NETWORK TOPOLOGY CHANGE-POINT DETECTION FROM GRAPH SIGNALS WITH PRIOR SPECTRAL SIGNATURES

Chiraag Kaushik, T. Mitchell Roddenberry, and Santiago Segarra

Rice University, Electrical and Computer Engineering, Houston, TX

ABSTRACT

We consider the problem of sequential graph topology change-point detection from graph signals. We assume that signals on the nodes of the graph are regularized by the underlying graph structure via a graph filtering model, which we then leverage to distill the graph topology change-point detection problem to a subspace detection problem. We demonstrate how prior information on the spectral signature of the post-change graph can be incorporated to implicitly denoise the observed sequential data, thus leading to a natural CUSUM-based algorithm for change-point detection. Numerical experiments illustrate the performance of our proposed approach, particularly underscoring the benefits of (potentially noisy) prior information.

Index Terms— Graph signal processing, change-point detection, network topology.

1. INTRODUCTION

Networks or graphs have emerged as effective tools to understand and summarize complex systems across multiple domains of knowledge [1–3]. Representing interconnected systems as graphs, where nodes correspond to agents and edges correspond to pairwise interactions between these agents, allows us to apply tools from graph theory to reveal key properties of the underlying systems such as the emergence of community structure [4, 5] or the existence of influential agents [6, 7].

In practice, one rarely has immediate access to the whole network of interest and rather must sample it, either by directly querying the connections between pairs of nodes in a strategic fashion [8], or performing network topology inference [9–13]. Once a network has been constructed, it can be used to gain insights about the system being modeled through a variety of downstream tasks. However, if the network is subject to change, *e.g.*, due to some external force influencing the propensity for connections between particular nodes, then it is in the user's interest to resample the network. Since constructing a network can be a burdensome task [9], it is important to minimize the resampling frequency.

Although it is often infeasible to constantly resample a network to assess whether its topology has changed, we frequently have access to data supported on the nodes of the graph. For instance, we might not be able to frequently query the connections within a social group but we can observe the preferences and actions of individuals. Moreover, this data is typically regularized by the underlying graph structure, *e.g.*, the data might be smooth on the underlying graph [13]. In our social group example, the preferences of individuals are partially driven by the unobserved social connections. Thus,

if a nominal graph is known, a sequence of sampled data on the nodes should reflect this nominal topology up to the point where the structure changes. At that point, the relationship between the nodal data and the nominal graph structure will deteriorate.

Working in the framework of graph signal processing, we aim to detect changes in the graph structure underlying a sequence of graph signals. By modeling nodal data as graph signals output by graph filters, we can extract dominant features of the underlying graph solely by observing the signals. We then leverage domain knowledge on the graph change model to propose a CUSUM-based algorithm for change-point detection.

Related work. The problem of detecting changes in a sequence of graphs has been studied from several perspectives. When a sequence of random graphs is directly available for computation, spectral methods have been designed for detecting changes in the underlying random graph model [14, 15]. However, these methods do not consider the case where only data on the nodes of a graph is available. A different body of work has focused on the detection of changes in the distribution of signals on a graph. For instance, [16, 17] study the detection of a mean change in a sequence of graph signals and [18] focuses on the change of the distribution of graph signals by leveraging the community structure of the underlying graph. However, these approaches do not aim to detect changes in graph topology but, rather, they detect changes in the model of the data supported on a static graph.

Most related to our work, detecting changes in graph topology from graph signals has also been studied. In [19], matched subspace detectors are considered for the non-sequential case, where assumptions of bandlimitedness in the graph Fourier domain are leveraged. Similarly, in [20], the generating process for the data is assumed to be known, unlike our work here which only leverages the principal eigenvectors of the data, instead of a precise distribution.

Contributions. Our contributions in this paper are twofold:
i) We present a CUSUM algorithm for network topology change-point detection that only depends on the observation of graph signals and that can incorporate prior knowledge about the nature of the change.
ii) Through numerical experiments, we demonstrate the utility of even crude knowledge of the post-change graph for denoising observed data and improving detection performance.

2. NOTATION AND BACKGROUND

Notation. The notation $[n]$ refers to the set of positive integers $\{1, 2, \ldots, n\}$. We refer to matrices using bold uppercase letters, *e.g.* $\mathbf{A}, \mathbf{B}, \mathbf{C}$, and to (column) vectors with bold lowercase letters, *e.g.* $\mathbf{v}, \mathbf{w}, \mathbf{x}$. Entries of matrix $\mathbf{A}$ are indicated by A_{ij} while those of vector $\mathbf{x}$ are denoted by x_i. For clarity, we alternatively use the notation $[\mathbf{x}]_i = x_i$. The ℓ_2-norm of a vector is denoted by $\|\cdot\|_2$.

