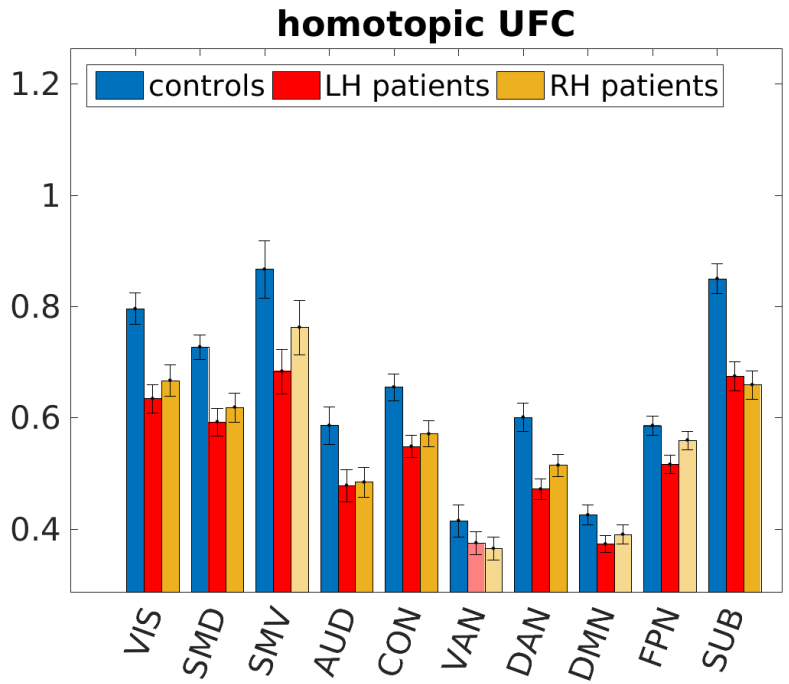
- 324 regions from cortical atlas (Gordonn-Laumann)
- 19 subcortical and cerebellar regions (from Harvard-Oxford and AAL)
- Cortical regions can be assigned to different resting state networks

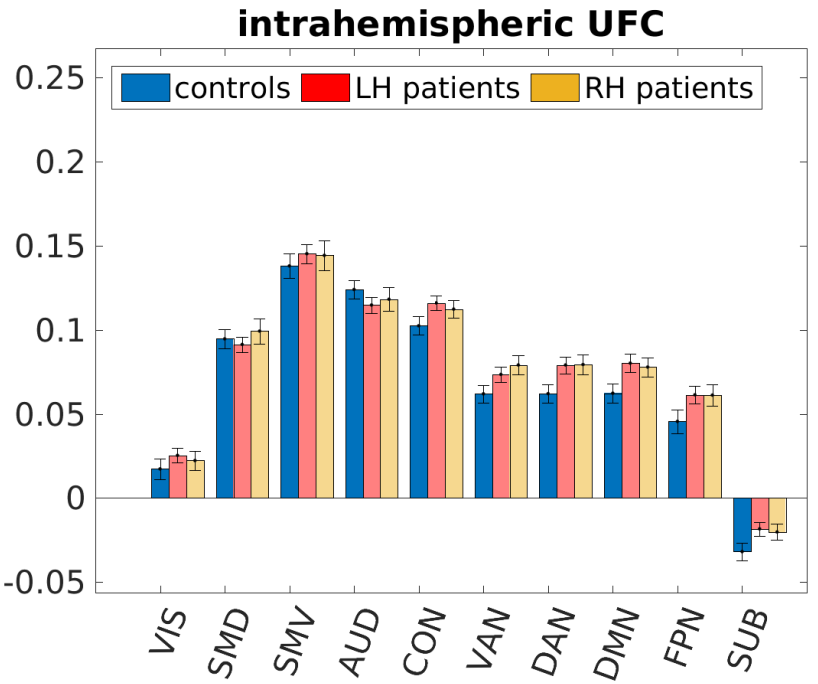


- We compute GC between all pairs of regions
- We compare the results of healthy controls, patients with left hemisphere (LH) lesions and patients with right hemisphere (RH) lesions

