Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Bayesian Graph Signal Recovery and Prediction Using an AR Model

Razieh Torkamani

r.torkamani@basu.ac.ir

Bu-Ali Sina University

Hadi Zayyani

Qom University of Technology

Mohammad Salman

American University of the Middle East

Research Article

Keywords:

Posted Date: November 25th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-11745809/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: No competing interests reported.

Bayesian Graph Signal Recovery and Prediction Using an AR Model

Razieh Torkamani[1*], Hadi Zayyani[2] and Mohammad Salman[3]

[1] Department of Electrical Engineering, Faculty of Engineering, Bu-Ali Sina University, Hamedan, Iran .

[2] Department of Electrical and Computer Engineering, Qom University of Technology, Qom, Iran .

[3] College of Engineering and Technology, American University of the Middle East, Egaila, 54200, Kuwait .

*Corresponding author(s). E-mail(s): r.torkamani@basu.ac.ir;
Contributing authors: zayyani@qut.ac.ir; mohammad.salman@aum.edu.kw;

Abstract

In this paper, a statistical spatio-temporal model is employed to characterize graph signals by jointly capturing both temporal and spatial dependencies. In the proposed model, the temporal dependency between consecutive graph signals is described using a state-space autoregressive (SS-AR) model, while the spatial dependency among graph nodes is modeled by a Gaussian Markov random field (GMRF). Furthermore, the graph structure is assumed to be unknown, and a generalized Gamma distribution (G$\mathbf{\Gamma}$D) is adopted as the prior distribution for the elements of the adjacency matrix. The graph signal, the SS-AR matrix, and the adjacency matrix are jointly estimated using the variational Bayes (VB) framework. Simulation results on both synthetic and real-world datasets demonstrate that the proposed algorithm, which jointly exploits spatial and temporal dependencies, achieves more accurate graph signal reconstruction and graph structure estimation than existing methods that consider only spatial dependencies among nodes.

Keywords: Graph Signal Recovery, Graph Learning, State-Space AR Model, Generalized Gamma Distribution, Gaussian Markov Random Field, Variational Bayes

1 Introduction

Graph Signal Processing (GSP) is a paradigm alternative to classical signal processing in which the signal is defined over an irregular domain such as a graph [1], [2]. Many problems are arisen within the framework of GSP which Graph Signal Recovery (GSR), Graph Learning (GL), and graph sampling are to name a few. In this paper, we focus on the two GSR and GL problems simultaneously. In GSR [3]-[5], the graph signal is to be recovered from only a partial known graph signal samples. In GL [6]-[7], the structure or the Laplacian matrix of the graph is to be estimated from the graph signal. But, the more recent problem is to simultaneously estimate the graph signal and learning the graph from incomplete measurements of the graph signal [8]-[11]. We nominate this problem as joint GSR-GL problem.

On the other hand, the graph signal model used in the literature and mainly in the joint problem of GSR and GL are diverse. A number of papers consider the bandlimited or smooth graph signal models [12], [13], while others utilize Gaussian Markov Random Field (GMRF) [8]-[10], and some others use Auto-Regressive (AR) models [14]-[17] or some variants of them such as Vector AutoRegressive Moving Average (VARMA) models [18]. Moreover, rare models used some unique models such as Generalized Autoregressive Conditional Heteroscedasticity (GARCH) model [19].

In this paper, due to higher capability of modeling the time dependence of graph signal samples, we focuss on the AR graph signal models. In pioneering paper of [14], it is suggested to use a AR model for causal modeling of unstructured data such as graph signal. [15] presented a topological Kalman filter for joint GSR-GL. Moreover, in [16], Structural Vector AutoRegressive Models (SVARMs) are utilized for the joint GSR-GL problem and some optimization techniques are proposed to solve the problem. In addition, [17] minimizes the smoothness with autoregressive coefficients. Also, in [18], the time-series of a graph signal is forecasted using a VARMA model by some fitting approaches. Besides, [19] utilized three different version of GARCH models for time-series prediction of a graph signal.

In this paper, due to success of Bayesian approaches in GSP problems [8]-[10], [20], [21], we extend using the Bayesian methods in joint GSR-GL problem using the AR graph signal model. Especially, to overcome the difficulty of calculating the Minimum Mean Square Estimation (MMSE) solution, we utilize the Variational Bayes (VB) algorithm to solve the problem. Hence, a near state-space model is defined for solving the problem and the posteriors are calculated in closed-form by some prevalent statistical assumptions. Then, the VB algorithm is derived for joint GSR-GL problem. Simulation results show the effectiveness of the proposed algorithm both in synthetic data scenario and also in recovering the real-world graph signals.

The rest of the paper is organized as follows: In Section 2, the preliminaries are explained. In Section 3, the derivation of the proposed VB-based algorithm is presented. The graph learning of the algorithm are discussed in Section 4. In Section 5, simulation results and their discussion are introduced. Finally, conclusions and potential future research directions are provided in Section 6 .

