

SEMI-SUPERVISED TRACKING OF DYNAMIC PROCESSES OVER SWITCHING GRAPHS

Qin Lu, Vassilis N. Ioannidis, Georgios B. Giannakis

Dept. of ECE and Digital Technology Center, University of Minnesota, USA

ABSTRACT

Several network science applications involve nodal processes with dynamics dependent on the underlying graph topology that can possibly jump over discrete states. The connectivity in dynamic brain networks for instance, switches among candidate topologies, each corresponding to a different emotional state. In this context, the present work relies on limited nodal observations to perform semi-supervised tracking of dynamic processes over switching graphs. To this end, leveraging what is termed interacting multi-graph model (IMGM), a scalable online Bayesian approach is developed to track the active graph topology and dynamic nodal process. Numerical tests with synthetic and real datasets demonstrate the merits of the novel approach.

Index Terms— Dynamic graphs, Bayesian tracking

1. INTRODUCTION

Given limited data at a subset of nodes, various applications deal with inference of processes across all network nodes. Such a semi-supervised learning (SSL) task over networks can be addressed thanks to the underlying graph topology that captures nodal inter-dependencies [5, 7, 14]. The scarcity of nodal observations can be due to e.g., cost, and computational or privacy constraints. To name a couple, individuals in social networks may be reluctant to share personal information, while acquiring nodal samples in brain networks may require invasive procedures such as electrocorticography.

This inference task becomes more challenging when nodal processes are *nonstationary*, and the graph topology is also *time-varying*. In a brain network for example, where nodes correspond to brain regions and edges capture dependencies among them, one may be interested in predicting the dynamic processes as well as the varying interconnections. An interesting time-varying topology model switches over a set of connectivity patterns, also known as “network modes” [1]. The connectivity among human brain regions varies as the humans’ emotional, mental or physical activities change [18]. Coupled with the topology, the dynamics of nodal processes can also switch among different modes. A similar *switching model* has been employed to capture the kinematics of maneuvering targets such as drones [2].

Methods for inference (or reconstruction) of nodal processes typically assume that the network topology is known

and undirected, while the processes are smooth, in the sense that neighboring vertices have similar values [15]. Inference of slow-varying functions over graphs has been pursued using the so-termed graph bandlimited model in [4, 17]. On the other hand, [6, 13] employ graph kernel-based estimators for reconstructing general dynamic processes. All these contemporary approaches rely on a *known* graph topology. However, the dynamic graph can change or switch in an *unknown* fashion among a set of possibly known topologies, which may reflect sudden changes in the partially observed signals.

The present paper relies on a known set of candidate topologies to put forth an approach for semi-supervised tracking and extrapolation of dynamic nodal processes over switching graphs. Rather than the kinematics in [2], the nodal processes here evolve in accordance with a switching dynamical model that depends on the active graph topology. Given partially observed nodal samples and the candidate graph topologies, a scalable Bayesian algorithm is developed to *jointly* track the dynamic graph processes and classify the active graph topology (or network mode) on-the-fly.

If observations were available at all nodes, it would have been possible to identify the active topology per slot without explicitly modeling the nodal process dynamics [1]. Relative to [1], this work accounts for dynamics to reconstruct unavailable nodal data, while at the same time identifying the active mode and tracking the nodal processes. Not necessarily graph related yet similar to that of [1] is the goal of subspace clustering [16], but different from the work here mode dynamics are not leveraged to reconstruct unavailable nodal processes.

2. PROBLEM FORMULATION

Consider a graph with N nodes and the vertex set $\mathcal{V} := \{v_1, \dots, v_N\}$, whose connectivity switches among S discrete modes. Each mode corresponds to a unique connectivity pattern captured by the $N \times N$ adjacency matrix $\mathbf{A}_t^s$, whose (n, n') th entry is the nonnegative weight of the edge connecting v_n with $v_{n'}$. The graph is considered undirected with no self-loops, that is, $\{\mathbf{A}_t^s\}_{s=1}^S$ are symmetric and $\mathbf{A}_t^s(n, n) = 0$. The Laplacian matrix of mode s is $\mathbf{L}_t^s := \text{diag}\{\mathbf{A}_t^s \mathbf{1}_N\} - \mathbf{A}_t^s$. Per time slot t only one network mode $\sigma_t = s$ is active. Switching topologies emerge in several interconnected systems. Besides brain networks [18], the email network switches from work-based connections on

weekdays to friends-and-family ones over weekends.

