Correlation-based Graph Smoothness Measures In Graph Signal Processing

Jari Miettinen, Sergiy A. Vorobyov, Esa Ollila, Xinjue Wang
Department of Information and Communications Engineering
Aalto University
Finland
{jari.p.miettinen, sergiy.vorobyov, esa.ollila, xinjue.wang}@aalto.fi

***Abstract*—Graph smoothness is an important prior used for designing sampling strategies for graph signals as well as for regularizing the problem of graph learning. Additionally, smoothness is an appropriate assumption for graph signal processing (GSP) tasks such as filtering or signal recovery from samples. The most popular measure of smoothness is the quadratic form of the Laplacian, which naturally follows from the factor analysis approach. This paper presents a novel smoothness measure based on the graph correlation. The proposed measure enhances the applicability of graph smoothness measures across a variety of GSP tasks, by facilitating interoperability and generalizing across shift operators.**

***Index Terms*—Graph signal processing, graph smoothness measures, graph autocovariance, graph autocorrelation**

I. Introduction

Graph signal processing (GSP) is an area of signal processing and data analysis that deals with structured data represented on graphs [1]–[3]. The application areas in which signals represented on graphs need to be processed are broad and include sensor networks, brain networks, gene regulatory networks, and social networks, to name a few [4]–[7]. By encoding the interaction between samples of data, graphs provide natural representations of data in irregular domains. This helps to improve the accuracy of data analysis. The structure of real-world data graphs varies and can take different forms, such as adjacency matrices, Laplacian matrices, or their normalized equivalents. Regardless of the above differences, specific signal processing tasks are addressed by algorithms of a unified nature. Therefore, in order to implement data-specific algorithms, an accurate understanding of how data structures are physically represented by graphs is critical.

Some graph structures, such as citation networks, are explicitly collected before processing; others, such as weather stations and electric power networks, are not known a priori. Therefore, such data topologies have to be pre-defined and learned in practice based on vertex's geometry, data similarity and association, and other criteria [8], [9], where the optimal sampling strategies also need to be found [10], [11]. The underlying graph structure for data can be fully specified by a graph adjacency matrix, denoted hereafter as $\mathbf{A}$, in which the (i,j)th element is nonzero if the ith and the jth data vertices are connected, and the corresponding value represents the strength of the relationship.

The problem of learning graphs from data samples is generally ill-posed. Pre-defined properties are usually required to associate signals with topologies. In a temperature-sensing network, the vertex distances encode physical properties, thus a k-nearest neighbor graph is constructed to detect malfunctioning sensors [12]; while [13] assumes the sparsity of the target graph and then learns the graph topology by optimization methods. In practice, one of the most important priors to impose on the data is the smoothness prior. Moreover, in GSP tasks such as filtering or signal recovery from samples, smoothness with respect to the chosen shift matrix is a key assumption as well because it acts as a regularizer in the corresponding problems [14]–[16]. In the vertex domain of a graph, smoothness can be described by the absence of abrupt changes in signal values. Specifically, a graph signal is smooth if the values in ith and jth vertices, which are connected by the edge (i,j), tend to be similar. It is often observed in real-world graph signals. For example, the records at close stations tend to be similar in a temperature-sensing network across a region [9].

The typical metric for signal smoothness is the quadratic form of Laplacian, $\mathbf{z}^{\mathsf{T}}\mathbf{L}\mathbf{z}$ [17], where $\mathbf{z}$ is the signal on a graph, which needs to be modeled such that it is close to actual observations, and $\mathbf{L} = \mathbf{D} - \mathbf{A}$ is the Laplacian matrix of a graph, which needs to be learned, $\mathbf{D}$ is the diagonal degree matrix with $d_{ii} = \sum_j a_{ij}$, and $(\cdot)^{\mathsf{T}}$ stands for the transpose. The quadratic form of Laplacian as a smoothness prior has been argued for in [9] as the one that naturally follows from the factor analysis approach, while applying Bayes' rule for finding a maximum a posteriori probability (MAP) estimate of the latent variables that control the graph signal. It has been demonstrated to work well in practice, but it is defined specifically for the Laplacian matrix shift operator.

