# Semi-Supervised Learning for Graph to Signal Mapping: a Graph Signal Wiener Filter Interpretation

Benjamin Girault, Paulo Gonçalves, Eric Fleury, Arashpreet Singh Mor

## HAL Id: hal-00942695
## https://hal.inria.fr/hal-00942695v2

Submitted on 14 May 2014

# SEMI-SUPERVISED LEARNING FOR GRAPH TO SIGNAL MAPPING: A GRAPH SIGNAL WIENER FILTER INTERPRETATION

*Benjamin Girault*[*†]     *Paulo Gonçalves*[◇†]     *Éric Fleury*[*†]     *Arashpreet Singh Mor*[◇†‡]

[*] École Normale Supérieure de Lyon
[◇] Inria
[†] LIP (UMR CNRS – ENS Lyon – UCB Lyon 1 – Inria 5668), Labex MILYON, Université de Lyon
[‡] Indian Institute of Technology Delhi, Department of Electrical Engineering

## ABSTRACT

In this contribution, we investigate a graph to signal mapping with the objective of analysing intricate structural properties of graphs with tools borrowed from signal processing. We successfully use a graph-based semi-supervised learning approach to map nodes of a graph to signal amplitudes such that the resulting time series is smooth and the procedure efficient and scalable.

Theoretical analysis of this method reveals that it essentially amounts to a linear graph-shift-invariant filter with the *a priori* knowledge put into the training set as input. Further analysis shows that we can interpret this filter as a Wiener filter on graphs. We finally build upon this interpretation to improve our results.

***Index Terms***— Signal processing on graphs, Semisupervised learning, Spectral analysis, Network science

## 1. INTRODUCTION

Network science is expanding rapidly with the availability of large datasets issued from complex systems. In a wide range of applications, the amount of collected data comes with an intricate organisation that is naturally or adequately captured by a graph structure. In that case, the relevant information often lies in the weighted edges that measure the strength binding entities viewed as nodes of a graph. Depending on the entities — computers in communication networks [1], individuals in social or peer-to-peer networks [2] or proteins in biological systems [3], to cite only a few — the nature of the interactions can be manifold: capacities, contacts, similarities... However, usual graph analytics do not always succeed in revealing or characterising the relational complexity of the underlying systems, such as communities or hidden patterns. One alternative, constructed on a duality principle, is then to associate to the graph structure a dual random walk time series whose states correspond to the nodes and where the transitions are encoded by the graph edges [4]. Thanks to this duality, graph analytics transposes to a signal analysis problem that can benefit from a rich signal processing toolkit.

In other scenarios, each node of a graph is itself the support of a measure (scalar or vector) that can be interpreted as one sample of a graph signal. This is for instance the case for temperatures delivered by a sensor network [5], activity of brain functional regions [6], or bacteria strains carried by interacting people [7]. The analysis of these high-dimensional signals supported by complex structures, as arbitrary graphs, generates a crescent interest [8] but faces the difficulty that they do not necessarily live in a Euclidean vector space. Therefore, extending the standard algorithms of signal (or image) processing to these structured objects is a topical challenge that already has marked successes. That is notably the case with the diffusion wavelets [9], the wavelets on graphs [10] and more generally the emerging field of discrete signal processing on graphs [8, 11].

In this contribution, we first tackle the problem of finding a sensible mapping between the nodes of a graph and the states of a dual time series. To this end, we propose to use graph-based semi-supervised learning methods [12] and we show that the results outperform standard approaches in many ways: *(i)* the resulting mapping entails dual signals that are smoother; *(ii)* the solution is unique; and *(iii)* algebraically scalable to massive datasets.

In a second step, we give this graph-based semi-supervised learning mapping an interpretation in terms of spectral filtering. Notably, we demonstrate that it corresponds to a linear graph-shift invariant filter, in accordance with the theoretical framework developed in [11]. In addition, we introduce the notion of Wiener filter in the context of graph signals, thus contributing to enrich the toolbox of graph signal processing.

