Blind Graph Topology Change Detection

Isufi, Elvin; Mahabir, Ashvant S.U.; Leus, Geert

DOI
10.1109/LSP.2018.2819127

Publication date
2018

Document Version
Final published version

Published in
IEEE Signal Processing Letters

Blind Graph Topology Change Detection

Elvin Isufi, Ashvant S. U. Mahabir, and Geert Leus

***Abstract*—This letter investigates methods to detect graph topological changes without making any assumption on the nature of the change itself. To accomplish this, we merge recently developed tools in graph signal processing with matched subspace detection theory and propose two blind topology change detectors. The first detector exploits the prior information that the observed signal is sparse w.r.t. the graph Fourier transform of the nominal graph, while the second makes use of the smoothness prior w.r.t. the nominal graph to detect topological changes. Both detectors are compared with their respective nonblind counterparts in a synthetic scenario that mimics brain networks. The absence of information about the alternative graph, in some cases, might heavily influence the blind detector's performance. However, in cases where the observed signal deviates slightly from the nonblind model, the information about the alternative graph turns out to be not useful.**

***Index Terms*—Anomalous subgraph detection, brain networks, graph detection, graph signal processing, matched subspace detection.**

I. Introduction

MATCHED subspace detection (MSD) [1] is a standard tool used for outlier detection in multidimensional data. By representing the data in their intrinsic low-dimensional subspace, MSD has achieved promising results in applications such as radar [2], communication [3], classification [4], and imaging [5].

In [6]–[8], the authors exploit similar ideas for outlier detection in structured data, mathematically represented by a graph. In these works, the goal is to detect an anomalous subgraph within the network, which is useful in cybersecurity attacks or suspicious network behavior. A similar analysis is also considered in [9] for detecting topological changes within the class of so-called stationary graph signals [10]–[12].

The work in [13] extends MSD to brain networks for Alzheimer's disease (AD) classification. Here, the authors consider two classes of subjects, the normal control (NC) subjects, and the AD subjects to perform patient classification based on fMRI measurements. Specifically, from a set of training data *a nominal* brain graph $\mathcal{G}_0$ and *an alternative* brain graph $\mathcal{G}_1$ are built for the NC and the AD classes, respectively. Then, [13] addresses the research question:

Is the data under consideration better matched with the graph $\mathcal{G}_0$, or with the graph $\mathcal{G}_1$?

Manuscript received January 11, 2018; revised March 15, 2018; accepted March 20, 2018. Date of publication March 26, 2018; date of current version April 4, 2018. The associate editor coordinating the review of this manuscript and approving it for publication was Dr. David I Shuman. *(Corresponding author: Elvin Isufi.)*

The authors are with the Faculty of Electrical Engineering, Mathematics and Computer Science, Delft University of Technology, Delft, 2628, CD, The Netherlands (e-mail: e.isufi-1@tudelft.nl; ashvantm@hotmail.com; g.j.t.leus@tudelft.nl).

Color versions of one or more of the figures in this letter are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/LSP.2018.2819127

Despite its classification success over contenders, the approach of [13] necessitates the knowledge of both $\mathcal{G}_0$ and $\mathcal{G}_1$. The latter presents the following challenges.

1) A "sufficient" number of training subjects is needed to build both $\mathcal{G}_0$ and $\mathcal{G}_1$.
2) A well-established graph model $\mathcal{G}_1$ needs to be investigated for a different category of alternative control (AC) subjects that we might be interested to test.
3) The alternative graph $\mathcal{G}_1$ may be hard to determine for AC subjects affected by multiple brain diseases.

While for the NC subjects we might leverage the extensive research of the past decades and the availability of a large set of available training subjects, this is not always the case for the AC subjects of rare brain diseases.

To solve the above issues, this letter explores ways in which topological changes can be detected without having specific knowledge of $\mathcal{G}_1$. We refer to these detectors as blind detectors. Concretely, we explore tools from graph signal processing (GSP) [14], [15] and MSD theory to answer the research question:

Is the data under consideration matched with the graph $\mathcal{G}_0$, or not?

The answer to the above question avoids the issues related to $\mathcal{G}_1$ and, therefore, sends only subjects that do not belong to $\mathcal{G}_0$ for further examinations. Moreover, since the analysis is not restricted to particular AC subjects, it covers a wider range of alternatives for which a graph-based analysis might be unknown. A final benefit worth mentioning for the introduced blind detectors is their ability to detect link losses within the nominal graph $\mathcal{G}_0$ directly from the observed signal.

The difference with [6]–[8] is that we do not consider any distribution of the graph topology, while, differently from [9], we do not restrict our analysis to stationary graph signals. Furthermore, we also provide closed-form expressions for the detectors' performance.

The remainder of the letter proceeds as follows: Section II covers the background information. Section III introduces the proposed blind topology change detectors, and Section IV evaluates their performance. Section V concludes the letter.

