Online Temperature Estimation using Graph Signals

Marcelo J. M. Spelta and Wallace A. Martins

***Abstract*— This article investigates the application of adaptive graph signal processing on real-world data. Using temperature measurements obtained from Brazilian weather stations, we construct a graph signal and verify that it can be approximated by a sparse frequency representation. Considering the properties of bandlimited graph signals, we analyze the conditions for perfect reconstruction and describe estimation methods based on adaptive strategies, such as the LMS and RLS algorithms. Numerical analyses suggest that these adaptive estimation algorithms provide a smaller mean-square deviation when compared to the optimal instantaneous linear estimator in noisy scenarios, for both constant and slowly time-varying graph signals.**

***Keywords*— Graph Signal Processing, Graph Signal Estimation, Adaptive Filtering, Sensor Network.**

I. INTRODUCTION

The desire to extend traditional signal processing techniques to handle data naturally defined on irregular discrete domains has led to the development of an emerging research field known as *graph signal processing* (GSP) [1]. Relying on graph structures with distributed vertices and weighted edges that indicate the connection and similarity between neighboring nodes, a graph signal consists in a vector formed by values corresponding to the graph vertices in a specific time instant. Recent papers point out the potential of the GSP framework by proposing its use in various areas such as sensor networks [2], classification tasks [2], and image processing [1].

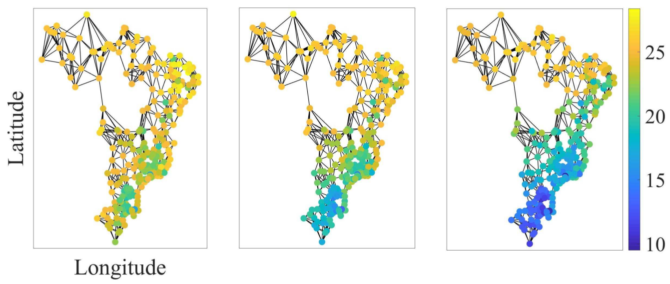
This work focuses on the recovery of a bandlimited, or approximately bandlimited, graph signal from a reduced number of sampled values. This problem is usually solved by a simple least-squares interpolation procedure [3] that allows perfect recovery of the original signal, if some requirements are met [4]. Alternatively, an adaptive approach for estimating graph signals has already been proposed [5], [6], [7]. This adaptive estimation is based on the well-known least-mean-squares (LMS) and recursive least-squares (RLS) algorithms [8] and is expected to benefit from its characteristics of enabling online reconstruction and tracking of time-varying signals. Besides that, another reason for using adaptive strategies is that we consider the reconstruction problem in noisy environments.

Inspired by the GSP signal compression application in [2], one realizes that the signal extracted from a spatial distribution of weather stations measuring local temperatures is usually a smooth signal, allowing its approximation by a bandlimited graph signal. The advantage of representing a temperature graph signal by a reduced number of its samples is notorious since it allows estimating temperature values in certain regions without having a sensor at that place, reducing the amount of sensor devices and information collected. However, one must know how to solve the inverse problem of evaluating the complete graph signal from its sampled representation, assuming that the graph signal is bandlimited.

Marcelo J. M. Spelta and Wallace A. Martins, Electrical Engineering Programme (PEE), Federal University of Rio de Janeiro (UFRJ), Rio de Janeiro-RJ, Brazil, E-mails: marcelo.spelta@smt.ufrj.br, wallace.martins@smt.ufrj.br.
This work was supported by CAPES, CNPq, and FAPERJ.

Thus, in this work we analyze the performance of the adaptive algorithms for graph signal estimation proposed in [5], [6] and compare their behavior with the instantaneous version of a standard strategy [3], [4], based on an instantaneous least-squares interpolation. In order to do so, we construct a graph signal using the monthly average temperature measurements obtained from Brazilian weather stations [9], as illustrated in Fig. 1, and verify that this signal is approximately bandlimited. Then, considering a sampled representation of the bandlimited graph signal, we assess the estimation algorithms in both static and time-varying scenarios, and verify that the adaptive strategies outperform the traditional interpolation technique in terms of mean-square deviation (MSD) for slowly time-varying graph signals.

This paper is organized as follows: Section II introduces some basic GSP concepts and Section III describes the adaptive strategies for graph signal reconstruction. In Section IV we obtain an approximately bandlimited graph signal from Brazilian temperature data and compare the performance of the standard and adaptive algorithms for graph signal estimation. At last, some conclusions are drawn in Section V.



(a) January (b) April (c) July

Fig. 1: Graph signals representing the monthly average temperatures from Brazilian weather stations during the period of 1961-1990 [9].

II. GRAPH SIGNAL PROCESSING

Intended for modeling relations among objects, a graph is a generic mathematical structure made up of vertices v_i connected by edges $\widehat{v_i v_j}$ with respective weights a_{ij}. Basically, a graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ consists of N nodes or vertices $\mathcal{V} = \{v_1, v_2, ..., v_N\}$ linked by a set of edges $\mathcal{E} = \{\widehat{v_i v_j}\}$. Every edge $\widehat{v_i v_j} \in \mathcal{E}$ is associated with a weight a_{ij} indicating a similarity or proximity measure between nodes v_i and $v_j \in \mathcal{V}$. Considering a zero value for a_{ij} when $\widehat{v_i v_j} \notin \mathcal{E}$, we can define the adjacency matrix $\mathbf{A} \in \mathbb{R}^{N \times N}$ as the matrix formed by elements a_{ij}. This adjacency matrix $\mathbf{A}$ is commonly used for graph structure representation and plays an important role in the emerging field of GSP [1], [10].