## 2. DUALITY BETWEEN TIME SERIES AND GRAPHS

### 2.1. Notations and State of the Art

We consider a discrete graph structure that we note $G = (V, E)$, with $V$ the set of nodes ($|V| = N$) and $E \subseteq V \times V$ the set of edges between any two nodes. We call $W$ the adjacency matrix whose entries $\{w_{ij}\}_{1 \leq i,j \leq N}$ are the weights of each existing edge $e = (i, j) \in E$ and zero if the nodes $i$ and $j$ are not connected. Let $D = \{d_{ii} = \sum_{j=1}^{N} w_{ij}\}_{1 \leq i \leq N}$ be the diagonal matrix of the nodes' degree. Finally, we denote $\mathcal{L} = D^{-1/2}(D - W)D^{-1/2}$ the *normalised graph Laplacian*.

It is common practice to map time series into graph objects with the aim of using the relational or the geometric structures of the latter to better characterise certain hidden dynamics of the former (see [13] for a survey of a number of proposed mappings). Although the converse transformation is less frequent, it raised a recent and strong interest where the rich arsenal of signal processing techniques could be exploited to reveal interesting structural or dynamic properties of complex networks.

Schematically, we can distinguish two possible approaches to perform this mapping: one relies on a decomposition of the matrix form $W$ of the graph into *node-series* that can be interpreted as many dual time series of the graph. A particularly interesting case uses *multidimensional scaling* methods and shows very promising results [14, 15]. On the other hand, a more direct approach consists to establish a one-to-one correspondence between the nodes of the
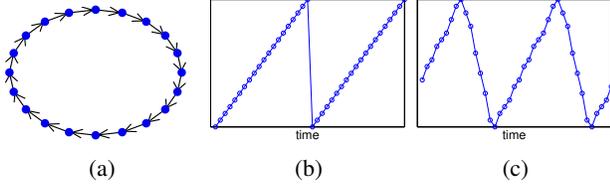
**Fig. 1**. (a) Cyclic oriented graph ($N = 20$). (b)-(c) Periodic time series, dual of the cyclic graph and corresponding to two distinct permutations $p^*$, solutions of Eq. (1)

graph and a set of values (integers, real intervals, sequences...) that the time series can take on. Then, after proper normalisation, the weights $w_{ij}$ are interpreted as transition probabilities between the different amplitude levels of a random walk.

In this direction, the certainly most accomplished work is presented in [4] and proposes to find a permutation (or renumbering) $p : \{1, \ldots, N\} \rightarrow \{1, \ldots, N\}$ so as to reorder the $N$ nodes of the graph according to their affinity as measured by the weights. More precisely, the sought permutation $p^*$ is solution of the following optimisation problem:

$$p^* = \operatorname*{argmin}_p \sum_{i=1}^{N} \sum_{j=1}^{N} w_{ij} |p(i) - p(j)|, \qquad (1)$$

imposing $p(i)$ and $p(j)$ to be close in the mapped space when $w_{ij}$ is large, that is for strongly tied nodes $i$ and $j$. The dual time series are then generated as random walks visiting the different mapped amplitude states. One motivation behind this procedure is to get smooth time series, by precluding too frequent strong discontinuities in the signal due to transitions between distant amplitudes.

As a striking illustration, we consider the oriented cyclic graph of Fig. 1a to which we apply the procedure described in [4]. Here, assuming an initial sequential numbering of the nodes, and owing to the cyclic structure of the graph, one trivial permutation $p^*$ is the identity yielding the periodic dual time series of the plot 1b. However, the optimisation problem of Eq. (1) is a NP-hard quadratic assignment problem, which does not systematically have a unique solution. To circumvent this issue, the authors of [4] propose to use the simulated annealing heuristic that leads to an approximated solution of $p^*$ with no closed-form expression. In the case of the cyclic graph, we can easily verify that other permutations satisfying (1) exist, yielding different periodic time series as the one plotted in Fig. 1c.

Regarding smoothness, the sawtooth solution shows one strong discontinuity per period. Intuitively though, one would expect the cyclic graph to be dual of a discretised pure sine wave, a $C^\infty$ periodic function. However, we can easily convince ourselves that no mapping of $\{1, \ldots, N\}$ onto itself can lead to a pure harmonic signal as solution of the dual problem stated in [4].