ICASSP 2021

Graphs and graph matrices. A *graph* is a finite set of nodes, coupled with a set of edges connecting those nodes. That is, for a set of n nodes denoted by $\mathcal{V}$, the set of edges $\mathcal{E} \subseteq \mathcal{V} \times \mathcal{V}$ forms the graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$. Typically, we endow the set of nodes with arbitrarily ordered integer labels, saying that $\mathcal{V} = [n]$. This representation allows us to represent the graph with the *adjacency matrix* $\mathbf{A} \in \mathbb{R}^{n \times n}$, where $A_{ij} = 1$ if $(i, j) \in \mathcal{E}$, taking value 0 otherwise.

In this paper, we focus on *undirected* graphs where $(i, j) \in \mathcal{E}$ implies that $(j, i) \in \mathcal{E}$, *i.e.*, $\mathcal{E}$ is composed of unordered pairs of nodes. Under this condition, the adjacency matrix is symmetric. With such an adjacency matrix, we define the *Laplacian matrix* as $\mathbf{L} = \mathbf{D} - \mathbf{A}$, where $\mathbf{D} = \mathrm{diag}(\mathbf{A1})$ is the diagonal matrix of node degrees. The adjacency matrix and the Laplacian are two examples of *graph shift operators* (generically denoted by $\mathbf{S}$), which are matrices whose sparsity and symmetry patterns correspond to those of the underlying graph [21].

Graph signals and graph filters. We model data on the nodes of a graph as a *graph signal*. A graph signal is a real-valued function supported on the nodes of a graph $x : \mathcal{V} \to \mathbb{R}$ that can be conveniently represented as a vector $\mathbf{x}$ in $\mathbb{R}^n$, where $[\mathbf{x}]_i = x(i)$ for each $i \in [n]$. We now define the notion of *graph filters* as linear maps between graph signals. Assuming that a given graph shift operator $\mathbf{S}$ has the eigenvalue decomposition $\mathbf{S} = \sum_i \lambda_i \mathbf{v}_i \mathbf{v}_i^\top$, a graph filter $\mathcal{H}(\mathbf{S})$ is a real polynomial of $\mathbf{S}$, *e.g.* for coefficients α_k,

$$\mathcal{H}(\mathbf{S}) = \sum_{k=0}^{T} \alpha_k \mathbf{S}^k = \sum_{i=1}^{n} h(\lambda_i) \mathbf{v}_i \mathbf{v}_i^\top, \tag{1}$$

where $h : \mathbb{R} \to \mathbb{R}$ is the extension of $\mathcal{H}$ to the real numbers. Graph filters have shown to be versatile tools in modeling linear network processes [22]. Notice that graph filters preserve the eigenvectors of the underlying graph shift operator, while distorting its eigenvalues.

3. NETWORK CHANGE-POINT DETECTION

Consider a streaming sequence of graphs with a single change point. More precisely, we have a nominal graph $\mathcal{G}_0 = (\mathcal{V}, \mathcal{E}_0)$, a post-change graph $\mathcal{G}_1 = (\mathcal{V}, \mathcal{E}_1)$, and a sequence of graphs $\{\mathcal{G}^{(t)}\}_{t=1}^{\infty}$ modeled as

$$\begin{aligned} \mathcal{G}^{(t)} &= \mathcal{G}_0, \quad t < \tau, \\ \mathcal{G}^{(t)} &= \mathcal{G}_1, \quad t \geq \tau. \end{aligned} \tag{2}$$

In (2), the graphs in the given sequence are equal to $\mathcal{G}_0$ before the unknown change point $t = \tau$ and, after the change point, they are equal to $\mathcal{G}_1$.

Our objective is to develop a method for detecting this change without direct observations of the graphs $\mathcal{G}^{(t)}$ or precise knowledge of the post-change graph $\mathcal{G}_1$. Specifically, we model the observations available to us as a sequence of graph signals given by

$$\mathbf{y}^{(t)} = \begin{cases} \mathcal{H}_0(\mathbf{S}_0)\, \mathbf{w}^{(t)} & \text{for } t < \tau, \\ \mathcal{H}_1(\mathbf{S}_1)\, \mathbf{w}^{(t)} & \text{for } t \geq \tau, \end{cases} \tag{3}$$

where $\mathbf{S}_0$ and $\mathbf{S}_1$ are graph shift operators of the graphs $\mathcal{G}_0$ and $\mathcal{G}_1$ in (2), respectively. Furthermore, we consider the challenging case where the graph filters $\mathcal{H}_0$ and $\mathcal{H}_1$ in (3) have unknown coefficients, and the unknown inputs $\mathbf{w}^{(t)}$ are independent random vectors drawn from a common centered distribution with identity covariance. With this model in place, we now formally state the problem of interest.

