DISTRIBUTED WIENER-BASED RECONSTRUCTION OF GRAPH SIGNALS

Elvin Isufi[†,‡]*, Paolo Di Lorenzo*[‡]*, Paolo Banelli*[‡] *and Geert Leus*[†]

[†]Faculty of EEMCS, Delft University of Technology, Delft, The Netherlands
[‡] Department of Engineering, University of Perugia, Perugia, Italy
E-mails: {e.isufi-1;g.j.t.leus}@tudelft.nl, {paolo.banelli; paolodilorenzo}@unipg.it

ABSTRACT

This paper proposes strategies for distributed Wiener-based reconstruction of graph signals from subsampled measurements. Given a stationary signal on a graph, we fit a distributed autoregressive moving average graph filter to a Wiener graph frequency response and propose two reconstruction strategies: i) reconstruction from a single temporal snapshot; ii) recursive signal reconstruction from a stream of noisy measurements. For both strategies, a mean square error analysis is performed to highlight the role played by the filter response and the sampled nodes, and to propose a graph sampling strategy. Our findings are validated with numerical results, which illustrate the potential of the proposed algorithms for distributed reconstruction of graph signals.

***Index Terms*—** Graph signal processing, stationary graph signals, Wiener regularization, ARMA graph filters.

1. INTRODUCTION

Recent advances in graph signal processing (GSP) have shown promising results in reconstructing missing values from signals that reside on top of networks [1–3]. By exploiting the expansion of a graph signal into a graph Fourier basis, [2, 4] propose reconstruction strategies when the signal of interest is sparse in this dual domain. This idea is then extended by [5–7] for adaptive graph signal reconstruction methods from a stream of data. On the other hand [8, 9] exploit regularization priors, such as the Tikhonov prior, to interpolate the missing values from subsampled measurements. A slightly different approach is taken by [10, 11], which interpolate stationary graph signals through a Wiener graph filter. Yet, the sampling set is built uniformly at random.

A common point of almost all the above works is their centralized processing. However, in many practical systems, a centralized implementation may be either infeasible or not efficient. The need for a distributed implementation is also acknowledged in [6,7,12] for graph signal reconstruction with a bandlimitedness prior.

Motivated by the benefits of distributed processing, in this work we propose distributed strategies for stationary graph signal reconstruction. Specifically, we make use of the ARMA graph filter [13] to approximate a Wiener graph frequency response, and then we employ the ARMA recursion to distributively reconstruct the missing values in the non-sampled nodes. Moreover, we extend our algorithm to enable recursive graph signal reconstruction when a stream of noisy data is available. For both approaches, we perform a mean square error (MSE) analysis to highlight the role played by the filter response and the sampling set. Finally, we adopt sparse sensing techniques [14] to sample the graph such that the mean square reconstruction error is minimized.

The work of Di Lorenzo was supported by the "Fondazione Cassa di Risparmio di Perugia".

2. BACKGROUND

This section recalls some background information about GSP, ARMA graph filtering and Wiener graph regularization.

GSP basics. We consider an undirected graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ consisting of a node set $\mathcal{V} = \{v_1, \ldots, v_N\}$ and an edge set $\mathcal{E}$. $\mathbf{W}$ indicates the weighted adjacency matrix, such that $W_{i,j}$ is the edge weight connecting the tuple $(v_i, v_j) \in \mathcal{E}$, or $W_{i,j} = 0$ otherwise. The discrete graph Laplacian is $\mathbf{L} = \text{diag}(\mathbf{1}^{\mathsf{T}}\mathbf{W}) - \mathbf{W}$. We indicate with $\mathbf{S}$ the graph shift operator, which has as plausible candidates $\mathbf{W}$, $\mathbf{L}$, or any of their generalization [15]. Since the graph is undirected, $\mathbf{S}$ can be eigendecomposed as $\mathbf{S} = \mathbf{U}\boldsymbol{\Lambda}\mathbf{U}^{\mathsf{H}}$, with eigenvector matrix $\mathbf{U}$ and eigenvalues $\boldsymbol{\Lambda} = \text{diag}(\lambda_1, \ldots, \lambda_N)$.

A graph signal $\mathbf{x}$ consists of a mapping from the node set to the set of complex numbers, i.e., the ith entry of $\mathbf{x}$ is the signal x_i at node v_i. The graph Fourier transform (GFT) of $\mathbf{x}$ and its inverse are $\hat{\mathbf{x}} = \mathbf{U}^{\mathsf{H}}\mathbf{x}$ and $\mathbf{x} = \mathbf{U}\hat{\mathbf{x}}$, respectively. The eigenvalues in $\boldsymbol{\Lambda}$ form the spectral support of $\hat{\mathbf{x}}$ and are referred to as the graph frequencies [15, 16].