2 Preliminaries

2.1 Graph Signal Recovery Problem

In this paper, we focus on recovering a graph signal from under-sampled and noisy observations. Let us consider an undirected graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ with N vertices, where $\mathcal{V} = \{\sqsubseteq_\infty, \sqsubseteq_\in, \ldots, \sqsubseteq_\mathcal{N}\}$ denotes the vertex set, and $\mathcal{E} \subset \mathcal{V} \times \mathcal{V}$ represents the edges. The adjacency matrix is denoted by $\mathbf{W}$, whose entries encode edge weights. From this, the Laplacian matrix is defined as $\mathbf{L} = \mathbf{D} - \mathbf{W}$, with $\mathbf{D}$ being the diagonal degree matrix such that $D_m = \sum_n W_{mn}$. A noisy under-sampled observation is then obtained by sampling the graph signal on M randomly selected vertices, leading to the model

$$\mathbf{y} = \Psi\mathbf{x} + \mathbf{n}, \tag{1}$$

Here, $\Psi \in \mathbf{R}^{M\times N}$ is the sampling matrix with a single "one" per row at the sampled node index and zeros elsewhere, $\mathbf{x} \in \mathbf{R}^{N\times 1}$ is the original graph signal, and $\mathbf{n} \in \mathbf{R}^{M\times 1}$ represents measurement noise.

The objective of this paper is to reconstruct the original graph signal $\mathbf{x}$ from observations that are limited to a subset of nodes and corrupted by noise. To achieve this, we employ a statistical graph signal recovery (GSR) algorithm based on variational Bayesian inference. This approach estimates the graph signal from the given measurements, effectively handling both the undersampling and the noise present in the data.

2.2 The Proposed Graph Signal Model

The state-space model for the observations in the graph can be represented as

$$\begin{aligned} \mathbf{y}(t) &= \mathbf{\Psi}\mathbf{x}(t) + \mathbf{n}(t) \\ \mathbf{x}(t) &= \mathbf{G}\mathbf{x}(t-1) + \mathbf{e}(t) \end{aligned} \tag{2}$$

where $\Psi \in \mathbf{R}^{M\times N}$ denotes the known sensing matrix, $\mathbf{n}$ is the measurement noise, $\mathbf{G} \in \mathbf{R}^{N\times N}$ is the diagonal matrix, and $\mathbf{e}$ is the error vector. In this paper, we assume a state-space AR (SS-AR)

model for the observations as

$$\mathbf{y}(t) = \mathbf{\Psi}\mathbf{x}(t) + \mathbf{n}(t)$$
$$\mathbf{x}(t) = \sum_{k=1}^{P} \mathbf{G}^{(k)}\mathbf{x}(t-k) + \mathbf{e}(t) \quad (3)$$

where $\mathbf{G}^{(k)}(t)$ is an $N \times N$ diagonal matrix which relates the graph signal at time t to the graph signal from k time steps before it. For the measurement noise and the error vector, we assume zero-mean Gaussian distributions with precisions α_n and $\mathbf{B}^{-1}$, respectively, i.e.,

$$p(\mathbf{n}) = \mathcal{N}(0, \alpha_{\mathrm{n}}^{-1}\mathbf{I}_{\mathrm{M}}) \quad (4)$$
$$p(\mathbf{e}) = \mathcal{N}(0, \mathbf{B}^{-1}) \quad (5)$$

where $\mathbf{I}_M$ denotes the $M \times M$ identity matrix. The precision matrix for the error vector is calculated as $\mathbf{B} = \mathbf{L} + \epsilon\mathbf{I}_N$, and the precision matrix for the measurement noise vector is drawn from a Gamma distribution with parameters d_n and f_n as

$$p(\alpha_n) = \mathrm{Gamma}(d_n, f_n) \quad (6)$$

Thus, the Likelihood function can be written as

$$p(\mathbf{y}|\mathbf{x}, \alpha_n^{-1}) = \mathcal{N}(\mathbf{\Psi}\mathbf{x}(\mathrm{t}), \alpha_{\mathrm{n}}^{-1}\mathbf{I}_{\mathrm{M}}) \quad (7)$$

Moreover, we assume a GMRF model for the graph signal, which reflects the spatial dependency between nodes,

$$p(\mathbf{x}) = \frac{|\mathbf{B}|^{\frac{1}{2}}}{(2\pi)^{\frac{N}{2}}} \exp\left(\frac{-\mathbf{x}^T\mathbf{B}\mathbf{x}}{2}\right) \quad (8)$$

where $|.|$ denotes the matrix determinant. In this paper, we employed an AR model of order P to model the temporal dependency among time instances. To this end, we use a Gaussian distribution to denote this relation, statistically:

$$p\left(\mathbf{x}(t)|\mathbf{x}(t-i)\right)_{i=1}^{P} = \mathcal{N}\left(\sum_{\mathrm{k}=1}^{\mathrm{P}} \mathbf{G}^{(\mathrm{k})}\mathbf{x}(\mathrm{t}-\mathrm{k}), \mathbf{B}^{-1}\right) \quad (9)$$

For the elements of the matrix $\mathbf{G}^{(m)}$, $m = 1 : P$, we assume a Gaussain distribution as

$$p(g_i^{(m)}) = \mathcal{N}\left(\rho_{\mathrm{i}}^{(\mathrm{m})}, \left(\beta_{\mathrm{i}}^{(\mathrm{m})}\right)^{-1}\right) \quad (10)$$

where $\rho_i^{(m)}$ is the weight which relates the i'th elements of the graph signal at time t to m time step before, and $\beta_i^{(m)}$ is the precision of the $g_i^{(m)}$, which we assume that follows from a Gamma distribution with parameters d_β and f_β as

$$p(\beta_i^{(m)}) = \mathrm{Gamma}(d_\beta, f_\beta) \quad (11)$$

Finally, we assume a Gaussian distribution for the weight parameters $\rho_i^{(m)}$

$$p(\rho_i^{(m)}) = \mathcal{N}\left(0, \mathrm{f}_{\rho,\mathrm{i}}^{(\mathrm{m})}\right) \quad (12)$$