In this paper, we develop smoothness measures for graph signals using a correlation-based approach as an alternative to the factor analysis-based approach. We consider the graph moving average model (GMA) [7], where the sampled signal is linearly parameterized by moving average coefficients and the set of powers of shift operators up to order m, that is, $\{\mathbf{A}, \ldots, \mathbf{A}^m\}$. Additionally, we assume that the underlying signal topology follows a certain probabilistic model. Thus, the topological statistical priors are inherited by the signal, thereby enhancing overall consistency. Our correlation-based

metric measures smoothness by *averaging* the weighted sum of covariances between the graph signal and its k-lagged counterpart. This average considers all possible graphs constrained by assumed model. For comparison, the quadratic Laplacian is based on only one graph representation. The proposed smoothness measure is also generic in the sense that it permits extension for any shift operator which can be expressed as a function of $\mathbf{A}$, including the Laplacian used in [9]. By specifying graph structure, this smoothness measure can be parametrized, and thus, it naturally leads to a simple interpretation (*cf.* Section 3). Our new smoothness measure is shown to be effective by applying mismatched shift operator to data and showing that the graph signal smoothness increases gradually as the mismatched shift operator approaches the true one. In other words, we show that as a mismatched shift operator deviates from the true one, the graph signal smoothness decreases gradually, which is captured by our smoothness measure.

II. Signal Model and Covariance Matrix

Let $\mathcal{G} = (\mathcal{N}, \mathcal{E})$ be a directed graph that represents the basis of a graph signal where $\mathcal{N}$ is the set of N vertices and $\mathcal{E}$ is the set of edges. The true adjacency matrix of the graph $\mathcal{G}$ is $\mathbf{A}$.

Let us consider the graph moving average (GMA) model as a more generic graph signal model as compared to just straightforward signal measurements. The GMA signal model of order m, denoted as GMA(m), is given as

$$\mathbf{z} = \mathbf{y} + \sum_{l=1}^{m} \theta_l \mathbf{A}^l \mathbf{y} \tag{1}$$

where $\mathbf{y} \triangleq [y_1, \ldots, y_N]^\mathsf{T}$ with $y_1, \ldots, y_N \sim N(0, \sigma^2)$ being mutually independent Gaussian random variables with zero mean and variance σ^2, and $\theta_1, \ldots, \theta_m$ are MA coefficients.

The same GMA model can be written in a number of forms, such as by adjusting the scales of θ's and $\mathbf{A}$ accordingly. Any GMA(m) model can be written as a GMA(1) model with $\sum_{l=1}^{m} \theta_l \mathbf{A}^l$ as its adjacency matrix. However, even if the adjacency matrix in GMA(m) presentation is unweighted, the adjacency matrix in the corresponding GMA(1) representation is weighted in general. Model (1) is a generalization of the traditional time series MA model, as the MA is obtained when $\mathbf{A}$ is the cycle graph that satisfies $a_{ij} = 1$, if $j = i - 1$, and $a_{ij} = 0$, otherwise.

Let us now derive some statistics of the graph signal given by (1). For simplicity and analytical tractability of the later studies, we consider the GMA(1) model given as

$$\mathbf{z} = \mathbf{y} + \theta \mathbf{A}\mathbf{y} = \tilde{\mathbf{A}}\mathbf{y} \tag{2}$$

where $\tilde{\mathbf{A}} \triangleq \mathbf{I}_{N\times N} + \theta \mathbf{A}$ and the matrix $\mathbf{I}_{N\times N}$ is an $N \times N$ identity matrix. Note that the value of the ith element of the graph signal $\mathbf{z}$ in (2) is given by

$$z_i = y_i + \theta \sum_{j \in \mathcal{N}_i} a_{ij} y_j \tag{3}$$

where $\mathcal{N}_i$ denotes the set of incoming neighbors of vertex i. Thus, if $\theta \neq 0$, two vertices are correlated if they are neighbors or if they have shared incoming neighbors.

The covariance matrix of the graph signal (2) as a function of $\tilde{\mathbf{A}}$ can be expressed as

$$\mathbf{C}_z(\tilde{\mathbf{A}}) \triangleq \mathbb{E}\left\{\mathbf{z}\mathbf{z}^\mathsf{T}\right\} = \mathbb{E}\left\{\tilde{\mathbf{A}}\mathbf{y}\mathbf{y}^\mathsf{T}\tilde{\mathbf{A}}^\mathsf{T}\right\} = \tilde{\mathbf{A}}\mathbb{E}\left\{\mathbf{y}\mathbf{y}^\mathsf{T}\right\}\tilde{\mathbf{A}}^\mathsf{T}$$