A graph filter consists of a function $h(\lambda_n)$ and acts as a point-wise multiplication in the GFT domain, yielding the output $\hat{\mathbf{y}} = h(\boldsymbol{\Lambda})\hat{\mathbf{x}}$. The diagonal of $h(\boldsymbol{\Lambda})$ is referred to as the filter frequency response. By means of the GFT, the filter output in the vertex domain is $\mathbf{y} = \mathbf{U}h(\boldsymbol{\Lambda})\mathbf{U}^{\mathsf{H}}\mathbf{x} \triangleq \mathbf{H}\mathbf{x}$. Depending on the form of $h(\lambda_n)$ we can distinguish between:

- polynomial graph filters [15, 16], characterized by a frequency response $h(\lambda_n) = \sum_{k=0}^{K} \varphi_k \lambda_n^k$ for some order K and coefficients φ_k. In the vertex domain, these filters yield the output $\mathbf{y} = \sum_{k=0}^{K} \varphi_k \mathbf{S}^k \mathbf{x}$.
- rational graph filters [13], with frequency response

$$h(\lambda_n) = \frac{\sum_{q=0}^{Q} \varphi_q \lambda_n^q}{1 + \sum_{p=1}^{P} \psi_p \lambda_n^p}, \tag{1}$$

978-1-5386-1570-3/18/$31.00 ©2018 IEEE

for some orders P, Q, and coefficients ψ_p, φ_q. These filters can be implemented in the vertex domain with conjugate gradients [17], or as we show next, distributively via ARMA recursions on graphs.

Distributed ARMA graph filters. From [18], the ARMA$_{P,Q}$ recursion

$$\mathbf{y}_t = -\overbrace{\sum_{p=1}^{P}}^{\mathbf{P}} \psi_p \mathbf{S}^p \mathbf{y}_{t-1} + \overbrace{\sum_{q=0}^{Q}}^{\mathbf{Q}} \varphi_q \mathbf{S}^q \mathbf{x}, \tag{2}$$

implements distributively the steady-state output

$$\mathbf{y} \triangleq \lim_{t\to\infty} \mathbf{y}_t = (\mathbf{I}-\mathbf{P})^{-1}\mathbf{Q}\mathbf{x}, \tag{3}$$

for any initialization $\mathbf{y}_0$ given the coefficients ψ_p are designed to satisfy $\|\mathbf{P}\| < 1$. By means of the GFT, it can be observed that (3) consists of a rational filter with frequency response (1). We observe that in computing the output $\mathbf{y}_t$, (2) performs P graph shifts of the previous output $\mathbf{y}_{t-1}$ and Q graph shifts of the input $\mathbf{x}$. Because of the locality of the shift operator, and since $\mathbf{S}^k\mathbf{x}$ can be recursively computed as $\mathbf{S}(\mathbf{S}^{k-1}\mathbf{x})$, the ARMA recursion on graphs (2) enjoys a distributed implementation in the vertex domain (see [13,18] for more details).

Wiener regularization on graphs. Let $\mathbf{x}_\mathrm{d} \sim \mathcal{P}(\bar{\boldsymbol{\mu}}_\mathrm{d}, \boldsymbol{\Sigma}_\mathrm{d})$ be an $N \times 1$ random vector and consider the task of recovering $\mathbf{x}_\mathrm{d}$ from noisy measurements $\mathbf{x} = \mathbf{x}_\mathrm{d} + \mathbf{n}$ with $\mathbf{n}$ being zero-mean noise with covariance matrix $\boldsymbol{\Sigma}_\mathrm{n}$. Under the assumption that $\mathbf{x}_\mathrm{d}$ and $\mathbf{n}$ are mutually independent, the Wiener solution to this problem consists of finding a filter $\mathbf{H}$ that minimizes the MSE, i.e.,

$$\begin{aligned}\mathbf{H}^\star &= \operatorname*{argmin}_{\mathbf{H}} \ \mathbb{E}\|\mathbf{H}(\mathbf{x}_\mathrm{d}-\bar{\boldsymbol{\mu}}_\mathrm{d}+\mathbf{n}) - \mathbf{x}_\mathrm{d} - \bar{\boldsymbol{\mu}}_\mathrm{d}\|_2^2 \\ &= \boldsymbol{\Sigma}_\mathrm{d}(\boldsymbol{\Sigma}_\mathrm{d}+\boldsymbol{\Sigma}_\mathrm{n})^{-1},\end{aligned} \tag{4}$$

and then setting $\mathbf{x}_\mathrm{d}^\star = \mathbf{H}^\star(\mathbf{x} - \bar{\boldsymbol{\mu}}_\mathrm{d})$.