Problem 1. *Given a streaming sequence of graph signals modeled by (3), detect at time t whether or not $t \geq \tau$.*

Problem 1 prompts us to decide, in real time, whether the change in the underlying graph has already occurred or not. This is motivated by the desire to detect graph topology changes when the correlation structure of the observed data is determined by interactions between neighboring nodes, *e.g.* to detect structural changes in social networks from opinion data over time, or anomalies in sensor networks from nodal measurements.

In order to solve Problem 1, we must be able to extract key features of the underlying graphs from the observed signals $\mathbf{y}^{(t)}$. In particular, a direct computation of the covariance $\mathbf{C}_{\mathbf{y}}^{(t)}$ of $\mathbf{y}^{(t)}$ – leveraging (1) and the identity covariance of the inputs $\mathbf{w}^{(t)}$ – reveals that $\mathbf{C}_{\mathbf{y}}^{(t)}$ preserves the eigenvectors of the underlying graph shift operators at time t. More precisely, if the shift operators have eigenvalue decompositions $\mathbf{S}_0 = \sum_i \lambda_i \mathbf{v}_i \mathbf{v}_i^\top$ and $\mathbf{S}_1 = \sum_i \mu_i \mathbf{u}_i \mathbf{u}_i^\top$, then

$$\mathbf{C}_{\mathbf{y}}^{(t)} = \begin{cases} \mathbf{C}_{\mathbf{y}}^0 := \sum_{i=1}^{n} h_0^2(\lambda_i) \mathbf{v}_i \mathbf{v}_i^\top & \text{for } t < \tau, \\ \mathbf{C}_{\mathbf{y}}^1 := \sum_{i=1}^{n} h_1^2(\mu_i) \mathbf{u}_i \mathbf{u}_i^\top & \text{for } t \geq \tau, \end{cases} \tag{4}$$

where h_0 and h_1 are the respective extensions of the matrix polynomials $\mathcal{H}_0$ and $\mathcal{H}_1$ to the real numbers. It can be seen from (4) that we can estimate the eigenvectors of the shift operator at time t by taking the sample covariance matrix of graph signals near that time.

Following this direction and in order to formulate a tractable hypothesis test for the solution of Problem 1, we restrict our view to the *dominant subspaces* of the covariance matrices $\mathbf{C}_{\mathbf{y}}^0$ and $\mathbf{C}_{\mathbf{y}}^1$ in (4). Thus, we translate our assumed knowledge of $\mathcal{G}_0$ and lack of precise knowledge of $\mathcal{G}_1$ into statements about these dominant subspaces. In particular, we assume that the leading subspace of $\mathbf{C}_{\mathbf{Y}}^0$ is known and denoted by $\mathbf{U}_0$, and that the leading subspace of $\mathbf{C}_{\mathbf{y}}^1$ is known to be an element of *a family of subspaces* $\{\mathbf{U}_1(\gamma)\}_{\gamma \in \Gamma}$, for some set Γ. Notice that this assumption of a parameterized post-change subspace is realistic, *e.g.*, in a social network under a polarization effect, γ could dictate how the nodes split into communities, where $\mathbf{U}_1(\gamma)$ is the dominant subspace of a stochastic block model parameterized by γ [5]. For simplicity, we assume that all the dominant subspaces considered are k-dimensional. Then, by partitioning the observed sequence of graph signals $\{\mathbf{y}^{(t)}\}_{t=1}^{\infty}$ into consecutive blocks of size b, we form estimates of the current dominant subspace by taking the top k eigenvectors of the sample covariance matrix of each block. We denote these subspaces by $\widehat{\mathbf{V}}^{(\ell)}$, where ℓ indicates the block under consideration. This naturally leads to the following hypothesis test for change-point detection

$$\begin{aligned} H_0 : & \quad \mathbf{V}^{(t)} = \mathbf{U}_0, \\ H_1 : & \quad \mathbf{V}^{(t)} = \mathbf{U}_1(\gamma) \text{ for some } \gamma \in \Gamma, \end{aligned} \tag{5}$$

where $\mathbf{V}^{(t)}$ is the true dominant subspace of the covariance matrix at time t, estimated by the corresponding $\widehat{\mathbf{V}}^{(\ell)}$. The form of (5) is tractable in the setting where the graphs are only observed via data on their nodes, allowing us to leverage sequential subspace detection methods [23, 24].