A. Graph Signal and Fourier Transform

A signal $x : \mathcal{V} \to \mathbb{R}$ defined on the vertices of a graph $\mathcal{G}$ with N nodes can be described as a vector $\mathbf{x} \in \mathbb{R}^N$, where its n-th component x_n represents the function value at vertex

$v_n \in \mathcal{V}$ [1]. An example of a graph signal representation can be seen in Fig. 1, where the color of each vertex v_n indicates its value x_n according to the colorbar scale.

A useful toolset for traditional signal processing is the Fourier transform, which expands an original time-domain signal into a Fourier basis of signals that are linear-filtering invariant. Likewise, the *graph Fourier transform* (GFT) of a graph signal $\mathbf{x} \in \mathbb{R}^N$ can be defined as its projection onto a set of orthonormal vectors $\{\mathbf{u}_n\} \subset \mathbb{R}^N$, where $n \in \mathcal{N} \triangleq \{1, 2, \dots, N\}$. Thus, the GFT of $\mathbf{x}$ is represented as [1], [2]

$$\mathbf{s} = \mathbf{U}^T \mathbf{x}, \tag{1}$$

where $\mathbf{U} \in \mathbb{R}^{N \times N}$ is the eigenvector matrix corresponding to the spectral decomposition of $\mathbf{A} = \mathbf{U}\mathbf{\Lambda}\mathbf{U}^T$.

Although $\mathbf{U}$ in (1) diagonalizes the graph adjacency matrix $\mathbf{A}$ as in [2], [11], alternative approaches use the spectral decomposition of the Laplacian matrix $\mathbf{L} = \mathbf{K} - \mathbf{A}$, where $\mathbf{K} \in \mathbb{R}^{N \times N}$ is a diagonal matrix with diagonal entries $k_i := \sum_{j=1}^{N} a_{ij}$, instead [1]. Moreover, for recovering the original signal $\mathbf{x}$ from its frequency decomposition, we can also define the *inverse graph Fourier transform* (IGFT) of $\mathbf{s}$ as

$$\mathbf{x} = \mathbf{U}\mathbf{s}. \tag{2}$$

B. Sampling and Reconstruction of Graph Signals

Similarly to the definition used for time signals, we can extend the idea of bandlimited signals to graph structures and say that a graph signal $\mathbf{x}_\text{o} \in \mathbb{R}^N$ is *bandlimited* or *spectrally sparse (ssparse)* when its frequency representation $\mathbf{s}$ given by the GFT in (1) is sparse. Taking $\mathcal{F}$ as an index subset of $\mathcal{N}$ ($\mathcal{F} \subseteq \mathcal{N}$), a graph signal $\mathbf{x}_\text{o}$ is defined as $\mathcal{F}$-ssparse if $\mathbf{s}$ is such that $\mathbf{s}_{\mathcal{N} \setminus \mathcal{F}}$ is a zero vector [12], i.e., the components of $\mathbf{s}$ with indices in $\mathcal{N} \setminus \mathcal{F}$ are equal to zero, where $\mathcal{N} \setminus \mathcal{F}$ denotes the difference between sets $\mathcal{N}$ and $\mathcal{F}$. In other words, this *support* or *frequency set* of $\mathcal{F}$ can be described as $\mathcal{F} = \{f \in \mathcal{N} \mid s_f \neq 0\}$ [5]. Thus, if we consider that $\mathbf{U}_\mathcal{F} \in \mathbb{R}^{N \times |\mathcal{F}|}$ and $\mathbf{s}_\mathcal{F} \in \mathbb{R}^{|\mathcal{F}|}$ are, respectively, the matrix formed by eigenvectors and the frequency representation indexed by the elements in $\mathcal{F}$, from (2) we can write that

$$\mathbf{x}_\text{o} = \mathbf{U}_\mathcal{F} \mathbf{s}_\mathcal{F}. \tag{3}$$

Bandlimited time signals can be sampled and reconstructed with no loss of information as long as the Nyquist criterion is satisfied; we describe now a similar result for graph signals. First, consider that the sampling and reconstruction operations over a graph signal $\mathbf{x}_\text{o}$ can be seen as the result of pre-multiplying the original signal by a sampling matrix $\mathbf{D} \in \mathbb{R}^{N \times N}$ and an interpolation matrix $\mathbf{\Phi} \in \mathbb{R}^{N \times N}$. For perfect recovery of the original graph signal, we intend to design matrices $\mathbf{D}$ and $\mathbf{\Phi}$ such that $\mathbf{x}_\text{o} = \mathbf{\Phi}\mathbf{D}\mathbf{x}_\text{o}$ for any bandlimited $\mathbf{x}_\text{o}$ described by (3) with a specific $\mathbf{U}_\mathcal{F}$, i.e., for ssparse graph signals represented by the same frequency components.