2.3 The Proposed Model for Adjacency Matrix

In this work, the entries of the adjacency matrix are modeled with the Generalized Gamma Distribution (GΓD), which can be formulated as:

$$p(w_{ij}) = \mathrm{G\Gamma D}(d_w, f_w, \gamma_w) \quad (13)$$

where w_{ij} represents the weight assigned to the edge between nodes i and j. The above GΓD is formulated as

$$p(w_{ij}) = \frac{\gamma_w}{(d_w)^{f_w}} \frac{w_{ij}^{(f_w-1)}}{\Gamma\left(\frac{f_w}{\gamma_w}\right)} \exp\left(-\left(\frac{w_{ij}}{d_w}\right)^{f_w}\right) \quad (14)$$

3 The proposed VB inference algorithm for GSR

In this section, we employ VB-inference method to derive the posterior distributions based on the priors presented in section 2.2. Let $\theta = \{\mathbf{x}, \alpha_n, \beta_i^{(m)}, g_i^{(m)}, \rho_i^{(m)}\}$ denote the set of all unknown variables. In the VB method, we estimate the posterior distribution of the parameter θ_i, i.e., $p(\theta_i|\mathbf{y})$, by a function $q(\theta_i)$, in the way that the Kullback-Leibler divergence (KLD) between the true posterior $p(\theta_i|\mathbf{y})$ and the estimated posterior $q(\theta_i)$ is minimized. The main assumption in VB is the factorization of the parameters as

$$q(\theta) = \prod_i q(\theta_i) \quad (15)$$

The posterior distribution for the graph signal $\mathbf{x}$, based on the prior distribution described in (9)

and the Likelihood function presented in (7), can be derived as

$$\begin{aligned}
&\ln\left(q(\mathbf{x}(t))\right)\\
&=\left\langle \ln p(\mathbf{y}(t),\mathbf{x}(t),\alpha_n,\beta_i^{(m)},g_i^{(m)},\rho_i^{(m)})\right\rangle_{\alpha_n,\alpha_i,g_i^{(m)},\rho_i^{(m)}}\\
&\quad+\text{const.}\\
&=\left\langle \ln p(\mathbf{y}(t)|\mathbf{x}(t))\right\rangle_{\alpha_n}+\left\langle \ln p(\mathbf{x}(t)|\mathbf{x}(t-m)\right\rangle_{g_i^{(m)}}+\text{const.}\\
&=-\left\langle \frac{\alpha_n}{2}\left(\mathbf{y}(t)-\Psi\mathbf{x}(t)\right)^T\left(\mathbf{y}(t)-\Psi\mathbf{x}(t)\right)\right\rangle_{\alpha_n}\\
&\quad-\frac{1}{2}\left\langle \mathbf{r}(t)^T\mathbf{B}\mathbf{r}(t)\right\rangle_{g_i^{(m)}}+\text{const.}
\end{aligned}\tag{16}$$

where we have

$$\mathbf{r}(t)=\mathbf{x}(t)-\sum_{k=1}^{P}\mathbf{G}^{(k)}\mathbf{x}(t-k)$$

It is obvious that the above posterior belongs to the Gaussian distribution with parameters as follows

$$q(\mathbf{x}(t))=\mathcal{N}(\mu_{\mathrm{x}},\Sigma_{\mathrm{x}}) \tag{17}$$

$$\Sigma_x=\left(\mathbf{B}+\alpha_n\Phi^T\Phi\right)^{-1} \tag{18}$$

$$\mu_x=\Sigma_x\left(\alpha_n\Phi^T\mathbf{y}(t)+\mathbf{B}\sum_{k=1}^{P}\mathbf{G}^{(k)}\mathbf{x}(t-k)\right) \tag{19}$$

For the diagonal elements of the matrix $\mathbf{G}^{(m)}$, i.e., $g_i^{(m)}, i=1:N, m=1:P$, the VB procedure for deriving the posterior distribution is as follows

$$\begin{aligned}
&\ln\left(q(g_i^{(m)})\right)\\
&=\left\langle \ln p(\mathbf{y}(t),\mathbf{x}(t),\alpha_n,\alpha_i,g_i^{(m)},\rho_i^{(m)})\right\rangle_{\mathbf{x},\alpha_n,\beta_i^{(m)},\rho_i^{(m)}}\\
&\quad+\text{const.}=\left\langle \ln p(\mathbf{x}(t)|g_i^{(m)})\right\rangle_{\mathbf{x}(t)}\\
&\quad+\left\langle \ln p(g_i^{(m)}|\rho_i^{(m)},\beta_i^{(m)})\right\rangle_{\rho_i^{(m)},\beta_i^{(m)}}+\text{const.}\\
&=-\frac{1}{2}\left\langle -2g_i^{(m)}h_i^{(m)}+(g_i^{(m)})^2(x_i(t-m))^2b_i\right\rangle_{\mathbf{x}}\\
&\quad-\left\langle \frac{\beta_i^{(m)}}{2}\left(g_i^{(m)}-\rho_i^{(m)}\right)^2\right\rangle_{\rho_i^{(m)},\beta_i^{(m)}}+\text{const.}
\end{aligned}\tag{20}$$