II. Background

A. Graphs and Graph Signals

A graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ consists of a set of N nodes $\mathcal{V}$ along with an edge set $\mathcal{E}$. The graph connectivity is captured by the weighted graph adjacency matrix $\mathbf{W}$. We assume $\mathcal{G}$ to be connected, undirected, and positive weighted, such that $W_{i,j} = W_{j,i} > 0$ if there is an edge connecting nodes (v_i, v_j), or $W_{i,j} = 0$ otherwise. In addition to $\mathbf{W}$, another graph descriptor matrix is the graph Laplacian matrix $\mathbf{L} = \text{diag}(\mathbf{1}^\mathsf{T}\mathbf{W}) - \mathbf{W}$. A graph signal is defined as a mapping from the vertex set to the set of complex numbers, i.e., $x : \mathcal{V} \to \mathbb{C}$. For convenience, the vector $\mathbf{x} = [x_1, \ldots, x_N]^\mathsf{T}$ collects all node signals. fMRI

measurements at a particular time instance in a brain network is an illustration of a graph signal.

B. Frequency Analysis on Graphs

Since $\mathbf{L}$ is real and symmetric, it can be eigendecomposed as $\mathbf{L} = \mathbf{U}\mathbf{\Lambda}\mathbf{U}^{\mathsf{H}}$, where $\mathbf{U} = [\mathbf{u}_1, \ldots, \mathbf{u}_N]$ is a complete set of orthonormal eigenvectors, with associated eigenvalues $0 = \lambda_1 = [\mathbf{\Lambda}]_{1,1} \leq \lambda_2 = [\mathbf{\Lambda}]_{2,2} \leq \ldots \leq \lambda_N = [\mathbf{\Lambda}]_{N,N}$. As stated in [14], the eigendecomposition of $\mathbf{L}$ extends the Fourier analysis from the structured temporal domain to the irregular graph domain. Specifically, the graph Fourier transform (GFT) $\hat{\mathbf{x}}$ of $\mathbf{x}$ is defined as the projection of $\mathbf{x}$ onto the eigenvectors of the graph Laplacian

$$\hat{\mathbf{x}} = \mathbf{U}^{\mathsf{H}}\mathbf{x} \tag{1}$$

and similarly, the inverse GFT is defined as $\mathbf{x} = \mathbf{U}\hat{\mathbf{x}}$. Matrix $\mathbf{U}$ contains the oscillating modes of the graph and $\{\lambda_1, \ldots, \lambda_N\}$ represent the graph frequencies, where $\lambda_i < \lambda_j$ is related to a graph mode $\mathbf{u}_i$ that varies more slowly over $\mathcal{G}$ than $\mathbf{u}_j$.

C. Bandlimited and Smooth Graph Signals

A graph signal $\mathbf{x}$ is said to be bandlimited if its graph frequency content $\hat{\mathbf{x}}$ is different from zero only on a limited set $\mathcal{F}$ of graph frequencies. The cardinality of $\mathcal{F}$, $|\mathcal{F}|$, denotes the graph signal bandwidth. Without loss of generality, let us consider that $\mathcal{F}$ consists of the first $|\mathcal{F}|$ graph frequencies and write $\hat{\mathbf{x}} = [\hat{\mathbf{x}}_{\mathcal{F}}^{\mathsf{T}}, \mathbf{0}_{|\overline{\mathcal{F}}|}^{\mathsf{T}}]^{\mathsf{T}}$, where $\overline{\mathcal{F}}$ denotes the complementary set of $\mathcal{F}$ with $|\overline{\mathcal{F}}| = N - |\mathcal{F}|$, and $\mathbf{0}_n$ is the $n \times 1$ vector of all zeros. Thus, the signal in the vertex domain can be written as

$$\mathbf{x} = \mathbf{U}_{\mathcal{F}}\hat{\mathbf{x}}_{\mathcal{F}} \tag{2}$$

where $\mathbf{U}_{\mathcal{F}} \in \mathbb{R}^{N \times |\mathcal{F}|}$ is the matrix containing the first $|\mathcal{F}|$ columns of $\mathbf{U}$. Similarly, we denote with $\mathbf{U}_{\overline{\mathcal{F}}} \in \mathbb{R}^{N \times |\overline{\mathcal{F}}|}$ the matrix containing the columns of $\mathbf{U}$ related to $\overline{\mathcal{F}}$.

Finally, from (2) we observe that the bandlimitedness of a graph signal is a low-rank representation of the latter and, thus, it represents a subspace projection of $\mathbf{x}$ where the MSD theory can be used. Nevertheless, it is a special subspace, since its basis carries the notion of frequency in the graph setting.

Beside bandlimitedness, another useful property of graph signals is their amount of variation over the graph. One way to quantify this is through the Laplacian quadratic form [16]

$$S_2(\mathbf{x}) = \mathbf{x}^{\mathsf{T}}\mathbf{L}\mathbf{x}, \tag{3}$$

a.k.a. the 2-Dirichlet form [14]. Then, for two graph signals $\mathbf{x}_1$ and $\mathbf{x}_2$ we say that $\mathbf{x}_1$ is smoother (or varies more slowly) than $\mathbf{x}_2$ w.r.t. $\mathcal{G}$ if $S_2(\mathbf{x}_1) < S_2(\mathbf{x}_2)$.

In the sequel, we show how hypothesis testing on graphs can be expressed in terms of GSP and MSD, and introduce the simple-MSD (SMSD) and constrained-MSD (CMSD) detectors from [13].