With this formulation in place, we propose Algorithm 1 to test the hypothesis (5) by applying a CUSUM-based [25, 26] method to our sequence of estimated subspaces $\{\widehat{\mathbf{V}}^{(\ell)}\}_{\ell=1}^{\infty}$. For each ℓ we determine the subspace in $\{\mathbf{U}_1(\gamma)\}_{\gamma \in \Gamma}$ that is closest to $\widehat{\mathbf{V}}^{(\ell)}$, where we measure the distance between subspaces using the Frobenius sin Θ distance [27] (see line 4 in Algorithm 1). We denote this closest subspace as $\mathbf{U}_1(\widehat{\gamma}^{(\ell)})$. We then compare the distances from the observed $\widehat{\mathbf{V}}^{(\ell)}$ to $\mathbf{U}_1(\widehat{\gamma}^{(\ell)})$ and from the nominal $\mathbf{U}_0$ to $\mathbf{U}_1(\widehat{\gamma}^{(\ell)})$ within a conventional CUSUM framework. In a nutshell, if the distance between the nominal subspace and $\mathbf{U}_1(\widehat{\gamma}^{(\ell)})$ is significantly

Algorithm 1 Network change-point detection with prior parameterized post-change subspaces

---

1: **Input:** Sequence of estimated dominant subspaces $\{\widehat{\mathbf{V}}^{(\ell)}\}_{\ell=1}^{\infty}$
2: **Parameters:** Nominal subspace $\mathbf{U}_0$, family of post-change subspaces $\{\mathbf{U}_1(\gamma)\}_{\gamma \in \Gamma}$
3: **Parameters:** CUSUM correction parameter c, decision threshold η
4: Let $d(\mathbf{V}, \mathbf{U}) = \|\sin\Theta(\mathbf{V}, \mathbf{U})\|_F$ for subspaces $\mathbf{V}, \mathbf{U}$
5: $S^{(0)} \leftarrow 0, \ell \leftarrow 1$
6: **while** $S^{(\ell-1)} < \eta$ **do**
7: $\quad \widehat{\gamma}^{(\ell)} \leftarrow \arg\min_{\gamma \in \Gamma} d(\mathbf{U}_1(\gamma), \widehat{\mathbf{V}}^{(\ell)})$
8: $\quad L^{(\ell)} \leftarrow d(\mathbf{U}_0, \mathbf{U}_1(\widehat{\gamma}^{(\ell)})) - d(\mathbf{U}_1(\widehat{\gamma}^{(\ell)}), \widehat{\mathbf{V}}^{(\ell)})) - c$
9: $\quad S^{(\ell)} \leftarrow \max(0, S^{(\ell-1)} + L^{(\ell)})$
10: $\quad \ell \leftarrow \ell + 1$
11: **end while**
12: **Output:** change detected

---

larger – where the level of significance is determined by the correction parameter c – than the distance between the observed subspace and $\mathbf{U}_1(\widehat{\gamma}^{(\ell)})$, we have reason to suspect that a topology change has occurred and we increase the value of our CUSUM statistic. This approach is inspired by the approach of [28], where the shift in the mean of a multivariate Gaussian random variable is assumed to be in a known direction, and the most likely scale of that shift is estimated at each step.

Remark 1. The choice of post-change subspaces $\{\mathbf{U}_1(\gamma)\}_{\gamma \in \Gamma}$ confers a high level of flexibility to Algorithm 1. Indeed, in the absence of any information about the post-change graph $\mathcal{G}_1$, one might adopt the most flexible setting, where $\{\mathbf{U}_1(\gamma)\}_{\gamma \in \Gamma}$ contains *all* k-dimensional subspaces of $\mathbb{R}^n$. Under this setting, $\widehat{\mathbf{V}}^{(\ell)} = \mathbf{U}_1(\widehat{\gamma}^{(\ell)})$ for all ℓ, so that Algorithm 1 only relies on the distance between $\mathbf{U}_0$ and $\widehat{\mathbf{V}}^{(\ell)}$ at each step. That is, $L^{(\ell)} = d(\mathbf{U}_0, \widehat{\mathbf{V}}^{(\ell)}) - c$, thus boiling down to a more conventional CUSUM statistic for detecting deviations in any direction from an expected behavior.

The effect of the correction parameter c. The performance of Algorithm 1 is dictated by c and the quality of the estimated dominant subspaces $\{\widehat{\mathbf{V}}^{(\ell)}\}_{\ell=1}^{\infty}$. Given the difficulty of precisely characterizing the distribution of the estimated subspaces, Algorithm 1 deviates from the standard formulation of CUSUM for Gaussian random variables based on likelihood (or generalized likelihood) ratios [29]. Nonetheless, we follow the provably-valid strategy of substituting the likelihood ratio with a general function with negative expected value in the nominal state and positive expected value after the change [30]. For this to hold, we require the following condition relating our estimates, our candidate post-change subspaces, and the correction parameter c,