A dynamic graph process is a mapping $x : \mathcal{V} \times \mathcal{T} \mapsto \mathbb{R}$, where $\mathcal{T} := \{1, 2, \dots\}$ is the set of slot indices. Specifically, $x_t(v_n)$ denotes the node v_n sample at time slot t . For example, $x_t(v_n)$ may denote the price of a stock at year t . The values of all nodes at time t will be collected in $\mathbf{x}_t := [x_t(v_1) \dots x_t(v_N)]^\top$, where $^\top$ stands for transposition.

In many applications, only a subset of the nodal samples are observed, yielding the observation model

$$\mathbf{z}_t = \mathbf{H}_t \mathbf{x}_t + \mathbf{e}_t \quad (1)$$

where $\mathbf{H}_t \in \{0, 1\}^{M \times N}$ is the $M \times N$ ($M < N$) sampling matrix, whose rows sum to 1, and $\mathbf{e}_t$ is zero-mean, temporally independent, Gaussian noise with covariance matrix $\mathbf{R}$.

To capture the spatio-temporal dynamics of the nodal processes that are connected through the mode-conditioned topology, we model the evolution from $\mathbf{x}_{t-1}$ to $\mathbf{x}_t$ as the first-order Markovian process

$$\mathbf{x}_t = \mathbf{F}_t^{\sigma_t} \mathbf{x}_{t-1} + \boldsymbol{\eta}_t^{\sigma_t} \quad (2)$$

where the state transition matrix $\mathbf{F}_t^{\sigma_t} := f(\mathbf{A}_t^{\sigma_t})$ is a function f of the active adjacency matrix $\mathbf{A}_t^{\sigma_t}$ at network mode $\sigma_t \in \{1, \dots, S\}$; and the mode-conditioned noise $\boldsymbol{\eta}_t^{\sigma_t}$ is Gaussian with zero mean and covariance $\mathbf{K}_t^{\sigma_t}$, which is taken from the so-termed family of Laplacian kernels described by [8] $\mathbf{K}_t^{\sigma_t} = r^\dagger(\mathbf{L}_t^{\sigma_t})$, where $r(\cdot)$ is a scalar decreasing function that promotes properties such as diffusion, smoothness, or graph bandlimitedness; and † denotes pseudo-inverse.

The dynamic model in (2) describes what is also known as a switching linear dynamical system (SLDS) [12], and it is widely employed in the tracking community to capture the kinematic state evolution of maneuvering targets [2].

Problem statement. Given T observations $\mathbf{Z}_T := [\mathbf{z}_1 \dots \mathbf{z}_T]$ as in (1), and candidate models $\{\{\mathbf{F}_t^s, \mathbf{K}_t^s\}_{s=1}^S\}_{t=1}^T$ as in (2), the goal is to jointly track the dynamic graph processes $\mathbf{X}_T := [\mathbf{x}_1 \dots \mathbf{x}_T]$, and the discrete modes $\{\sigma_t\}_{t=1}^T$.