where $\mathbb{E}\{\cdot\}$ stands for the mathematical expectation. Using the fact that the elements of $\mathbf{y}$ are uncorrelated, i.e., $\mathbb{E}\{\mathbf{y}\mathbf{y}^\mathsf{T}\} = \sigma^2\mathbf{I}_{N\times N}$, and substituting the expression for $\tilde{\mathbf{A}}$, the covariance matrix $\mathbf{C}_z$ can be written as a function of the adjacency matrix $\mathbf{A}$ as

$$\mathbf{C}_z(\mathbf{A}) = \sigma^2\tilde{\mathbf{A}}\tilde{\mathbf{A}}^\mathsf{T} = \sigma^2\left(\mathbf{I}_{N\times N} + \theta\left(\mathbf{A} + \mathbf{A}^\mathsf{T}\right) + \theta^2\mathbf{A}\mathbf{A}^\mathsf{T}\right). \tag{4}$$

The covariance matrix for the general GMA(m) in (1) can be obtained by substituting the expression $\mathbf{I}_{N\times N} + \sum_{l=1}^{m}\theta_l\mathbf{A}^l$ instead of $\tilde{\mathbf{A}}$ into (4). It is important to note that for graph signals represented on unweighted graphs (that is when the entries of $\mathbf{A}$ take only zero and one values), the diagonal elements of the matrix $\mathbf{A}\mathbf{A}^\mathsf{T}$ in (4) give the numbers of outcoming neighbors for the vertices. Moreover, the (i,j)th element of $\mathbf{A}\mathbf{A}^\mathsf{T}$ for $i \neq j$, i.e., $[\mathbf{A}\mathbf{A}^\mathsf{T}]_{i,j}$, $i \neq j$, is the number of mutual outcoming neighbors of the ith and the jth vertices.

To enable a direct analysis of the correlation-based graph smoothness metrics that we introduce in the next section, let us introduce errors in the adjacency matrix as [18]

$$\mathbf{W} = \mathbf{A} + \mathbf{E} \tag{5}$$

where $\mathbf{W}$ presents the estimated (inaccurate) adjacency matrix and $\mathbf{E}$ is an unknown error matrix. The covariance matrix of the graph signal (2) with inaccurately learned graph adjacency matrix modeled as in (5) can then be derived by simply substituting the mismatched adjacency matrix $\mathbf{W}$ instead of the actual one $\mathbf{A}$, yielding

$$\mathbf{C}_z(\mathbf{W}) = \sigma^2\left(\mathbf{I}_{N\times N} + \theta\left(\mathbf{W} + \mathbf{W}^\mathsf{T}\right) + \theta^2\mathbf{W}\mathbf{W}^\mathsf{T}\right). \tag{6}$$

Using (5), the covariance matrix (6) can be now rewritten as a function of the true adjacency matrix $\mathbf{A}$ and the mismatched matrix $\mathbf{E}$, which is assumed to be fixed. Then we get the covariance matrix

$$\begin{aligned}\mathbf{C}_z(\mathbf{W}) &= \mathbf{C}_z(\mathbf{A}) + \sigma^2\theta(\mathbf{E} + \mathbf{E}^\mathsf{T} + \theta(\mathbf{A}\mathbf{E}^\mathsf{T} + \mathbf{E}\mathbf{A}^\mathsf{T} + \mathbf{E}\mathbf{E}^\mathsf{T})) \\ &= \mathbf{C}_z(\mathbf{A}) + \mathbf{C}_z(\mathbf{E}) + \sigma^2\left(\theta^2\left(\mathbf{A}\mathbf{E}^\mathsf{T} + \mathbf{E}\mathbf{A}^\mathsf{T}\right) - \mathbf{I}_{N\times N}\right).\end{aligned} \tag{7}$$

It can be seen from (7) that the covariance matrix of the graph signal (2) with inaccurately learned $\mathbf{A}$ is not only the summation of the covariance matrices of $\mathbf{A}$ and $\mathbf{E}$, but also contains the third cross-correlation component $\sigma^2\left(\theta^2\left(\mathbf{A}\mathbf{E}^\mathsf{T} + \mathbf{E}\mathbf{A}^\mathsf{T}\right) - \mathbf{I}_{N\times N}\right)$, which is scaled by the signal variance σ^2 and the square of coefficient θ.