D. Hypothesis Testing on Graphs

Consider two candidate graphs $\mathcal{G}_0 = (\mathcal{V}, \mathcal{E}_0)$ and $\mathcal{G}_1 = (\mathcal{V}, \mathcal{E}_1)$ with respective Laplacian matrices $\mathbf{L}_0 = \mathbf{U}_0\mathbf{\Lambda}_0\mathbf{U}_0^{\mathsf{H}}$ and $\mathbf{L}_1 = \mathbf{U}_1\mathbf{\Lambda}_1\mathbf{U}_1^{\mathsf{H}}$. Let also $\mathbf{y} = \mathbf{x} + \mathbf{n}$ denote the version of $\mathbf{x}$ corrupted by some zero-mean Gaussian noise $\mathbf{n} \sim \mathcal{N}(\mathbf{0}_N, \sigma^2\mathbf{I}_N)$. Subsequently, the hypothesis test is formulated as

$$\begin{aligned} \mathcal{H}_0: &\quad \mathbf{y} \text{ is attributed to } \mathcal{G}_0 \\ \mathcal{H}_1: &\quad \mathbf{y} \text{ is attributed to } \mathcal{G}_1. \end{aligned} \tag{4}$$

By exploiting the GFT (1), (4) can be rephrased as

$$\begin{aligned} \mathcal{H}_0: &\quad \mathbf{y} = \mathbf{U}_0\hat{\mathbf{x}}_{\mathcal{G}_0} + \mathbf{n} \\ \mathcal{H}_1: &\quad \mathbf{y} = \mathbf{U}_1\hat{\mathbf{x}}_{\mathcal{G}_1} + \mathbf{n} \end{aligned} \tag{5}$$

with $\hat{\mathbf{x}}_{\mathcal{G}_j}$ the GFT of $\mathbf{x}$ on $\mathcal{G}_j$ for $j = 0, 1$. Then, the Neyman–Pearson detector $T(\mathbf{y})$ discriminates between the nominal graph $\mathcal{G}_0$ and the alternative graph $\mathcal{G}_1$ if [1]

$$T(\mathbf{y}) = \frac{p(\mathbf{y}; \mathcal{H}_1)}{p(\mathbf{y}; \mathcal{H}_0)} \underset{\mathcal{H}_0}{\overset{\mathcal{H}_1}{\gtrless}} \gamma \tag{6}$$

where $p(\mathbf{y}; \mathcal{H}_j)$ is the probability density function of $\mathbf{y}$ given hypothesis $\mathcal{H}_j$ and γ denotes the decision threshold that guarantees a target probability of false alarm $P_{\text{FA}} \triangleq \Pr\{T(\mathbf{y}) > \gamma; \mathcal{H}_0\}$.

Based on the properties of $\mathbf{x}$ w.r.t. $\mathcal{G}_0$ or $\mathcal{G}_1$, [13] identifies the following MSD detectors on graphs:

1) SMSD: From the bandlimited assumption of $\mathbf{x}$ w.r.t. both graphs $\mathcal{G}_0$ and $\mathcal{G}_1$, $\mathbf{y}$ can be projected on the respective low-rank subspaces, i.e., $\hat{\mathbf{y}}_{\mathcal{F}_0} = \mathbf{U}_{0,\mathcal{F}_0}^{\mathsf{H}}\mathbf{y}$ and $\hat{\mathbf{y}}_{\mathcal{F}_1} = \mathbf{U}_{1,\mathcal{F}_1}^{\mathsf{H}}\mathbf{y}$. Then, the SMSD detector can be written as

$$T_{\text{SMSD}}(\mathbf{y}) = \|\hat{\mathbf{y}}_{\mathcal{F}_0}\|_2^2 - \|\hat{\mathbf{y}}_{\mathcal{F}_1}\|_2^2 \underset{\mathcal{H}_0}{\overset{\mathcal{H}_1}{\gtrless}} \gamma \tag{7}$$

which consists of the in-band energy difference of the GFT of $\mathbf{y}$ w.r.t. $\mathcal{G}_0$ and $\mathcal{G}_1$, respectively.

Detector (7) consists of the difference of two Chi-squared random variables. While in [13] its performance is only assessed numerically, the distribution of the SMSD can be approximated as normal [17, Chapter 7]. This approximation may often be inaccurate, however, we have observed satisfactory results even for moderate values of $|\mathcal{F}_0|$ and $|\mathcal{F}_1|$. Accordingly, we can write

$$T_{\text{SMSD}}(\mathbf{y}) \sim \begin{cases} \mathcal{N}(\mu_0, \sigma_0^2) & \text{under } \mathcal{H}_0 \\ \mathcal{N}(\mu_1, \sigma_1^2) & \text{under } \mathcal{H}_1 \end{cases} \tag{8}$$

where μ_0, μ_1, σ_0, and σ_1 can be estimated from the training data; for instance the data used for building the graphs. Therefore, by using (8) we can assess the SMSD performance theoretically as

$$\gamma = \sigma_0 Q^{-1}(P_{\text{FA}}) + \mu_0 \tag{9}$$

$$P_D = Q\left(\sigma_1{}^{-1}(\gamma - \mu_1)\right) \tag{10}$$

where (9) sets the threshold γ to guarantee a target P_{FA} and (10) relates the probability of detection P_D to γ (P_{FA}) by the Q-function. Here $Q^{-1}(\cdot)$ stands for the inverse Q-function. In Section IV we show that these theoretical derivations match the numerical performance of the SMSD.