For $\mathbf{x}_\mathrm{d}$ representing a signal on a graph, $\mathbf{x}_\mathrm{d}$ is said to be stationary on $\mathcal{G}$ if: $i)$ $\bar{\boldsymbol{\mu}}_\mathrm{d} = \mathbf{0}$, and $ii)$ $\boldsymbol{\Sigma}_{x_\mathrm{d}} = \mathbf{U}\mathrm{diag}(\boldsymbol{\sigma}_\mathrm{d}^2)\mathbf{U}^\mathsf{H}$ [10, 11, 19]. These conditions impose respectively $\mathbf{x}_\mathrm{d}$ to have a zero mean and the covariance matrix to be diagonalizable by graph shift eigenvectors.

Therefore, under the above stationary setting, the optimal Wiener filter (4) consists of a graph filter with a rational frequency response similar to (1) (not necessarily with a polynomial structure). In the sequel, we make use of recursion (2) to approximate $\mathbf{H}^\star$ and distributively reconstruct the graph signal $\mathbf{x}_\mathrm{d}$ from subsampled measurements. To ease the exposition, let us denote with $\mathbf{H}_\mathrm{W} = (\mathbf{I}-\mathbf{P}_\mathrm{W})^{-1}\mathbf{Q}_\mathrm{W}$ the ARMA filter response (3) that approximates (4).

3. GRAPH WIENER RECONSTRUCTION

Now, given $\mathbf{H}_\mathrm{W}$, we are interested in designing a sampling set $\mathcal{S} = \{v_{s_1}, \ldots, v_{s_M}\}$ of cardinality $|\mathcal{S}| = M \leq N$ such that it guarantees the minimum steady-state MSE reconstruction error. To this aim, let $\mathbf{D} = \mathrm{diag}(d_1, \ldots, d_N)$ denote the selection matrix with $d_i = 1$ if $v_i \in \mathcal{S}$ and $d_i = 0$, otherwise. Likewise, let $\mathbf{z} = \mathbf{D}\mathbf{x}$ be the $N \times 1$ vector of collected measurements with $z_i = x_i$ if $v_i \in \mathcal{S}$ and $z_i = 0$ otherwise.

Algorithm 1: Distributed computation of the ARMA output

Initialize $\mathbf{y}_0 = \mathbf{z}$, $\mathbf{w}^{(0)} = \mathbf{z}$, and ψ_p, φ_q to approximate (4)
for $q = 1, \ldots, Q$ compute the MA terms as:
  Collect $w^{(q-1)}$ from all neighbors $m \in \mathcal{N}_n$
  Compute $w_n^{(q)} = \sum_{m\in\mathcal{N}_n} L_{n,m}\left(w_n^{(q-1)} - w_m^{(q-1)}\right)$
Set $z_{Q,n} = \sum_{q=0}^{Q} \varphi_q w_n^{(q)}$
Set $\mathbf{y}_1 = \mathbf{z}_Q$, and compute $\mathbf{y}_t$ for $t \geq 2$ as:
  Set $\mathbf{w}^{(0)} = \mathbf{y}_{t-1}$
    for p = 1, ..., P
    Collect $w^{(p-1)}$ from all neighbors $m \in \mathcal{N}_n$
    Compute $w_n^{(p)} = \sum_{m\in\mathcal{N}_n} L_{n,m}\left(w_n^{(p-1)} - w_m^{(p-1)}\right)$
  Set $y_{t,n} = -\sum_{p=1}^{P} \psi_p z_n^{(p)} + z_{Q,n}$

With $\mathbf{z}$ being the new ARMA input, (3) is rewritten as

$$\mathbf{y}_\mathrm{W} = \mathbf{H}_\mathrm{W}\mathbf{z} = (\mathbf{I}-\mathbf{P}_\mathrm{W})^{-1}\mathbf{Q}_\mathrm{W}\mathbf{D}\mathbf{x}, \tag{5}$$

i.e., the steady-state reconstructed output is a Wiener spreading version of the sampled noisy measurements. Algorithm 1 illustrates the distributed implementation of this filter for $\mathbf{S} = \mathbf{L}$. The following proposition quantifies the mean square deviation (MSD) of the reconstructed signal $\mathbf{y}_\mathrm{W}$.