Stroke: functional connectivity changes







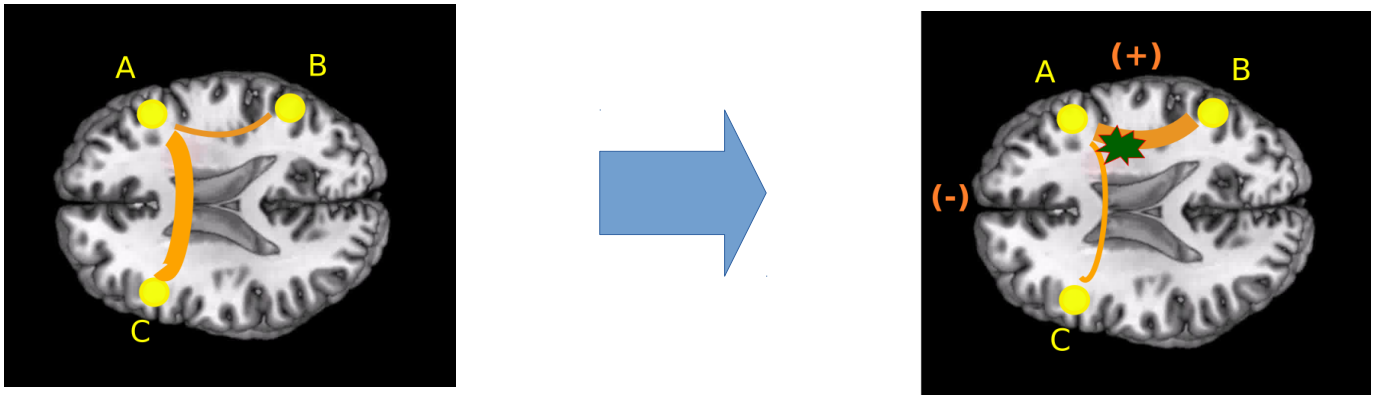
VIS
AUD
CON
DAN
sub

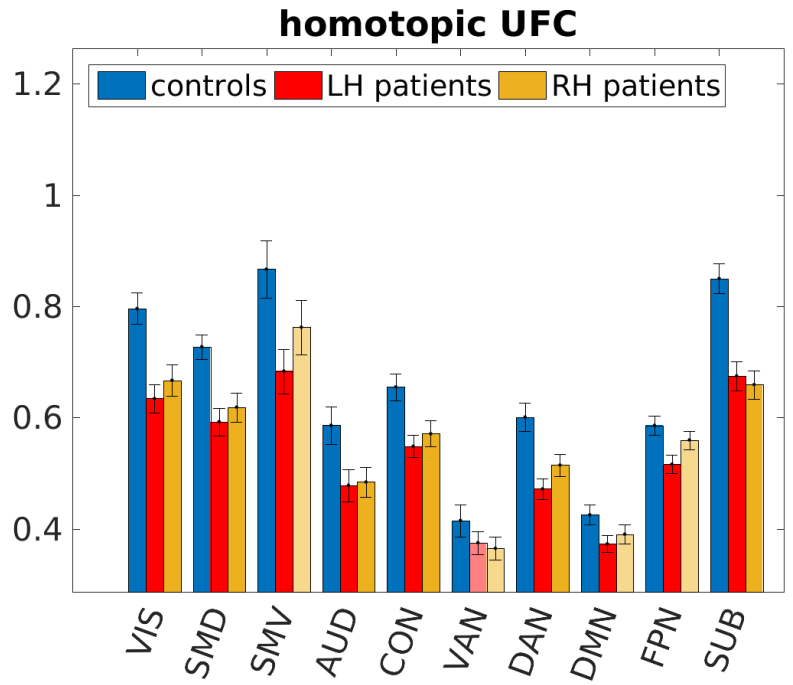
DAN
DMN

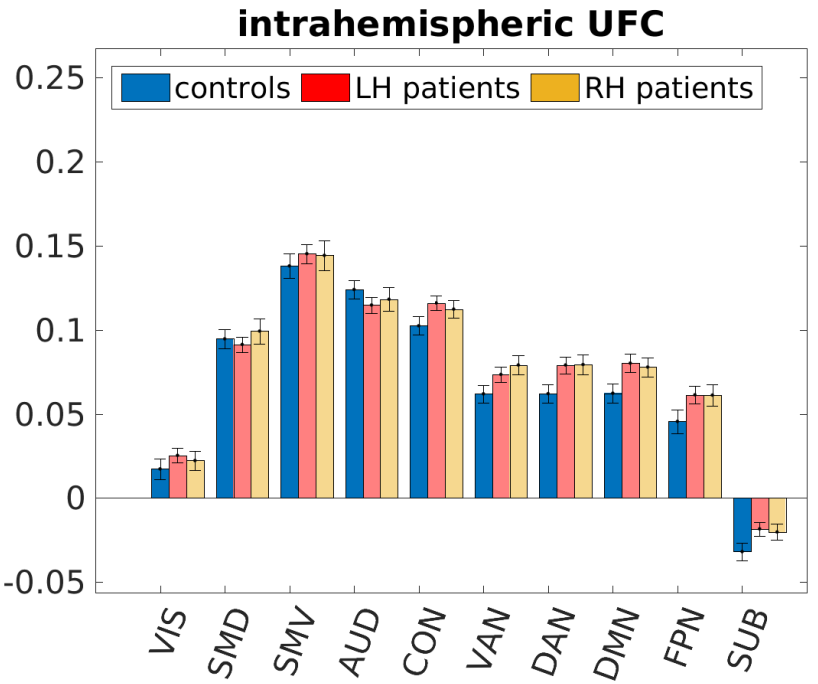
Interhemispheric (homotopic) FC is reduced in patients

