MODELING SIGNALS OVER DIRECTED GRAPHS THROUGH FILTERING

Pierre Borgnat

Senior Research Fellow CNRS – Physics Laboratory, ENS de Lyon, France
Sisyphe group (Signals, Systems and Physics)
and IXXI (Complex System Institute Lyon)

GlobalSIP, Anaheim, CA, 11/2018

Scope of the work

- Motivation: tasks of signal modeling on **directed graphs**
- Extension of Graph Signal Processing to digraph
- Numerical explorations about parametric signal modeling

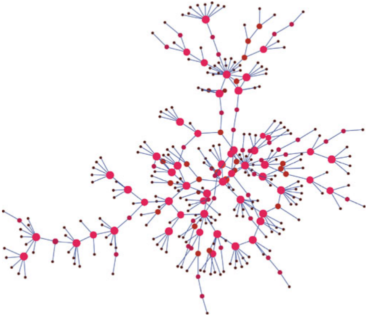
- Joint work with Harry Sevi (PhD defended last week) and Gabriel Rilling (CEA List)

- Work supported by:
  - ANR-14-CE27-0001 GRAPHSIP grant
  - ACADEMICS grant in the Scientific Breakthrough Program IDEXLYON

A mandatory slide on Graph Signal Processing

- Given a graph G, let's consider a signal s on the nodes V
- How to apply signal processing on this data / signal ?

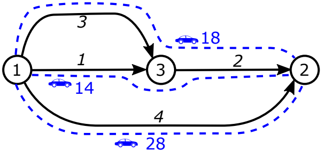
Epidemiological network



Undirected graph

[G.Ghoshal (2009), Potterat et al. (2002)]

Road network



Directed graph

[G. Michau, PB et al., 2017]

A Fundamental analogy for undirected graphs

[Shuman et al., *IEEE SP Mag*, 2013]

A fundamental analogy

On *any* graph, the **eigenvectors χ_i of the Laplacian matrix $\mathbf{L} = \mathbf{D} - \mathbf{A}$ will be considered as the Fourier modes**, and its eigenvalues λ_i the associated (squared) frequencies.

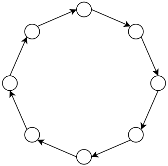
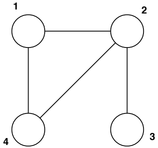
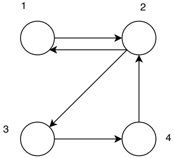
$$\boxed{\hat{\mathbf{s}} = \chi^\top \mathbf{s}}$$

where $\chi = (\chi_1 | \chi_2 | \dots | \chi_N)$

- Two ingredients:
  - **Fourier modes** = Eigenvectors χ_i
  - **Frequencies** = Eigenvalues related to variation: $\frac{||\nabla \chi_i||^2}{||\chi_i||^2} = \lambda_i$, because

$$\forall \mathbf{x} \in \mathbb{R}^N \quad \sum_{e=(i,j)\in E} A_{ij}(\mathbf{x}_i - \mathbf{x}_j)^2 = \mathbf{x}^\top \mathbf{L} \mathbf{x}$$

What about directed graphs ?

Graph	cyclic	undirected	directed
			
Fourier Modes	$e^{i\omega t}$	χ	?
Operator		$\mathbf{L}$	?
Frequency	ω	λ	?
Variation		$\langle \chi, \mathbf{L}\chi \rangle$	?

Measure of Variations

Undirected:

$$\mathrm{VQ}(\boldsymbol{f}) = \sum_{i,j} a_{ij} |f_i - f_j|^2$$
$$= \langle \boldsymbol{f}, \mathbf{L}\boldsymbol{f} \rangle$$

with

$$\mathbf{L} = \mathbf{D} - \mathbf{A}.$$

Directed:

$$\mathcal{D}^2_{\pi,\mathbf{P}}(\boldsymbol{f}) = \frac{1}{2} \sum_{i,j} \pi_i p_{ij} |f_i - f_j|^2.$$
$$= \langle \boldsymbol{f}, \mathbf{L}_{dir} \boldsymbol{f} \rangle.$$

with

$$\mathbf{L}_{dir} = \boldsymbol{\Pi} - \frac{\boldsymbol{\Pi}\mathbf{P} + \mathbf{P}^\top \boldsymbol{\Pi}}{2}$$

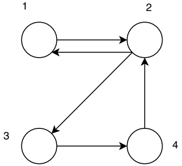
- Directed case
  - use of $\mathbf{P} = \mathbf{D}^{-1}\mathbf{A}$ the random walk operator
  - and its associated stationary distribution $\boldsymbol{\pi}$
- Undirected case : $\boldsymbol{\Pi} \propto \mathbf{D} \Rightarrow \mathbf{L}_{dir} \propto \mathbf{L}$.