3. GRAPH-ADAPTIVE BAYESIAN TRACKER

Here we develop a Bayesian approach, starting from the joint probability density function (pdf) of the nodal processes in $\mathbf{X}_T$ that can be expressed as $p(\mathbf{X}_T) = p(\mathbf{x}_T | \mathbf{x}_{T-1}) p(\mathbf{x}_{T-1}) = \dots = \prod_{t=1}^T p(\mathbf{x}_t | \mathbf{x}_{t-1})$, due to the Markovian model in (2). Because $\mathbf{e}_t$ is temporally white, the conditional data pdf also factorizes as $p(\mathbf{Z}_T | \mathbf{X}_T) = \prod_{t=1}^T p(\mathbf{z}_t | \mathbf{x}_t)$. Hence, Bayes' rule yields the posterior pdf proportional to

$$\begin{aligned} p(\mathbf{X}_T | \mathbf{Z}_T) &\propto p(\mathbf{Z}_T | \mathbf{X}_T) p(\mathbf{X}_T) = \prod_{t=1}^T p(\mathbf{z}_t | \mathbf{x}_t) p(\mathbf{x}_t | \mathbf{x}_{t-1}) \\ &= \prod_{t=1}^T p(\mathbf{z}_t | \mathbf{x}_t) \left(\sum_{s=1}^S w_t^s p(\mathbf{x}_t | \mathbf{x}_{t-1}; \sigma_t = s) \right) \end{aligned} \quad (3)$$

where $\sum_{s=1}^S w_t^s = 1$, and $w_t^s \in \{0, 1\}$, with $w_t^s = 1$ indicating that topology s is active at time slot t . To stress the active

topology present, we abused notation by explicitly incorporating $\sigma_t = s$ in $p(\mathbf{x}_t | \mathbf{x}_{t-1})$. The conditional likelihood $p(\mathbf{z}_t | \mathbf{x}_t)$ and the transition pdf $p(\mathbf{x}_t | \mathbf{x}_{t-1}; \sigma_t = s)$ are Gaussian; that is, $p(\mathbf{z}_t | \mathbf{x}_t) = \mathcal{N}(\mathbf{z}_t; \mathbf{H}_t \mathbf{x}_t, \mathbf{R})$ and $p(\mathbf{x}_t | \mathbf{x}_{t-1}; \sigma_t = s) = \mathcal{N}(\mathbf{x}_t; \mathbf{F}_t^s \mathbf{x}_{t-1}, \mathbf{K}_t^s)$. Thus, the maximum a posteriori estimate of the state is given in batch form by (cf. (3))

$$\begin{aligned} \arg \min_{\substack{\{\mathbf{x}_t\}_{t=1}^T \\ \{\{w_t^s\}_{s=1}^S\}_{t=1}^T}} & \frac{1}{2} \sum_{t=1}^T [\|\mathbf{z}_t - \mathbf{H}_t \mathbf{x}_t\|_{\mathbf{R}}^2 + \sum_{s=1}^S w_t^s \|\mathbf{x}_t - \mathbf{F}_t^s \mathbf{x}_{t-1}\|_{\mathbf{K}_t^s}^2] \\ \text{s.to } & w_t^s \in \{0, 1\}, \quad \sum_{s=1}^S w_t^s = 1. \end{aligned} \quad (4)$$

Unfortunately, (4) is a mixed integer program and thus computationally prohibitive to solve, especially in an online setup, where estimates of $\mathbf{x}_t$ and σ_t are sought on-the-fly.

Aiming at a computationally efficient online scheme, we will innovate the interacting multi-model (IMM) algorithm [3] that has been applied to target tracking [11] and air traffic control [10], but without graph-related information. Our graph-aware algorithm is naturally termed interacting multi-graph model (IMGM), and takes into account dynamically switching topologies. Given partially observed nodal samples $\mathbf{z}_t$, IMGM offers a scalable Bayesian scheme for tracking not only all nodal processes in $\mathbf{x}_t$, but also the discrete network mode σ_t per slot t online.

Our IMGM replaces the hard constraint $w_t^s \in \{0, 1\}$ with the soft one $w_t^s \in [0, 1]$. This allows one to think of w_t^s as the posterior probability of mode s being active at slot t given $\mathbf{Z}_t$, namely $w_t^s = \Pr(\sigma_t = s | \mathbf{Z}_t)$. We further model the evolving mode σ_t as a first-order Markov chain parameterized by the $S \times S$ mode transition matrix $\boldsymbol{\Pi}$, whose (i, j) th entry $\pi_{ij} = \Pr(\sigma_t = i | \sigma_{t-1} = j)$ denotes the transition probability from mode j at slot $t-1$ to mode i at slot t .