Consider the unweighted graph error model [18]

$$\mathbf{W} = \mathbf{A} - \mathbf{\Delta}_{\epsilon_1} \odot \mathbf{A} + \mathbf{\Delta}_{\epsilon_2} \odot (\mathbf{1}_{N\times N} - \mathbf{A}) \tag{8}$$

where $\mathbf{\Delta}_\epsilon$ is the adjacency matrix of the Erdös–Rényi random graph, $\odot$ stands for the Hadamard product, $\mathbf{1}_{N\times N}$ is the $N\times N$ matrix full of ones. The Erdös–Rényi graph is $\mathcal{G}=(\mathcal{N},\epsilon)$ with random $\mathbf{\Delta}_\epsilon$ such that $\mathbb{P}([\mathbf{\Delta}_\epsilon]_{i,j}=1)=\epsilon$ and $\mathbb{P}([\mathbf{\Delta}_\epsilon]_{i,j}=0)=1-\epsilon$ for all $i\neq j$, and $[\mathbf{\Delta}_\epsilon]_{i,i}=0$ for $i=1,\ldots,N$, where each element of the matrix is generated independently of the other elements [19], and $\mathbb{P}(\cdot)$ denotes the probability. The parameter ϵ_1 (resp. ϵ_2) then describes the probability of edge removal (resp. addition).

III. Correlation-Based Graph Smoothness Metrics

First, we introduce graph autocorrelation as a measure of smoothness, and then we generalize it for multivariate signals to introduce graph autocorrelation matrix.

We start with the autocovariance of a graph signal. Let us first define the centering matrix $\mathbf{H}\triangleq\mathbf{I}_{N\times N}-\mathbf{1}_{N\times N}/N$. The matrix $\mathbf{H}$ is symmetric and idempotent. Centering the graph signal $\mathbf{Hz}$ removes the signal's mean, and thus, forces the first moment to be 0 [20]. The graph signal autocovariance of lag k with respect to $\mathbf{W}$ is then defined as

$$s_{\mathbf{z},k}(\mathbf{W})\triangleq\mathbb{E}\left\{\frac{1}{N-k}(\mathbf{Hz})^{\mathsf{T}}\mathbf{W}^k\mathbf{Hz}\right\}. \tag{9}$$

Substituting the graph signal model (2) into (9), we obtain

$$\begin{aligned}s_{\mathbf{z},k}(\mathbf{W})&=\frac{1}{N-k}\mathbb{E}\left\{\mathbf{y}^{\mathsf{T}}\tilde{\mathbf{A}}^{\mathsf{T}}\mathbf{H}\mathbf{W}^k\mathbf{H}\tilde{\mathbf{A}}\mathbf{y}\right\}\\&=\frac{\sigma^2}{N-k}\mathrm{tr}\left\{\left(\mathbf{H}+\theta\mathbf{A}^{\mathsf{T}}\mathbf{H}\right)\mathbf{W}^k\left(\mathbf{H}+\theta\mathbf{H}\mathbf{A}\right)\right\},\end{aligned} \tag{10}$$

where $\mathrm{tr}\{\cdot\}$ stands for the trace of a matrix. Here the property $\mathrm{tr}\{\mathbf{AB}\}=\mathrm{tr}\{\mathbf{BA}\}$ has been used. Thus, graph autocovariance is a weighted sum of the covariances between the vertices of a graph signal.

The graph autocorrelation of lag k as function of $\mathbf{W}$ is defined as

$$r_{\mathbf{z},k}(\mathbf{W})=\frac{\mathbb{E}\left\{(\mathbf{Hz})^{\mathsf{T}}\mathbf{W}^k\mathbf{Hz}\right\}}{\left(\mathbb{E}\{\|\mathbf{Hz}\|^2\}\mathbb{E}\{\|\mathbf{HW}^k\mathbf{z}\|^2\}\right)^{1/2}} \tag{11}$$

where $\|\cdot\|$ stands for the Euclidean norm of a vector.

A. Autocorrelation of Inaccurate Adjacency Matrix

If model (8) is adopted for inaccurate adjacency matrix $\mathbf{W}$, then the graph autocorrelation (11) naturally depends on the parameters ϵ_1 and ϵ_2. For simplicity, assume that $\mathbf{A}=\mathbf{\Delta}_\alpha$, i.e., true adjacency matrix $\mathbf{A}$ also follows the Erdös-Rényi model with the probability parameter α. The nonzero values of $\mathbf{A}$ and $\mathbf{W}$ are denoted by a and w, respectively.