Sampling is the operation of observing the value of a graph signal on a *sampling set* $\mathcal{S} \subseteq \mathcal{V}$. In this context, we can define $\mathbf{D}_\mathcal{S} \in \mathbb{R}^{N \times N}$ as a diagonal matrix with entries d_i, in which $d_i = 1$ if $v_i \in \mathcal{S}$ and $d_i = 0$ otherwise. Thus, we obtain the sampled vector $\mathbf{x}_\mathcal{S} \in \mathbb{R}^N$ by sampling $|\mathcal{S}|$ components of a generic graph signal $\mathbf{x}$ as

$$\mathbf{x}_\mathcal{S} = \mathbf{D}_\mathcal{S} \mathbf{x}. \tag{4}$$

In order to recover a original bandlimited signal $\mathbf{x}_\text{o}$ from its sampled version $\mathbf{x}_\mathcal{S}$, we remember the IGFT expression of an $\mathcal{F}$-ssparse signal in (3) and verify that, when $\left(\mathbf{U}_\mathcal{F}^T \mathbf{D}_\mathcal{S} \mathbf{U}_\mathcal{F}\right)$ has full rank, the interpolation matrix $\mathbf{\Phi}$ can be chosen as [4]

$$\mathbf{\Phi} = \mathbf{U}_\mathcal{F} \left(\mathbf{U}_\mathcal{F}^T \mathbf{D}_\mathcal{S} \mathbf{U}_\mathcal{F}\right)^{-1} \mathbf{U}_\mathcal{F}^T. \tag{5}$$

Thus, by considering expressions (4) and (5), it is clear that the sampling and interpolation procedure described by $\mathbf{\Phi}\mathbf{D}_\mathcal{S}\mathbf{x}_\text{o}$ results in the original graph signal $\mathbf{x}_\text{o} = \mathbf{U}_\mathcal{F}\mathbf{s}_\mathcal{F}$ [3], [4].

Therefore, we conclude that perfect reconstruction of an $\mathcal{F}$-ssparse graph signal from its sampled version $\mathbf{x}_\mathcal{S}$ is possible as long as the chosen sampling set $\mathcal{S}$ guarantees that [4], [13]

$$\text{rank}(\mathbf{D}_\mathcal{S} \mathbf{U}_\mathcal{F}) = |\mathcal{F}|. \tag{6}$$

As $\text{rank}(\mathbf{D}_\mathcal{S}) = |\mathcal{S}|$, from (6) we conclude that a necessary condition for perfect recovery of a sampled graph signal is that $|\mathcal{S}| \geq |\mathcal{F}|$, i.e., the number of samples retained must be at least the amount of non-zero frequency components of $\mathbf{s}$.

Moreover, as Section III describes adaptive reconstruction algorithms which allow the graph signal to vary in time and might be corrupted by noise, it is of interest to use the standard interpolation ideas presented so far for comparison purposes. In fact, employing the interpolation strategy for handling time varying bandlimited graph signals is straightforward, as presented in Algorithm 1. This algorithm basically computes an instantaneous least-squares linear estimate $\hat{\mathbf{x}}_\text{o}[k]$ for the original bandlimited signal $\mathbf{x}_\text{o}[k]$ based on the current sampled signal $\mathbf{D}_\mathcal{S}\mathbf{x}_w[k]$ at time instant $k \in \mathbb{Z}$, where $\mathbf{x}_w[k]$ is the measured graph signal vector, possibly corrupted by noise. Although Algorithm 1 is a fast method that allows perfect reconstruction for bandlimited graph signals in ideal scenarios, its use in noisy environments might be limited due to its inherent instantaneous nature.

Algorithm 1 Instantaneous linear interpolation

1: $k \leftarrow 0$
2: $\mathbf{\Phi} = \mathbf{U}_\mathcal{F} \left(\mathbf{U}_\mathcal{F}^T \mathbf{D}_\mathcal{S} \mathbf{U}_\mathcal{F}\right)^{-1} \mathbf{U}_\mathcal{F}^T$
3: **while** (true) **do**
4: $\quad \hat{\mathbf{x}}_\text{o}[k] = \mathbf{\Phi}\mathbf{D}_\mathcal{S}\mathbf{x}_w[k]$
5: $\quad k \leftarrow k + 1$
6: **end**

C. Sampling Set Selection

As so far we have considered recovering a bandlimited graph signal $\mathbf{x}_\text{o}[k]$ from sampled values $\mathbf{x}_\mathcal{S}[k]$, any choice of *sampling set* $\mathcal{S}$ respecting condition (6) results in a perfect recovery of the original graph signal when applying the interpolation method in Algorithm 1. However, in a practical situation in which one acquires data from distributed sensors, it is expected that the resulting measurements are corrupted by noise.

Based on this assumption, a more adequate modeling for a practical graph signal is represented by

$$\boldsymbol{x}_w[k] = \mathbf{x}_\text{o}[k] + \boldsymbol{w}[k], \tag{7}$$

where $\boldsymbol{x}_w[k]$ is the noisy random signal[1] available at the vertices of the graph, $\mathbf{x}_\text{o}[k] \in \mathbb{R}^N$ is the original bandlimited graph signal, and $\boldsymbol{w}[k]$ is a zero-mean noise vector with covariance matrix $\mathbf{C}_w[k] \in \mathbb{R}^{N \times N}$.