where b_i is the i'th diagonal element of the matrix $\mathbf{B}$, and we have

$$\begin{aligned}
h_i^{(m)}=&x_i(t-m)\mathbf{x}(t)^T\mathbf{B}_i\\
&-x_i(t-m)\mathbf{B}\sum_{\substack{k=1\\k\neq m}}^{P}\mathbf{G}^{(k)}\mathbf{x}(t-k)\\
&-\frac{1}{2}x_i(t-m)\mathbf{x}^T(t-m)\mathbf{G}^{(m)}\mathbf{B}_i
\end{aligned}\tag{21}$$

where $\mathbf{B}_i$ denotes the i'th column of the matrix $\mathbf{B}$. It is evident that the posterior for the $g_i^{(m)}$ is Gaussian distribution as

$$q(g_i^{(m)})=\mathcal{N}(\mu_{g_i^{(m)}},\sigma^2_{g_i^{(m)}}) \tag{22}$$

$$\sigma^2_{g_i^{(m)}}=\left(x_i(t-m)^2b_i+\beta_i^{(m)}\right)^{-1} \tag{23}$$

$$\mu_{g_i^{(m)}}=\sigma^2_{g_i^{(m)}}\left(\rho_i^{(m)}\beta_i^{(m)}+h_i^{(m)}\right) \tag{24}$$

We can derive the posterior for the noise precision,α_n in the same manner as

$$\begin{aligned}
&\ln\left(q(\alpha_n)\right)\\
&=\left\langle \ln p(\mathbf{y}(t),\mathbf{x}(t),\alpha_n,\beta_i^{(m)},g_i^{(m)},\rho_i^{(m)})\right\rangle_{\mathbf{x},\beta_i^{(m)},g_i^{(m)},\rho_i^{(m)}}\\
&\quad+\text{const.}=\left\langle \ln p(\mathbf{y}(t)|\mathbf{x},\alpha_n)+\ln p(\alpha_n)\right\rangle_{\mathbf{x}(t)}+\text{const.}\\
&=-\frac{M}{2}\ln(\alpha_n)-\frac{\alpha_n}{2}\left\langle \left(\mathbf{y}(t)-\Psi\mathbf{x}(t)\right)^T\left(\mathbf{y}(t)-\Psi\mathbf{x}(t)\right)\right\rangle_{\mathbf{x}}\\
&\quad+(d_n-1)\ln(\alpha_n)-f_n\alpha_n+\text{const.}
\end{aligned}\tag{25}$$

which belongs to the Gamma distribution as

$$q(\alpha_n)=\text{Gamma}(d_n',f_n') \tag{26}$$

$$d_n'=d_n+\frac{M}{2} \tag{27}$$

$$\begin{aligned}
f_n'=&f_n+\frac{1}{2}(\mathbf{y}(t)-\Psi\mu_x)^T(\mathbf{y}(t)-\Psi\mu_x)\\
&+\frac{1}{2}\text{Tr}(\Psi\Sigma_x\Psi^T)
\end{aligned}\tag{28}$$

Moreover, the posterior distribution for the weights $\rho_i^{(m)}$ can be derived as

$$\begin{aligned}
&\ln\left(q(\rho_i^{(m)})\right)\\
&=\left\langle \ln p(\mathbf{y}(t),\mathbf{x}(t),\alpha_n,\beta_i^{(m)},g_i^{(m)},\rho_i^{(m)})\right\rangle_{\mathbf{x},\alpha_n,\beta_i^{(m)},g_i^{(m)}}
\end{aligned}$$

$$+\text{const.} = -\frac{1}{2}\frac{\left(\rho_i^{(m)}\right)^2}{f_{\rho_i}^{(m)}} - \frac{\left(\rho_i^{(m)}\right)^2}{2}\beta_i^{(m)} + \rho_i^{(m)}\beta_i^{(m)}g_i^{(m)} + \text{const.} \quad (29)$$

It is evident that $\rho_i^{(m)}$ has Gaussian distribution given by

$$q(\rho_i^{(m)}) = \mathcal{N}(\mu_{\rho_i^{(m)}}, \sigma^2_{\rho_i^{(m)}}) \quad (30)$$

$$\sigma^2_{\rho_i^{(m)}} = \left(\frac{1}{f_{\rho,i}^{(m)}} + \beta_i^{(m)}\right)^{-1} \quad (31)$$

$$\mu_{\rho_i^{(m)}} = \sigma^2_{\rho_i^{(m)}}\beta_i^{(m)}g_i^{(m)} \quad (32)$$

The posterior of the $\beta_i^{(m)}$ can be calculated as follows:

$$\ln\left(q(\beta_i^{(m)})\right) = \left\langle \ln p(\mathbf{y}(t), \mathbf{x}(t), \alpha_n, \beta_i^{(m)}, g_i^{(m)}, \rho_i^{(m)})\right\rangle_{\mathbf{x},\alpha_n,g_i^{(m)},\rho_i^{(m)}} + \text{const.} = (d_\beta - 1)\ln(\beta_i^{(m)}) - f_\beta\beta_i^{(m)} + \frac{1}{2}\ln(\beta_i^{(m)}) - \frac{1}{2}\beta_i^{(m)}\left(g_i^{(m)} - \rho_i^{(m)}\right)^2 + \text{const.} \quad (33)$$

It is obvious that the above function belongs to a Gaussian distribution with parameters

$$q(\beta_i^{(m)}) = \text{Gamma}(d'_\beta, f'_{\beta_i^{(m)}}) \quad (34)$$

$$d'_\beta = d_\beta + \frac{1}{2} \quad (35)$$

$$f'_{\beta_i^{(m)}} = f_\beta + \frac{1}{2}\left(g_i^{(m)} - \rho_i^{(m)}\right)^2 \quad (36)$$

