# Graph signal processing for the study of multivariate physiological signals

Laurent Oudre
laurent.oudre@ens-paris-saclay.fr

May, 26th 2023

# The Centre Borelli

**The Centre Borelli**

Fusion of two labs :

- The Centre de mathématiques et de leurs applications (CMLA) : applied mathematics for the study of complex phenomena and data
- The Cognition & Action Group (CognacG) : quantification and study of human and animal behavior

# Main scientific questions

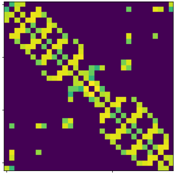
**How to quantify the human behavior**

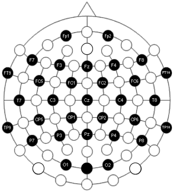
- Adventure launched since 2012 : interdisciplinary collaboration between mathematicians, physicians, neuroscientists, engineers, biologists, etc...
- Implementation of measurement chains “pipelines”, platforms and intelligent tools but also of procedures for analysis, measurement and processing of data
- Creation of tools for diagnostic assistance, inter-individual comparison and longitudinal follow-up
- Integration into a clinical environment and interaction between algorithms and medical/neuroscience experts

# First usecase

## Study of EEG data

- Study of EEG recorded during general anesthesia
- 32 sensors at 256 Hz
- How can we learn a structure and use it to process the signals ?

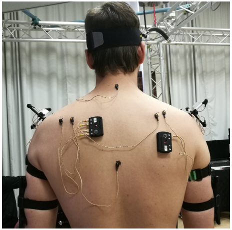


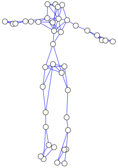


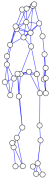

# Second usecase

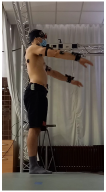
## Study of 3D upper-limb movements

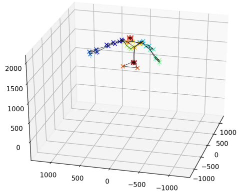
- Study of upper-limb movements with 3D markers
- Around 30 sensors recording the 3D positions over time (100 Hz)
- How can we take into account the skeleton structure for the study of these time series?



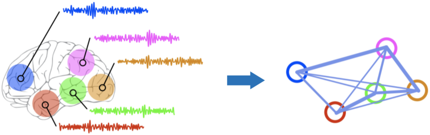








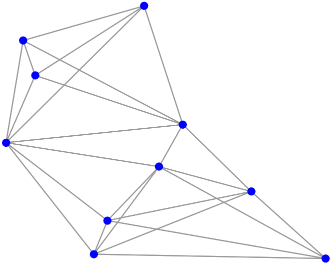
# Graph Signal Processing



- In most practical applications, the different dimensions of a multivariate signal $\mathbf{x}[t]$ are linked
  - Notion of correlation between recorded variables (ex: pressure/temperature/precipitation)
  - Sensor networks, body sensors, social networks...: spatial proximity, interactions...
- These links can be explicitly be modeled through a graph structure: **Graph Signal Processing** [Ortega et al., 2018]
- Each multivariate sample $\mathbf{x}[t]$ is assumed to be carried on the graph

# Contents

# What is a graph ?



A graph G is a triplet $(\mathcal{V}, \mathcal{E}, \mathcal{W})$

- $\mathcal{V}$ is a finite set of $D$ nodes or vertices (usually $\{1, 2, ..., N\}$)
- $\mathcal{E} \subset \mathcal{V} \times \mathcal{V}$ is a set of edges
- $\mathcal{W} : \mathcal{E} \to \mathbb{R}$ is a map from the set of edges to scalar values

Here : undirected graph, positive weights

$\mathcal{W}(i, j)$ encodes the strength of the relationship between dimensions $i$ and $j$

# Laplacian of a graph

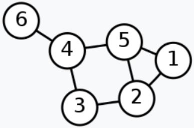
$$\boxed{\mathbf{L} = \mathbf{D} - \mathbf{W}}$$