Proposition 1. *The steady-state* MSD $= \mathbb{E}\|\mathbf{H}_\mathrm{W}\mathbf{D}\mathbf{x} - \mathbf{x}_\mathrm{d}\|_2^2$ *of the reconstructed signal from* (5) *is*

$$\mathrm{MSD} = \mathrm{tr}\Big((\mathbf{H}_\mathrm{W}\mathbf{D}-\mathbf{I})\boldsymbol{\Sigma}_\mathrm{d}(\mathbf{H}_\mathrm{W}\mathbf{D}-\mathbf{I})^\mathsf{H} + \mathbf{H}_\mathrm{W}\mathbf{D}\boldsymbol{\Sigma}_\mathrm{n}\mathbf{D}\mathbf{H}_\mathrm{W}^\mathsf{H}\Big), \tag{6}$$

where $\mathbf{H}_\mathrm{W}$ *is the ARMA filter response that approximates the Wiener filter* (4)*,* $\mathbf{D}$ *is a diagonal sampling matrix, and* $\boldsymbol{\Sigma}_\mathrm{d}$ *and* $\boldsymbol{\Sigma}_\mathrm{n}$ *are the covariance matrices of* $\mathbf{x}_\mathrm{d}$ *and* $\mathbf{n}$*, respectively.*

Proof. (Sketch.) The claim can be proven by considering the norm property $\|\mathbf{a}\|_2^2 = \mathrm{tr}(\mathbf{a}\mathbf{a}^\mathsf{T})$ along with the the linearity of the trace and expectation and the mutual independence between $\mathbf{x}_\mathrm{d}$ and $\mathbf{n}$. □

Proposition 1 shows the impact of the filter accuracy $\mathbf{H}_\mathrm{W}$ and the sampling set $\mathbf{D}$ in the reconstruction MSD. Therefore, these parameters represent our handle to amend the steady-state reconstruction MSD. In the sequel, we propose a sampling strategy that acts on $\mathbf{D}$ such that the minimum reconstruction MSD (6) is attained.

Sampling strategy. Although the conventional approach to reduce the MSD involves a joint design of filtering coefficients and sampling nodes, this approach results in a challenging non-convex problem. To tackle this issue, we consider a non-joint design of these parameters. Specifically, we consider (ψ_p, φ_q) being designed in the filter design phase,

and we then tune $\mathbf{D}$ to minimize the reconstruction MSD. This idea is not totally new and it has been formerly used in [20, 21] for graph sampling for sketching and control, respectively. Pursuing then a sparse-sensing approach [14], the sampling set can be obtained from the optimization problem

$$\begin{aligned} \min_{\mathbf{d}} \quad & \mathrm{tr}\Big((\mathbf{H}_{\mathrm{W}}\mathbf{D}-\mathbf{I})\boldsymbol{\Sigma}_{\mathrm{d}}(\mathbf{H}_{\mathrm{W}}\mathbf{D}-\mathbf{I})^{\mathsf{H}}+\mathbf{H}_{\mathrm{W}}\mathbf{D}\boldsymbol{\Sigma}_{\mathrm{n}}\mathbf{D}\mathbf{H}_{\mathrm{W}}^{\mathsf{H}}\Big) \\ \text{s.t.} \quad & \mathbf{1}^{\mathsf{T}}\mathbf{d}=M,\ \ \mathbf{D}=\mathrm{diag}(\mathbf{d}),\ \ \mathbf{d}\in[0,1]^N. \end{aligned} \tag{7}$$

Problem (7) consists already of the relaxed convex version of the selection strategy. The approach designs $\mathbf{d}$ (thus the sampling set $\mathcal{S}$) that yields the minimum MSD value. The first constraint is the standard l_1-norm surrogate of the l_0-norm and imposes a maximum cardinality on $\mathcal{S}$. The last constraint, on the other hand, is a relaxation of the Boolean constraint $\mathbf{d}\in\{0,1\}^N$ to the box one. Conversely, one can also consider the opposite problem, i.e., finding the sparsest sampling pattern that guarantees a target MSD reconstruction error. The latter can as well be formulated as a convex problem.

In Section 6, we analyze the impact of $\mathbf{d}$ on the reconstruction error, and show that the position of the sampled nodes w.r.t. the graph is as important as the cardinality of the sampling set.

4. RECURSIVE WIENER RECONSTRUCTION

In this section we focus on Wiener-based reconstruction from a stream of independent noisy data $\mathbf{x}_t=\mathbf{x}_{\mathrm{d}}+\mathbf{n}_t$, for $t\geq 1$. This is a common situation in sensor networks, and to avoid the challenges of batch processing we consider a recursive implementation to incorporate the current data on the fly. Following the joint Tikhonov denoising in [22], we write the recursive Wiener reconstruction (RWR) algorithm as

$$\mathbf{y}_t=\mathbf{P}_{\mathrm{W}}\mathbf{y}_{t-1}+\mathbf{Q}_{\mathrm{W}}\mathbf{D}\mathbf{x}_t, \tag{8a}$$

$$\tilde{\mathbf{y}}_t=\frac{1}{t}\big((t-1)\tilde{\mathbf{y}}_{t-1}+\mathbf{y}_t\big), \tag{8b}$$

where (8a) incorporates the time-varying measurements into the ARMA recursion and (8b) consists of the running average of the ARMA output.