Fourier modes on directed graphs

Random walk operator

- Random walk X_n : position X at time n.
- $\mathbf{P}_{ij} = \mathbb{P}(X_n = j | X_{n-1} = i)$ is its transition probability

$$\mathbf{P} = \begin{pmatrix} 0 & 1 & 0 & 0 \\ \frac{1}{2} & 0 & \frac{1}{2} & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \end{pmatrix} = \mathbf{D}^{-1}\mathbf{A}$$



Proposition of Fourier Modes

- Eigenvectors $\mathbf{P}\boldsymbol{\xi}_k = \theta_k \boldsymbol{\xi}_k$

$\boldsymbol{\Xi} = [\boldsymbol{\xi}_1, \dots, \boldsymbol{\xi}_N]$ the basis

- Fourier representation of $\boldsymbol{s}$

$$\boldsymbol{s} = \sum_k \hat{s}_k \boldsymbol{\xi}_k = \boldsymbol{\Xi}\hat{\boldsymbol{s}}.$$

where $\hat{\boldsymbol{s}} = [\hat{s}_1, \dots, \hat{s}_N]^\top$ are the Fourier coefficients

- *Digraph Fourier Transform* :

$$\hat{\boldsymbol{s}} = \boldsymbol{\Xi}^{-1}\boldsymbol{s}.$$

- Beware : complex eigenvalues : $\theta = \alpha + \boldsymbol{i}\beta$, $|\theta| \leq 1$.

Frequency analysis of modes of **P**

Fourier Modes:

$$[\xi_1, \ldots, \xi_N]$$

Variations:

$$\mathcal{D}^2_{\pi,\mathbf{P}}(\boldsymbol{f}) = \langle \boldsymbol{f}, \mathbf{L}_{dir} \boldsymbol{f} \rangle$$

Frequency analysis:

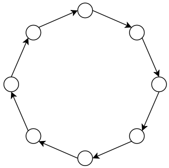
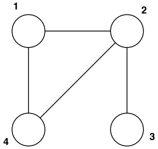
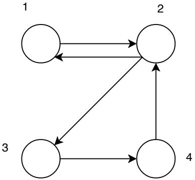
$$\frac{\mathcal{D}^2_{\pi,\mathbf{P}}(\xi)}{\langle \xi, \mathbf{\Pi} \xi \rangle} = 1 - \mathfrak{Re}(\theta)$$

- Let's define the **frequency** of ξ from its complex eigenvalue θ :

$$\omega = 1 - \mathfrak{Re}(\theta), \quad \omega \in [0, 2]$$

["Analyse fréquentielle et filtrage sur graphes dirigés", Sevi et al., GRETSI, 2017]

Summary of the proposed framework

Graphe	cyclic	undirected	directed
			
Fourier Mode	$e^{i\omega t}$	φ	ξ
Operator		$\mathbf{L}$	$\mathbf{P}$
Frequency	ω	λ_k	$\omega = 1 - \Re\mathfrak{e}(\theta)$
Variation		$\langle \varphi, \mathbf{L}\varphi_k \rangle$	$\langle \xi, \mathbf{L}_{dir}\xi \rangle$

Parametric modeling

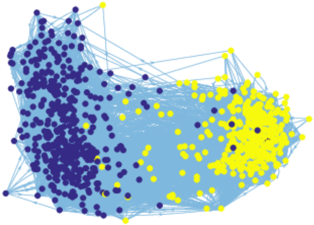
Problem formulation

- Model a signal, eg., for compression or inpainting
- Assumption: a partial observation $\boldsymbol{y}$ of $\boldsymbol{f}$

Objective

- Estimate the missing data points from $\boldsymbol{y}$

Dataset of test: Political Blogs US 2004 [Adamic et al., 2004]



- **Node** : a blog
- **Edge** : hyperlink from a blog to another
- **Signal** : political side (Republican / Democrat).