Intrahemispheric FC is increased in patients

Stroke: functional connectivity changes







VIS
AUD
CON
DAN
sub

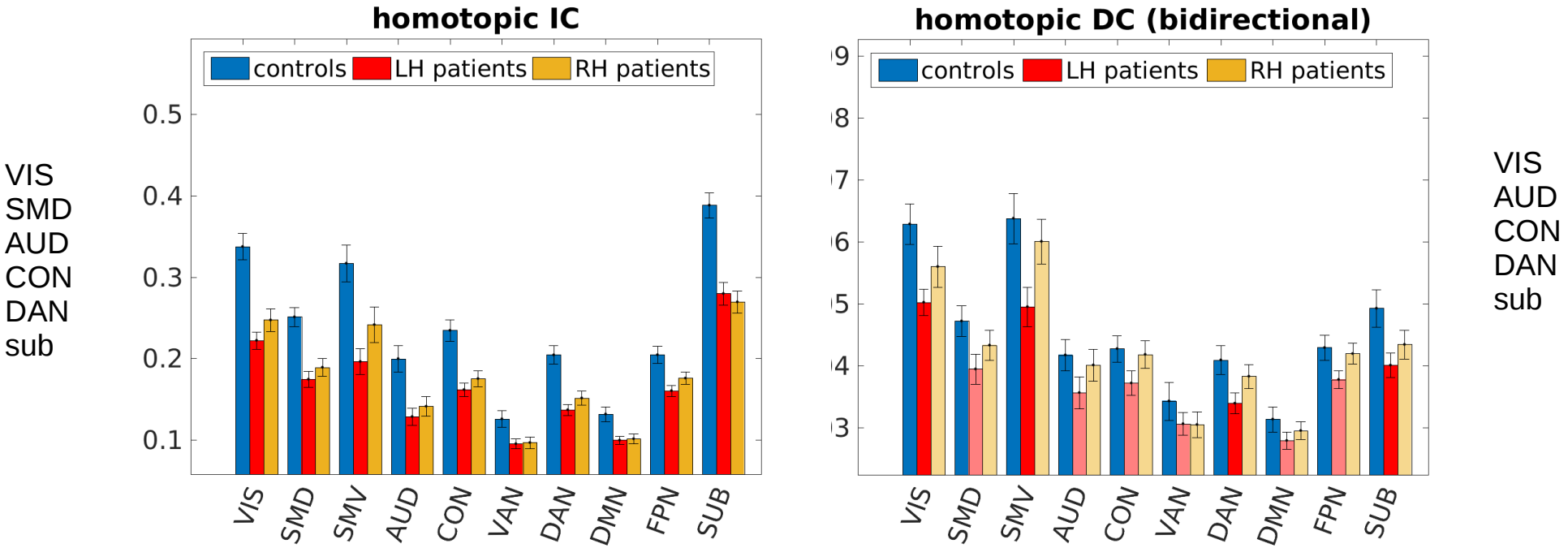
DAN
DMN

Interhemispheric (homotopic) FC is reduced in patients

Intrahemispheric FC is increased in patients

Stroke: homotopic GC changes

- covariance-based GC using a window of L=5 time points to define X_{past}, Y_{past} (slow flows on the scale of ~10s)
- Instantaneous Causality (IC) for homologous regions
- Directed Causality (DC) for homologous regions: «bidirectional flow» $F_{X\to Y}+F_{Y\to X}$



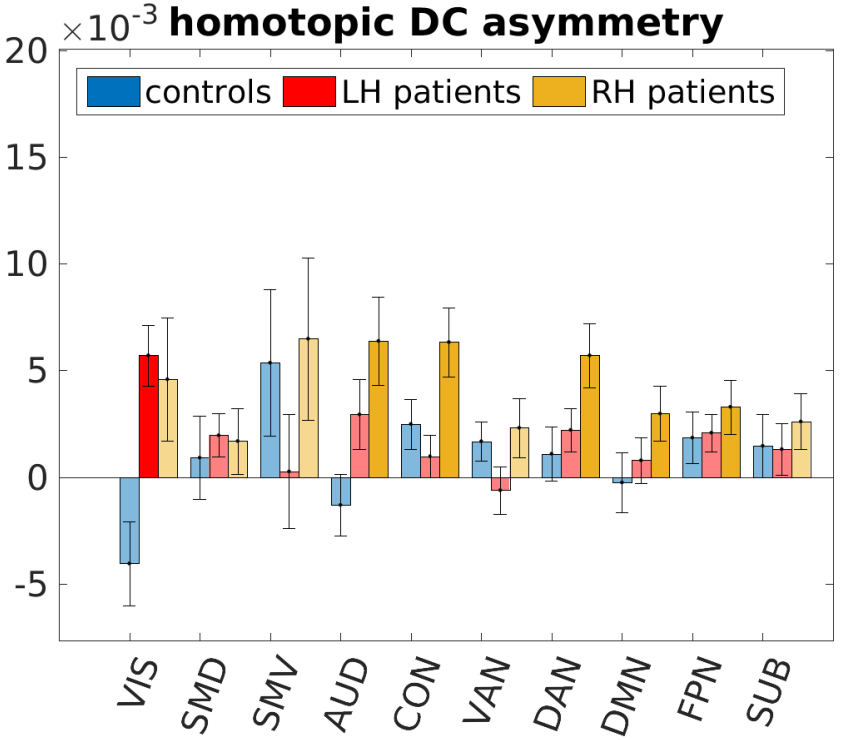
Interhemispheric (homotopic) IC is reduced in patients

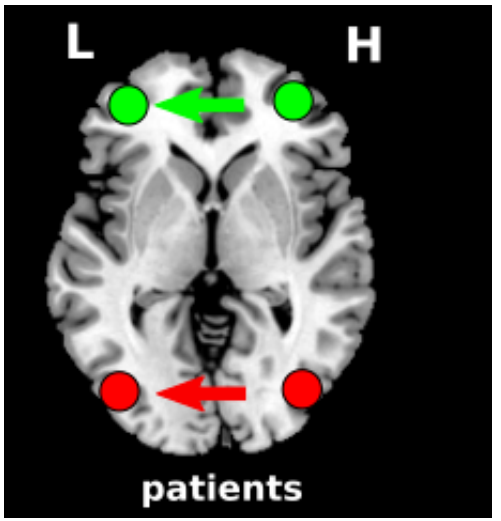
Stroke: homotopic GC changes

- Directed Causality (DC) for homologous regions: «net flow» $F_{X \to Y} - F_{Y \to X}$

  X is in the healthy hemisphere, Y in the lesioned hemisphere (left/right for controls)

VIS
AUD
CON
DAN
FPN
sub



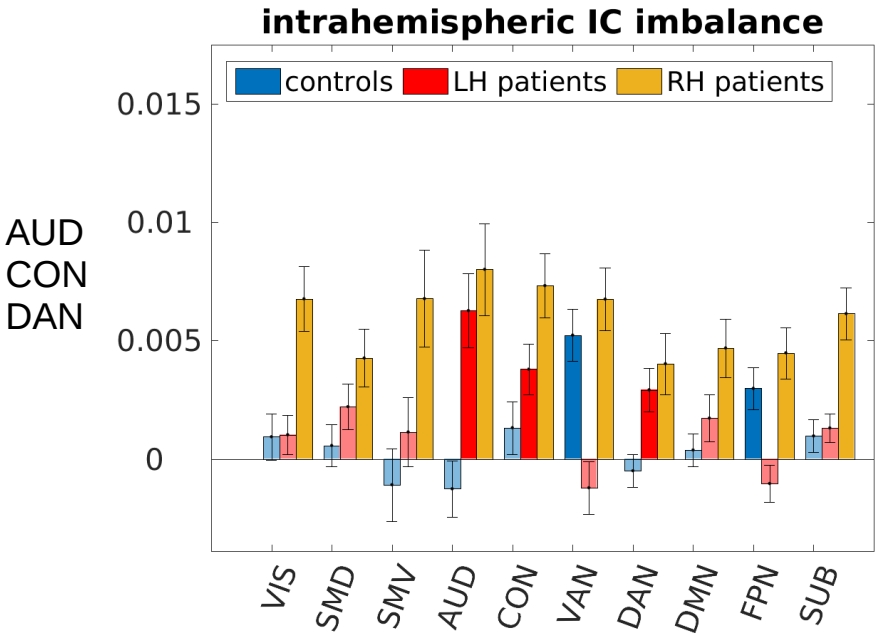


The interhemispheric (homotopic) DC from the healthy to the lesioned hemisphere is higher than the reverse in patients