Then for graph autocorrelation $r_{\mathbf{z},k}(\mathbf{W})$ of signal (2), we would like to derive the expected value $\mathbb{E}_{\mathbf{A},\mathbf{W},\mathbf{z}}\{r_{\mathbf{z},1}(\mathbf{W})\}$ for the lag $k=1$, where both $\mathbf{A}$ and $\mathbf{W}$ are considered to be random. Even though this expected value is an average over multiple implementations of Erdös-Rényi graphs $\mathbf{A}$, we have verified by extensive simulations that the graph autocorrelation values are very similar for any specific and fixed realization of $\mathbf{A}$. The expected graph autocorrelation value can be then expressed in terms of the GMA coefficient θ, the probability parameter α in the graph model $\mathbf{A}=\mathbf{\Delta}_\alpha$, and the parameters ϵ_1 and ϵ_2 in the graph error model in (8). The expected value $\mathbb{E}_{\mathbf{A},\mathbf{W},\mathbf{z}}\{r_{\mathbf{z},1}(\mathbf{W})\}$ can be used for example to verify that GMA signal is smoother with respect to its adjacency matrix than the mismatched adjacency matrix. Specifically, the larger the graph autocorrelation term in (11), the smoother the graph signal. More details are illustrated in Section IV.

Now we can derive the expected value of the graph autocorrelation for graph signal (2) and graph error model (8). Let us start with rescaling the matrices $\mathbf{A}$ and $\mathbf{W}$ and the variance σ^2 so that

$$\begin{aligned}\tfrac{1}{N}\mathbb{E}\{\|\mathbf{Hz}\|^2\}&=1\\\tfrac{1}{N}\mathbb{E}\{\|\mathbf{HAz}\|^2\}&=1\\\tfrac{1}{N}\mathbb{E}\{\|\mathbf{HWz}\|^2\}&=1.\end{aligned} \tag{12}$$

For given values of N, α, ϵ_1 and ϵ_2, the only unknowns in (12) are a, w and σ^2. Hence, we can find a, w and σ^2 by simply solving the system of three equations (12). To calculate expected values in (12), we need the fact that $\mathbb{E}\{y_iy_j\}=\sigma^2$, if $i=j$ and $\mathbb{E}\{y_iy_j\}=0$, otherwise; in conjunction with the assumption that the elements of the matrices $\mathbf{\Delta}_\alpha$, $\mathbf{\Delta}_{\epsilon_1}$ and $\mathbf{\Delta}_{\epsilon_2}$ are statistically independent both within each matrix and between them.

The expectation in the first equation in (12) is approximated as follows. We start by substituting (2) into the expectation in the first equation in (12), and opening the square

$$\begin{aligned}&\mathbb{E}\{\|\mathbf{Hz}\|^2\}=\mathbb{E}\{\|\mathbf{H}\left(\theta\mathbf{Ay}+\mathbf{y}\right)\|^2\}\\&=\mathbb{E}\left\{\sum_{i=1}^N\left(\theta\sum_{k=1}^N a_{ik}y_k+y_i-\frac{1}{N}\sum_{j=1}^N\left(\theta\sum_{l=1}^N a_{jl}y_l-y_j\right)\right)^2\right\}\\&=\mathbb{E}\left\{\sum_{i=1}^N\left(\theta\sum_{k=1}^N a_{ik}y_k+y_i-\frac{1}{N}\sum_{j=1}^N\left(\theta\sum_{l=1}^N a_{jl}y_l-y_j\right)\right)\right.\\&\left.\times\left(\theta\sum_{k'=1}^N a_{ik'}y_{k'}+y_i-\frac{1}{N}\sum_{j'=1}^N\left(\theta\sum_{l'=1}^N a_{j'l'}y_{l'}-y_j\right)\right)\right\}.\end{aligned} \tag{13}$$