$$\mathbb{E}_0\left[d_{\mathbf{U}_0, \mathbf{U}_1(\widehat{\gamma}^{(\ell)}), \widehat{\mathbf{V}}^{(\ell)}}\right] < c < \mathbb{E}_1\left[d_{\mathbf{U}_0, \mathbf{U}_1(\widehat{\gamma}^{(\ell)}), \widehat{\mathbf{V}}^{(\ell)}}\right], \quad (6)$$

where $d_{\mathbf{U}_0, \mathbf{U}_1(\widehat{\gamma}^{(\ell)}), \widehat{\mathbf{V}}^{(\ell)}} := d(\mathbf{U}_0, \mathbf{U}_1(\widehat{\gamma}^{(\ell)})) - d(\widehat{\mathbf{V}}^{(\ell)}, \mathbf{U}_1(\widehat{\gamma}^{(\ell)}))$, and $\mathbb{E}_0$ and $\mathbb{E}_1$ refer to the nominal and post-change expectations, respectively. We illustrate the geometry of this requirement in Fig. 1 for the $k = 1$ dimensional case. The red circle around $\mathbf{U}_0$ indicates the set of 1-dimensional estimated subspaces such that $d(\mathbf{U}_0, \widehat{\mathbf{V}}^{(\ell)}) - c < 0$, as in the case where $\{\mathbf{U}_1(\gamma)\}_{\gamma \in \Gamma}$ contains all 1-dimensional subspaces of $\mathbb{R}^n$ (cf. Remark 1). In this setting, all estimated subspaces beyond this radius obey $d(\mathbf{U}_0, \widehat{\mathbf{V}}^{(\ell)}) - c > 0$.
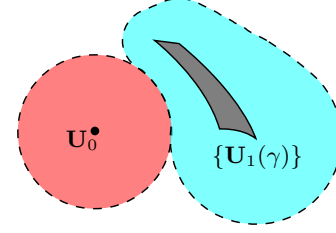
Fig. 1. Geometry of $L^{(\ell)}$ for two different choices of the post-change subspaces $\{\mathbf{U}_1(\gamma)\}_{\gamma \in \Gamma}$.

However, when there is prior information about the post-change subspaces (as in the gray region $\{\mathbf{U}_1(\gamma)\}$), only the cyan region yields $d(\mathbf{U}_0, \mathbf{U}_1(\widehat{\gamma}^{(\ell)})) - d(\widehat{\mathbf{V}}^{(\ell)}, \mathbf{U}_1(\widehat{\gamma}^{(\ell)})) - c > 0$ in order to contribute to the CUSUM statistic. In a sense, only subspace estimates that are pointed in the right direction from $\mathbf{U}_0$ to $\{\mathbf{U}_1(\gamma)\}_{\gamma \in \Gamma}$ contribute to the CUSUM statistic, rather than all subspaces beyond the open ball drawn in red. Intuitively, fulfilling the requirement in (6) amounts to choosing the parameter c such that most of the estimated nominal subspaces lie in the open ball around $\mathbf{U}_0$ and most of the estimated post-change subspaces lie in the region around $\{\mathbf{U}_1(\gamma)\}_{\gamma \in \Gamma}$.

4. EXPERIMENTS

Algorithm 1 relies on a parameterization of the dominant subspace of the post-change graph in terms of $\gamma \in \Gamma$. In practice, however, the spectral properties of the post-change graph may only be modeled roughly. In this section, we demonstrate how even the use of an *approximate* parameterization of the post-change subspace can improve CUSUM-based subspace change-point detection algorithms[1]. In the synthetic experiments, both graph filters in (3) are equal to the adjacency matrix squared, *i.e.* $\mathcal{H}_0(\mathbf{A}_0) = \mathbf{A}_0^2, \mathcal{H}_1(\mathbf{A}_1) = \mathbf{A}_1^2$, with $\mathbf{w}$ being drawn from a standard normal multivariate Gaussian distribution. Under this model, the dominant subspaces of the nominal and post-change covariance matrices respectively correspond to the eigenvectors of $\mathbf{A}_0$ and $\mathbf{A}_1$ with the highest-magnitude eigenvalues.

Detecting concentrations of centrality. We consider a model where the nominal graph is an Erdős-Rényi (ER) graph on $n = 100$ nodes with density parameter $p = 2\log n/n$. The post-change graph is a randomly drawn Barabási-Albert (BA) network on the same set of nodes with parameters $m_0 = m = 1$. In terms of eigenvector centralities, one expects the pre-change centrality to be flat across nodes, in accordance with the expectation of the ER graph, and the post-change centrality to be concentrated in a few nodes, in accordance with the power-law structure of the BA random graph model [31].