2) CMSD: The CMSD detector exploits a constraint on the smoothness (3) to discriminate between $\mathcal{G}_0$ and $\mathcal{G}_1$. Specifically, the hypothesis test (5) is first expressed in the graph Fourier domain by multiplying both sides of align $\mathcal{H}_j$ with $\mathbf{U}_j^{\mathsf{H}}$, for $j = 0, 1$, leading to $\hat{\mathbf{y}}_{\mathcal{G}_j} = \mathbf{U}_j^{\mathsf{H}}\mathbf{y}$. Then, the constrained maximum likelihood estimator (C-MLE) $\hat{\mathbf{x}}_{\mathcal{G}_j}^*$ is found from

$$\begin{aligned} \hat{\mathbf{x}}_{\mathcal{G}_j}^* = &\underset{\hat{\mathbf{x}}_{\mathcal{G}_j}}{\text{argmax}} \quad p(\hat{\mathbf{y}}_{\mathcal{G}_j}; \mathcal{H}_j) \\ &\text{subject to} \quad \hat{\mathbf{x}}_{\mathcal{G}_j}^{\mathsf{T}}\mathbf{\Lambda}_j\hat{\mathbf{x}}_{\mathcal{G}_j} \leq r\|\hat{\mathbf{x}}_{\mathcal{G}_j}\|_2^2. \end{aligned} \tag{11}$$

The objective function $p(\hat{\mathbf{y}}_{\mathcal{G}_j};\mathcal{H}_j) \sim \mathcal{N}(\hat{\mathbf{x}}_{\mathcal{G}_j}, \sigma^2\mathbf{I}_N)$ is the likelihood of $\hat{\mathbf{y}}_{\mathcal{G}_j}$ given hypothesis $\mathcal{H}_j$. The quadratic constraint imposes an upper bound r on $S_2(\mathbf{x})$ (3). The role of $\|\hat{\mathbf{x}}_{\mathcal{G}_j}\|_2^2$ is simply for normalization purposes. With this in place, the CMSD detector can be written as

$$T_{\text{CMSD}}(\mathbf{y}) = \left\|\left(\hat{\mathbf{y}}_{\mathcal{G}_0} - \hat{\mathbf{x}}^*_{\mathcal{G}_0}\right)\right\|_2^2 - \left\|\left(\hat{\mathbf{y}}_{\mathcal{G}_1} - \hat{\mathbf{x}}^*_{\mathcal{G}_1}\right)\right\|_2^2 \underset{\mathcal{H}_0}{\overset{\mathcal{H}_1}{\gtrless}} \gamma \quad (12)$$

which, similar to the SMSD, consists of the energy difference between two GFTs, but now the respective MLE spectrum is subtracted (i.e., $\hat{\mathbf{y}}_{\mathcal{G}_j} - \hat{\mathbf{x}}^*_{\mathcal{G}_j}$).

Next, we show how these detectors extend to the case where no information about $\mathcal{G}_1$ is available.

III. Blind Graph Topology Change Detectors

We reformulate the blind hypothesis testing on graphs as

$$\begin{aligned} \mathcal{H}_0: &\quad \mathbf{y} = \mathbf{U}_0\hat{\mathbf{x}}_{\mathcal{G}_0} + \mathbf{n} \\ \mathcal{H}_1: &\quad \mathbf{y} \neq \mathbf{U}_0\hat{\mathbf{x}}_{\mathcal{G}_0} + \mathbf{n} \end{aligned} \quad (13)$$

which is the equivalent of (5) and aims to quantify if $\mathbf{y}$ is attributed to $\mathcal{G}_0$, or not. In the sequel, we approach the problem (13) by extending both the SMSD and CMSD detectors to their blind counterparts.

Blind-SMSD (B-SMSD): In the B-SMSD detector, we exploit the fact that, on $\mathcal{G}_0$, $\mathbf{x}$ has a graph frequency content different from zero only on the set $\mathcal{F}_0$. Therefore, $\hat{\mathbf{x}}_{\overline{\mathcal{F}}_0} = \mathbf{U}^{\mathsf{H}}_{\overline{\mathcal{F}}_0}\mathbf{x} = \mathbf{0}_{|\overline{\mathcal{F}}_0|}$ under hypothesis $\mathcal{H}_0$ and $\hat{\mathbf{x}}_{\overline{\mathcal{F}}_0} \neq \mathbf{0}_{|\overline{\mathcal{F}}_0|}$ under $\mathcal{H}_1$. With this in place, the observation $\mathbf{y}$ w.r.t. $\mathcal{G}_0$ is transformed as

$$\begin{aligned} \hat{\mathbf{y}}_{\overline{\mathcal{F}}_0} &= \mathbf{U}^{\mathsf{H}}_{\overline{\mathcal{F}}_0}\mathbf{y} = \mathbf{U}^{\mathsf{H}}_{\overline{\mathcal{F}}_0}\mathbf{x} + \mathbf{U}^{\mathsf{H}}_{\overline{\mathcal{F}}_0}\mathbf{n} \\ &\triangleq \hat{\mathbf{x}}_{\overline{\mathcal{F}}_0} + \hat{\mathbf{n}}_{\overline{\mathcal{F}}_0} \end{aligned} \quad (14)$$

with $\hat{\mathbf{n}}_{\overline{\mathcal{F}}_0} \sim \mathcal{N}(\mathbf{0}_{|\overline{\mathcal{F}}_0|}, \sigma^2\mathbf{I}_{|\overline{\mathcal{F}}_0|})$. Subsequently, the blind hypothesis test (13) can be rephrased as