- $\mathbf{W}$ (weight or adjacency matrix) :

$$W_{i,j} = \begin{cases} \mathcal{W}(i,j) & \text{if } (i,j) \in \mathcal{E} \\ 0 & \text{Otherwise} \end{cases}$$

- $\mathbf{D}$ (degree matrix) : diagonal matrix with

$$D_{i,i} = \sum_j W_{i,j}$$

| Labeled graph | Degree matrix | Adjacency matrix | Laplacian matrix |
|---|---|---|---|
|   | $\begin{pmatrix} 2&0&0&0&0&0 \\ 0&3&0&0&0&0 \\ 0&0&2&0&0&0 \\ 0&0&0&3&0&0 \\ 0&0&0&0&3&0 \\ 0&0&0&0&0&1 \end{pmatrix}$ | $\begin{pmatrix} 0&1&0&0&1&0 \\ 1&0&1&0&1&0 \\ 0&1&0&1&0&0 \\ 0&0&1&0&1&1 \\ 1&1&0&1&0&0 \\ 0&0&0&1&0&0 \end{pmatrix}$ | $\begin{pmatrix} 2&-1&0&0&-1&0 \\ -1&3&-1&0&-1&0 \\ 0&-1&2&-1&0&0 \\ 0&0&-1&3&-1&-1 \\ -1&-1&0&-1&3&0 \\ 0&0&0&-1&0&1 \end{pmatrix}$ |

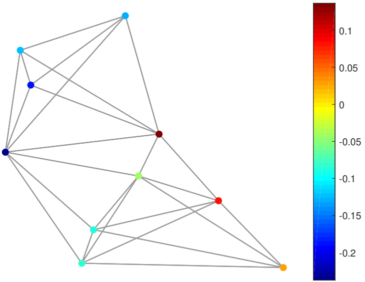
Symmetric, positive semi-definite

# Laplacian of a graph

By construction of the Laplacian matrix

- $\forall \mathbf{x} \in \mathbb{R}^N, \quad \mathbf{x}^T \mathbf{L} \mathbf{x} = \frac{1}{2} \sum_{(i,j) \in \mathcal{E}} W_{i,j} (x_i - x_j)^2 \geqslant 0$
- The constant vector $\mathbf{1}_N$ is an eigenvector for matrix $\mathbf{L}$ associated to eigenvalue $\lambda_1 = 0$
  Obvious as the sum of the matrix along the rows/column is equal to zero
- The number of connected components in the graph is equal to the number of eigenvalues equal to zero

# What is a graph signal ?



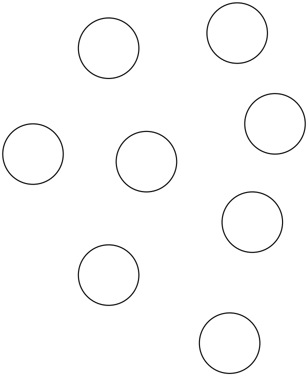

## Graph signal

Given a graph $\mathcal{G}$ of $N$ nodes, a **graph signal** is an array $\mathbf{x} \in \Omega^N$ that associates an element of $\Omega$ to each node of $\mathcal{G}$.

- $\Omega = \mathbb{R}$ : Simple signal
- $\Omega = \mathbb{R}^T$ : Time signal
- $\Omega = \mathbb{R}^d$ : Multivariate signal
- $\Omega = \mathbb{R}^{d \times T}$ : Multivariate temporal signal

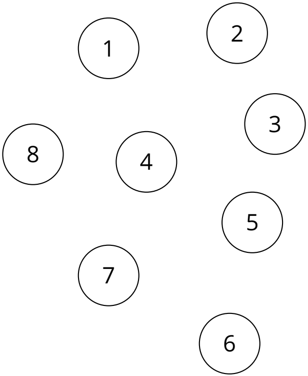
In most illustrations we will consider a single sample $\mathbf{x}[t]$ that belongs to $\mathbb{R}^N$ and will therefore define ONE graph signal