4 Graph Learning Algorithm

In this section, we derive the posterior distribution for the adjacency matrix, based on the prior distribution presented in 13. After that the posterior of the adjacency matrix is determined, we can calculate the Laplacian matrix using $\mathbf{L} = \mathbf{D} - \mathbf{W}$. So, we have

$$\ln\left(q(w_{ij})\right) = \left\langle \ln p(\mathbf{x}(t)|\mathbf{L}) + \ln(p(w_{ij}))\right\rangle_{\mathbf{x},g_i^{(m)}} + \text{const.} = \left\langle \frac{1}{2}\ln(|\mathbf{B}|) - \frac{1}{2}\mathbf{r}(t)^T\mathbf{B}\mathbf{r}\right\rangle_{\mathbf{x}} + \ln(p(w_{ij})) + \text{const.}$$

Algorithm 1 Proposed VB-based graph signal inference algorithm

Input Observations y; **Sampling matrix** Ψ; max_iter; ϵ.

Initialize $\hat{\mathbf{x}} = \Psi^T\mathbf{y}, d_n, f_n, \rho_i^{(m)}, \beta_i^{(m)}, d_\beta, f_\beta, f_{\rho,i}^{(m)}, d_w, f_w, \gamma_w, t = 0$.

repeat

- Update the elements of Laplacian matrix w_{ij} $(1 \le i, j \le N)$ using (44);
- Update the expectation of posterior of the noise precision $q(\alpha_n)$ via $\langle\alpha_n\rangle = \frac{f'_n}{d'_n}$ and then set $\sigma_n^2 = \alpha_n^{-1}$
- Update the mean and covariance of signal using (17) and update the expectation of posterior of signal $q(\mathbf{x})$ by its mean
- Set $\hat{\mathbf{x}} = \mu$ and update $error = \|\mathbf{x} - \hat{\mathbf{x}}\|_2^2/\|\mathbf{x}\|_2^2$ and $t \leftarrow t + 1$

until $(error < \epsilon)$ or $(t \le max_iter)$

It has been shown in [10] that the determinant term is obtained from the following relation.

$$\det(\mathbf{B}) = \det(\mathbf{L} + \epsilon\mathbf{I}) = c_{ij}\left(w_{ij} - u_{ij}\right)^2 + s_{ij} \quad (37)$$

where c_{ij}, u_{ij} and s_{ij} are some constants. By substituting the above relation in $q(w_{ij})$, we have

$$\ln\left(q(w_{ij})\right) = \frac{1}{2}\ln\left(c_{ij}\left(w_{ij} - u_{ij}\right)^2 + s_{ij}\right) - \frac{1}{2}\left\langle \mathbf{r}(t)^T\mathbf{B}\mathbf{r}\right\rangle_{\mathbf{x}} + \ln\left(p(w_{ij})\right) + \text{const.} \quad (38)$$

Moreover, it can be proven that

$$\mathbf{r}(t)^T\mathbf{B}\mathbf{r} = \sum_{v_1=1}^{N}\sum_{v_2=1}^{N} r_{v_1}r_{v_2}\ell_{v_1v_2} = 2r_ir_j\ell_{ij} + C = -2r_ir_jw_{ij} + C \quad (39)$$

where r_i is the i-th element of the vector $\mathbf{r}$ and C is the part of the above term that is independent of ℓ_{ij}.

Using equation 13 with $\gamma_w = 2$, we have

$$q\left(w_{ij}\right) \propto \left(c_{ij}\left(w_{ij} - u_{ij}\right)^2 + s_{ij}\right)^{\frac{1}{2}} \times w_{ij}^{f_w - 1}\exp\left(r_ir_jw_{ij} - \frac{w_{ij}^2}{d_w^2}\right)$$

$$\propto \sqrt{c_{ij}} d_w \left(\left(\frac{w_{ij} - u_{ij}}{d_w} \right)^2 + \frac{s_{ij}}{c_{ij} d_w^2} \right)^{\frac{1}{2}} \times w_{ij}^{f_w - 1} \exp\left(- \left(\frac{w_{ij} - \frac{1}{2} r_i r_j d_w^2}{d_w} \right)^2 \right) \tag{40}$$

Assuming $d_w = \sqrt{\frac{2u_{ij}}{r_i r_j}}$ and $f_w = 1$, the above posterior distribution for w_{ij} belongs to a generalization of the modified generalized Gamma distribution (MGΓ) [22]. The MGΓ distribution is given by

$$p(w; \alpha, k, b, \lambda, \theta, \beta) = \frac{\beta}{\theta \Gamma_\lambda(\alpha, k, b)} \left(\frac{w}{\theta} \right)^{\alpha\beta - 1} \times \left(\left(\frac{w}{\theta} \right)^\beta + k \right)^{-\lambda} e^{-b\left(\frac{w}{\theta}\right)^\beta} \tag{41}$$

where $\Gamma(.)$ is the modified generalized Gamma function. In this work, we define a generalization for GMGΓ distribution as

$$p(w; u, \alpha, k, b, \lambda, \theta, \beta) = \frac{\beta}{\theta \Gamma_\lambda(\alpha, k, b)} \left(\frac{w - u}{\theta} \right)^{\alpha\beta - 1} \times \left(\left(\frac{w - u}{\theta} \right)^\beta + k \right)^{-\lambda} e^{-b\left(\frac{w-u}{\theta}\right)^\beta} \tag{42}$$

which is the same as the GMGΓ distribution introduced in [23] with $m = 1$. Thus the posterior distribution for w_{ij} is as follows

$$q\left(w_{ij}\right) = \mathrm{GMGG}(u_{ij}, 0.5, \frac{s_{ij}}{c_{ij} d_w^2}, 1, -0.5, d_w, 2) \tag{43}$$

with the mean value calculated as

$$< w_{ij} >= u_{ij} + \frac{d_w \Gamma_{-0.5}\left(1, \frac{s_{ij}}{c_{ij} d_w^2}, 1\right)}{\Gamma_{-0.5}\left(0.5, \frac{s_{ij}}{c_{ij} d_w^2}, 1\right)} \tag{44}$$