$$\begin{aligned} \mathcal{H}_0: &\quad \hat{\mathbf{y}}_{\overline{\mathcal{F}}_0} = \hat{\mathbf{n}}_{\overline{\mathcal{F}}_0} \\ \mathcal{H}_1: &\quad \hat{\mathbf{y}}_{\overline{\mathcal{F}}_0} = \hat{\mathbf{x}}_{\overline{\mathcal{F}}_0} + \hat{\mathbf{n}}_{\overline{\mathcal{F}}_0} \end{aligned} \quad (15)$$

which is the well-known problem of detecting an unknown deterministic signal ($\hat{\mathbf{x}}_{\overline{\mathcal{F}}_0}$) embedded in noise [1]. Then, the B-SMSD detector decides $\mathcal{H}_1$ if the generalized LRT (GLRT)

$$L(\mathbf{y}) = \frac{p(\mathbf{y};\hat{\mathbf{x}}^*_{\overline{\mathcal{F}}_0},\mathcal{H}_1)}{p(\mathbf{y};\mathcal{H}_0)} \quad (16)$$

exceeds a threshold, where $\hat{\mathbf{x}}^*_{\overline{\mathcal{F}}_0}$ is the MLE of $\hat{\mathbf{x}}_{\overline{\mathcal{F}}_0}$ assuming $\mathcal{H}_1$ is true (it maximizes $p(\mathbf{y};\hat{\mathbf{x}}_{\overline{\mathcal{F}}_0},\mathcal{H}_1)$), i.e., $\hat{\mathbf{x}}^*_{\overline{\mathcal{F}}_0} = \hat{\mathbf{y}}_{\overline{\mathcal{F}}_0}$.[1] Therefore, the B-SMSD detector can be written as

$$T_{\text{B-SMSD}}(\mathbf{y}) = \|\hat{\mathbf{y}}_{\overline{\mathcal{F}}_0}\|_2^2/\sigma^2 \underset{\mathcal{H}_0}{\overset{\mathcal{H}_1}{\gtrless}} \gamma \quad (17)$$

which differently from (7) decides $\mathcal{H}_1$ if the out-of-band signal-to-noise ratio (SNR) exceeds a threshold γ.[2]

A notable benefit of the B-SMSD detector (17) is that its detection performance can be computed in closed form [18]. Specifically,

$$T_{\text{B-SMSD}}(\mathbf{y}) \sim \begin{cases} \chi^2_{|\overline{\mathcal{F}}_0|} & \text{under } \mathcal{H}_0 \\ \chi^2_{|\overline{\mathcal{F}}_0|}(\delta) & \text{under } \mathcal{H}_1 \end{cases} \quad (18)$$

where $\chi^2_{|\overline{\mathcal{F}}_0|}$ stands for a Chi-squared distribution with $|\overline{\mathcal{F}}_0|$ degrees of freedom, and $\delta = \|\hat{\mathbf{x}}_{\overline{\mathcal{F}}_0}\|_2^2/\sigma^2$ indicates the noncentrality parameter. The detection threshold γ and P_D can be found by making use of the right-tail probability function $Q_{\chi^2_{|\overline{\mathcal{F}}_0|}}(\mathbf{y})$ for a $\chi^2_{|\overline{\mathcal{F}}_0|}$ random variable, i.e.,

$$\gamma = Q^{-1}_{\chi^2_{|\overline{\mathcal{F}}_0|}}(P_{\text{FA}}) \quad (19)$$

$$P_D = Q_{\chi^2_{|\overline{\mathcal{F}}_0|}(\delta)}(\gamma) \quad (20)$$

where $Q^{-1}_{\chi^2_{|\overline{\mathcal{F}}_0|}}(\cdot)$ indicates the inverse of $Q_{\chi^2_{|\overline{\mathcal{F}}_0|}}(\cdot)$, and $Q_{\chi^2_{|\overline{\mathcal{F}}_0|}(\delta)}$ is the Q-function of the noncentral Chi-squared random variable $\chi^2_{|\overline{\mathcal{F}}_0|}(\delta)$.

Blind-CMSD (B-CMSD): In the B-CMSD detector, we exploit the smoothness constraint (3) solely on $\mathcal{G}_0$, to decide whether the observation $\mathbf{y}$ can be attributed to that graph, or not. Proceeding similarly as in the CMSD, the C-MLE $\hat{\mathbf{x}}^*_{\mathcal{G}_0|\mathcal{H}_0}$ is obtained from (11) for $j = 0$, while the MLE of $\hat{\mathbf{x}}_{\mathcal{G}_0|\mathcal{H}_1}$ given $\mathcal{H}_1$ (it maximizes $p(\hat{\mathbf{y}};\hat{\mathbf{x}}_{\mathcal{G}_0|\mathcal{H}_1},\mathcal{H}_1)$) is $\hat{\mathbf{x}}^*_{\mathcal{G}_0|\mathcal{H}_1} = \hat{\mathbf{y}}_{\mathcal{G}_0}$ (i.e., the GFT of $\mathbf{y}|\mathcal{H}_1$ w.r.t. $\mathcal{G}_0$). Then, from the GLRT