# Example



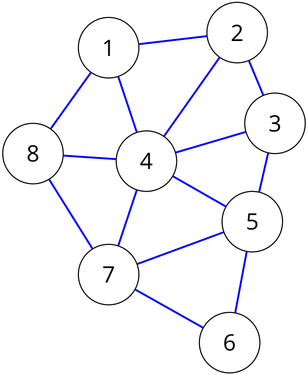
$N$ nodes: one per signal dimension

# Example



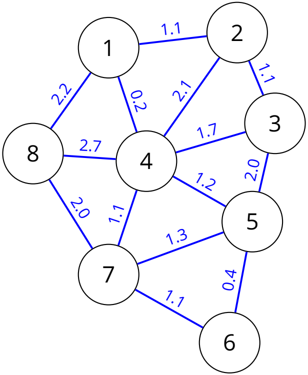

Dimension 1 lies on node 1, 2 on node 2, etc.

# Example



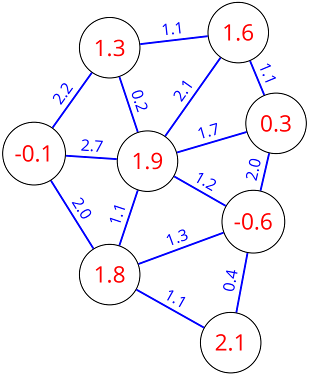

Edges model links between dimensions

# Example



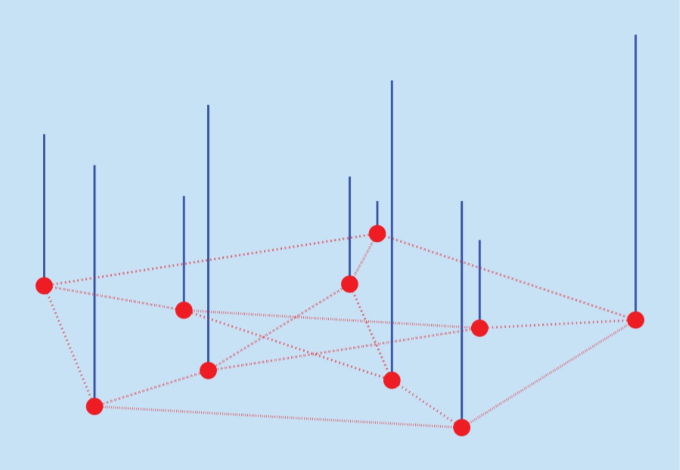

Weights model the strengths of these links

# Example

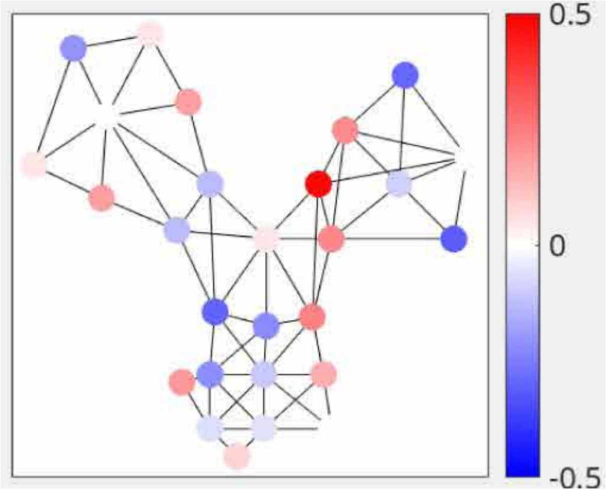



Visualization of one multivariate sample $\mathbf{x}[t]$ on the graph

# How to visualize graph signals?



# How to visualize graph signals?



# Examples

Different kinds of signals :

- Sensor networks (meteorology, population flux, energy consumption)
- Interaction networks (social networks, communication networks, ...)
- Economy based signals (market dependencies, stocks)
- Image processing (intensity, color)
- Cloud points (position, color)