We focus on single dimensional leading subspaces for our change-point detection and we do not assume that the high-centrality nodes in the post-change graph are known. Rather, we crudely approximate the post-change structure of the leading eigenvector (centrality) with a spike: a discrete delta function $\mathbf{u}_1(\gamma) = \delta_\gamma$, for $\gamma \in [n]$. Evaluating the distances in Algorithm 1 boils down to finding the maximum magnitude element of the estimated leading eigenvector, and then comparing to the same element in the nominal eigenvector. Formally, if we have that $\widehat{\gamma}^{(\ell)} = \arg\max_{\gamma \in [n]} |[\widehat{\mathbf{v}}^{(\ell)}]_\gamma|$ then

$$
\begin{aligned}
d(\widehat{\mathbf{v}}^{(\ell)}, \mathbf{u}_1(\widehat{\gamma}^{(\ell)})) &= \sqrt{1 - \|\widehat{\mathbf{v}}^{(\ell)}\|_\infty^2}, \\
d(\mathbf{u}_0, \mathbf{u}_1(\widehat{\gamma}^{(\ell)})) &= \sqrt{1 - [\mathbf{u}_0]_{\widehat{\gamma}^{(\ell)}}^2},
\end{aligned}
\quad (7)
$$

---

[1]Code at github.com/chiraagk7/graph-cusum
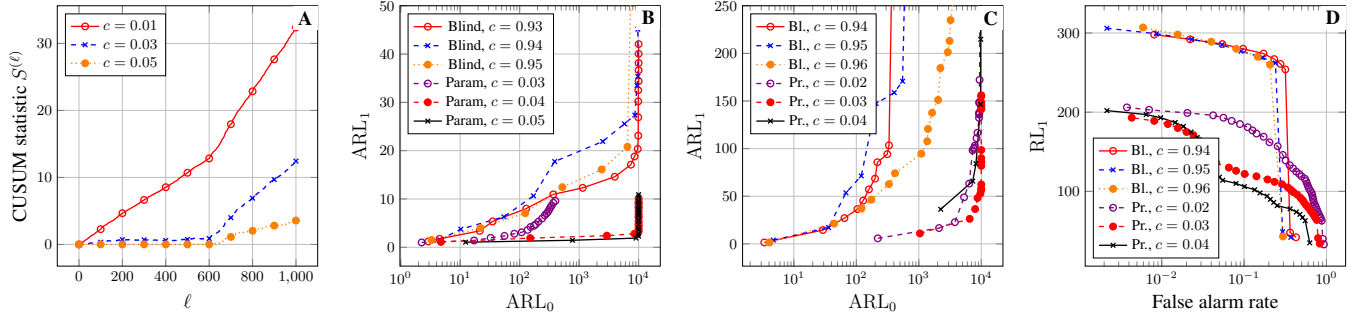
Fig. 2. **(A)** CUSUM statistic for detecting the change of an ER graph into a BA graph. **(B)** Average run length before (ARL_0) and after (ARL_1) the change for detection of a BA graph from an ER graph. **(C)** Average run length before and after the change for detection of an emerging community. **(D)** Detection of change points in the market from mentions on Twitter. False alarm rate for detection of changes before March 9th 2015 compared to the run length after the morning of March 9th.

where $\widehat{\mathbf{v}}^{(\ell)}$ is the estimated leading eigenvector, $\mathbf{u}_1(\gamma)$ is a vector taking value 1 at index γ and 0 elsewhere, and $\mathbf{u}_0$ is the leading eigenvector of the nominal ER graph.

We illustrate the case where the block size is $b = 1$, *i.e.*, the estimate of the leading eigenvector $\widehat{\mathbf{v}}^{(\ell)}$ is given simply by the observed graph signal at time $t = \ell$, scaled to have unit norm. The CUSUM statistics $S^{(\ell)}$ of this experiment are shown in Fig. 2A, for varying correction parameters c on a sequence of 1000 signals with a change point at $\tau = 600$. When the correction parameter is small ($c = 0.01$), the left inequality in (6) is not satisfied. This leads to an increase in the CUSUM statistic even before the change point. However, when c is sufficiently large, the nature of the CUSUM statistic is apparent. It remains close to zero in the nominal state, and then increases with slope inversely proportional to c after the change point.

We compare our method using the parameterization of the post-change subspace with a delta function with a method using no knowledge of the post-change subspace, where $\{\mathbf{u}_1(\gamma)\}_{\gamma \in \Gamma}$ is the set of all 1-dimensional subspaces of $\mathbb{R}^n$, as discussed in Remark 1. We generate 10000 signals in the nominal state, as well as 10000 post-change signals, and compare both methods for varying correction parameters c, plotting the expected nominal and post-change run lengths for a range of decision thresholds η, as in [24, Section 5]. As illustrated in Fig. 2B, for a fixed average run length in the nominal state (ARL_0), the parameterized approach detects the change more quickly (ARL_1) than the blind approach, where no prior information is used. Although the post-change subspace will never take the form of a delta function, this result shows that even a rough estimate of how the post-change graph is structured can be leveraged to improve change-point detection algorithms.