IMGM leverages the current observation $\mathbf{z}_t$ to propagate the posterior $p(\mathbf{x}_{t-1} | \mathbf{Z}_{t-1})$ to $p(\mathbf{x}_t | \mathbf{Z}_t)$. Based on Bayes' rule and the total probability theorem (TPT), the posterior pdf is

$$\begin{aligned} p(\mathbf{x}_t | \mathbf{Z}_t) &= \sum_{s=1}^S \Pr(\sigma_t = s | \mathbf{Z}_t) p(\mathbf{x}_t | \sigma_t = s, \mathbf{Z}_t) \\ &\approx \sum_{s=1}^S w_t^s \mathcal{N}(\mathbf{x}_t; \hat{\mathbf{x}}_{t|t}^s, \mathbf{P}_{t|t}^s) \end{aligned} \quad (5)$$

where we approximated the mode-conditional posterior of $\mathbf{x}_t$ with a Gaussian pdf having mean $\hat{\mathbf{x}}_{t|t}^s$ and covariance matrix $\mathbf{P}_{t|t}^s$. We will henceforth suppose that $p(\mathbf{x}_t | \mathbf{Z}_t)$ adheres to an exact Gaussian mixture (GM) pdf parameterized by the set $\mathcal{P}_t := \{w_t^s, \hat{\mathbf{x}}_{t|t}^s, \mathbf{P}_{t|t}^s, s = 1, \dots, S\}$. This GM model facilitates the propagation from $p(\mathbf{x}_{t-1} | \mathbf{Z}_{t-1})$ to $p(\mathbf{x}_t | \mathbf{Z}_t)$ through closed-form updates of the elements in $\mathcal{P}_{t-1}$ to those in $\mathcal{P}_t$. These updates are implemented using the prediction and correction steps described next.

Prediction. Given $\mathcal{P}_{t-1}$, the mode-conditioned predicted

state pdf at slot t can be expressed using the TPT as

$$p(\mathbf{x}_t|\sigma_t = s', \mathbf{Z}_{t-1}) = \quad (6)$$

$$\sum_{s=1}^S \Pr(\sigma_{t-1} = s|\sigma_t = s', \mathbf{Z}_{t-1}) p(\mathbf{x}_t|\sigma_t = s', \sigma_{t-1} = s, \mathbf{Z}_{t-1})$$

where the first factor $\Pr(\sigma_{t-1} = s|\sigma_t = s', \mathbf{Z}_{t-1}) := w_{t-1|t}^{s|s'}$ can be viewed as a backward mode transition probability, while the second factor is the predicted state pdf conditioned on mode s' at slot t and mode s at slot $t-1$. Upon appealing to Bayes' rule and the TPT, the first factor boils down to

$$\begin{aligned} w_{t-1|t}^{s|s'} &= \frac{\Pr(\sigma_{t-1} = s|\mathbf{Z}_{t-1}) \Pr(\sigma_t = s'|\sigma_{t-1} = s, \mathbf{Z}_{t-1})}{\sum_{s=1}^S \Pr(\sigma_{t-1} = s|\mathbf{Z}_{t-1}) \Pr(\sigma_t = s'|\sigma_{t-1} = s, \mathbf{Z}_{t-1})} \\ &= \frac{w_{t-1}^s \pi_{s's}}{\sum_{s=1}^S w_{t-1}^s \pi_{s's}}. \end{aligned} \quad (7)$$