$$L(\mathbf{y}) = \frac{p(\hat{\mathbf{y}};\hat{\mathbf{x}}^*_{\mathcal{G}_0|\mathcal{H}_1},\mathcal{H}_1)}{p(\hat{\mathbf{y}};\hat{\mathbf{x}}^*_{\mathcal{G}_0|\mathcal{H}_0},\mathcal{H}_0)} \quad (21)$$

with $p(\hat{\mathbf{y}};\hat{\mathbf{x}}^*_{\mathcal{G}_0|\mathcal{H}_0},\mathcal{H}_0) \sim \mathcal{N}(\hat{\mathbf{x}}^*_{\mathcal{G}_0|\mathcal{H}_0}, \sigma^2\mathbf{I})$ and $p(\hat{\mathbf{y}};\hat{\mathbf{x}}^*_{\mathcal{G}_0|\mathcal{H}_1},\mathcal{H}_1) \sim \mathcal{N}(\hat{\mathbf{x}}^*_{\mathcal{G}_0|\mathcal{H}_1}, \sigma^2\mathbf{I})$, the B-CMSD can be written as

$$T_{\text{B-CMSD}}(\mathbf{y}) = \left(\|\hat{\mathbf{y}}_{\mathcal{G}_0}\|_2^2 - \hat{\mathbf{y}}^{\mathsf{H}}_{\mathcal{G}_0}\hat{\mathbf{x}}^*_{\mathcal{G}_0|\mathcal{H}_0}\right)/\sigma^2 \underset{\mathcal{H}_0}{\overset{\mathcal{H}_1}{\gtrless}} \gamma. \quad (22)$$

Detector (22) decides $\mathcal{H}_1$ if the *nonmatched*[3] SNR w.r.t. $\mathcal{G}_0$ exceeds a threshold γ.

To conclude this section, we remark that the meaning of probability of detection (P_D) in the blind detectors is different from that of the nonblind ones. In fact, for the nonblind detectors a detection is made if the observed signal is attributed to *the specific* $\mathcal{G}_1$. On the contrary, blind detectors count a detection if the observed signal belongs to any other $\mathcal{G}_1$ that is not $\mathcal{G}_0$. As we will show next, this difference plays a central role when evaluating the detector performances.

IV. Numerical Results

To validate the detectors' performance, we first consider the synthetic brain scenario from [13], and then study topological changes that follow the random edge sampling (RES) model [19]. In our simulations, we make use of the GSP box [20].

A. Results for the Brain Scenario of [13]

This scenario considers two candidate graphs $\mathcal{G}_0$ and $\mathcal{G}_1$, both modeled as small-world (SW) networks [21]. As shown in [22]–[24] SW networks are effective models for brain connectivity structures. Concisely, we indicate a SW graph of

[1] Note that the MLE does not invoke the knowledge of $\mathcal{G}_1$.

[2] In practice, σ^2 can be incorporated into the threshold γ, but is left here as a normalization term to derive compactly the theoretical performance [see (18)]. For consistency, it is also kept in the other blind detector [see (22)].

[3] $\|\hat{\mathbf{y}}_{\mathcal{G}_0}\|_2^2$ consists of the signal energy, while $\hat{\mathbf{y}}^{\mathsf{H}}_{\mathcal{G}_0}\hat{\mathbf{x}}^*_{\mathcal{G}_0|\mathcal{H}_0}$ can be interpreted as the matched filter output of $\hat{\mathbf{y}}_{\mathcal{G}_0}$ onto the C-MLE $\hat{\mathbf{x}}^*_{\mathcal{G}_0|\mathcal{H}_0}$.

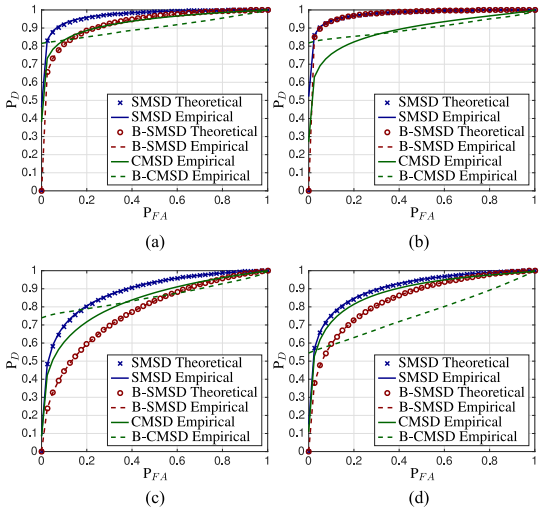

Fig. 1. P_D versus P_{FA} for the proposed blind detectors and their nonblind counterparts in the brain scenario of [13]. (a) $\mathcal{G}_1 \in \mathcal{G}_{SW}(40, 20, 0.4)$, exp. decay, $\sigma = 0.3$. (b) $\mathcal{G}_1 \in \mathcal{G}_{SW}(40, 20, 0.4)$, step, $\sigma = 0.5$. (c) $\mathcal{G}_1 \in \mathcal{G}_{SW}(40, 20, 0.4)$, exp. decay, $\sigma = 0.5$. (d) $\mathcal{G}_1 \in \mathcal{G}_{SW}(40, 12, 0.1)$, exp. decay, $\sigma = 0.3$.