Detecting emerging communities. To illustrate our approach on a more complex subspace problem, let the nominal graph again be an ER graph of $n = 100$ nodes with density parameter $p = 2 \log n/n$. For some unknown cutoff $1 \leq n_0 \leq n$, we model the post-change graph as identical to the nominal graph but having the edges (i, j) for $i, j \leq n_0$ redrawn with density parameter $q = 5p$.

This models the emergence of a community dictated by strong adherence to a particular ideology, *e.g.*, a faction of the most politically extreme individuals in a social network forming a dense community, where the sorted labeling of the nodes implicitly places nodes on the political spectrum. In this setting, our nominal subspace hypothesis compares the estimated subspace to the leading $k = 2$ eigenvectors of the nominal graph, while the post-change hypothesis compares the estimated subspace to the $k = 2$ leading eigenvectors of the expected post-change adjacency matrix, param-

eterized by $\gamma = n_0$. We consider signal blocks of size $b = 50$ in estimating the leading subspaces. As shown in Fig. 2C, using the parameterized (Pr.) model of the post-change subspace is effective in denoising the observed signals, yielding $\mathrm{ARL}_0/\mathrm{ARL}_1$ curves that are below those of the case where no knowledge of the post-change subspace is assumed (Bl.).

Detecting events in Twitter data. To illustrate the utility of this approach on real data, we consider a dataset recording daily mentions of 10 companies on Twitter, where the signals count the number of mentions of each company every 5 minutes. We determine the nominal dominant eigenvector from the sample covariance of the first 4 days of data, where we have implicitly assumed the existence of an underlying network structure between these companies. We aim to detect sudden events related to public activity of a company over the next 10 days. Our post-change hypothesis is similar to our first experiment, where we model the post-change dominant subspace as a discrete delta function. To estimate the subspaces $\widehat{\mathbf{v}}^{(\ell)}$, we use a sliding window of width $b = 36$ samples, corresponding to 3-hour periods of time. As a ground-truth change point, we consider the morning of March 9th, 2015, which corresponds to a special event from Apple. We compare our parameterization with a delta function to the case where the post-change subspace is not parameterized. In Fig. 2D, we compare the false alarm rate, defined as the proportion of time before the ground-truth change point that the CUSUM statistic $S^{(\ell)}$ is above the threshold, to the number of samples after the change point needed for detection. By leveraging our model of the post-change subspace, we achieve earlier detection (RL_1) than the case where no model is used. This shows that Algorithm 1 has applications beyond the case where a precise network is known, since it only relies on having a model for the shape of the post-change dominant eigenvectors in the data.

5. CONCLUSION

We have considered a graph change-point detection problem, where we detect a change in the graph without directly observing its topology. Instead, we have access to data on the nodes, which is shaped by the underlying graph via some unknown network process. Through the lens of graph signal processing, we frame this as a subspace detection problem, and then propose an algorithm allowing one to incorporate prior knowledge on the structure of the post-change graph. We demonstrate how this prior domain knowledge can improve CUSUM-based detection algorithms on synthetic and real data, even when the precise post-change structure is not known.

6. REFERENCES

[1] S. H. Strogatz, "Exploring complex networks," *Nature*, vol. 410, no. 6825, pp. 268–276, Mar. 2001.

[2] M. E. J. Newman, *Networks: An Introduction*, Oxford University Press, USA, Mar. 2010.

[3] M. O. Jackson, *Social and Economic Networks*, Princeton university press, 2010.

[4] S. Fortunato, "Community detection in graphs," *Phys. Rep.*, vol. 486, no. 3, pp. 75 – 174, 2010.

[5] E. Abbe, "Community detection and stochastic block models: recent developments," *The Journal of Machine Learning Research*, vol. 18, no. 1, pp. 6446–6531, 2017.

[6] L. Page, S. Brin, R. Motwani, and T. Winograd, "The PageRank citation ranking: Bringing order to the web," Technical Report 1999-66, Stanford InfoLab, Nov. 1999.

[7] S. Segarra and A. Ribeiro, "Stability and continuity of centrality measures in weighted graphs," *IEEE Trans. Signal Process.*, vol. 64, no. 3, pp. 543–555, 2016.

[8] N. K. Ahmed, J. Neville, and R. Kompella, "Network sampling: From static to streaming graphs," *ACM Trans. Disc. Knowl. Data (TKDD)*, vol. 8, no. 2, Jun. 2013.