# Tasks

- Several machine learning tasks can be extended to graph signals [Ortega et al., 2018]:
  - Sampling/compression: choose the most relevant nodes (i.e. dimensions) to reconstruct the whole data
  - Graph inference: learn the graph structure from data [Mateos et al., 2019]
  - Denoising/filtering: use the graph structure to remove noise, outliers... [Chen et al., 2014]
  - Interpolation: use the graph structure to reconstruct missing data [Narang et al., 2013]
  - Classification, event detection, anomaly detection, prediction...
- Use the structure to improve performances on multivariate time series

# Graph Fourier Transform

Given a signal $\mathcal{G}$ with only one connex component, we compute the eigen-decomposition of its Laplacian $\mathbf{L}$ :

$$\mathbf{L} = \mathbf{U} \begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_N \end{bmatrix} \mathbf{U}^T, \quad 0 = \lambda_1 < \lambda_2 \leqslant \ldots \leqslant \lambda_N$$

- $\lambda_i$ are interpretable as **frequencies** (see later for a more intuitive definition) $\lambda_1 = 0$ : DC component
- $\mathbf{u}_i$ is the eigenvector associated to frequency $\lambda_i$

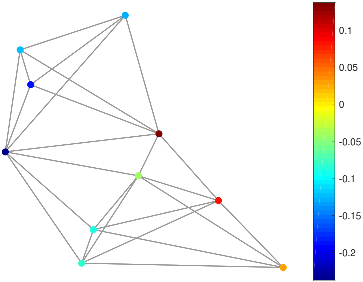
## Graph Fourier Transform

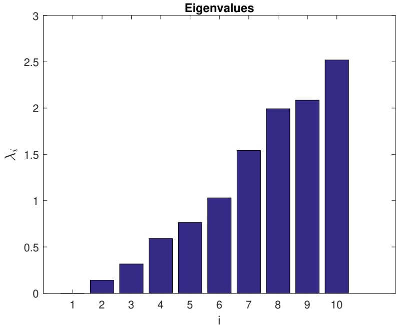
The Graph Fourier Transform (GFT) $\hat{\mathbf{x}}$ of a graph signal $\mathbf{x} \in \mathbb{R}^N$ is defined as