Solution of the problem

- We observe $\boldsymbol{Y}_k = \varepsilon_k \boldsymbol{f}_k$, where the $\varepsilon_k = 1$ if known, else 0
- Decide upon a **reference operator**, noted $\mathbf{R}$, first $\mathbf{R} = \mathbf{P}$ or $\mathbf{A}$
- Model the signal thanks to a parametric graph filter $\mathbf{H}$:

$$\mathbf{H}(\boldsymbol{\theta}) = \sum_{k=0}^{K} \theta_k \mathbf{R}^k, \quad \theta_k \in \mathbb{R}. \tag{1}$$

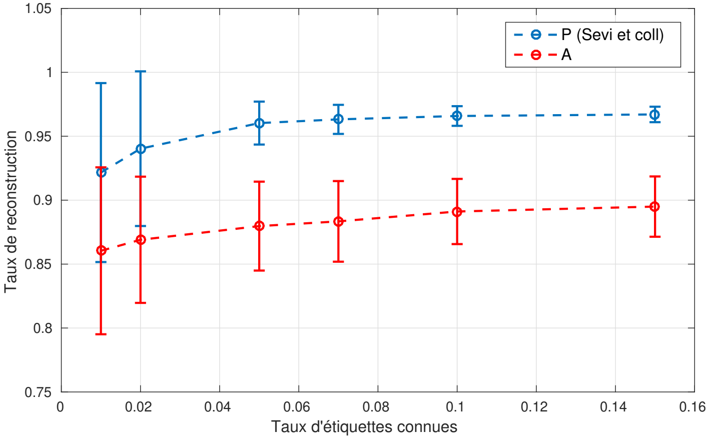
- Parameter estimation

$$\widehat{\boldsymbol{\theta}} = \underset{\boldsymbol{\theta}=\{\theta_k\}_{k=0}^{K} \in \mathbb{R}^{K+1}}{\operatorname{argmin}} \mathbb{E}\left[\|\boldsymbol{f} - \sum_{k=0}^{K} \theta_k \mathbf{R}^k \boldsymbol{Y}\|_{\mu}^2 \right], \tag{2}$$

- (that has well-known solution)
- **Signal model** :

$$\hat{\boldsymbol{f}}(\theta) = \sum_{k=0}^{K} \hat{\theta}_k \mathbf{R}^k \boldsymbol{Y}$$

Experimental results (1)



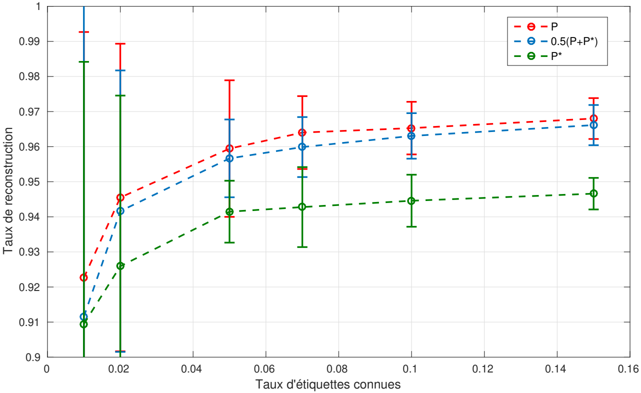
Alternative Reference Operator (1)

Other Reference operators R could be used :

- $\mathbf{P}^*$, associated to the time reversed random walk: $\mathbf{P}^* = \mathbf{\Pi}^{-1}\mathbf{P}^\top\mathbf{\Pi}$.
- $\bar{\mathbf{P}}$, the additive reversibilization of $\mathbf{P}$: $\bar{\mathbf{P}} = \dfrac{\mathbf{P}+\mathbf{P}^*}{2}$.