Considering the model in (7), if one evaluates some common figures of merit (FoM) for an estimated graph signal $\hat{\mathbf{x}}_\text{o}[k]$, such as the *mean-square deviation* (MSD) defined as

$$\text{MSD} = \mathbb{E}\left\{\|\hat{\boldsymbol{x}}_\text{o}[k] - \mathbf{x}_\text{o}[k]\|_2^2\right\}, \tag{8}$$

[1]In this work, $\boldsymbol{x}$ denotes a random vector with realizations denoted as $\mathbf{x}$.

or the *squared deviation* (SD) given by

$$\mathrm{SD} = \|\hat{\mathbf{x}}_{\mathrm{o}}[k] - \mathbf{x}_{\mathrm{o}}[k]\|_2^2 \,, \tag{9}$$

it is simple to verify that an estimate $\hat{\mathbf{x}}_{\mathrm{o}}[k]$ based on the linear interpolation procedure $\mathbf{\Phi}\mathbf{D}_{\mathcal{S}}$ results in

$$\begin{aligned} \mathrm{MSD}(\mathcal{S}) &= \mathbb{E}\left\{\|\mathbf{U}_{\mathcal{F}}\big(\mathbf{U}_{\mathcal{F}}^T\mathbf{D}_{\mathcal{S}}\mathbf{U}_{\mathcal{F}}\big)^{-1}\mathbf{U}_{\mathcal{F}}^T\mathbf{D}_{\mathcal{S}}\boldsymbol{w}[k]\|_2^2\right\} \,, \\ \mathrm{SD}(\mathcal{S}) &= \|(\mathbf{U}_{\mathcal{F}}^T\mathbf{D}_{\mathcal{S}}\mathbf{U}_{\mathcal{F}})^{-1}\mathbf{U}_{\mathcal{F}}^T\mathbf{D}_{\mathcal{S}}\mathbf{w}[k]\|_2^2 \,. \end{aligned} \tag{10}$$

Both expressions in (10) present an explict dependency on $\mathbf{D}_{\mathcal{S}}$, which indicates that the sampling set $\mathcal{S}$ could be properly chosen to reduce a desired FoM. Nonetheless, the resulting optimization problem is combinatorial in nature.

In order to obtain an interesting trade-off between reasonable reconstruction performance and time required for calculating the sampling set $\mathcal{S}$, a common approach is to employ a greedy algorithm for minimizing a specific FoM. A greedy algorithm basically reduces the overall computational complexity by searching for an optimal selection at each stage and expecting to find a near optimal final value. Detailed information regarding the reconstruction performance of greedy strategies in GSP is presented in [12].

Particularly, in this work we use the sampling set method presented in Algorithm 2, which employs at each iteration a greedy search for the index $n \in \mathcal{N}$ to be added to the current set $\mathcal{S}$ in order to maximize the minimum non-negative eigenvalue of $(\mathbf{U}_{\mathcal{F}}^T\mathbf{D}_{\mathcal{S}\cup\{n\}}\mathbf{U}_{\mathcal{F}})$. The number of indexes in the final set $\mathcal{S}$ is denoted as M. This method has been initially suggested in [4] and takes the same form as one of the sampling strategies described in [5]. In fact, following an idea similar to [4], it can be shown that Algorithm 2 uses a greedy scheme for minimizing the SD in (10).

Algorithm 2 Greedy algorithm for selection of $\mathcal{S}$

$\mathcal{S} \leftarrow \emptyset$
while $|\mathcal{S}| < M$ **do**
  $m = \underset{n\in\mathcal{N}}{\mathrm{argmax}}\ \lambda_{\min}^{+}(\mathbf{U}_{\mathcal{F}}^T\mathbf{D}_{\mathcal{S}\cup\{n\}}\mathbf{U}_{\mathcal{F}})$
  $\mathcal{S} \leftarrow \mathcal{S} + \{m\}$
end
return $\mathcal{S}$

Although more general selection approaches allow adaptive graph sampling [5], in which $\mathbf{D}_{\mathcal{S}}[k]$ might change at each instant k, this work considers a static sampling matrix $\mathbf{D}_{\mathcal{S}}$ for the sake of simplicity. This assumption implies there is prior knowledge about the signal representation in the frequency domain, which defines $\mathbf{U}_{\mathcal{F}}$, while the adaptive sampling idea is more suitable for the cases where the support $\mathcal{F}$ is unknown.

III. Adaptive Estimation of Graph Signals

A first attempt to merge the traditional area of adaptive filtering [8] with the brand-new field of GSP is done in [5], where the authors suggest LMS-based strategies for handling the problem of graph signal reconstruction. Additionally, an RLS-based algorithm is proposed in [6] for an identical estimation task.

The reason for using an adaptive strategy in a graph estimation context comes from the possibility of robust online estimation and tracking of time-varying graph signals. The robustness of the adaptive approach is handy to work in noisy scenarios such as (7), when we expect an adaptive strategy to provide a smaller MSD in comparison to an instantaneous signal interpolation, as in Algorithm 1.

A. LMS Estimation

The most widely used algorithm in adaptive filtering is the LMS [14], [15]. Among many reasons, such as its stable behavior for finite-precision implementation, the popularity of the LMS algorithm is mainly due to its simplicity and low computational complexity [8].