[9] G. Mateos, S. Segarra, A. G. Marques, and A. Ribeiro, "Connecting the dots: Identifying network structure via graph signal processing," *IEEE Signal Process. Mag.*, vol. 36, no. 3, pp. 16–43, 2019.

[10] S. Segarra, A. G. Marques, G. Mateos, and A. Ribeiro, "Network topology inference from spectral templates," *IEEE Trans. Signal Inf. Process. Netw.*, vol. 3, no. 3, pp. 467–483, 2017.

[11] R. Shafipour, S. Segarra, A. G. Marques, and G. Mateos, "Network topology inference from non-stationary graph signals," in *IEEE Intl. Conf. Acoust., Speech and Signal Process. (ICASSP)*. IEEE, 2017, pp. 5870–5874.

[12] 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, 2016.

[13] V. Kalofolias, "How to learn a graph from smooth signals," in *Intl. Conf. Artif. Intel. Stat. (AISTATS)*, 2016, pp. 920–929.

[14] M. Zhang, L. Xie, and Y. Xie, "Online community detection by spectral CUSUM," in *IEEE Intl. Conf. Acoust., Speech and Signal Process. (ICASSP)*. IEEE, 2020, pp. 3402–3406.

[15] X. He, Y. Xie, S.-M. Wu, and F.-C. Lin, "Sequential graph scanning statistic for change-point detection," in *Asilomar Conf. on Signals, Systems, and Computers*. IEEE, 2018, pp. 1317–1321.

[16] J. Sharpnack, A. Rinaldo, and A. Singh, "Detecting anomalous activity on networks with the graph fourier scan statistic," *IEEE Trans. Signal Process.*, vol. 64, no. 2, pp. 364–379, 2015.

[17] A. Ferrari, C. Richard, and L. Verduci, "Distributed change detection in streaming graph signals," in *IEEE Intl. Wrksp. Computat. Advances Multi-Sensor Adaptive Process. (CAMSAP)*. IEEE, 2019, pp. 166–170.

[18] A. Ferrari and C. Richard, "Non-parametric community change-points detection in streaming graph signals," in *IEEE Intl. Conf. Acoust., Speech and Signal Process. (ICASSP)*. IEEE, 2020, pp. 5545–5549.

[19] E. Isufi, A. S. Mahabir, and G. Leus, "Blind graph topology change detection," *IEEE Signal Processing Letters*, vol. 25, no. 5, pp. 655–659, 2018.

[20] S. P. Chepuri and G. Leus, "Subgraph detection using graph signals," in *2016 50th Asilomar Conference on Signals, Systems and Computers*. IEEE, 2016, pp. 532–534.

[21] 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.

[22] S. Segarra, A. G. Marques, and A. Ribeiro, "Optimal graph-filter design and applications to distributed linear network operators," *IEEE Trans. Signal Process.*, vol. 65, no. 15, pp. 4117–4131, 2017.

[23] L. Xie, G. V. Moustakides, and Y. Xie, "First-order optimal sequential subspace change-point detection," in *IEEE Global Conf. Signal and Info. Process. (GlobalSIP)*. IEEE, 2018, pp. 111–115.

[24] Y. Jiao, Y. Chen, and Y. Gu, "Subspace change-point detection: A new model and solution," *IEEE J. Sel. Topics Signal Process.*, vol. 12, no. 6, pp. 1224–1239, 2018.

[25] E. S. Page, "Continuous inspection schemes," *Biometrika*, vol. 41, no. 1, pp. 100–115, 1954.

[26] J. J. Pignatiello Jr and G. C. Runger, "Comparisons of multivariate CUSUM charts," *Journal of quality technology*, vol. 22, no. 3, pp. 173–186, 1990.

[27] Y.-C. Wong, "Differential geometry of grassmann manifolds," *Proc. of the National Academy of Sciences*, vol. 57, no. 3, pp. 589, 1967.

[28] J. Chen, H. Yang, and J. Yao, "A new multivariate CUSUM chart using principal components with a revision of crosier's chart," *Communications in Statistics-Simulation and Computation*, vol. 47, no. 2, pp. 464–476, 2018.

[29] J. D. Healy, "A note on multivariate CUSUM procedures," *Technometrics*, vol. 29, no. 4, pp. 409–412, 1987.

[30] T. Oskiper and H. V. Poor, "Quickest detection of a random signal in background noise using a sensor array," *EURASIP Journal on Adv. in Signal Process.*, vol. 2005, no. 1, pp. 360150, 2005.

[31] I. J. Farkas, I. Derényi, A.-L. Barabási, and T. Vicsek, "Spectra of "real-world" graphs: Beyond the semicircle law," *Physical Review E*, vol. 64, no. 2, pp. 026704, 2001.