N nodes, average vertex degree d , and rewiring probability p_r as $\mathcal{G}_{SW}(N, d, p_r)$. Then, $\mathcal{G}_0 \in \mathcal{G}_{SW}(40, 12, 0.1)$ and $\mathcal{G}_1 \in \mathcal{G}_{SW}(40, 20, 0.4)$. There is also one case (Fig. 1(d)) where $\mathcal{G}_1 \in \mathcal{G}_{SW}(40, 12, 0.1)$, i.e., it is drawn from the same distribution as $\mathcal{G}_0$ but it is a different realization. The graph signal is defined in the graph Fourier domain as $\hat{\mathbf{x}}_{\mathcal{G}_j} \sim \mathcal{N}(\mathbf{0}, \mathbf{\Sigma})$ for $j = 0, 1$. The signal covariance matrix is considered diagonal and defined in two ways: $1)$ as an exponential decay $[\mathbf{\Sigma}]_{i,i} = \exp(-i/5)$ for all i, and $2)$ as a step function $[\mathbf{\Sigma}]_{i,i} = 1$ for $i = 1, \ldots, 12$, and $[\mathbf{\Sigma}]_{i,i} = 0.1$, otherwise. Furthermore, two different values of σ are assumed for $\mathbf{n}$, i.e., $a)$ $\sigma = 0.3$, and $b)$ $\sigma = 0.5$.

Our results are averaged over 100 different graphs and graph signal realizations, and 10^4 realizations of the noise. $|\mathcal{F}|$ is set to 12 for the B-SMSD detector, r is the 12th eigenvalue of the respective Laplacian for CMSD, while for B-CMSD r is chosen data dependent as

$$r(\mathbf{y}) = \min\left\{[\mathbf{\Lambda}_0]_{i,i} \mid \mathbf{y}^{\mathsf{H}}\mathbf{L}_0\mathbf{y} \leq [\mathbf{\Lambda}_0]_{i,i}\|\mathbf{y}\|_2^2\right\} \qquad (23)$$

i.e., as the minimum eigenvalue of $\mathbf{L}_0$ that bounds the normalized $S_2(\mathbf{y})$ over $\mathcal{G}_0$.

Fig. 1(a)–(d) show the corresponding results. We remark the matching of the proposed theoretical derivations in (9), (10) and (19), (20) with their numerical counterparts. For the B-SMSD detector, we note that the lack of information about $\mathcal{G}_1$ influences the performance, especially when the out-of-band signal spectrum is significantly different from zero, e.g., exp. decay in Fig. 1(a), (c), and (d). This result is not totally surprising as no information about $\mathcal{G}_1$ is exploited. However, this difference in P_D is around 10% in Fig. 1(a) ($P_{\text{FA}} = 0.1$), and is nearly indistinguishable in Fig. 1(b) (i.e., when the signal spectrum deviates only a little from the bandlimitedness model). The latter suggests that the knowledge about $\mathcal{G}_1$ *is not necessary* for this case. The B-CMSD detector, on the other hand, compares well with the respective nonblind detector (CMSD), especially for low values of P_{FA} (in Fig. 1(b) and (c) it performs better). B-CMSD is in general more sensitive to small topological changes in low noise regimes (e.g., Fig. 1(d)); a behavior that is observed also in the upcoming scenario.

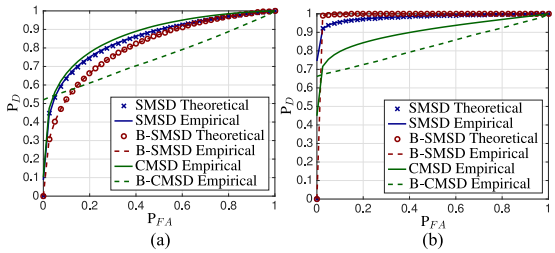

Fig. 2. P_D versus P_{FA} for the proposed blind detectors and their nonblind counterparts in the RES change model [19].

B. Results for the RES Change Model [19]

In this second scenario, we assess the discriminative power of the detectors when the alternative graph $\mathcal{G}_1$ is a RES version of $\mathcal{G}_0$ denoted as $\mathcal{G}_1 \in \text{RES}(\mathcal{G}_0, p_l)$, where p_l indicates the probability that an edge of $\mathcal{G}_0$ is removed. Our justification for this is that SW graphs are also suitable models for social and sensor networks, where edges can go down randomly from the original graph. In this case, information about the specific $\mathcal{G}_1$ realization might not even be obtained.

Fig. 2(a) shows the detectors' performance for $\sigma = 0.3$ and $p_l = 0.25$, when the graph signal is characterized by an exponential decay spectrum. Similarly to the results in Fig. 1(a), the lack of information about $\mathcal{G}_1$ deteriorates the P_D of the B-SMSD detector by approximately 10% ($P_{\text{FA}} = 0.1$).

Fig. 2(b) compares the performance when the covariance matrix of the signal follows the step function. As the signal is now approximately bandlimited, the B-SMSD performs the best. We observe that B-SMSD performs better than SMSD (enforcing the result in Fig. 1(b)). A similar performance was also observed for $\sigma = 0.5$, not reported here due to space limitations. The latter suggests once again that the knowledge of $\mathcal{G}_1$ is not necessary. We attribute these results to the small model deviations, and to differences in the meaning of P_D.

For the B-CMSD detector, we observe that it is more sensitive to small topological changes than the B-SMSD detector, and it performs worse than CMSD.