### 2.2. Duality and Graph-Based Semi-Supervised Learning

We propose to reformulate the mapping problem exposed in the previous section, as a classification problem using *graph-based semi-supervised learning* (SSL) methods [12]. Before adapting it to our specific problem, we first recall the main lines of SSL. The goal of graph-based SSL is to assign each node $i \in V$ of a non-oriented graph $G = (V, E)$ of size $N$, to a class $k(i) \in \{1, \ldots, K\}$, taking into account relations of dependence between nodes, via the weights $\{w_{ij}\}_{1 \le i, j \le N}$, and using a few labelled nodes as training set.

Formally, we write this training set as a label matrix $Y$ of size $N$-by-$K$ where each row corresponds to a node in the graph, and each column corresponds to a class $k \in \{1, \ldots, K\}$. For a labelled node $i$, for which a ground truth $k(i) \in \{1, \ldots, K\}$ is known, only the entry $Y_{i,k(i)} = 1$ while $Y_{i,k'} = 0$, $\forall k' \neq k(i)$. For unlabelled nodes, the corresponding rows of $Y$ are identically set to zero. The objective of the generalised graph-based SSL derived in [12] is to find the classification matrix $F$ of the same size as $Y$ and solution of the following optimisation problem:

$$\min_F \Big\{ \sum_{k=1}^{K} \underbrace{\|D^{\sigma-\frac{1}{2}}(F_{.k} - Y_{.k})\|_2^2}_{\text{fitting term}} + \alpha \underbrace{\|D^{\sigma-\frac{1}{2}} F_{.k}\|_{\mathcal{L}}^2}_{\text{penalty}} \Big\}. \qquad (2)$$

We readily recognise in this setting the expression of a regularised inverse problem where the fitting term forces the sought solution $F$ to match the training set $Y$ and where the penalising term constrains the solution $F$ to be smooth with respect to the Laplacian semi-norm of the underlying graph. In this expression also, the regularisation factor $\alpha$ allows to tune the trade-off between these two effects and $\sigma$ is a free unifying parameter that was introduced in [16] to get a canonical formulation of different semi-supervised machine learning methods: standard ($\sigma = 1$) or normalised ($\sigma = 1/2$) Laplacian, PageRank ($\sigma = 0$). Then, according to [16, 17], the solution of (2) has a closed-form expression that reads:

$$
\begin{aligned}
F &= C_{\sigma, \alpha} Y \\
\text{s.t.} \quad C_{\sigma, \alpha} &= \big[(1 + \alpha) I_N - \alpha D^{-\sigma} W D^{\sigma-1}\big]^{-1}. \quad (3)
\end{aligned}
$$

The classification task is then to assign to each node $i$ the class $k(i) = \operatorname*{argmax}_{k'} F_{i,k'}$. As it is thoroughly commented in [18], SSL methods show to perform remarkably well when only a few labelled points are known.

Let us now apply this SSL scheme to our mapping problem, seeking for a more relevant and easier to determine mapping between the nodes of the graph and the amplitudes set of the corresponding time series. To start with, we must identify sensible classes of the nodes and a corresponding training set. Actually, this choice is relatively free, and therefore we adopt the most obvious one that amounts to identify two most distant nodes in the graph (*w.r.t.* the shortest path) and to label those as the centroids of two distinct classes. With this choice, the label matrix $Y$ encompasses only two columns, each of them having only one non-zero entry for its associated centroid. Then, computing straightforwardly the solution given by expression (3), yields for a given $\sigma$ and $\alpha$, a classification matrix $F$ with only two columns and real entries comprised in the interval $[0, 1]$. Now, instead of sorting the coefficients $F_{i,1}$ and $F_{i,2}$ to determine the class of each node $i$, we form the function defined on each node:

$$X(i) = \frac{F_{i,1} - F_{i,2}}{F_{i,1} + F_{i,2}}, \qquad (4)$$

corresponding to its relative position with respect to the two classes. It is this function that we directly associate to a mapping from the nodes indices $\{1, \ldots, N\}$ onto the time series amplitudes lying in the interval $[-1, 1]$.