As for the second factor in (6), state equation (2) implies that

$$p(\mathbf{x}_t|\sigma_t = s', \sigma_{t-1} = s, \mathbf{Z}_{t-1}) = \mathcal{N}(\mathbf{x}_t; \hat{\mathbf{x}}_{t|t-1}^{s',s}, \mathbf{P}_{t|t-1}^{s',s})$$

where the mean and the covariance for the predicted state are respectively obtained as

$$\hat{\mathbf{x}}_{t|t-1}^{s',s} = \mathbf{F}_t^{s'} \hat{\mathbf{x}}_{t-1|t-1}^s \quad (8a)$$

$$\mathbf{P}_{t|t-1}^{s',s} = \mathbf{F}_t^{s'} \mathbf{P}_{t-1|t-1}^s (\mathbf{F}_t^{s'})^\top + \mathbf{K}_t^{s'}. \quad (8b)$$

Although (7) and (8) yield the predicted GM pdf $p(\mathbf{x}_t|\sigma_t = s', \mathbf{Z}_{t-1})$, evolving it to its posterior in (6) is challenging, simply because a GM pdf is a non-Gaussian pdf. To arrive at a computationally tractable mode-conditioned Gaussian posterior, we will approximate (6) by the following Gaussian pdf

$$p(\mathbf{x}_t|\sigma_t = s', \mathbf{Z}_{t-1}) \approx \mathcal{N}(\mathbf{x}_t; \hat{\mathbf{x}}_{t|t-1}^{s'}, \mathbf{P}_{t|t-1}^{s'}) \quad (9)$$

where $\hat{\mathbf{x}}_{t|t-1}^{s'}$ and $\mathbf{P}_{t|t-1}^{s'}$ are chosen to minimize the Kullback-Leibler divergence between the pdfs in (9) and (6) that gives

$$\hat{\mathbf{x}}_{t|t-1}^{s'} = \sum_{s=1}^S w_{t-1|t}^{s|s'} \hat{\mathbf{x}}_{t|t-1}^{s',s} \quad (10a)$$

$$\begin{aligned} \mathbf{P}_{t|t-1}^{s'} &= \sum_{s=1}^S w_{t-1|t}^{s|s'} \left(\mathbf{P}_{t|t-1}^{s',s} \right. \\ &\quad \left. + (\hat{\mathbf{x}}_{t|t-1}^{s',s} - \hat{\mathbf{x}}_{t|t-1}^{s'}) (\hat{\mathbf{x}}_{t|t-1}^{s',s} - \hat{\mathbf{x}}_{t|t-1}^{s'})^\top \right). \end{aligned} \quad (10b)$$

Correction. Given the new observation $\mathbf{z}_t$, the predicted state pdf specified by (10) is propagated via Bayes rule as

$$\begin{aligned} p(\mathbf{x}_t|\sigma_t = s', \mathbf{Z}_t) &= p(\mathbf{x}_t|\sigma_t = s', \mathbf{z}_t, \mathbf{Z}_{t-1}) \\ &= \frac{p(\mathbf{x}_t|\sigma_t = s', \mathbf{Z}_{t-1}) p(\mathbf{z}_t|\mathbf{x}_t, \sigma_t = s', \mathbf{Z}_{t-1})}{p(\mathbf{z}_t|\sigma_t = s', \mathbf{Z}_{t-1})} \end{aligned} \quad (11)$$

where $p(\mathbf{z}_t|\mathbf{x}_t, \sigma_t = s', \mathbf{Z}_{t-1}) = p(\mathbf{z}_t|\mathbf{x}_t)$ by independence. Since $p(\mathbf{x}_t|\sigma_t = s', \mathbf{Z}_{t-1})$ and $p(\mathbf{z}_t|\mathbf{x}_t)$ are Gaussian, $p(\mathbf{x}_t|\sigma_t = s', \mathbf{Z}_t)$ will also be Gaussian with the first two

moments in (12d) and (12e) given by Kalman updates [2]

$$\hat{\mathbf{z}}_{t|t-1}^{s'} = \mathbf{H}_t \hat{\mathbf{x}}_{t|t-1}^{s'} \quad (12a)$$

$$\Phi_t^{s'} = \mathbf{H}_t \mathbf{P}_{t|t-1}^{s'} (\mathbf{H}_t)^\top + \mathbf{R} \quad (12b)$$