The complete variational Bayes (VB) procedure for graph signal recovery is summarized in Algorithm 1. As can be seen from the pseudocode, the proposed method operates by iteratively updating the model parameters and latent variables. At each iteration, the variational distributions are refined based on the current estimates, following the standard mean-field approximation framework. This iterative process continues until one of two stopping criteria is met: either the algorithm converges to a fixed point, meaning that the changes in the variational lower bound or the estimated parameters fall below a predefined tolerance, or the number of iterations surpasses a preset maximum. This design ensures both convergence stability and computational efficiency in practice.

5 Simulation results

In this section, we test the performance of our proposed algorithm for GSR problem. To evaluate the performance of the proposed algorithm, we examine two types of data: synthetic data and real data. For both data types, we compare the normalized mean square error (NMSE) metric for the proposed algorithm and three recently introduced GSR algorithms. This metric is defined as follows:

$$\mathrm{NMSE} = \frac{||\hat{\mathbf{x}} - \mathbf{x}||_2^2}{||\mathbf{x}||_2^2}.$$

5.1 Experiments on synthetic data

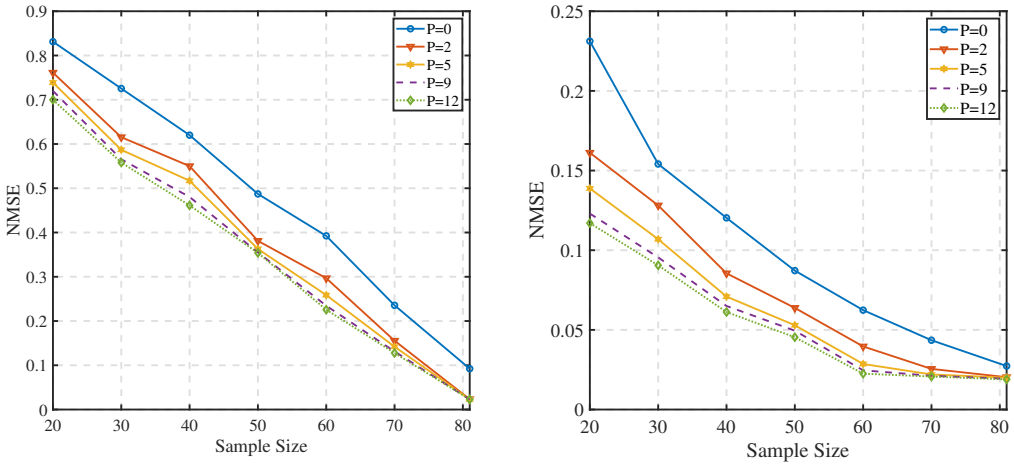
We first evaluate the performance of the proposed variational Bayes (VB)-based algorithm using a synthetic graph under a configuration similar to that in [10]. A network comprising $N = 81$ nodes is considered, constructed from the seed matrix

$$\mathbf{P}_0 = \begin{bmatrix} 0.6 & 0.1 & 0.7 \\ 0.3 & 0.1 & 0.5 \\ 0 & 1 & 0.1 \end{bmatrix}$$

and expanded to an $N \times N$ matrix via the Kronecker product $\mathbf{P} = \mathbf{P}_0 \otimes \mathbf{P}_0 \otimes \mathbf{P}_0 \otimes \mathbf{P}_0$. The weighted adjacency matrix $\mathbf{W}$ is generated by drawing each entry as $W_{ij} \sim \mathrm{Bernoulli}(P_{ij})$,, followed by symmetrization through adding its transpose. The Laplacian matrix is then computed as $\mathbf{L} = \mathrm{diag}(\mathbf{W}.\mathbf{1}_N) - \mathbf{W}$. Graph signals are synthesized using a bandlimited model:

$$\mathbf{y} = \sum_{i=1}^{\omega} \alpha^{(i)} \mathbf{r}^{(i)}$$

where $\alpha_k^{(i)} \sim \mathcal{N}(0, 1)$, ω denotes the bandwidth, and $\{\mathbf{r}^{(i)}\}_{i=1}^{\omega}$ are the eigenvectors associated with



(a) NMSE for graph signal (b) NMSE for Laplacian

Fig. 1: The NMSE of the GSR and graph learning for synthetic data ($N = 81$ and $SNR \sim 15dB$).