$$\boxed{\hat{\mathbf{x}} = \mathbf{U}^T \mathbf{x}}$$

# Example

$$\hat{\mathbf{x}} = \mathbf{U}^T \mathbf{x}$$



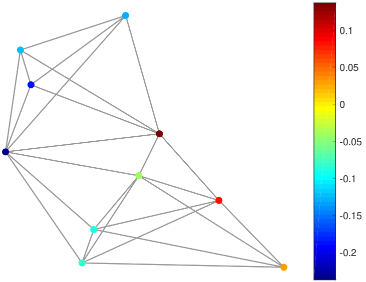



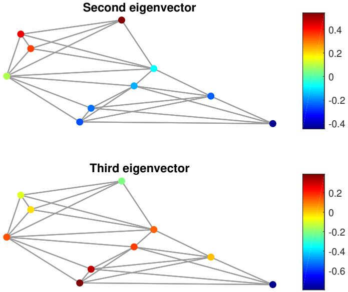

Eigenvalues $\lambda_i$ can be interpreted as *spatial frequencies*

Low frequencies : global phenomena, high frequencies : local phenomena

# Example

$$\hat{\mathbf{x}} = \mathbf{U}^T \mathbf{x}$$



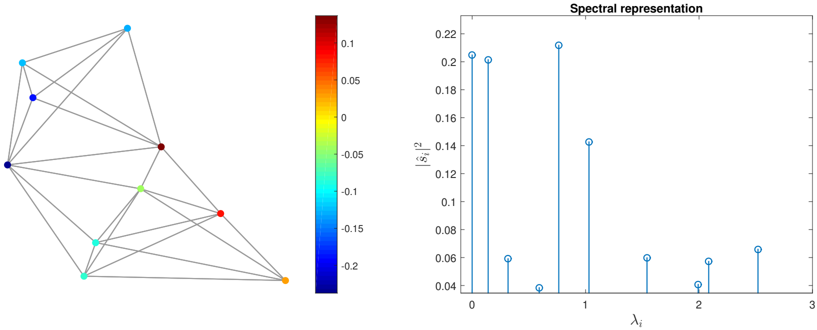




Eigenvectors $\mathbf{u}_2$ and $\mathbf{u}_3$

Can model symmetries, anti-symmetries, spatial phenomena

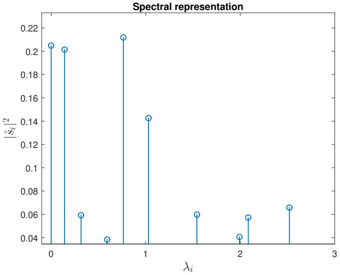
# Example

$$\hat{\mathbf{x}} = \mathbf{U}^T\mathbf{x}$$




Graph spectrum

# Bandlimitedness

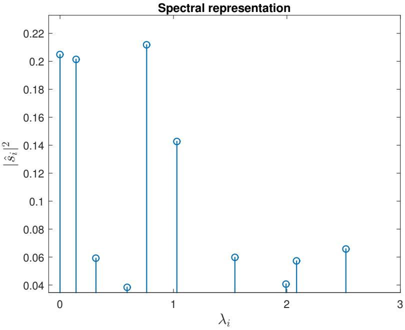


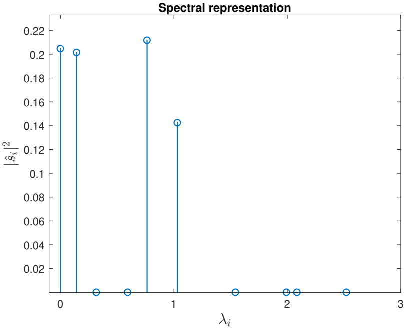

## Bandlimitedness

- Common assumption in signal processing : sparsity of the spectrum
- In SP : bandlimitedness of signals (baseband, wideband...) is used for sampling, denoising
- In GSP : same notion but on the graph spectrum

$\mathbf{x}$ is $K$-bandlimited iff. $\|\hat{\mathbf{x}}\|_0 = K$

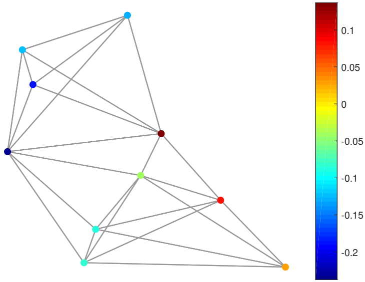
# Example

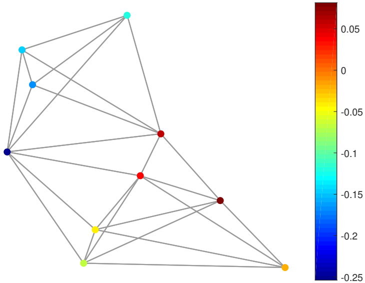



Filtering by removing all frequencies except for the 4 most dominant frequencies

# Example




4-bandlimited approximation

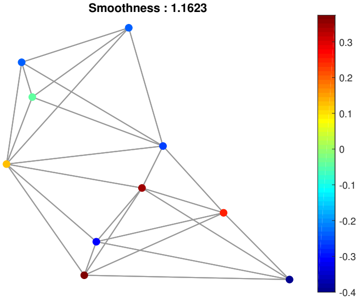
# Smoothness

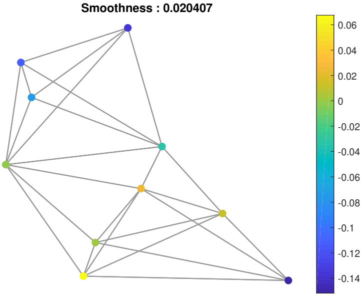
- Intuitively, signal values taken on adjacent nodes should be quite similar
- Notion of **smoothness** for a graph signal $\mathbf{x}$ :

$$\boxed{S(\mathbf{x}) = \mathbf{x}^T \mathbf{L} \mathbf{x} = \frac{1}{2} \sum_{(i,j)\in\mathcal{E}} W_{i,j}(x_i - x_j)^2}$$

- $S(\mathbf{x})$ is small if $(x_i - x_j)^2$ is small for large $W_{i,j}$
- Careful! This quantity in counterintuitive: large smoothness is achieved for non-smooth signals and vice-versa!