Particularly, for graph signal reconstruction [5], [16] the LMS algorithm is designed to minimize the mean-square sampled deviation, i.e.,

$$\min_{\mathbf{s}_{\mathcal{F}}[k]} \mathbb{E}\left\{\, \|\mathbf{D}_{\mathcal{S}}(\boldsymbol{x}_w[k] - \mathbf{U}_{\mathcal{F}}\mathbf{s}_{\mathcal{F}}[k])\|_2^2 \,\right\} . \tag{11}$$

By using a stochastic gradient approach for the minimization problem in (11), one finds an update expression for $\mathbf{s}_{\mathcal{F}}[k+1]$. Moreover, using the IGFT in (2), one can easily find an estimate $\hat{\mathbf{x}}_{\mathrm{o}}[k]$ for the bandlimited graph signal $\mathbf{x}_{\mathrm{o}}[k]$ in (3), which corresponds to the LMS update equation [16]

$$\hat{\mathbf{x}}_{\mathrm{o}}[k+1] = \hat{\mathbf{x}}_{\mathrm{o}}[k] + \mu\mathbf{U}_{\mathcal{F}}\mathbf{U}_{\mathcal{F}}^T\mathbf{D}_{\mathcal{S}}(\mathbf{x}_w[k] - \hat{\mathbf{x}}_{\mathrm{o}}[k]) \,, \tag{12}$$

where $\mu \in \mathbb{R}_+$ is the so-called convergence factor [8], a parameter that controls the algorithm behavior to either improve its convergence speed or reduce the steady-state error.

Thus, the procedure for finding an online estimation of $\mathbf{x}_{\mathrm{o}}[k]$ based on the LMS update equation (12) is presented in Algorithm 3. The initial value $\hat{\mathbf{x}}_{\mathrm{o}}[0]$ must be a bandlimited signal (such as $\hat{\mathbf{x}}_{\mathrm{o}}[0] = \mathbf{0}$) and the step-size parameter μ must be sufficiently small to promote stability [5].

Algorithm 3 LMS estimation of graph signals

$k \leftarrow 0$
while (true) **do**
  $\hat{\mathbf{x}}_{\mathrm{o}}[k+1] = \hat{\mathbf{x}}_{\mathrm{o}}[k] + \mu\,\mathbf{U}_{\mathcal{F}}\mathbf{U}_{\mathcal{F}}^T\mathbf{D}_{\mathcal{S}}(\mathbf{x}_w[k] - \hat{\mathbf{x}}_{\mathrm{o}}[k])$
  $k \leftarrow k+1$
end

B. RLS Estimation

Another well-known technique in the adaptive filtering area is an iterative procedure that solves the least-squares problem for each incoming signal in a recursive form, called the RLS algorithm [8]. The LMS strategy takes a long time until reaching its steady-state, calling for alternatives, such as the RLS, a much faster approach. However, the trade-off for using a faster convergence algorithm like the RLS is a considerable increase in computational complexity, which might be a limiting factor, depending on the application.

Similarly to the LMS idea, the RLS algorithm for estimating graph signals [7] intends to minimize an error function, which in this case is given by

$$\min_{\mathbf{s}_{\mathcal{F}}} \ \sum_{l=1}^{k} \beta^{k-l}\|\mathbf{D}_{\mathcal{S}}(\mathbf{x}_w[l] - \mathbf{U}_{\mathcal{F}}\mathbf{s}_{\mathcal{F}})\|_{\mathbf{C}_w^{-1}}^2 + \beta^k\|\mathbf{s}_{\mathcal{F}}\|_{\mathbf{\Pi}}^2 \,, \tag{13}$$

where $0 \ll \beta \leq 1$ is the forgetting factor, and $\mathbf{\Pi}$ is a regularization matrix, usually taken as $\mathbf{\Pi} = \delta\mathbf{I}$, for a small $\delta > 0$.

When solving the convex problem in (13), one finds that the estimate for the bandlimited signal $\mathbf{x}_{\mathrm{o}}[k]$ is equal to

$$\hat{\mathbf{x}}_{\mathrm{o}}[k] = \mathbf{U}_{\mathcal{F}}\mathbf{\Psi}^{-1}[k]\boldsymbol{\psi}[k] \,, \tag{14}$$

where $\mathbf{\Psi}[k] \in \mathbb{R}^{|\mathcal{F}|\times|\mathcal{F}|}$ and $\boldsymbol{\psi}[k] \in \mathbb{R}^{|\mathcal{F}|}$ are

$$\begin{aligned} \mathbf{\Psi}[k] &= \beta\mathbf{\Psi}[k-1] + \mathbf{U}_{\mathcal{F}}^T\mathbf{D}_{\mathcal{S}}\mathbf{C}_w^{-1}[k]\,\mathbf{U}_{\mathcal{F}} \,, \\ \boldsymbol{\psi}[k] &= \beta\boldsymbol{\psi}[k-1] + \mathbf{U}_{\mathcal{F}}^T\mathbf{D}_{\mathcal{S}}\mathbf{C}_w^{-1}[k]\,\mathbf{x}_w[k]. \end{aligned} \tag{15}$$

Then, based on the update equations (14) and (15), we present the complete RLS method for graph signal estimation [7] in Algorithm 4. The initial values of this algorithm can be taken as $\boldsymbol{\Psi}[0] = \boldsymbol{\Pi}$ and a random $\boldsymbol{\psi}[0]$, such as $\boldsymbol{\psi}[0] = \mathbf{0}$.