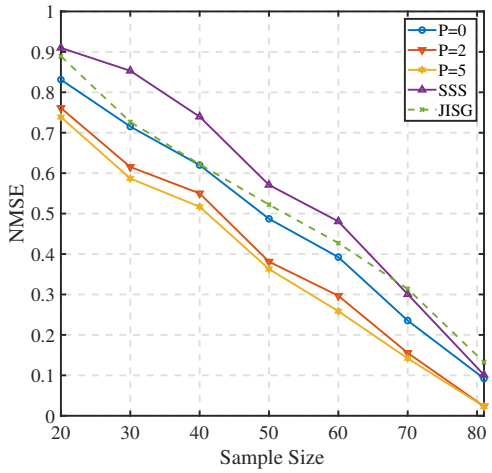


Fig. 2: The NMSE comparison for synthetic data.

the ω smallest eigenvalues of $\mathbf{L}$. Finally, Gaussian noise $\mathbf{n}$ with entries drawn from $\mathcal{N}(0, \sigma_n^2)$ is added to each signal such that the signal-to-noise ratio (SNR) is approximately 12dB.

The NMSE results for graph signal and Laplacian matrix recovery are shown in Fig. 1. In this figure, different numbers of time sequences are used for signal recovery and graph learning. $P = 0$ corresponds to the case where no recent graph signals are used to estimate the current signal, and the graph signal estimation is based solely on noisy, sub-sampled observations at the present time. In the other cases, $P = 2, 5, 9$, and 12 recent signals are used, along with noisy, sub-sampled observations, to estimate the signal at the current time. As can be seen by comparing the estimation error for $P = 0$ with those for other values of P, exploiting the temporal dependencies among signals to estimate the unknown signal significantly reduces the error compared to the case where only current observations are used. It is also evident that using a larger number of past signals leads to a lower estimation error. These results demonstrate that employing a model that simultaneously accounts for both temporal and spatial dependencies among samples enables more accurate recovery of the graph signal and the Laplacian matrix with lower error.

We have also compared the proposed algorithm with two other algorithms: JISG [16] and ESSS [24]. The results are ahown in Fig. 2. As can be observed, the proposed method, which exploits both temporal and spatial dependencies, can recover the graph signal with lower NMSE compared to the other methods.

5.2 Experiments on real-valued data

In this subsection, the performance of the proposed AR-VB-GSR algorithm is evaluated using real-valued data. This scenario involves large-scale graph signal data derived from global mean sea level pressure (MSLP) fields obtained from the

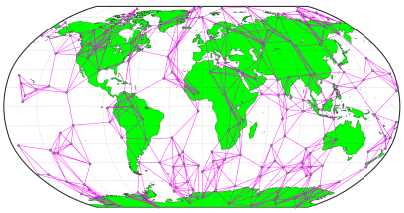


Fig. 3: The graph constructed by 6-nearest neighbors algorithm for the global mean sea level pressure dataset.

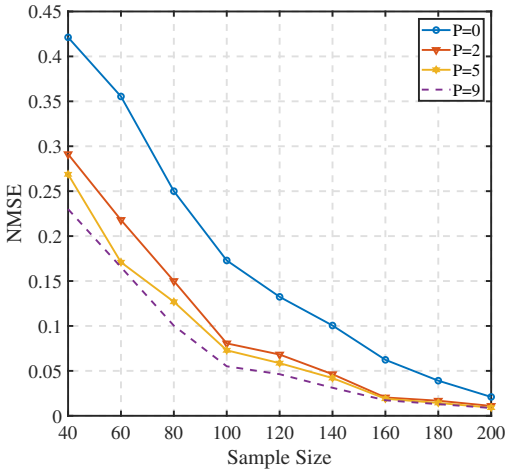


Fig. 4: The NMSE under different sampling sizes for the global mean sea level pressure dataset with $N = 200$ nodes.

Coupled Model Intercomparison Project (CMIP) archive [25] and distributed via the Earth System Grid Federation (ESGF) in NetCDF format. The original dataset comprises regularly spaced latitude–longitude grid measurements covering the entire globe. To construct the graph, $N = 500$ spatial locations were randomly subsampled from the global grid and treated as graph nodes. The graph topology was defined using a 6-nearest neighbors (kNN with $k = 6$) criterion based on the geodesic distances between node coordinates, as illustrated in Fig. 3. The atmospheric pressure values at the selected locations serve as graph signals, where each node represents a spatial sensor and the associated signal value denotes the sea level pressure at that location at a given time instant.

To examine the influence of sample size, the number of sampled nodes, M, was varied from 40 to 200 across the global sensor network. In all experiments, the signal-to-noise ratio (SNR) was about $9dB$. The results, presented in Fig. 4, consistently demonstrate that employing an appropriate spatio-temporal model leads to a reduction in the NMSE value, while underscoring its robustness and superior scalability for large-scale graph signal recovery tasks. Furthermore, since the time-dependency between consecutive time instants in real data is stronger than that in synthetic data, the reconstruction error in this case is significantly lower than that observed for synthetic data.

6 Conclusion

In this paper, a spatio-temporal statistical model was proposed for modeling graph signals by jointly capturing both temporal and spatial dependencies. Specifically, the temporal dependency between consecutive graph signals was modeled using a state-space autoregressive (SS-AR) model, while the spatial dependency among graph nodes was characterized by a Gaussian Markov random field (GMRF). The graph structure was assumed to be unknown, and a generalized Gamma prior was imposed on the elements of the adjacency matrix. A variational Bayes (VB) framework was developed to jointly reconstruct the graph signal and estimate both the SS-AR parameters and the adjacency matrix. Simulation results on both synthetic and real-world datasets demonstrate that the proposed algorithm consistently achieves lower reconstruction errors for both the graph signal and the Laplacian matrix than existing methods that rely solely on spatial dependencies.

Declarations

Funding

The authors received no financial support for the research, authorship, and/or publication of this article.

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 Trans. Signal Processing Magazine*, vol. 30, no. 3, pp. 83–98, May 2013.