After algebraic manipulations of (13), we obtain

$$\begin{aligned}&\mathbb{E}\{\|\mathbf{Hz}\|^2\}\\&=\mathbb{E}\left\{\sum_{i=1}^N\left(\theta^2\sum_{k=1}^N a_{ik}^2y_k^2-\frac{2}{N}\sum_{j=1}^N\left(\theta^2\sum_{k=1}^N a_{ik}a_{jk}y_k^2-\theta a_{ij}y_j^2\right)\right.\right.\\&+y_i^2-\frac{2}{N}\sum_{j=1}^N\left(\theta a_{ji}y_i^2-y_iy_j\right)+\frac{\theta^2}{N^2}\sum_{j=1,j'=1,l=1}^N a_{jl}a_{j'l}y_l^2\\&\left.\left.+\frac{2\theta}{N^2}\sum_{j=1,l=1}^N a_{jl}y_l^2+\frac{1}{N^2}\sum_{j=1}^N y_j^2\right)\right\}\\&=\sigma^2\left(N^2\theta^2\alpha a^2-2N\theta^2\alpha a^2-2N^2\theta^2\alpha^2a^2-2N\theta\alpha a+N\right.\\&\left.-2N\theta\alpha a-2+N\theta^2\alpha a^2+N^2\theta^2\alpha^2a^2+2N\theta\alpha a+1\right).\end{aligned} \tag{14}$$

Dropping the negligible terms in (14), i.e., the zero-order terms of N and first-order terms of N that also include a^2, gives the approximation

$$\frac{1}{N}\mathbb{E}\{\|\mathbf{Hz}\|^2\} \approx \sigma^2\left((\alpha-\alpha^2)\theta^2a^2N - 2\alpha\theta a + 1\right). \quad (15)$$

The derivations of the expected values of the latter two equations in (12) are tedious, but otherwise follow the same steps. Thus, we skip these derivations for the sake of brevity. The first two equations in (15) do not depend on w. Therefore, a and σ^2 can be found from the first two equations,

$$\begin{cases}\sigma^2((\alpha-\alpha^2)\theta^2a^2N - 2\alpha\theta a + 1) = 1,\\ \sigma^2((\alpha^2-\alpha^3)\theta^2a^4N^2 - 2\alpha^2\theta a^3N + (\alpha-\alpha^2)a^2N) = 1.\end{cases}$$

After solving for a and σ^2, the parameter w can be found from the last equation in (12) as

$$\begin{aligned} w = \big(&\sigma^2\left(-\alpha^3N^2\theta^2a^2(\epsilon_1+\epsilon_2-1)^2 + \alpha N^2\theta^2a^2(\epsilon_2-\epsilon_2^2)\right.\\ &+\alpha^2N^2\theta^2a^2(1-\epsilon_1+\epsilon_2(2(\epsilon_1+\epsilon_2)-3))\\ &+\alpha^2N(\theta^2a^2(\epsilon_1+\epsilon_2-1)-(\epsilon_1+\epsilon_2-1)^2)\\ &+\alpha N(1-\epsilon_1+\epsilon_2(2(\epsilon_1+\epsilon_2)-\theta^2a^2-3))\\ &\left.+\ N(\epsilon_2-\epsilon_2^2)+\alpha(\epsilon_1+\epsilon_2-1)-\epsilon_2\big)\right)^{-1/2}. \end{aligned}$$

Finally, for the mismatched matrix $\mathbf{W}$ in (8), a closed-form approximate expression for the expected autocorrelation value of the graph signal $\mathbf{z}$ as a function of α, θ, ϵ_1 and ϵ_2 is

$$\begin{aligned}&\mathbb{E}_{\mathbf{A},\mathbf{W},\mathbf{z}}\{r_{\mathbf{z},1}(\mathbf{W})\} \approx\\ &-w\epsilon_2 + \sigma^2(\theta\, a\, w\, N(\alpha-\alpha^2) - \alpha\, w)(1-\epsilon_1-\epsilon_2).\end{aligned} \quad (16)$$

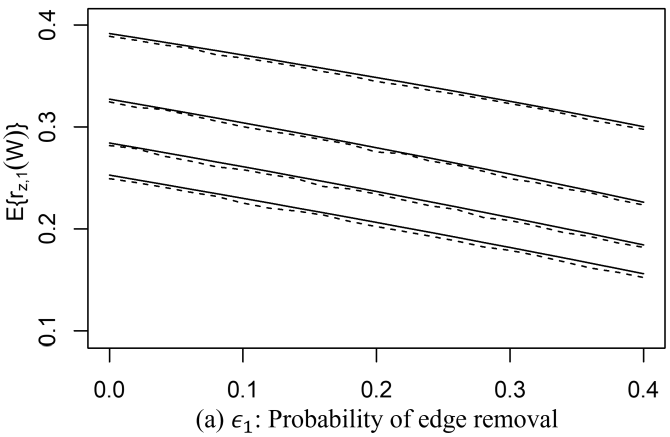
IV. Simulation Example

In our simulation example, we examine how the expected graph autocorrelation derived in (16) changes when the matrix with respect to which it is computed is gradually changing from the adjacency matrix of the signal according to model (8). We generate graph signal $\mathbf{z}$ as a GMA(1) signal with unweighted Erdös-Rényi adjacency matrix $\mathbf{A} = \mathbf{\Delta}_\alpha$ with probability $\alpha = 0.05$, and the GMA coefficient is $\theta = 0.5$. Furthermore, we compute 2000 graph autocorrelations. For each repetition, a new Erdös-Rényi adjacency matrix $\mathbf{A}$ is generated. The graph size is $N = 500$.