# Example




Smoothness decreases as the graph signal becomes more smooth

# Interpretation of the eigenvectors/eigenvalues

- For the eigenvectors of the Laplacian $\mathbf{u}_i$, we have

$$S(\mathbf{u}_i) = \mathbf{u}_i^T \mathbf{L} \mathbf{u}_i = \lambda_i$$

- New interpretation of the eigenvalues $\lambda_i$: **smoothness of the associated eigenvector**
- For one connex component, $\lambda_1 = 0$ and $\mathbf{u}_1 = \mathbf{1}_D$
  Constant eigenvector: perfect smoothness!

# How to interpret the graph spectrum

Parallels can be drawn between standard signal processing and graph signal processing:

- Notion of smoothness and low-frequency approximation
  Useful for denoising, interpolation...
- Notion of sparsity and bandlimitedness
  Useful for subsampling and reconstruction

# Outline

1. **Graph learning** from graph signals, based on the bandlimitedness and smoothness assumptions (with application to EEG/brain data)

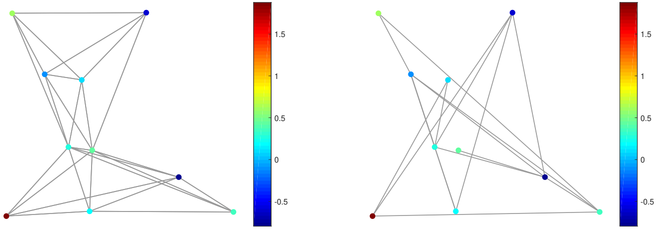
   - P. Humbert, B. Le Bars, L. Oudre, A. Kalogeratos, and N. Vayatis. Learning Laplacian Matrix from Graph Signals with Sparse Spectral Representation. Journal of Machine Learning Research, 22(195):1-47, 2021.
   - P. Humbert, L. Oudre, and C. Dubost. Learning spatial filters from EEG signals with Graph Signal Processing methods. In Proceedings of the International Conference of the IEEE Engineering in Medecine and Biology Society (EMBC), Guadalajara, Mexico, 2021.
   - B. Le Bars, P. Humbert, L. Oudre, and A. Kalogeratos. Learning laplacian matrix from bandlimited graph signals. In Proceedings of the International Conference on Acoustics, Speech, and Signal Processing (ICASSP), pages 2937-2941, Brighton, UK, 2019.

2. **Graph signal interpolation**, based on “non-smooth” assumption (with application to 3D movement analysis)

   - A. Mazarguil, L. Oudre, and N. Vayatis. Non-smooth interpolation of graph signals. Signal Processing, 196:108480, 2022.
   - A. Mazarguil, L. Oudre, and N. Vayatis. Localized interpolation for graph signals. In Proceedings of the European Signal Processing Conference (EUSIPCO), Amsterdam, The Netherlands, 2020.

# Contents

# Graph inference




- **Aim :** Given a collection of $n$ observed graph signals $\{\mathbf{y}^{(k)}\}_{k=1}^{n}$ of size $N$, learn the graph $\mathcal{G}$ that best explains the structure observed in the signals
- **Assumption :** The graph signals $\mathbf{Y}$ should be bandlimited and smooth for $\mathcal{G}$
- **Inputs :**

$$\mathbf{Y} = [\mathbf{y}^{(1)}, \cdots, \mathbf{y}^{(n)}] \in \mathbb{R}^{N \times n} : \text{input graph signals}$$