Applied to our toy example of an oriented cyclic graph[1], the resulting dual time series corresponding to different choices for the

---

[1] To apply SSL methods on oriented graphs, we first symmetrize the weight matrix $(W^T + W)/2$. This operation does not alter the proximity relations between nodes, and therefore does not hinder the pertinence of the mapping. The orientation of the links is taken into account when generating the states transitions in the dual random walk.
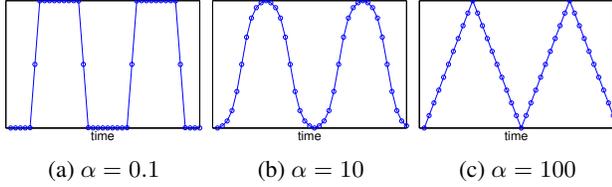
(a) $\alpha = 0.1$      (b) $\alpha = 10$      (c) $\alpha = 100$

**Fig. 2**. Dual signals of the non-oriented cyclic graph where the mappings stem from graph-based semi-supervised learning methods ($\sigma = 1/2$) with different regularisation parameters $\alpha$.

regularisation parameter $\alpha$ are plotted in Fig. 2. Let us notice that in this particular case, the matrix $D$ of degrees is equal to the identity matrix and therefore the choice of $\sigma$ has no effect on the solution $F$. In contrast to the permutation solution of Eq. (1), each of these mappings is uniquely derived from the closed-form expression (3) and scalable to very large datasets [19]. Moreover, regarding our original motivation about smoothness of the dual time series, a pertinent choice of $\alpha$ yields a harmonic-like solution (see Fig. 2b) as the paradigm that one would naturally expect for a smooth periodic function.

## 3. SEMI-SUPERVISED LEARNING AS A GRAPH SIGNAL FILTER

As we saw in the previous section, the SSL classifier can be viewed as an operator whose input is a $K$-dimensional signal $Y$ defined on the graph and the output, a transformed version $F$ of it. We now show in a more general setting, that the SSL solution of Eq. (2) is the output of a linear, shift-invariant filter consistent with the discrete signal processing on graphs theory developed in [11]. Furthermore, we identify the frequency response of this regularising SSL operator to a Wiener filter.

Thereafter, we consider only non-oriented graphs with doubly stochastic adjacency matrices $W$, leading in particular to identity matrices $D$ and to Laplacians of the form $\mathcal{L} = I_N - W$.

### 3.1. Graph filters

In [11], the authors generalised the *linear shift-invariant* filter theory on discrete signals to graph signals. In their framework, the usual time-shift operator is replaced by a *graph-shift* operator simply defined as the multiplication by the adjacency matrix. Then, a filter on a graph signal is said to be *linear shift-invariant* if it verifies the superposition principle and commutes with the graph-shift operator.

We use two important results from [11]: *(i)* a filter $H$ is linear shift invariant on a graph signal if and only if it can be written as a polynomial in the graph shift operator $h(W)^2$. *(ii)* Given the diagonalisation $W = U\Lambda U^T$ and identifying eigenvectors of the graph shift to eigenmodes, by analogy to the standard case, the Fourier operator on graphs is defined by $\mathcal{F} = U^{-1} = U^T$. Formally then, the analogue of the convolution of a graph signal $\mathbf{s}$ with a filter $H$ can be expressed as $\widehat{H\mathbf{s}} = h(\Lambda)\widehat{\mathbf{s}}$, where $\widehat{\cdot}$ denotes the Fourier transform and $h$ is the polynomial form of $H$.

---

[2]Additionally, a general condition is assumed on $W$ which in our case amounts to say that all the eigenvalues of $W$ must have a one-multiplicity. This condition is circumvented, as suggested in [11], using an equivalent filter on a graph shift operator $W'$ constructed by slightly changing the eigenvalues of $W$.