As a final remark, we advocate the use of the nonblind detectors when the true signal matches well the model. Meanwhile, a blind approach is more recommended for those cases where there are uncertainties about the alternative graph structure and the signal model.

V. Conclusion and Future Work

The primary goal of this letter was to develop detection strategies that could discriminate whether a graph topological change has happened without making any assumptions on the change itself. By making use of MSD and GSP theories, we proposed two strategies that could achieve this goal. The evidence from the obtained results yields two main observations: $1)$ the absence of information about the topological change may, in some situations, play a crucial role in the detection performance; and $2)$ when the topological change deviates "slightly" from the assumed model, such information may deteriorate the detectors that incorporate it.

As a further study, we will consider assessing the detectors' performance in classifying subjects as healthy and nonhealthy, without being restricted to AD. Furthermore, it would also be interesting to analyze the performance of blind detectors that use jointly the bandlimitedness and smoothness priors.

REFERENCES

[1] L. L. Scharf and B. Friedlander, "Matched subspace detectors," *IEEE Trans. Signal Process.*, vol. 42, no. 8, pp. 2146–2157, Aug. 1994.

[2] M. A. Richards, J. Scheer, W. A. Holm, and W. L. Melvin, *Principles of Modern Radar*. Citeseer, 2010.

[3] B. Sklar, *Digital communications*. Upper Saddle River, NJ, USA: Prentice-Hall, 2001, vol. 2.

[4] G. B. Giannakis and M. K. Tsatsanis, "Signal detection and classification using matched filtering and higher order statistics," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 38, no. 7, pp. 1284–1296, Jul. 1990.

[5] Z. Li, Q. Li, X. Yu, P. S. Conti, and R. M. Leahy, "Lesion detection in dynamic fdg-pet using matched subspace detection," *IEEE Trans. Med. Imag.*, vol. 28, no. 2, pp. 230–240, Feb. 2009.

[6] B. A. Miller, N. T. Bliss, P. J. Wolfe, and M. S. Beard, "Detection theory for graphs," *Lincoln Laboratory J.*, vol. 20, no. 1, pp. 10–30, 2013.

[7] N. Singh, B. A. Miller, N. T. Bliss, and P. J. Wolfe, "Anomalous subgraph detection via sparse principal component analysis," in *Proc. 2011 IEEE Statistical Signal Process. Workshop*, 2011, pp. 485–488.

[8] B. A. Miller, M. S. Beard, and N. T. Bliss, "Matched filtering for subgraph detection in dynamic networks," in *Proc. 2011 IEEE Statistical Signal Process. Workshop*, 2011, pp. 509–512.

[9] S. P. Chepuri and G. Leus, "Subgraph detection using graph signals," in Proc. Asilomar Conf. Signals, Syst., Comput., 2016, pp. 532–534.

[10] B. Girault, "Stationary graph signals using an isometric graph translation," in *Proc. 2015 23rd Eur. Signal Process. Conf.,* IEEE, 2015, pp. 1516–1520.

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

[12] A. G. Marques, S. Segarra, G. Leus, and A. Ribeiro, "Stationary graph processes and spectral estimation," *IEEE Trans. Signal Process.*, vol. 65, no. 22, pp. 5911–5926, Nov. 2017.

[13] C. Hu, J. Sepulcre, K. A. Johnson, G. E. Fakhri, Y. M. Lu, and Q. Li, "Matched signal detection on graphs: Theory and application to brain imaging data classification," *NeuroImage*, vol. 125, pp. 587–600, 2016.

[14] 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, May 2013.

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

[16] D. A. Spielman, "Spectral graph theory and its applications," in *Proc. 48th Annu. IEEE Symp. Foundations Comput. Sci.,* 2007, pp. 29–38.

[17] J. K. Patel and C. B. Read, *Handbook of the normal distribution*. Boca Raton, FL, USA: CRC Press, 1996, vol. 150.

[18] S. M. Kay, "Fundamentals of Statistical Signal Processing. Detection Theory, Volume II," Englewood Cliffs, NJ, USA: Prentice-Hall, 1998.

[19] E. Isufi, A. Loukas, A. Simonetto, and G. Leus, "Filtering random graph processes over random time-varying graphs," *IEEE Trans. Signal Process.*, vol. 65, no. 16, pp. 4406–4421, Aug. 2017.

[20] N. Perraudin *et al.*, "Gspbox: A toolbox for signal processing on graphs," arXiv:1408.5781, 2014.

[21] D. J. Watts and S. H. Strogatz, "Collective dynamics of ësmall-worldínetworks," *nature*, vol. 393, no. 6684, pp. 440–442, 1998.

[22] D. S. Bassett and E. Bullmore, "Small-world brain networks," *Neuroscientist*, vol. 12, no. 6, pp. 512–523, 2006.

[23] Y. He, Z. J. Chen, and A. C. Evans, "Small-world anatomical networks in the human brain revealed by cortical thickness from MRI," *Cerebral Cortex*, vol. 17, no. 10, pp. 2407–2419, 2007.

[24] C. J. Stam, B. Jones, G. Nolte, M. Breakspear, and P. Scheltens, "Small-world networks and functional connectivity in Alzheimer's disease," *Cerebral Cortex*, vol. 17, no. 1, pp. 92–99, 2006.