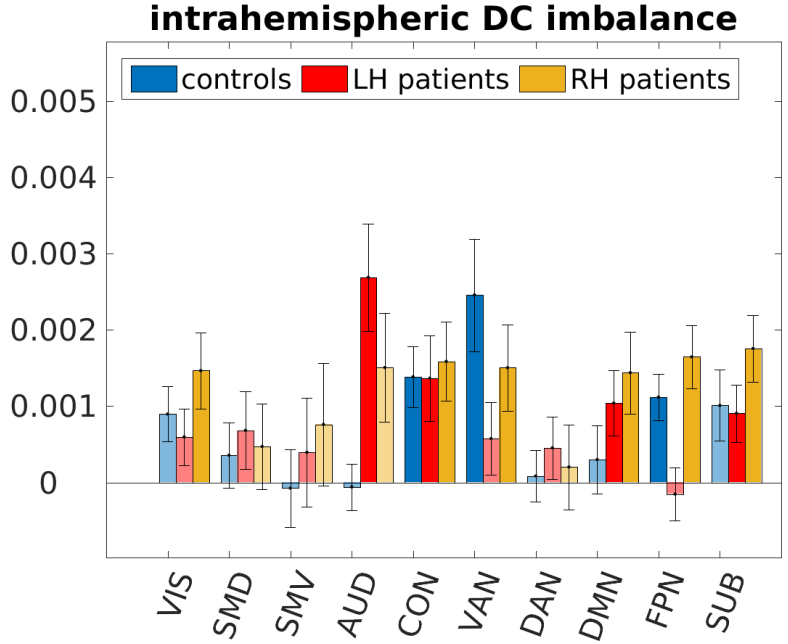
The net information flow is in the direction of the lesioned hemisphere

Stroke: intra-hemispheric GC changes

- We compute the average DC and IC in each hemisphere separately
- We compute an imbalance (difference) between the lesione and healthy hemisphere (left/right for controls)



AUD
CON
DAN



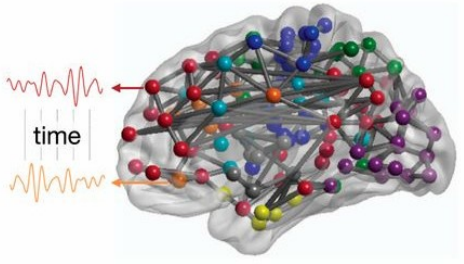
AUD
CON
DMN
sub

The intra-hemispheric IC is higher in the healthy hemisphere for patients

- **The intra-hemispheric DC**
- **is higher in the healthy hemisphere cor patients**

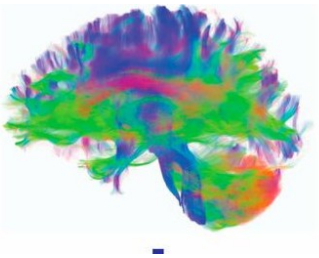
Stroke: FC and GC markers

We obtain seven functional bio-markers of stroke:



- UFC_{homo} (average homotopic functional connectivity)
- IC_{homo} (average homotopic instantaneous causality)
- ΣDC_{homo} (average homotopic directed causality)
- ΔDC_{homo} (lesioned/healthy asymmetry in homotopic directed causality)
- ΔIC_{intra} (average intrahemispheric functional connectivity)
- ΔDC_{intra} (lesione/healthy asymmetry in intra-hemispheric instantaneous causality)
- UFC_{intra} (lesione/healthy asymmetry in intra-hemispheric directed causality)

We also have structural markers of stroke:



- V (lesion volume),
- Dischomo (amount of homotopic fibers severed by lesion),
- Discintra (amount of intra-hemispheric fibers severed by lesion)

- **The homotopic measures UFC_{homo}, IC_{homo} , ΣDC_{homo} are reduced in patients, correlate among themselves, correlate negatively with V and with $Disc_{homo}$**

- **The «unbalance» measures ΔDC_{homo}, ΔIC_{intra} , ΔDC_{intra} are enhanced in patients, correlate among themselves, correlate positively with Discintra**

ID estimation: projective approach

Project *D*-dimensional data into lower dimension *d* : $\Pi^d : \mathbf{x}_i \in \mathbb{R}^D \mapsto \mathbf{y}_i \in \mathbb{R}^d$

- Try different *d* and evaluate for each a “loss function” $\mathcal{L}(\Pi^d)$
- $\mathcal{L}(\Pi^d)$ measures the “data loss” occurring in the projection. Examples:

$\mathcal{L}(\Pi^d) = \sum_i ||\mathbf{x}_i - \mathbf{y}_i||^2$ preservation of original distance relations