Let us now use the diagonalisation of $W$ and rewrite the semi-supervised learning operator $C_{\sigma,\alpha}$ defined in Eq. (3) as:

$$\begin{aligned}
C_{\sigma,\alpha} &= U\big[(1+\alpha)I_N - \alpha\Lambda\big]^{-1}U^T \\
&= \sum_{i=1}^{N} u_i \frac{1}{(1+\alpha) - \alpha\lambda_i} u_i^T \\
&= \sum_{i=1}^{N} u_i\, g_\alpha(\lambda_i)\, u_i^T.
\end{aligned} \tag{5}$$

Since the set of entries $\{\lambda_i : i \in \{1,\ldots,N\}\}$ of the diagonal matrix $\Lambda$ is finite, we can find an interpolating polynomial $h_\alpha(\lambda)$ verifying $\forall i \in \{1,\ldots,N\}, h_\alpha(\lambda_i) = g_\alpha(\lambda_i)$. Then, Eq. (5) reads:

$$C_{\sigma,\alpha} = U h_\alpha(\Lambda) U^T = h_\alpha(W), \tag{6}$$

thus proving that the SSL operator $C_{\sigma,\alpha}$ is a linear shift-invariant graph filter on $Y$.

In standard Fourier analysis on $L^2(\mathbb{R})$, the eigenvalues of the continuous Laplacian operator are the square of its eigenfrequencies. By identification and recalling that we considered graphs such that $\mathcal{L} = I_N - W$, we pose $\omega_i = \sqrt{1 - \lambda_i}, \forall i \in \{1,\ldots,N\}$, the eigenfrequencies of the graph. With this change of variables, the (eigen-)frequency response of the SSL filter becomes:

$$\widetilde{g}_\alpha(\omega_i) = g_\alpha(1 - \omega_i^2) = \frac{1}{1 + \alpha\omega_i^2}. \tag{7}$$

We recognise in the above expression (7), the frequency response of an optimal Wiener filter found in usual signal processing theory to solve the following problem: let $\mathbf{y} = \mathbf{x} + \eta$, be the observation of a signal $\mathbf{x}$ corrupted by an additive white Gaussian noise $\eta$ independent of $\mathbf{x}$. Estimating $\mathbf{x}$ faces an ill-posed inverse problem that we tackle imposing a Tikhonov regularisation on the least square regression. Choosing, as the penalty, to minimise the $L^2$-norm of the gradient, or equivalently the Laplacian, yields a similar form as in (7), where $\alpha$ is the inverse of the signal to noise ratio [20].

### 3.2. Graph signal Wiener filter and duality

Let us reinterpret the smooth mapping we obtained from semi-supervised learning applied to the oriented cyclic graph of Fig. 1a, in the light of the Wiener filtering of a graph signal. Using Eq. (4), we transformed the initial classification problem into the estimation of a one-dimensional function $X$ defined on the nodes of a graph. Following the same lines, we can consider the label matrix $Y$ of Eq. (3) to be the observation measured on the nodes of a graph and being the input of the Wiener-like SSL filter. For instance, we can assume that the time series plotted in Fig. 1c, is a rough and possibly noisy version of an underlying smooth signal that we want to recover using SSL filtering. Fig. 3 shows the results obtained for different values of the regularisation parameter $\alpha$. Plots in the left column represent the input $Y(i)$ and the output $X(i)$ of the SSL filter. The right column displays the corresponding Graph-Fourier transforms $\widehat{Y}(\lambda_i)$ and $\widehat{X}(\lambda_i)$ superimposed to the frequency response of the low-pass filters of expression (7).