Algorithm 4 RLS estimation of graph signals

$k \leftarrow 0$
while (true) **do**
  $\boldsymbol{\Psi}[k] = \beta\boldsymbol{\Psi}[k-1] + \mathbf{U}_{\mathcal{F}}^T\mathbf{D}_{\mathcal{S}}\mathbf{C}_w^{-1}[k]\,\mathbf{U}_{\mathcal{F}}$
  $\boldsymbol{\psi}[k] = \beta\boldsymbol{\psi}[k-1] + \mathbf{U}_{\mathcal{F}}^T\mathbf{D}_{\mathcal{S}}\mathbf{C}_w^{-1}[k]\,\mathbf{x}_w[k]$
  $\hat{\mathbf{x}}_{\mathrm{o}}[k] = \mathbf{U}_{\mathcal{F}}\boldsymbol{\Psi}^{-1}[k]\boldsymbol{\psi}[k]$
  $k \leftarrow k+1$
end

IV. Numerical Simulations

In order to compare the performance of the adaptive strategies presented in Section III with the standard method for reconstructing graph signals in Algorithm 1, we collect two datasets from the Instituto Nacional de Meteorologia (INMET) website [9]: the first one contains the latitude and longitude coordinates of active weather stations, while the second dataset presents a monthly average temperature recorded in some of these stations, during the 1961-1990 period. From these data we obtain a total of 299 nodes for our graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$, in which each of these vertices represents a weather station. Thus, a node v_n of the graph $\mathcal{G}$ and its signal value $x_n[k]$ are given, respectively, by the geographical coordinates and the average temperature of the associatied weather station in a given month. For reproducibility purposes and future experiments, this compiled location and temperature dataset has been made available at [17].

As we have no explicit connection between weather stations, we are free to design the set $\mathcal{E}$ and choose the weights $\{a_{ij}\}$. For simplicity, we construct the graph edges by connecting a vertex v_n to its 8 closest neighbor nodes [1], considering that the distance between two nodes is given by the Haversine formula [18], which evaluates the great-circle distance between points using their latitude and longitude coordinates. Thus, based on this procedure we obtain the graph structures displayed in Fig. 1.

As GSP techniques require the use of either the Laplacian or adjacency matrix, we need to define the edge weights $\{a_{ij}\}$ of $\mathbf{A}$, which can be seen as a similarity measure between neighboring vertices. Like [1], we evaluate the edge weights $\{a_{ij}\}$ based on the Gaussian kernel weighting function

$$a_{ij} = \begin{cases} \exp\left(-\frac{d_{\mathrm{H}}(i,j)^2}{2\theta^2}\right), & \text{if } \widehat{v_i v_j} \in \mathcal{E}, \\ 0 , & \text{otherwise}, \end{cases} \tag{16}$$

where $d_{\mathrm{H}}(i,j)$ is the Haversine distance between vertices v_i and v_j, θ is taken as $2 \cdot 10^3$ and the condition $\widehat{v_i v_j} \in \mathcal{E}$ checks if the edge connecting nodes v_i and v_j is part of the set $\mathcal{E}$.

A common assumption in GSP literature is that smooth signals on graphs present a bandlimited or approximately bandlimited frequency representation, in this case given by their low-frequency components or eigenvectors. As a graph signal $\mathbf{x}_w[k]$ obtained from temperature measurements across the country presents this smooth behavior (despite minor outlier points) at every instant k, we expect $\mathbf{x}_w[k]$ to be approximately bandlimited, which would allow us to test the sampling and reconstruction strategies mentioned in Sections II and III. However, before doing so we need to decide how many frequency components are necessary for representing the graph signal with an acceptable deviation error.

As this task consists in a signal compression problem, we take a similar procedure to [2] and evaluate the average reconstruction error (ARE) $\|\mathbf{x}_{\mathrm{o}} - \bar{\mathbf{x}}_{\mathrm{o}}^P\|_2 / \|\mathbf{x}_{\mathrm{o}}\|_2$ for different estimates $\bar{\mathbf{x}}_{\mathrm{o}}^P$ using only P frequency components of the original bandlimited graph signal $\mathbf{x}_{\mathrm{o}}$, taken as the signal from the July dataset depicted in Fig. 1(c). The signal $\bar{\mathbf{x}}_{\mathrm{o}}^P$ is obtained by sorting the absolute values of signal $\mathbf{s}$ obtained from (1) and selecting the indices p of the P-largest components $|s_n|$ to form the auxiliar set $\mathcal{F}_P$. Based on these indices $p \in \mathcal{F}_{\mathrm{P}} \subseteq \mathcal{N}$, we pick the p-th eigenvector of $\mathbf{U}$ and the p-th frequency component of $\mathbf{s}$ to define $\mathbf{U}_{\mathcal{F}_P}$ and $\mathbf{s}_{\mathcal{F}_P}$. Then, the estimate $\bar{\mathbf{x}}_{\mathrm{o}}^P$ using P components is given by $\mathbf{U}_{\mathcal{F}_P}\mathbf{s}_{\mathcal{F}_P}$.

Following this compression procedure, we compute the percentual ARE for different values of used components P and display the results for the range $[50, 250]$ in Fig. 2. Considering that a deviation error of 2.5% is acceptable in the current application, we approximate the original graph signal by its $P = 200$ largest frequency components. From this assumption we can define the bandlimited set $\mathcal{F} = \mathcal{F}_{\mathcal{P}}$, where $|\mathcal{F}| = P$.