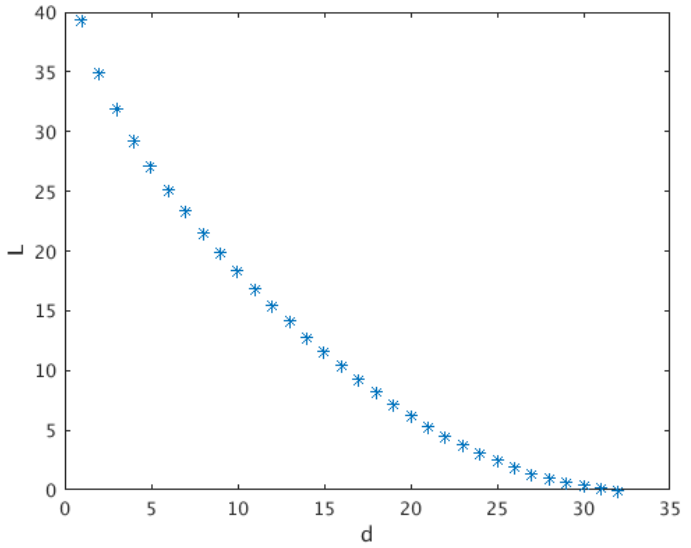
$\mathcal{L}(\Pi^d) = \sum_i \mathbf{x}_i \mathbf{x}_i^T - \mathbf{y}_i \mathbf{y}_i^T$ preservation of original covariance matrix

- **ID estimate based on tradeoff between dimensionality reduction and data loss**
- Problem (1): Computationally burdensome (search for optimal projection for each *d*)
- Problem (2): robust ID estimates only if $\mathcal{L}(\Pi^d)$ has large gap as a function of *d*

  if no gap, the estimation can be rather arbitrary

ID estimation: projective approach

- Example: Principal Component Analysis (PCA)
- Projects data onto linear subspace spanned by first *d* eigenvalues of covariance matrix. $X^T X$ Loss: $\mathcal{L}(\Pi^d) = || \sum_i \mathbf{x}_i \mathbf{x}_i^T - \mathbf{y}_i \mathbf{y}_i^T ||$
- On the villin headpiece simulation:



- How can one select an appropriate *d*?

ID estimation: TWO-NN (derivation)

- 1) **Poisson sampling (valid id N is large)**

  $P(n \text{ points in } A) = (\rho V_A)^n (1-\rho V_A)^{N-n} \sim (\rho V_A)^n/n! \exp(-\rho V_A)$

- 2) **Consider** first two neighbors of *i* and **hyperspherical shells S_{i1} and S_{i2}** with volumes V_{i1} and V_{i2}

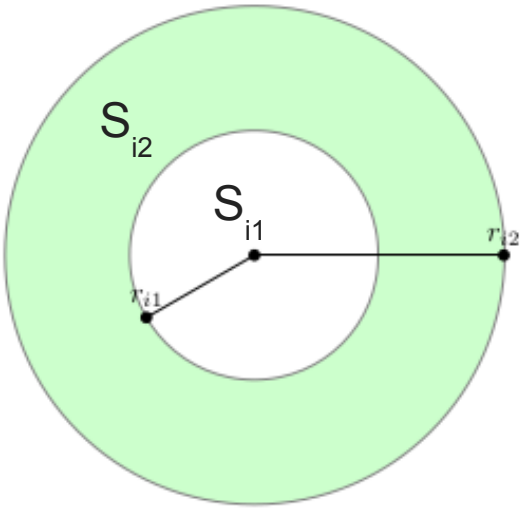
- 3) **Derive probability density of the shell volumes** from Poisson sampling

  $P(V_{i1}>V)=P(0 \text{ points in } S_{i1})= \exp(-\rho V) \rightarrow P(V_{i1} \leq V)=1-\exp(-\rho V) \rightarrow f(V_{i1}) = \rho \exp(-\rho V_{i1})$

- 4) **Derive probability of $\mu_i=r_{i2}/r_{i1}$**

  $f(V_{i1},V_{i2})=\rho^2 \exp(-\rho V_{i1}-\rho V_{i2}) \rightarrow f(V_{i2}/V_{i1})=1/(1+V_{i2/}V_{i1})^2$

  $(V_{i2}/V_{i1})=(r_{i2}/r_{i1})^d-1 \rightarrow f(r_{i2}/r_{i1})=d/(r_{i2}/r_{i1})^{d+1}$

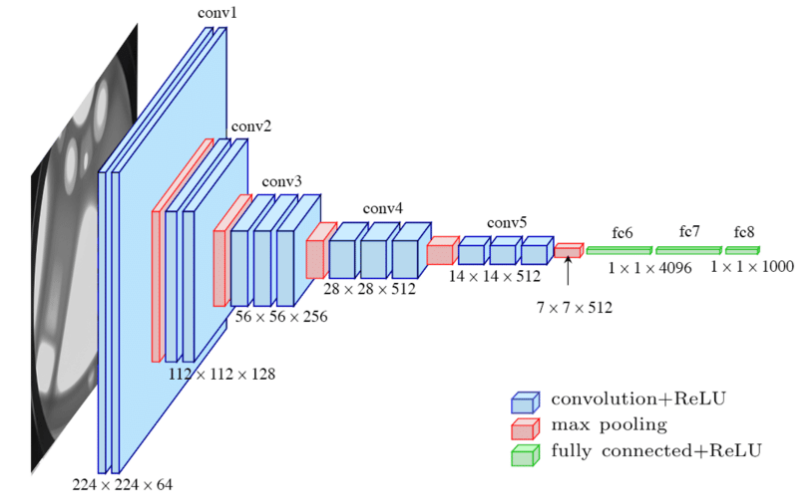


Application: dimension of data in neural networks

Ansuini, Alessio, et al. arXiv preprint arXiv:1905.12784 (2019).

Deep neural networks transform their inputs across m ultiple processing layers.

What are the geometrical properties of the representations learned?



Representations should be relatively low-dimensional

Intuitively **the dimension should decrease in successive layers**
(the NN progressively eliminates irrelevant features)

Application: dimension of data in neural networks

Ansuini, Alessio, et al. arXiv preprint arXiv:1905.12784 (2019).