$$\mathbf{G}_t^{s'} = \mathbf{P}_{t|t-1}^{s'} (\mathbf{H}_t)^\top (\Phi_t^{s'})^{-1} \quad (12c)$$

$$\hat{\mathbf{x}}_{t|t}^{s'} = \hat{\mathbf{x}}_{t|t-1}^{s'} + \mathbf{G}_t^{s'} (\mathbf{z}_t - \hat{\mathbf{z}}_{t|t-1}^{s'}) \quad (12d)$$

$$\mathbf{P}_{t|t}^{s'} = \mathbf{P}_{t|t-1}^{s'} - \mathbf{G}_t^{s'} \Phi_t^{s'} (\mathbf{G}_t^{s'})^\top. \quad (12e)$$

The mode probabilities are then updated as

$$\begin{aligned} w_t^{s'} &= \Pr(\sigma_t = s'|\mathbf{z}_t, \mathbf{Z}_{t-1}) \\ &= \frac{p(\mathbf{z}_t|\sigma_t = s', \mathbf{Z}_{t-1}) \Pr(\sigma_t = s'|\mathbf{Z}_{t-1})}{\sum_{s'=1}^S p(\mathbf{z}_t|\sigma_t = s', \mathbf{Z}_{t-1}) \Pr(\sigma_t = s'|\mathbf{Z}_{t-1})} \end{aligned} \quad (13)$$

where $p(\mathbf{z}_t|\sigma_t = s', \mathbf{Z}_{t-1}) = \mathcal{N}(\mathbf{z}_t; \hat{\mathbf{z}}_{t|t-1}^{s'}, \Phi_t^{s'})$ from (12a) and (12b), and the predicted mode probability is given by

$$\begin{aligned} \Pr(\sigma_t = s'|\mathbf{Z}_{t-1}) &= \sum_{s=1}^S \Pr(\sigma_t = s', \sigma_{t-1} = s|\mathbf{Z}_{t-1}) \\ &= \sum_{s=1}^S \Pr(\sigma_t = s'|\sigma_{t-1} = s, \mathbf{Z}_{t-1}) \Pr(\sigma_{t-1} = s|\mathbf{Z}_{t-1}) \\ &= \sum_{s=1}^S \pi_{s's} w_{t-1}^s. \end{aligned} \quad (14)$$

Finally, the wanted posterior pdf is

$$p(\mathbf{x}_t|\mathbf{Z}_t) = \sum_{s'=1}^S w_t^{s'} \mathcal{N}(\mathbf{x}_t; \hat{\mathbf{x}}_{t|t}^{s'}, \mathbf{P}_{t|t}^{s'}) \approx \mathcal{N}(\mathbf{x}_t; \hat{\mathbf{x}}_{t|t}, \mathbf{P}_{t|t})$$

where the single Gaussian approximant of GM has moments

$$\begin{aligned} \hat{\mathbf{x}}_{t|t} &= \sum_{s'=1}^S w_t^{s'} \hat{\mathbf{x}}_{t|t}^{s'} \\ \mathbf{P}_{t|t} &= \sum_{s'=1}^S w_t^{s'} \left(\mathbf{P}_{t|t}^{s'} + (\hat{\mathbf{x}}_{t|t}^{s'} - \hat{\mathbf{x}}_{t|t}) (\hat{\mathbf{x}}_{t|t}^{s'} - \hat{\mathbf{x}}_{t|t})^\top \right). \end{aligned}$$

4. NUMERICAL TESTS

In this section, we test the performance of IMGGM using synthetic and real dynamic graph processes. IMGGM is compared with existing algorithms including kernel Kalman filter (KKF) [13], the adaptive least mean-square algorithm [4], and distributed least-squares reconstruction (DLSR) [17], where the last two are adaptive algorithms to track slow-varying B-bandlimited graph signals. The competing algorithms know the active network mode per slot t , whereas IMGGM estimates σ_t on-the-fly. The performance metric is the normalized mean-square error (NMSE) over unobserved nodes, which is given by $\text{NMSE}(t) := \|\mathbf{H}_t^c (\hat{\mathbf{x}}_{t|t} - \mathbf{x}_t)\|_2^2 / \|\mathbf{H}_t^c \mathbf{x}_t\|_2^2$, where $\mathbf{H}_t^c$ is the sampling matrix for the unobserved nodes.