Recursion (8) can be expanded to all its terms as

$$\begin{aligned} \tilde{\mathbf{y}}_t &= \frac{1}{t}\sum_{\tau_1=1}^{t}\sum_{\tau_2=0}^{\tau_1-1}\mathbf{P}_{\mathrm{W}}^{\tau_2}\mathbf{Q}_{\mathrm{W}}\mathbf{D}\mathbf{x}_{\tau_1-\tau_2} \\ &= \underbrace{\frac{1}{t}\sum_{\tau_1=1}^{t}\sum_{\tau_2=0}^{\tau_1-1}\mathbf{P}_{\mathrm{W}}^{\tau_2}\mathbf{Q}_{\mathrm{W}}\mathbf{D}\mathbf{x}_{\mathrm{d}}}_{\boldsymbol{\chi}_t}+\underbrace{\frac{1}{t}\sum_{\tau_1=1}^{t}\sum_{\tau_2=0}^{\tau_1-1}\mathbf{P}_{\mathrm{W}}^{\tau_2}\mathbf{Q}_{\mathrm{W}}\mathbf{D}\mathbf{n}_{\tau_1-\tau_2}}_{\boldsymbol{\eta}_t}. \end{aligned} \tag{9}$$

Then, by working out the sums we rewrite $\boldsymbol{\chi}_t$ and $\boldsymbol{\eta}_t$ as:

$$\begin{aligned} \boldsymbol{\chi}_t &= \frac{1}{t}\sum_{\tau_1=1}^{t}(\mathbf{I}-\mathbf{P}_{\mathrm{W}}^{\tau_1})(\mathbf{I}-\mathbf{P}_{\mathrm{W}})^{-1}\mathbf{Q}_{\mathrm{W}}\mathbf{D}\mathbf{x}_{\mathrm{d}} \\ &= \frac{t+1}{t}(\mathbf{I}-\mathbf{P}_{\mathrm{W}})^{-1}\mathbf{Q}_{\mathrm{W}}\mathbf{D}\mathbf{x}_{\mathrm{d}} \\ &\quad -\frac{1}{t}(\mathbf{I}-\mathbf{P}_{\mathrm{W}}^{t+1})(\mathbf{I}-\mathbf{P}_{\mathrm{W}})^{-2}\mathbf{Q}_{\mathrm{W}}\mathbf{D}\mathbf{x}_{\mathrm{d}}\triangleq\mathbf{H}_{\mathrm{WD},t}\mathbf{x}_{\mathrm{d}}, \end{aligned} \tag{10}$$

$$\boldsymbol{\eta}_t=\frac{1}{t}\sum_{\tau_1=1}^{t}\sum_{\tau_2=0}^{t-\tau_1}\mathbf{P}_{\mathrm{W}}^{\tau_2}\mathbf{Q}_{\mathrm{W}}\mathbf{D}\mathbf{n}_{\tau_1}, \tag{11}$$

where (10) is obtained by expressing the geometric series in closed form since $\|\mathbf{P}_{\mathrm{W}}\|<1$, and (11) follows from simple algebra. With these in place, we next characterize the steady state performance and the MSD of the RWR algorithm.

Steady state performance. The main result is summarized in the following proposition.

Proposition 2. *The steady-state behavior of the RWR recursion* (8) *is*

$$\tilde{\mathbf{y}}\triangleq\lim_{t\to\infty}\tilde{\mathbf{y}}_t=(\mathbf{I}-\mathbf{P}_W)^{-1}\mathbf{Q}_W\mathbf{D}\mathbf{x}_{\mathrm{d}}. \tag{12}$$

Recursion (8) *converges to* (12) *with rate* t^{-1}.

Proof. (Sketch.) From (10) we can observe that $\lim_{t\to\infty}\boldsymbol{\chi}_t=(\mathbf{I}-\mathbf{P}_{\mathrm{W}})^{-1}\mathbf{Q}_{\mathrm{W}}\mathbf{D}\mathbf{x}_{\mathrm{d}}$ since all other terms vanish as t^{-1}. Similarly, by expressing all terms of (11) in closed form ($\sum_{\tau=0}^{a}\mathbf{A}^{\tau}=(\mathbf{I}-\mathbf{A})^{-1}(\mathbf{I}-\mathbf{A}^{a+1})$, if $\|\mathbf{A}\|<1$) it can be observed that $\lim_{t\to\infty}\boldsymbol{\eta}_t=\mathbf{0}$ with rate t^{-1}. □

Proposition 2 states that the asymptotic reconstructed signal of $\mathbf{x}_{\mathrm{d}}$ through the RWR algorithm (8) consists of a percolated version of the true sampled signal $\mathbf{x}_{\mathrm{d}}$ through a Wiener graph filter. Observe that the noise contribution is removed due to the running average in (8b).