Synthetic data-set of 1440 images of 40 objects

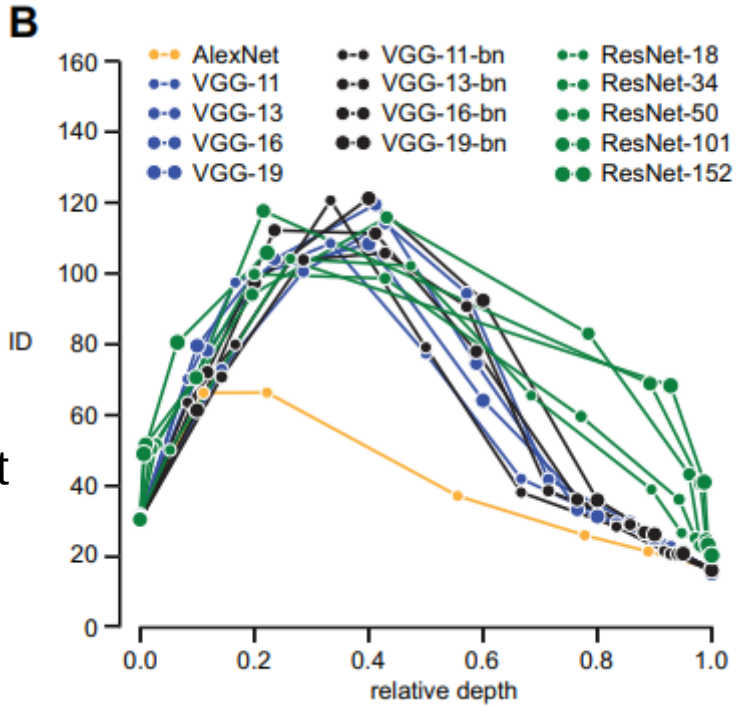
Train deep neural networks with different architectures (generally ~10^6 units)

Representations are low dimensional (ID <100)

The ID first increases, then decreases as function of the layer (depth)

The NN first eliminates “gross” features that effectively dominate the ID at a large scale

Then NN progressively reduces “finer” features that are irrelevant for prediction



Application: dimensionality reduction

TWO-NN yields only an estimate of the intrinsic dimension

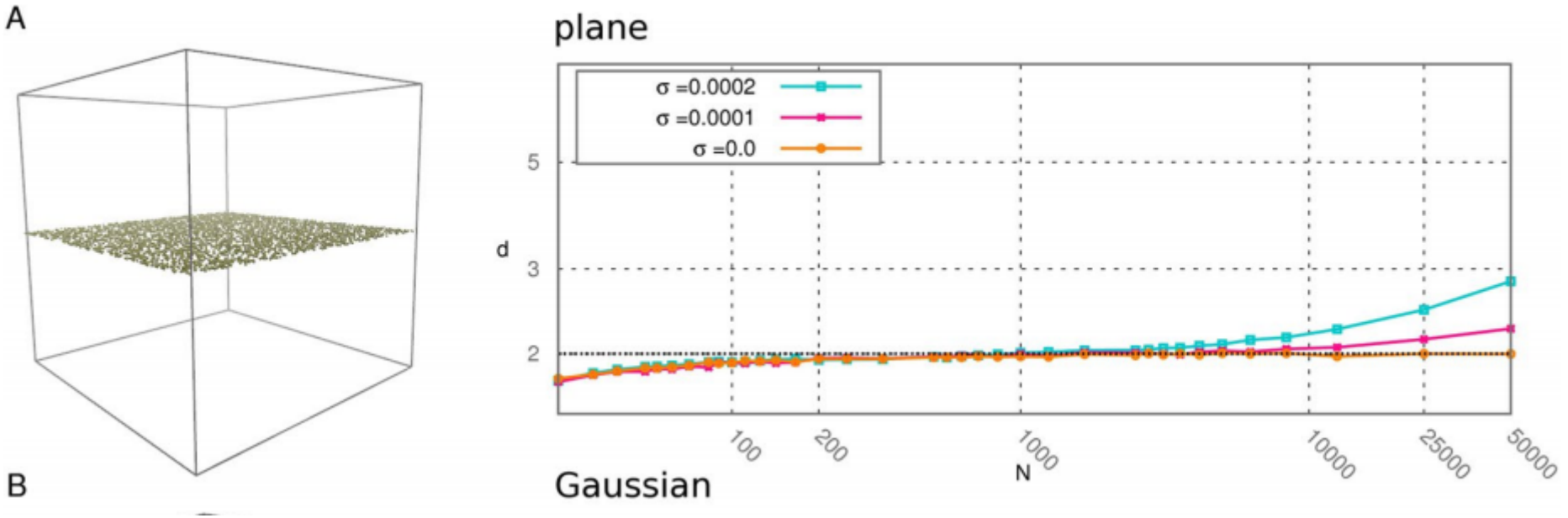
It does not yield a scheme for dimensionality reduction
(explicit parametrization in terms of low-dimensional set of coordinates

However, the ID estimate given by TWO-NN can be used
To select proper **target dimension in** dimensionality reduction methods

Project *D*-dimensional data into lower dimension *d*: $\Pi^d : \mathbf{x}_i \in \mathbb{R}^D \mapsto \mathbf{y}_i \in \mathbb{R}^d$

ID estimation: TWO-NN

- The ID can depend on the scale at which we look at the system!
- Explore different scales by decimating the data
- **The “effective” dimension is stable across wide range of scales**



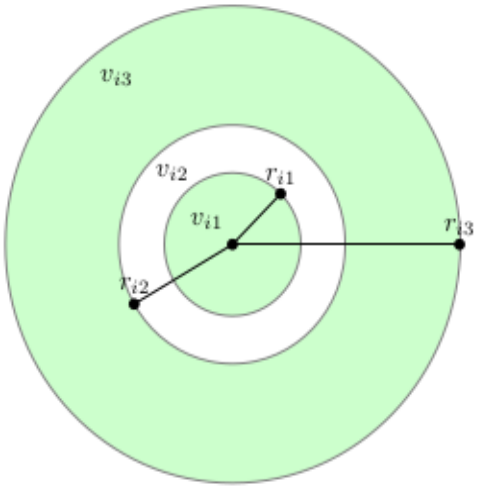
Application: density estimation

Rodriguez, A., d'Errico, M., Facco, E., & Laio, A. (2018) JCTC 14(3), 1206-1215.