- **Outputs :**

$$\mathbf{L} = \mathbf{U}\Lambda\mathbf{U}^T : \text{Laplacian matrix of } \mathcal{G}$$

$$\hat{\mathbf{Y}} : \text{GFT of signals } \mathbf{Y} \text{ on } \mathcal{G}$$

# Problem formulation

$$\min_{\hat{\mathbf{Y}},\mathbf{U},\Lambda} \|\mathbf{Y} - \mathbf{U}\hat{\mathbf{Y}}\|_F^2 + \alpha\|\Lambda^{1/2}\hat{\mathbf{Y}}\|_F^2 + \beta\|\hat{\mathbf{Y}}\|_S$$

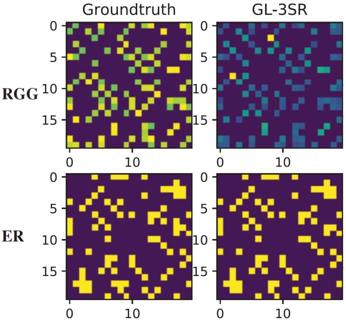
$$\text{s.t.} \quad \begin{cases} \mathbf{U}^T\mathbf{U} = \mathbf{I}_N, \ \mathbf{u}_1 = \frac{1}{\sqrt{N}}\mathbf{1}_N, & \text{(a)} \\ (\mathbf{U}\Lambda\mathbf{U}^T)_{k,l} \leqslant 0 \quad k \neq l, & \text{(b)} \\ \Lambda = \mathrm{diag}(0, \lambda_2, \ldots, \lambda_N) \succeq 0, & \text{(c)} \\ \mathrm{tr}(\Lambda) = N \in \mathbb{R}_*^+. & \text{(d)} \end{cases}$$

- $\|\mathbf{Y} - \mathbf{U}\hat{\mathbf{Y}}\|_F^2$ : $\mathbf{Y}$ should be close to the inverse GFT of its spectral representation in $\mathcal{G}$
- $\|\Lambda^{1/2}\hat{\mathbf{Y}}\|_F^2$ : $\mathbf{Y}$ should be smooth on $\mathcal{G}$
- $\|\hat{\mathbf{Y}}\|_S$ : sparsity constraint on the frequency representation of $Y$

Constraints : $\mathbf{L} = \mathbf{U}\Lambda\mathbf{U}^T$ should be a Laplacian matrix (symmetric, semi positive)
Resolution with alternate minimization

# Synthetic data

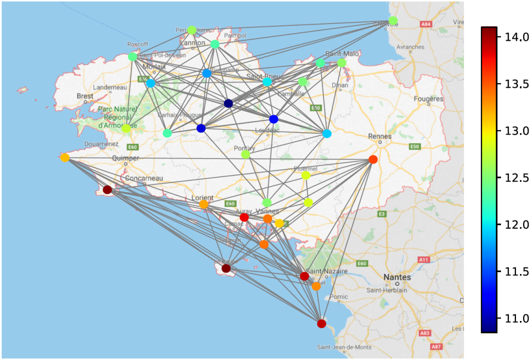
- Two synthetic graphs : Random Geometric Graph (RGG) and Erdos-Rényi Graph (ER)
- Noisy graph signals with $n = 1000$, $N = 20$ and 10-bandlimited




Comparison of the adjacency matrices

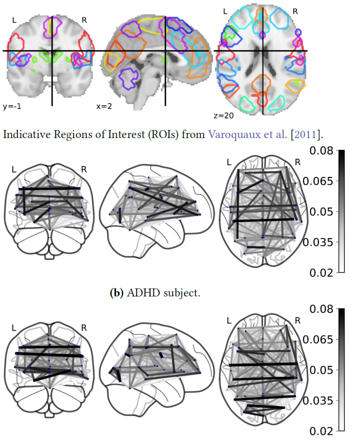
# Real data: temperature data

- Temperature in Brittany (32 weather stations, 747 graph signals)



# Real data: fMRI data

- 40 subjects (20 healthy, 20 ADHD), $N = 39$ regions of interest
- We also used the graph to classify the subjects : 65% (52.5% with standard correlation graphs)




**(a)** Indicative Regions of Interest (ROIs) from Varoquaux et al. [2011].

**(b)** ADHD subject.

**(c)** Healthy subject.

**Figure 2.14:** (a) Indicative ROIs from the Multi-Subject Dictionary Learning atlas extracted in Varoquaux et al. [2011] with sparse dictionary learning. Results: Graphs returned by FGL-3SR, separately for (b) an ADHD patient and (c) a healthy subject, where darker edges indicate larger weights of connection.