Mean square error analysis. To fully characterize the RWR performance, the following proposition provides a closed form expression for the MSD of the RWR recursion (8).

Proposition 3. *The* $\mathrm{MSD}_t=\mathbb{E}\|\tilde{\mathbf{y}}_t-\mathbf{x}_{\mathrm{d}}\|_2^2$ *of the reconstructed signal* $\mathbf{y}_t$ *from the RWR algorithm* (8) *is*

$$\begin{aligned} \mathrm{MSD}_t &= \mathrm{tr}\left[\Big(\mathbf{H}_{\mathrm{WD},t}-\mathbf{I}\Big)\boldsymbol{\Sigma}_{\mathrm{d}}\Big(\mathbf{H}_{\mathrm{WD},t}-\mathbf{I}\Big)^{\mathsf{H}}\right] \\ &+\frac{1}{t^2}\mathrm{tr}\left[\left(\sum_{\tau=0}^{t-1}\mathbf{P}_W^{2\tau}+\ldots+\sum_{\tau=0}^{0}\mathbf{P}_W^{2\tau}\right)\mathbf{Q}_W\mathbf{D}\boldsymbol{\Sigma}_n\mathbf{D}\mathbf{Q}_W^{\mathsf{H}}\right], \end{aligned} \tag{13}$$

with $\mathbf{H}_{\mathrm{WD},t}$ *defined in* (10)*. Moreover,* (13) *converges to the steady-state MSD*

$$\mathrm{MSD}_\infty=\lim_{t\to\infty}\mathrm{MSD}_t=\mathrm{tr}\Big((\mathbf{H}_{\mathrm{W}}\mathbf{D}-\mathbf{I})\boldsymbol{\Sigma}_{\mathrm{d}}(\mathbf{H}_{\mathrm{W}}\mathbf{D}-\mathbf{I})^{\mathsf{H}}\Big), \tag{14}$$

with rate t^{-1}

Proof. (Sketch.) From the mutual independence of $\mathbf{x}_{\mathrm{d}}$ and $\mathbf{n}_t$ the MSD at time t can be written as

$$\mathrm{MSD}_t=\mathrm{tr}\mathbb{E}\big[(\boldsymbol{\chi}_t-\mathbf{x}_{\mathrm{d}})(\boldsymbol{\chi}_t-\mathbf{x}_{\mathrm{d}})^{\mathsf{H}}\big]+\mathrm{tr}\mathbb{E}\big(\boldsymbol{\eta}_t\boldsymbol{\eta}_t^{\mathsf{H}}\big). \tag{15}$$

The first trace-term on the right-hand side of (13) can then be obtained from the definition of the covariance matrix in

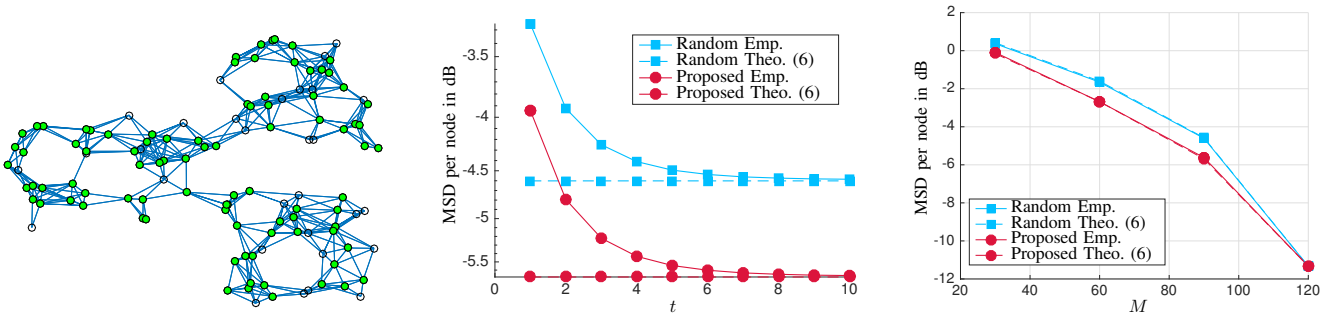

Fig. 1: (Left) Considered graph topology with $N = 120$. Green circles depict the selected $M = 90$ nodes from (7). (Center) Average MSD per node as a function of t for the sampling strategy (7) compared to a uniformly random sampling. (Right) Average MSD per node as a function of M, for $t = 10$.

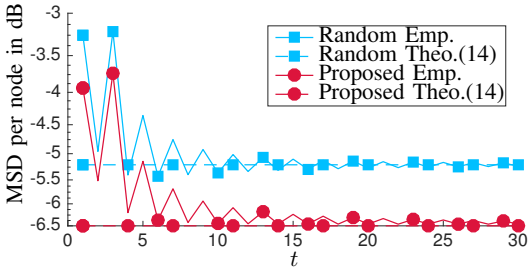

Fig. 2: Average MSD per node as a function of t for the RWR algorithm with $M = 90$ sampled nodes.