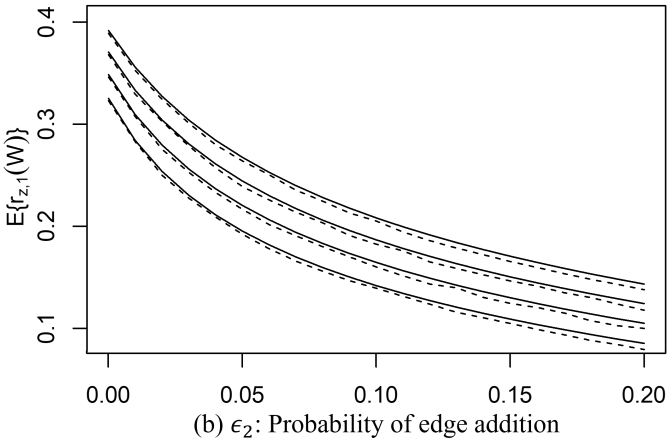
Fig. 1 displays the theoretical and simulated graph autocorrelation as functions of ϵ_1 and ϵ_2, respectively, while the other one remains fixed. The figures show that the graph autocorrelation is decreasing with respect to both ϵ_1 and ϵ_2, which hints that graph autocorrelation might be useful to estimate the structure of the graph even if we only have a single realization of the graph data.

V. Conclusion and Discussion

In this paper, a new correlation-based graph smoothness measure is introduced for graph signals as an alternative to the quadratic form of Laplacian that follows from the factor analysis approach. The usefulness of this correlation-based graph smoothness measure is also demonstrated by a synthetic example.



(a) ϵ_1: Probability of edge removal



(b) ϵ_2: Probability of edge addition

Fig. 1: (a). Theoretical (solid) and simulated (dash) values of graph autocorrelations. The lines from top to bottom are given by $\epsilon_2 = 0, 0.02, 0.04, 0.06$; (b). The lines from top to bottom are given by $\epsilon_1 = 0, 0.1, 0.2, 0.3$.

In terms of other examples, it is worth pointing out that similar, but preliminary ideas of graph autocorrelation has been used for multivariate graph signals in the context of blind source separation (BSS) of graph signals [7]. Namely, let $\mathbf{X} \in \mathbb{R}^{P\times N}$ denote centered P-dimensional graph signal generated as a mixture of independent components according to the model $\mathbf{X} = \mathbf{\Omega Z}$, where $\mathbf{\Omega} \in \mathbb{R}^{P\times P}$ is a full rank mixing matrix, $\mathbf{Z} \in \mathbb{R}^{P\times N}$ is the matrix of P mutually independent graph signals with zero means and unit variances. The goal of graph-BSS is to estimate the unmixing matrix $\mathbf{\Gamma} = \mathbf{\Omega}^{-1}$ using only the signal matrix $\mathbf{X}$. Then, letting $\mathbf{X}_\mathrm{w} = \hat{\mathbf{S}}_0^{-1/2}\mathbf{X}$ be the whitened signals, where $\hat{\mathbf{S}}_0$ is the sample covariance matrix of $\mathbf{X}$, the measure of smoothness used for estimating the unmixing matrix $\mathbf{\Omega}^{-1}$ is the autocorrelation matrix or the set of autocorrelation matrices given by

$$\hat{\mathbf{S}}_k(\mathbf{W}) = \frac{1}{N-k}(\mathbf{X}_\mathrm{w}\mathbf{W}^k\mathbf{X}_\mathrm{w}^\mathsf{T}), \quad k = 1,\dots,K. \quad (17)$$

Then (17) can be diagonalized or jointly diagonalized, for example, as in Grade method [7] (see also [18]) to find independent components. Thus, this measure can be considered as instantaneous (single sample) multivariate extension of the ideas presented in this paper.