# Contents

# Interpolation of missing samples

- **Aim :** Given a multivariate time series $\mathbf{Y}$ of $n$ samples recorded on $N$ sensors with missing values and a graph $\mathcal{G}$, reconstruct the missing values
- **Assumption :** The missing graph signal values can be reconstructed by using the neighborhood nodes
- **Inputs :**

$$\mathbf{Y} \in \mathbb{R}^{N \times n} : \text{input graph signals}$$

$$\mathcal{U} \text{ and } \mathcal{K} : \text{sets of unknown/known samples}$$

$$\mathcal{G} : \text{graph}$$

- **Outputs :**

$$\mathbf{Y}_{\mathcal{U}} : \text{missing samples imputation}$$

# Problem formulation

$$\min_{\mathbf{Y},\mathbf{A},\mathbf{b}} \|\mathbf{Y} - (\mathbf{AY} + \mathbf{b}\mathbf{1}_N^T)\|_F^2 + \mu \,\mathrm{Loc}(\mathbf{A})$$

where

- $\|\mathbf{Y} - (\mathbf{AY} + \mathbf{b}\mathbf{1}_N^T)\|_F^2$ : $\mathbf{Y}$ should follow a linear structural equation model
- $\mathrm{Loc}(\mathbf{A}) = \sum_{i,j} d_{\mathcal{G}}^2(i,j) a_{i,j}^2$ is a localization term. This penalty ensures that the contributions for signal reconstruction on node $i$ is mostly carried on a set of nodes that are close (according to the geodesic distance on the graph) to $i$

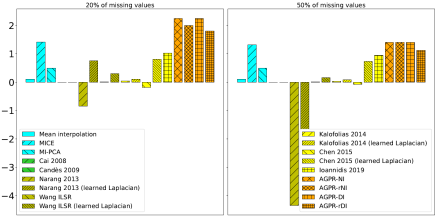
Biconvex problem according to $\mathbf{Y}$, $(\mathbf{A}, \mathbf{b})$

Resolution with alternate minimization

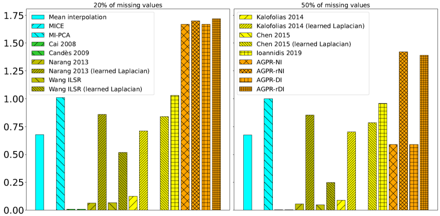
Two resolution methods: one based on closed form solution (heavy) and one relaxed iterative method

# Real data

- Normalized Root Mean Square Error (in dB) for several datasets



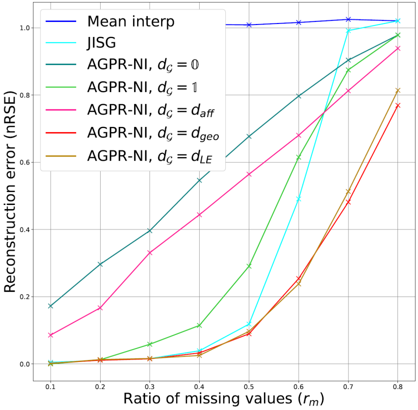

(e) nRSE-db for the **mocap** dataset.




(f) nRSE-db for the **walking dog** dataset.

# Is the graph important ?

- Comparison of several distances on synthetic data
- Using the graph information is especially relevant when the percentage of missing data is large



# Contents

# Conclusion

- Interpretable assumptions such as smoothness or bandlimitesness can be useful for several tasks (sampling, learning, interpolation...)
- Graph Signal Processing allows to take into account the relationships and correlation observed in multivariate data
- Versatile framework : graphs can model several types of interactions
- Various applications in healthcare : sensor networks, EEG, multisensors...