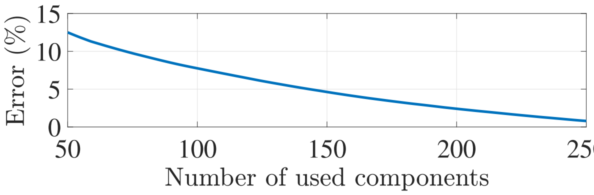


Fig. 2: Percentage of reconstruction error when the original signal is compressed using P frequency components.

Based on this approximately $\mathcal{F}$-ssparse signal $\mathbf{x}_{\mathrm{o}}$, we need to take a practical project decision and select both the amount $|\mathcal{S}|$ and which vertices $v_n \in \mathcal{V}$ of the graph signal should be sampled. As stated in [12], increasing the number of samples in $\mathcal{S}$ always decreases the MSD in (8). However, as we also want to reduce the amount of nodes to be measured, we consider that $|\mathcal{S}| = 210$ provides a reasonable trade-off and then find the *sampling set* $\mathcal{S}$ by using Algorithm 2, with $M = |\mathcal{S}| = 210$. Then, at this point we obtain the sampled bandlimited graph signal which will be used to test the linear interpolation and adaptive reconstruction strategies.

In the following simulations we consider the MSD in (8) as figure of merit for comparing algorithms and take a reference signal $\bar{\mathbf{x}}_{\mathrm{o}}$, defined as the 210-frequency component representation of the approximately $\mathcal{F}$-ssparse graph signal in Fig. 1(c). The simulation results were generated with each estimation algorithm running during 5000 iterations, and this procedure was repeated 10 times, forming the ensemble. From this ensemble, the MSD is estimated as the average for each iteration k. Parameters were set as: initial estimation $\hat{\mathbf{x}}_{\mathrm{o}}[0] = \mathbf{0} \in \mathbb{R}^{299}$, noise covariance matrix $\mathbf{C}_w = \sigma_w^2\mathbf{I}$, with $\sigma_w^2 = 0.01$, convergence factor $\mu = 0.25$, and forgetting factor $\beta = 0.95$.

Fig. 3 depicts the results for the first reconstruction simulation, where we represent the graph readings $\mathbf{x}_w[k]$ by the noisy scenario in (7) with a constant graph signal $\mathbf{x}_{\mathrm{o}}[k] = \bar{\mathbf{x}}_{\mathrm{o}}$ until time instant $k = 2500$, when the reference signal is scaled by a factor of 1.1. The MSD values in Fig. 3 are considerably large due to the high dimension of the reconstructed vectors and the high noise level. The initial iterations suffer even more, considering the initialization step with $\hat{\mathbf{x}}_{\mathrm{o}}[0] = \mathbf{0}$.

As expected, from Fig. 3 we observe that the standard online interpolation method in Algorithm 1 presents the fastest convergence speed, reaching its steady-state value after the first iteration. However, as this traditional strategy ignores

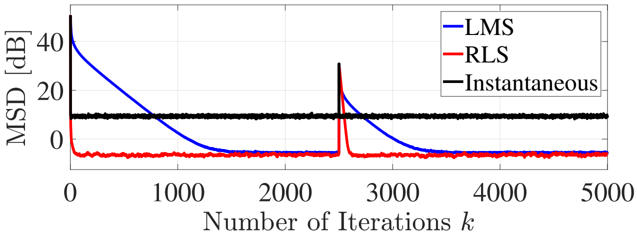


Fig. 3: MSD for different reconstruction algorithms using a constant signal $\mathbf{x}_o[k]$ with a step-transition at $k = 2500$.

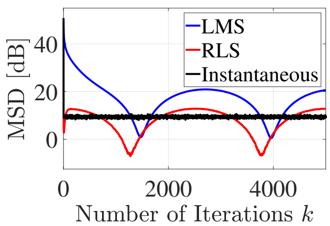
previous values, after a short time it is also noticeable that by using adaptive strategies such as the LMS and RLS algorithms one can achieve a more accurate estimation of the original graph signal $\mathbf{x}_o[k]$, because they provide a lower level of MSD. Additionally, it is clear that the RLS algorithm presents a much faster convergence than the LMS method.

For assessing the performance with time-varying signals, we remember that the temperature follows a regular daily behavior, reaching a peak around mid-day and decreasing to its lowest values in the night period. Then, we assume that a simplified temperature measurement changes in time according to the periodic formula

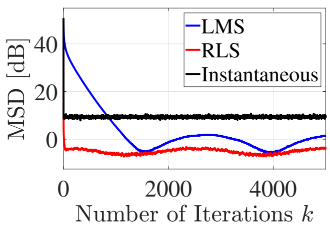
$$\mathbf{x}_o[k] = \left[1 + A \sin\left(\frac{2\pi k}{T}\right)\right] \cdot \bar{\mathbf{x}}_o \,, \tag{17}$$

where A is a scalar used to indicate the range of temperatures and T is a constant for defining the daily periodic behavior.