REFERENCES

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

[2] D. I. Shuman, S. K. Narang, P. Frossard, A. Ortega, and P. Vandergheynst, "The emerging field of signal processing on graphs: Extending high-dimensional data analysis to networks and other irregular domains," *IEEE Signal Process. Mag.*, vol. 30, no. 3, pp. 83–98, Apr. 2013.

[3] A. Ortega, P. Frossard, J. Kovacevic, J. M. F. Moura, and P. Vandergheynst, "Graph signal processing: Overview, challenges, and applications," *Proc. of the IEEE*, vol. 106, no. 5, pp. 808–828, Apr. 2018.

[4] S.-Y. Tu and A. H. Sayed, "Mobile adaptive networks," *IEEE J. Sel. Topics Signal Process.*, vol. 5, no. 4, pp. 649–664, Aug. 2011.

[5] D. Acemoglu, G. Como, F. Fagnani, and A. Ozdaglar, "Opinion fluctuations and disagreement in social networks," *Math. Oper. Res.*, vol. 38, no. 1, pp. 1–27, Jan. 2013.

[6] W. Huang, L. Goldsberry, N. F. Wymbs, S. T. Grafton, D. S. Bassett, and A. Ribeiro, "Graph frequency analysis of brain signals," *IEEE J. Sel. Topics Signal Process.*, vol. 10, no. 7, pp. 1189–1203, Oct. 2016.

[7] F. Blöchl, A. Kowarsch, and F. J. Theis, "Second-order source separation based on prior knowledge realized in a graph model," in *Proc. 9th Int. Conf. Latent Var. Anal. Signal Sep.*, St. Malo, France, Sep. 27-30, 2010, pp. 434–441.

[8] L. Stanković, D. P. Mandic, M. Daković, M. Brajović, B. Scalzo, S. Li, and A. G. Constantinides, "Data analytics on graphs part III: Machine learning on graphs, from graph topology to applications," *Foundations and Trends® in Mach. Learn.*, vol. 13, no. 4, pp. 332–530, Dec. 2020.

[9] X. Dong, D. Thanou, P. Frossard, and P. Vandergheynst, "Learning laplacian matrix in smooth graph signal representations," *IEEE Trans. Signal Process.*, vol. 64, no. 23, pp. 6160–6173, Dec. 2016.

[10] S. P. Chepuri and G. Leus, "Graph sampling for covariance estimation," *IEEE Trans. Signal Process.*, vol. 3, no. 3, pp. 451–466, July. 2017.

[11] P. A. Traganitis, Y. Shen, and G. B. Giannakis, "Network toplogy inference via elastic net structural equation models," in *Proc. 25th Eur. Signal Process. Conf.*, Kos island, Greece, Aug. 28 - Sep. 2, 2017, pp. 156–160.

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

[13] B. Lake and J. Tenenbaum, "Discovering structure by learning sparse graphs," in *Proc. 32nd Annu. Conf. Cognit. Sci. Soc.*, Portland, USA, Aug. 11-14, 2010, pp. 1–6.

[14] E. Isufi, A. Loukas, A. Simonetto, and G. Leus, "Autoregressive moving average graph filtering," *IEEE Trans. Signal Process.*, vol. 65, no. 2, pp. 274–288, Jan. 2017.

[15] ——, "Filtering random graph processes over random time-varying graphs," *IEEE Trans. Signal Process.*, vol. 65, no. 16, pp. 4406–4421, Aug. 2017.

[16] S. Segarra, A. G. Marques, G. Leus, and A. Ribeiro, "Reconstruction of graph signals through percolation from seeding nodes," *IEEE Trans. Signal Process.*, vol. 64, no. 16, pp. 4363–4378, Aug. 2016.

[17] D. Zhou and B. Schölkopf, "A regularization framework for learning from graph data," in *Proc. 21st Int. Conf. Mach. Learn. Workshop*, Banff, Alberta, Canada, July 4-8, 2004, pp. 132–137.

[18] J. Miettinen, S. A. Vorobyov, and E. Ollila, "Modelling and studying the effect of graph errors in graph signal processing," *Signal Process.*, vol. 189, 108256, pp. 1–8, Dec. 2021.

[19] P. Erdös and A. Rényi, "On random graphs I." *Publicationes Mathematicae (Debrecen)*, vol. 6, pp. 290–297, 1959.

[20] N. Perraudin and P. Vandergheynst, "Stationary signal processing on graphs," *IEEE Trans. Signal Process.*, vol. 65, no. 13, pp. 3462–3477, July. 2017.