[2] A. Ortega, P. Frossard, J. Kovacevic, J. F. Moura, and P. Vandergheynst, "Graph Signal Processing: Overview, Challanges, and Applications," *Proceedings of the IEEE*, vol. 106, pp. 808–828, 2018.

[3] S. Chen, A. Sandryhaila, J. M. F. Moura, and J. Kovacevic, "Signal Recovery on Graphs: Variation Minimization," *IEEE Transaction on Signal Processing.*, vol. 63, no. 17, pp. 4609–4624, Sep 2015.

[4] D. Romero, M. Ma, and G. B. Giannakis, "Kernel-based Reconstruction of Graph Signals," *IEEE Transaction on Signal Processing.*, vol. 65, no. 3, pp. 764–778, Feb 2017.

[5] X. P. Li, et al, "Robust Recovery for Graph Signal via l0-Norm Regularization," *IEEE Signal Processing Letters*, vol. 30, pp. 1322–1326, 2023.

[6] X. Dong, D. Thanou, P. Frossard, and P. Vandergheynst, "Learning Laplacian Matrix in Smooth Graph Signal Representations," *IEEE Transaction on Signal Processing.*, vol. 64, no. 23, pp. 6160–6173, Dec 2016.

[7] H. E. Egilmez, E. Pavez, and A. Ortega, "Graph Learning from Data under Laplacian and Structural Constraints," *IEEE Journal of Selected Topics in Signal Processing.*, vol. 11, no. 6, pp. 825–841, Sep 2017.

[8] R. Torkamani, H. Zayyani and M. Korki, "Proportionate adaptive graph signal recovery," *IEEE Trans. Signal Inf. Process. Netw.*, vol. 9, pp. 386–396, 2023.

[9] R. Torkamani, H. Zayyani, and F. Marvasti, "Joint topology learning and graph signal recovery using variational Bayes in non-gaussian noise," *IEEE Trans. Circuits Syst. II, Express Briefs*, vol. 69, no. 3, pp. 1887–1891, 2022.

[10] R. Torkamani, and H. Zayyani, "Statistical graph signal recovery using variational Bayes," *IEEE Trans. Circuit and Systems II: Express Briefs*, Vol. 68, no. 6, pp. 2232-2236, June 2021.

[11] A. H. Javaheri, et al, "Learning spatiotemporal graphical models from incomplete observations," *IEEE Trans. Signal Processing*, Vol. 72, pp. 1361-1374, 2024.

[12] X. Wang, J. Chen, and Y. Gu, "Local measurement and reconstruction for noisy bandlimited graph signals," *Elsevier Signal Processing.*, vol. 129, pp. 119–129, 2016.

[13] R. Torkamani, H. Zayyani, M. Korki, F. Marvasti "Robust Adaptive Generalized Correntropy-based Smoothed Graph Signal Recovery with a Kernel Width Learning", *Signal, Image, and Video Processing (SIVP).*, 2025.

[14] J. Mei, and J. F. Moura, "Signal Processing on Graphs: Causal Modeling of Unstructured Data," *IEEE Trans. Signal Processing*, Vol. 65, pp. 2077-2092, 2017.

[15] M. Ramezani-Mayiami, B. Beferull-Lozano, "Joint Topology Learning and Graph Signal Recovery via Kalman Filter in Causal Data Processes," *2018 IEEE 28th International Workshop on Machine Learning for Signal Processing (MLSP).*, Aalborg, Denmark, Sep 2018.

[16] V. N. Ioannidis, Y. Shen, and G. B Giannakis, "Semi-Blind Inference of Topologies and Dynamical Processes over Dynamic Graphs," *IEEE Transaction on Signal Processing.*, vol. 67, no. 9, pp. 2263–2274, May 2019.

[17] K. Donoghue, et al, "Learning Graph Structures With Autoregressive Graph Signal Models," *IEEE Open Journal of Signal Processing.*, vol. 6, pp. 838–855, 2025.

[18] E. Isufi, A. Loukas, N. Perraudin, and G. Leus, "Forecasting Time Series With VARMA Recursions on Graphs," *IEEE Transaction on Signal Processing.*, vol. 67, no. 18, pp. 4870–4885, Sep 2019.

[19] J. Hong, et al, "Multivariate Time Series Forecasting With GARCH Models on Graphs," *IEEE Trans. Signal Inf. Process. Netw.*, vol. 9, pp. 557–568, 2023.

[20] R. Torkamani, et al, "Bayesian Convolutional Neural Networks for Prior Learning in Graph Signal Recovery," *IEEE Trans. Signal Inf. Process. Netw.*, vol. 12, pp. 399–410, 2026.

[21] A. Kroizer, T. Rottenberg, and Y. Eldar, "Bayesian Estimation of Graph Signals," *IEEE Transaction on Signal Processing.*, vol. 70, pp. 2207–2223, 2012.

[22] M. Mead, M. Nassar, and S. Dey, "A Generalization of Generalized Gamma Distributions" *Pakistan Journal of Statistics and Operation Research*, February 2018.

[23] Z. Sadeghigol, et. al., "Simultaneous Spatiotemporal Modeling of Graph Signals with Bayesian Markov Models" *arxiv*

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

[25] Earth System Grid Federation, "Coupled Model Intercomparison Project Phase 5 (CMIP5) Data Archive," Lawrence Livermore National Laboratory, 2011–2023. [Online]. Available: https://cds.climate.copernicus.eu/datasets/projections-cmip5-daily-pressure-levels?tab=overview