Fig. 4 illustrates the tracking behavior when we consider a reference signal $\mathbf{x}_o[k]$ in (17) with $T = 5000$ and two different temperature ranges $A = 0.5$ and $A = 0.05$, and run the simulation procedure described previously. From these results we verify that once more the traditional interpolation scheme overcomes the adaptive algoritms in terms of tracking speed in the initial iterations. Furthermore, we observe that the reduction in the estimation deviation (MSD) brought by using adaptive strategies depends on the daily temperature range, where the implementation of adaptive algorithms seems more suitable to track signals in slowly time-varying environments.



(a) $A = 0.5$



(b) $A = 0.05$

Fig. 4: MSD for different reconstruction algorithms when $\mathbf{x}_o[k]$ varies as (17) with (a) $A = 0.5$ and (b) $A = 0.05$.

In fact, this idea of a slowly time-varying graph signal depends not only on the amplitude range A, but also on the parameter T in (17), which is related to the time interval t_k taken between consecutive signal measurements at instants k and $k+1$. If the time interval t_k is small enough, we can consider that the temperature graph signal changes slowly and then the use of the LMS and RLS adaptive algorithms for estimating the original signal from its samples is justified since they outperform the traditional interpolation in terms of MSD in these scenarios, as displayed in Fig. 4(b). Particularly, when complexity is not an issue the RLS algorithm can be seen as a more adequate choice since it presents a good trade-off between MSD and speed of convergence.

V. Conclusions

The recently developed GSP framework produces interesting insights about signals defined on irregular structures, allowing one to perform some processing on them. Based on this, we have verified that a graph signal can be estimated from a reduced set of vertex values in applications where the signal spatial variations is smooth, as is usual in a temperature measurement network. For a practical application this reduction implies budget and maintenance savings, since one can employ a smaller number of sensor devices in the network, but it also requires an adequate method for recovering the original graph signal.

In this paper we compared a standard online interpolation strategy to adaptive methods for signal reconstruction and, based on the obtained results, we concluded that the adaptive algorithms achieved a smaller estimation error for slowly time-varying graph signals.

References

[1] 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 Processing Magazine*, vol. 30, pp. 83–98, May 2013.

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

[3] S. K. Narang, A. Gadde, and A. Ortega, "Signal processing techniques for interpolation in graph structured data," in *2013 IEEE International Conference on Acoustics, Speech and Signal Processing*, pp. 5445–5449, May 2013.

[4] S. Chen, R. Varma, A. Sandryhaila, and J. Kovačević, "Discrete signal processing on graphs: Sampling theory," *IEEE Transactions on Signal Processing*, vol. 63, pp. 6510–6523, Dec 2015.

[5] P. D. Lorenzo, S. Barbarossa, P. Banelli, and S. Sardellitti, "Adaptive least mean squares estimation of graph signals," *IEEE Transactions on Signal and Information Processing over Networks*, vol. 2, no. 4, pp. 555–568, 2016.

[6] P. D. Lorenzo, E. Isufi, P. Banelli, S. Barbarossa, and G. Leus, "Distributed recursive least squares strategies for adaptive reconstruction of graph signals," in *2017 25th European Signal Processing Conference (EUSIPCO)*, pp. 2289–2293, Aug 2017.

[7] P. D. Lorenzo, P. Banelli, E. Isufi, S. Barbarossa, and G. Leus, "Adaptive graph signal processing: Algorithms and optimal sampling strategies," *CoRR*, vol. abs/1709.03726, 2017.

[8] P. S. R. Diniz, *Adaptive Filtering: Algorithms and Practical Implementation*. Springer, 2013.

[9] Instituto Nacional de Meteorologia (INMET) , "Normais Climatológicas do Brasil." `http://www.inmet.gov.br/portal/index.php?r=clima/normaisClimatologicas`.

[10] A. Sandryhaila and J. M. F. Moura, "Discrete signal processing on graphs," *IEEE Transactions on Signal Processing*, vol. 61, pp. 1644–1656, April 2013.

[11] A. Sandryhaila and J. M. F. Moura, "Discrete signal processing on graphs: Frequency analysis," *IEEE Transactions on Signal Processing*, vol. 62, no. 12, pp. 3042–3054, 2014.

[12] L. F. O. Chamon and A. Ribeiro, "Greedy sampling of graph signals," *IEEE Transactions on Signal Processing*, vol. 66, pp. 34–47, Jan 2018.

[13] S. Chen, A. Sandryhaila, and J. Kovačević, "Sampling theory for graph signals," in *2015 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pp. 3392–3396, April 2015.

[14] B. Widrow and M. E. Hoff, "Adaptive switching circuits," *WESCOM Conv. Rec.*, no. pt 4, pp. 96–140, 1960.

[15] B. Widrow, J. M. McCool, M. G. Larimore, and C. R. Johnson, "Stationary and nonstationary learning characteristics of the lms adaptive filter," *Proceedings of the IEEE*, vol. 64, pp. 1151–1162, Aug 1976.

[16] P. D. Lorenzo, P. Banelli, and S. Barbarossa, "Optimal sampling strategies for adaptive learning of graph signals," in *2017 25th European Signal Processing Conference (EUSIPCO)*, pp. 1684–1688, 2017.

[17] M. J. M. Spelta, "Brazilian weather stations." `https://github.com/mspelta/brazilian-weather-stations#brazilian-weather-stations`, 2018.

[18] H. Mahmoud and N. Akkari, "Shortest path calculation: A comparative study for location-based recommender system," in *2016 World Symposium on Computer Applications Research (WSCAR)*, pp. 1–5, March 2016.