A small value of the regularisation parameter ($\alpha = 0.1$ in the top plots) is not very stringent on the smoothness penalty in relation (2). By contrast, it gives a high confidence to the observation, which is reflected by a SSL filter with a all-pass frequency response that leaves the input almost unchanged. Conversely, an overstressed regularisation term ($\alpha = 1000$ in the bottom plots) overwrites the entire relevance of the input, with an equivalent SSL filter that zeroes all the
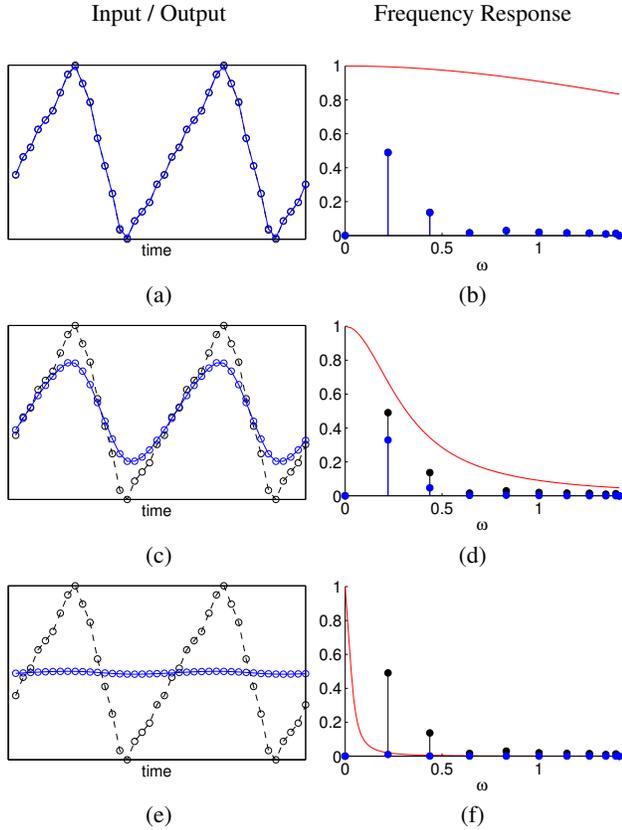
Input / Output | Frequency Response



(a)            (b)

(c)            (d)

(e)            (f)

**Fig. 3**. Semi-supervised learning as a graph signal filter. Left column: input (black dashed) and output (blue plain) of the SSL filter. Right column: frequency response (red curve) of the SSL filter (Eq. (7)) superimposed to the graph Fourier transforms of the input (black) and of the output (blue). Top row: $\alpha = 0.1$. Middle row: $\alpha = 10$. Bottom row: $\alpha = 1000$.

frequency modes in the spectral domain. Finally, $\alpha = 10$ (middle plots) yields a low-pass SSL filter with a cut-off frequency that preserves the first harmonic only. The corresponding output achieves a a good trade-off between smoothness and coherence with the input.

## 4. CONCLUSIONS

The main contribution of the current paper is twofold: on one hand, it introduces a graph-based semi-supervised learning technique to derive more adaptive and more efficient mappings between graphs and time series. Compared to the state-of-the-art, our method yields smoother time series that are easier to analyse in general. Moreover, the formulation we proposed leads to a unique solution, whose closed-form expression is algebraically straightforward and scalable to large datasets. On the other hand, our work bridges an original connection between graph-based semi-supervised learning and the theory of linear graph-shift invariant filters, with special emphasis on optimal Wiener filters.

Since SSL techniques form a wide class of statistical classifiers, they open further research directions. We first plan to benefit from key characteristics of SSL to investigate adaptive mappings that match the changes of dynamic graphs. We also conduct an ongoing work to investigate the use of $N$ classes instead of only 2, bearing in

mind to take advantage of the classifier to derive efficient algorithms for mining multiscale communities.

By revisiting graph-based semi-supervised learning in terms of linear graph-shift invariant signal filtering, we embedded the duality problem into a more general framework of signal recovery from noisy observations. Under this angle, we believe that we can leverage the assets of SSL to better deal with uncertainty. Here, uncertainty may relate to the signal observed on the nodes of the graph (*e.g.* the *a priori* knowledge of a training set), or to the graph structure itself which may be known only partially or unreliably.

Finally, in the present contribution, we focused on graphs corresponding to doubly stochastic weight matrices. An important question that is left open is to better understand the role of the canonical parameter $\sigma$ in the relation (2) and its impact on the (general) eigenvectors of the weight matrix $W$. We are currently considering general graphs beyond the case of Markov matrices.

## 5. REFERENCES

[1] D. Easley and J. Kleinberg, *Networks, Crowds, and Markets: Reasoning About a Highly Connected World*, Cambridge University Press, 2010.

[2] P. Boldi and S. Vigna, "The webgraph framework I: compression techniques," in *WWW '04: Proceedings of the 13th international conference on World Wide Web*, New York, NY, USA, 2004, pp. 595–602, ACM Press.

[3] P.F. Jonsson, T. Cavanna, D. Zicha, and P.A. Bates, "Cluster analysis of networks generated through homology: automatic identification of important protein communities involved in cancer metastasis," *BMC bioinformatics*, vol. 7, no. 1, pp. 2, 2006.