Prop.: $\mathbf{P}, \mathbf{P}^*, \bar{\mathbf{P}}$ lead all to DiGFT with frequency related to Variations

Experimental results (2)



Alternative Reference Operator (2)

- $\mathbf{P}^*$, associated to the time reversed random walk: $\mathbf{P}^* = \mathbf{\Pi}^{-1}\mathbf{P}^\top\mathbf{\Pi}$.
- $\bar{\mathbf{P}}$, the additive reversibilization of $\mathbf{P}$: $\bar{\mathbf{P}} = \dfrac{\mathbf{P}+\mathbf{P}^*}{2}$.

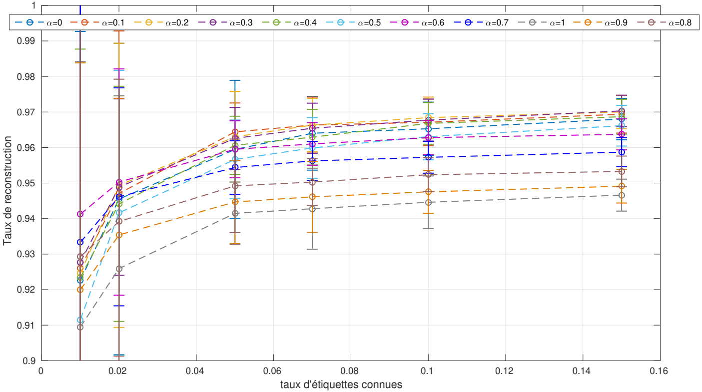
Generalization: convex combination between $\mathbf{P}$ and $\mathbf{P}^*$

$$\mathbf{P}_\alpha = (1-\alpha)\mathbf{P} + \alpha\mathbf{P}^*$$

for $\alpha \in [0,\ 1]$.

Prop.: $\mathbf{P}_\alpha$ leads all to DiGFT with frequency related to Variations

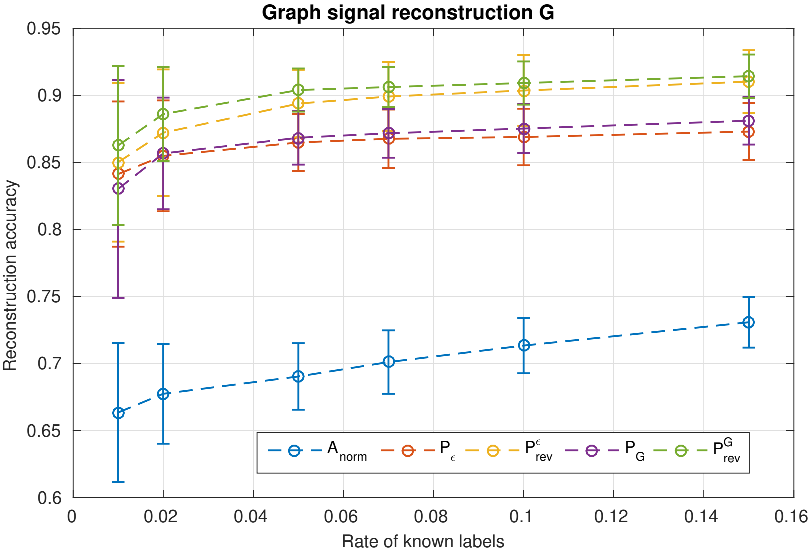
Experimental results (3)



Further numerical explorations

- Results depends on the sampling law for ε_k where $\boldsymbol{Y}_k = \varepsilon_k \boldsymbol{f}_k$
- A limit of choosing **P**: it requires a strongly connected graph...
- 1) use connected components,
- or 2) modify the graph
  - add a small rank-one perturbation (Cons: non-sparse)
  - construct the “google” matrix: complete dangling nodes (i.e., nodes with $d^{out} = 0$) and then add a probability of jumping anywhere

Experimental results (4)



Conclusion

- Use of GSP on directed graphs
- A full framework to generalize Laplacian-based approaches to digraphs,
  - using random walk (or generalisations $\mathbf{P}_\alpha$) as Reference operator
  - and $\mathbf{L}_{dir}$ to measure variations and define frequency
- A numerical exploration around the task of signal modeling
- More developments:
  spectral wavelets and diffusion wavelets with **P** on digraphs
- Contact and more information:

http://perso.ens-lyon.fr/pierre.borgnat