K-nearest-neighbor estimate: assume ρ ≈ const in small region around each point

- For each point *i*, consider its *k* nearest neighbors at distances $r_{i1}, r_{i2}, r_{i3}, \ldots$
- density= *k*/volume of sphere containing the *k* points

$$\rho = \frac{k}{V_{ik}} \qquad \delta\rho = \frac{\sqrt{k}}{V_{ik}} \qquad V_{ik} = \omega_d r_{ik}^d$$



What is right d? The intrinsic dimension of the manifold!

(… then optimize k locally)

Extending TWO-NN to multiple IDs

- TWO-NN assumptions:
  - **H1)** the data points x_i are **independent samples** from a density ρ(x).
  - **H2) local uniformity:** ρ(x) ~ const. in the region containing the first 2 neighbors of x_i

  Additional assumption:

- H3) the distribution ρ(x) has support on K manifolds with different IDs **d**=$d_1,\ldots,d_K$

- Under H1), H2), H3) the distribution of μ is simply a **mixture of Pareto distributions**

$$\mathcal{L}(\boldsymbol{\mu}|\mathbf{d},\mathbf{p}) = \prod_{i=1}^{N}\sum_{k=1}^{K} p_k d_k \mu_i^{-d_k-1}$$

Heterogeneous ID

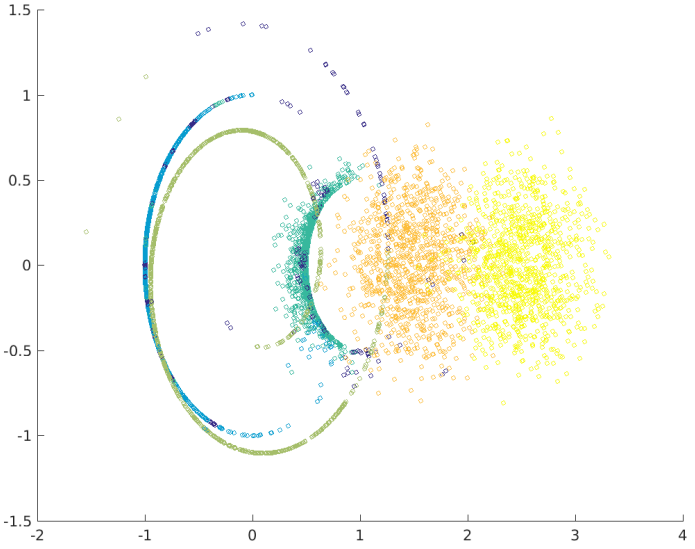
M Allegra, E Facco, A Laio and A Mira, arXiv:1902.10459 (2019)

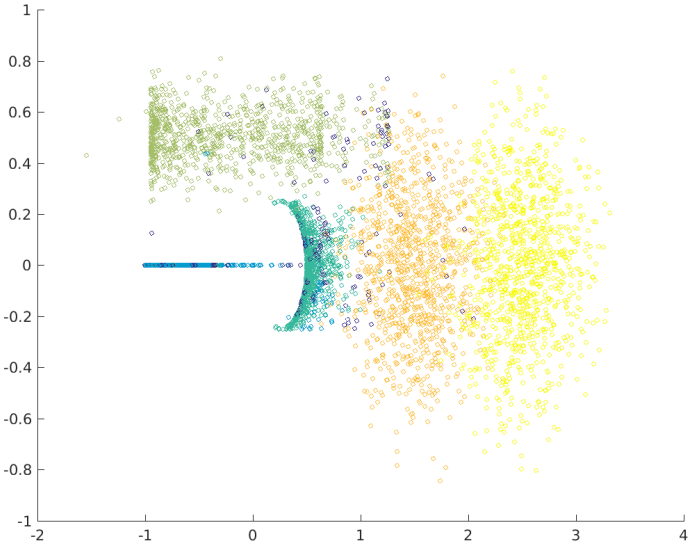
Find regions (manifolds) of different ID in the data

Works also for nonlinear and topologically complex manifolds

Circle d=1, swiss roll in d=4, torus d=2, sphere d=5, sphere d=9

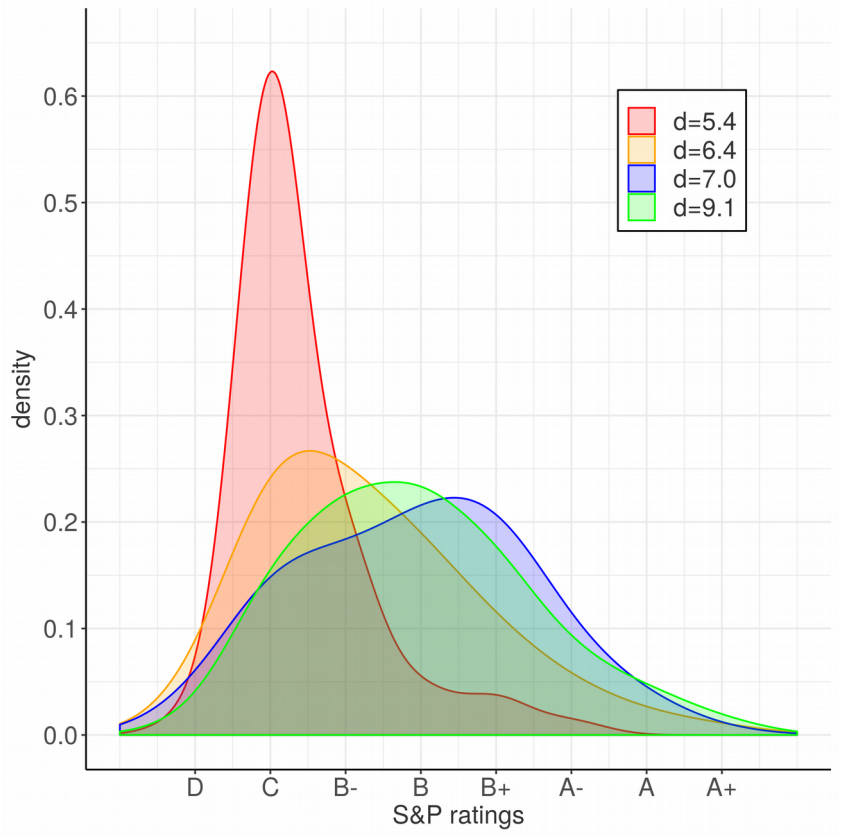
Estimated dimensions 0.9,2.0,4.1,5.2,8.5