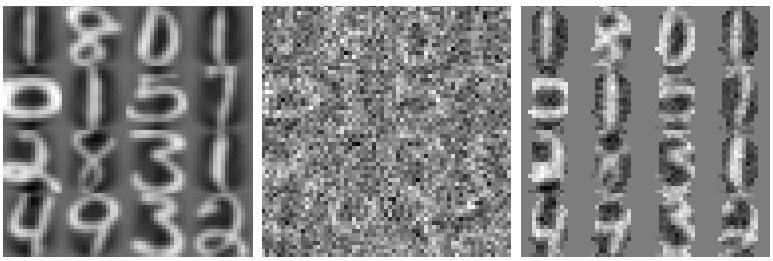

Fig. 3: Illustration for the USPS data set. (Left) Some digits from the data set. (Center) Noisy graph signal (digits) with 50% non-sampled nodes. (Right) Recovered digits after 30 observations.

$\text{tr}\mathbb{E}\left[(\boldsymbol{\chi}_t - \mathbf{x}_\text{d})(\boldsymbol{\chi}_t - \mathbf{x}_\text{d})^\text{H}\right]$. The second trace-term in the right hand-side of (13) is obtained by first substituting the expression for $\boldsymbol{\eta}_t$, and then by making use of the independence of the noise realizations in time and exploiting the trace property $\text{tr}(\mathbf{ABC}) = \text{tr}(\mathbf{CAB}) = \text{tr}(\mathbf{BCA})$.

The convergence of (13) to (14) can be proven by expressing all sums in closed form (recall $\|\mathbf{P}\| < 1$) and then taking the limit for $t \to \infty$. □

An important outcome of Proposition 3 is that (14) can be used to design the sampling strategy, similarly to (7), for sampling the graph such that the steady-state MSD is minimized.

5. NUMERICAL RESULTS

In this section, we evaluate the performance of the proposed algorithms first in a synthetic scenario and then in the USPS data set.

Synthetic scenario. We consider a graph composed of $N = 120$ nodes, with a topology depicted in Fig. 1 (Left). The graph signal power spectral density is $\sigma_\text{d}^2(\lambda_n) = \exp(-8\lambda_n)$ and the zero-mean Gaussian noise $\mathbf{n}$ has a covariance matrix $\boldsymbol{\Sigma}_\text{n} = 0.01\mathbf{I}$. The Wiener filter (4) is approximated with an ARMA recursion of $P = 3$ and $Q = 5$ [18] leading to an MSE approximation error of order 10^{-5}. The shift operator is $\mathbf{S} = \mathbf{L}_\text{n} - 0.5\lambda_\text{max}(\mathbf{L}_\text{n})\mathbf{I}$, where $\mathbf{L}_\text{n}$ denotes the normalized Laplacian. The empirical results are averaged over 10^3 different realizations of $\mathbf{x}_\text{d}$ and 10^6 realizations of the noise.

In Fig. 1 (Center) we show the MSD per node as a function of t for the proposed sampling strategy and random uniformly sampling. In this example, M is fixed to 90 and the selected nodes are shown in green in Fig. 1 (Left). We observe that both strategies attain their theoretical performance in just a few iterations with the sparse sensing sampling (7) performing 1.1 dB better. Fig. 1 (Right) shows the MSD per node as a function of M. In these results, the ARMA recursion is arrested after $t = 10$ iterations. We observe that the signal reconstruction is affected when more than 50% of the nodes are not sampled. From the latter findings sampling sets with $M \geq 90$ nodes seem a reasonable choice.

Next, in Fig. 2 we show the performance of the RWR algorithm as a function of t for $M = 90$. Once again, we observe the fast convergence of the ARMA to its asymptotic MSD (14), and the performance improvement of the proposed sampling strategy (minimizing (14)) over random sampling.

USPS data set. To further illustrate the reconstruction performance, we consider the USPS dataset which presents a high degree of graph stationarity [10]. Under the settings of [10], we additionally corrupted the data with zero-mean Gaussian noise of variance $\sigma_\text{n}^2(\lambda_n) = 1$, for all λ_n. The Wiener filter is approximated by an ARMA with $P = 3$ and $Q = 5$ leading to an MSE of order 10^{-7}.

Fig. 3 illustrates 16 digits from the dataset (Left), their noisy subsampled version (Center) and the reconstructed signal (Right). For this setup, N is 256, $M = 128$ pixels are selected using (14) and $t = 30$. We observe how the sampling approach focuses on reconstructing the most relevant part of the images (i.e., where the digits are) and leaves more errors in the borders (i.e., where there is no information).