[4] A.S.L.O. Campanharo, M.I. Sirer, R.D. Malmgren, F.M. Ramos, and L.A.N. Amaral, "Duality between Time Series and Networks," *PLoS one*, vol. 6, no. 8, pp. e23378, 2011.

[5] A. Sandryhaila and J.M.F. Moura, "Discrete Signal Processing on Graphs: Frequency Analysis," Preprint, July 2013.

[6] P. Hagmann, L. Cammoun, X. Gigandet, R. Meuli, C.J. Honey, V.J. Wedeen, and O. Sporns, "Mapping the Structural Core of Human Cerebral Cortex," *PLoS Biology*, vol. 6, no. 7, pp. e159, July 2008.

[7] B. Girault, P. Gonçalves, and E. Fleury, "Graphe de contacts et ondelettes : étude d'une diffusion bactérienne," in *Proceedings of Gretsi 2013*, Sept. 2013.

[8] D.I. Shuman, S.K. Narang, P. Frossard, A. Ortega, and P. Vandergheynst, "The Emerging Field of Signal Processing on Graphs: Extending High-Dimensional Data Analysis to Networks and Other Irregular Domains.," *IEEE Signal Process. Mag.*, vol. 30, no. 3, pp. 83–98, 2013.

[9] R.R. Coifman and M. Maggioni, "Diffusion wavelets," *Applied and Computational Harmonic Analysis*, vol. 21, no. 1, pp. 53–94, 2006.

[10] D.K. Hammond, P. Vandergheynst, and R. Gribonval, "Wavelets on graphs via spectral graph theory," *Applied and Computational Harmonic Analysis*, vol. 30, no. 2, pp. 129–150, 2011.

[11] A. Sandryhaila and J.M.F. Moura, "Discrete Signal Processing on Graphs," *Signal Processing, IEEE Transactions on*, vol. 61, no. 7, pp. 1644–1656, 2013.

[12] X. Zhu and A.B. Goldberg, *Introduction to Semi-Supervised Learning*, vol. 3 of *Synthesis Lectures on Artificial Intelligence and Machine Learning*, Morgan & Claypool, 2009.

[13] R.V. Donner, M. Small, J.F. Donges, N. Marwan, Y. Zou, R. Xiang, and J. Kurths, "Recurrence-Based Time Series Analysis by Means of Complex Network Methods.," *I. J. Bifurcation and Chaos*, vol. 21, no. 4, pp. 1019–1046, 2011.

[14] Y. Shimada, T. Ikeguchi, and T. Shigehara, "From Networks to Time Series," *Phys. Rev. Lett.*, vol. 109, pp. 158701, Oct 2012.

[15] R. Hamon, P. Borgnat, P. Flandrin, and C. Robardet, "Transformation de graphes dynamiques en signaux non stationnaires," in *Proceedings of Gretsi 2013*, Sept. 2013.

[16] M. Sokol, K. Avrachenkov, P. Gonçalves, and A. Mishenin, "Generalized Optimization Framework for Graph-based Semi-supervised Learning.," in *SDM*. 2012, pp. 966–974, SIAM / Omnipress.

[17] J. Varah, "A Practical Examination of Some Numerical Methods for Linear Discrete Ill-Posed Problems," *SIAM Review*, vol. 21, no. 1, pp. 100–111, 1979.

[18] K. Avrachenkov, P. Gonçalves, and M. Sokol, "On the Choice of Kernel and Labelled Data in Semi-supervised Learning Methods," in *10th Workshop on Algorithms and Models for the Web Graph*, Harvard University, Cambridge, MA, USA, December 14-15 2013.

[19] K. Avrachenkov, P. Gonçalves, A. Legout, and M. Sokol, "Classification of content and users in BitTorrent by semi-supervised learning methods.," in *IWCMC*. 2012, pp. 625–630, IEEE.

[20] M. Feigin and N. Sochen, "Anisotropic Regularization for Inverse Problems with Application to the Wiener Filter with Gaussian and Impulse Noise," in *Scale Space and Variational Methods in Computer Vision*, vol. 5567 of *Lecture Notes in Computer Science*, pp. 319–330. Springer Berlin Heidelberg, 2009.