Heterogeneous ID example: companies balance sheets

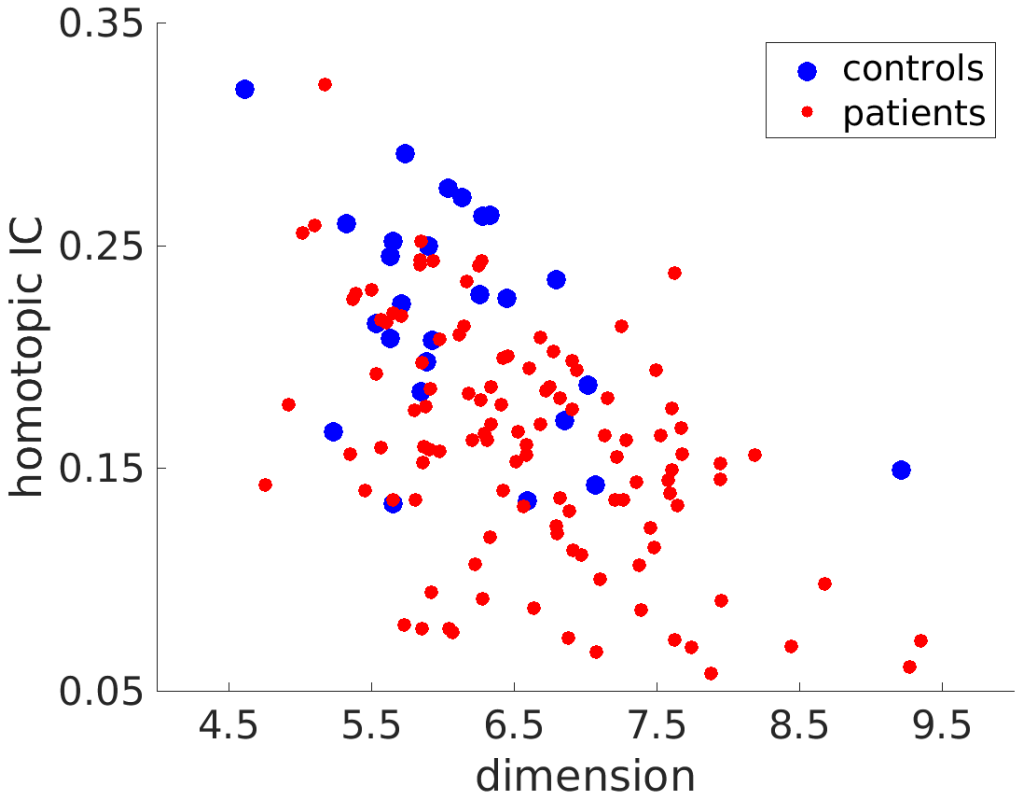
- consider D=38 balance sheet variables for N=8000 companies
- **We find four manifolds with dimensions** d=5.4,d=6.4,d=7.0, d=9.1
- Consider the financial risk of the companies assigned to different manifolds



Companies with higher risk are preferentially assigned to low dimensional manifolds!

ID and brain dynamics?

- **Idea 2: the dimension of dynamical « attractors» in brain dynamics**
- Compute the ID of the dynamics (# variables needed to define dynamical states)

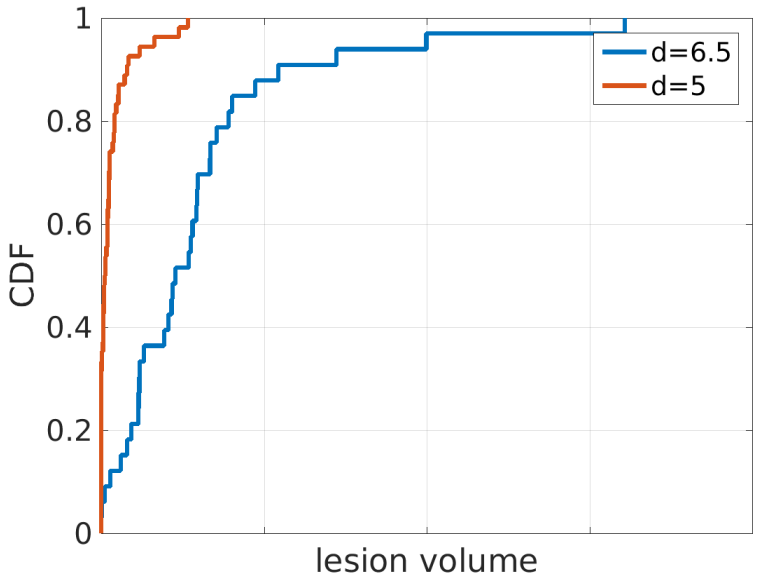


More homotopic disconnection

Higher dimension

ID and brain dynamics?

- **Idea 1: distinguish groups of subjects on the basis of dynamical features**
- We defined 7 markers of stroke UFC_{homo} , IC_{homo} , ΣDC_{homo} , ΔDC_{homo} , ΔIC_{intra} , ΔDC_{intra} , UFC_{intra}
- Are the data really 7-dimensional?



- We find a manifold of dimension 5 and a manifold of dimension 6.5
- The manifold of dimension 5 includes controls and patients with weak lesions; The manifold with dimension 6.5 patients with strong lesions
- Controls are less «variable» in this space