6. REFERENCES

[1] 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.

[2] Mikhail Tsitsvero, Sergio Barbarossa, and Paolo Di Lorenzo, "Signals on graphs: Uncertainty principle and sampling," *IEEE Transactions on Signal Processing*, vol. 64, no. 18, pp. 4845–4860, 2016.

[3] Antonio G Marques, Santiago Segarra, Geert Leus, and Alejandro Ribeiro, "Sampling of graph signals with successive local aggregations," *IEEE Transactions on Signal Processing*, vol. 64, no. 7, pp. 1832–1843, 2016.

[4] Aamir Anis, Akshay Gadde, and Antonio Ortega, "Efficient sampling set selection for bandlimited graph signals using graph spectral proxies," *IEEE Transactions on Signal Processing*, vol. 64, no. 14, pp. 3775–3789, 2016.

[5] Paolo Di Lorenzo, Sergio Barbarossa, Paolo Banelli, and Stefania 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] Paolo Di Lorenzo, Paolo Banelli, Sergio Barbarossa, and Stefania Sardellitti, "Distributed adaptive learning of graph signals," *IEEE Transactions on Signal Processing*, 2017.

[7] Paolo Di Lorenzo, Paolo Banelli, Elvin Isufi, Sergio Barbarossa, and Geert Leus, "Adaptive graph signal processing: Algorithms and optimal sampling strategies," *arXiv preprint arXiv:1709.03726*, 2017.

[8] Sunil K Narang, Akshay Gadde, and Antonio Ortega, "Signal processing techniques for interpolation in graph structured data," in *Acoustics, Speech and Signal Processing (ICASSP), 2013 IEEE International Conference on*. IEEE, 2013, pp. 5445–5449.

[9] Daniel Romero, Meng Ma, and Georgios B Giannakis, "Kernel-based reconstruction of graph signals," *IEEE Transactions on Signal Processing*, vol. 65, no. 3, pp. 764–778, 2017.

[10] Nathanaël Perraudin and Pierre Vandergheynst, "Stationary signal processing on graphs," *IEEE Transactions on Signal Processing*, vol. 65, no. 13, pp. 3462–3477, 2017.

[11] Antonio G Marques, Santiago Segarra, Geert Leus, and Alejandro Ribeiro, "Stationary graph processes and spectral estimation," *IEEE Transactions on Signal Processing*, 2017.

[12] Paolo Di Lorenzo, Elvin Isufi, Paolo Banelli, Sergio Barbarossa, and Geert Leus, "Distributed Recursive Least Squares Strategies for Adaptive Reconstruction of Graph Signals," in *EURASIP European Signal Processing Conference (EUSIPCO), Kos, Greece, Aug.-Sept. 2017*, 2017.

[13] Elvin Isufi, Andreas Loukas, Andrea Simonetto, and Geert Leus, "Autoregressive moving average graph filtering," *IEEE Transactions on Signal Processing*, vol. 65, no. 2, pp. 274–288, 2017.

[14] Sundeep Prabhakar Chepuri and Geert Leus, "Sparse sensing for statistical inference," *Foundations and Trends® in Signal Processing*, vol. 9, no. 3–4, pp. 233–368, 2016.

[15] David I Shuman, Sunil K Narang, Pascal Frossard, Antonio Ortega, and Pierre 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, no. 3, pp. 83–98, 2013.

[16] Aliaksei Sandryhaila and José MF Moura, "Discrete signal processing on graphs," *IEEE transactions on signal processing*, vol. 61, no. 7, pp. 1644–1656, 2013.

[17] Jiani Liu, Elvin Isufi, and Geert Leus, "Filter design for autoregressive moving average graph filters," *arXiv preprint arXiv:1711.09086*, 2017.

[18] Elvin Isufi, Andreas Loukas, and Geert Leus, "Autoregressive moving average graph filters a stable distributed implementation," in *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2017, number EPFL-CONF-223825.

[19] Benjamin Girault, "Stationary graph signals using an isometric graph translation," in *Signal Processing Conference (EUSIPCO), 2015 23rd European*. IEEE, 2015, pp. 1516–1520.

[20] Fernando Gama, Antonio G Marques, Gonzalo Mateos, and Alejandro Ribeiro, "Rethinking sketching as sampling: A graph signal processing approach," *arXiv preprint arXiv:1611.00119*, 2016.

[21] Fernando Gama, Elvin Isufi, Geert Leus, and Alejandro Ribeiro, "Control of Graph Signal Over Random Time-Varying Graphs," in *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2018, number to appear.

[22] Elvin Isufi, Andreas Loukas, Andrea Simonetto, and Geert Leus, "Filtering random graph processes over random time-varying graphs," *IEEE Transactions on Signal Processing*, 2017.