A dynamic process is generated over a graph having $N = 100$ nodes, and $S = 2$ modes corresponding to topologies

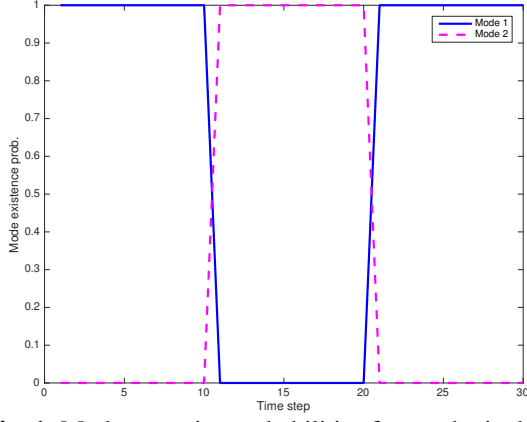


Fig. 1. Mode posterior probabilities for synthetic data.

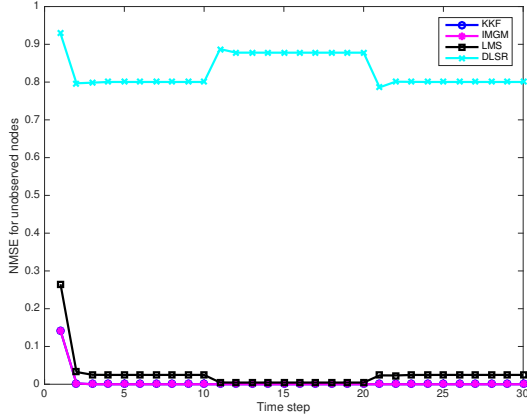


Fig. 2. NMSE for synthetic data ($\mu_{\text{LMS}} = 2$, $B_{\text{LMS}} = 2$, $\mu_{\text{DLSR}} = 2$, $B_{\text{DLSR}} = 50$, $\beta_{\text{DLSR}} = 0.2$).

obtained by two symmetric Erdős-Rényi random graphs with edge existence probabilities 0.3 and 0.7, respectively. Process $\mathbf{x}_t$ is generated according to (2) with $\mathbf{F}_t^{\sigma_t} = 0.4(\mathbf{A}^{\sigma_t} + \mathbf{I}_N)$, and $\mathbf{K}_t^{\sigma_t}$ a bandlimited kernel with $B = 50$ and $\beta = 100$ (see in [6, Table I]). The network switches from mode 1 to mode 2 at slot 11, and switches back to mode 1 at slot 21 over a total $T = 30$ slots. The observations adhere to (1) with $M = 50$, $\mathbf{H}_t = [\mathbf{I}_M, \mathbf{0}_{M, N-M}]$ and $\mathbf{R} = 3^2 \mathbf{I}_M$. To assess the average performance, 100 Monte-Carlo runs are conducted. Fig. 1 depicts $\{w_t^s\}_{s=1}^2$ found by IMGM, and demonstrates how efficiently IMGM tracks the active modes. Fig. 2 shows that IMGM's NMSE is comparable to that of KKF, which relies on extra information, while it outperforms LMS and DLSR.

ECoG brain data. Here we test the IMGM performance using ECoG data obtained from an epilepsy study [9]. The ECoG time series were obtained from $N = 76$ electrodes implanted in a patient's brain before and after a seizure, where the onset of the seizure was identified by a neurophysiologist. Therefore, there are $S = 2$ modes, the pre-ictal and ictal mode that correspond to before and after the seizure. We extract 250 samples from the dataset for each of the two modes, which are preprocessed by subtracting the sample mean and normalizing by the sample standard deviation. The preprocessed samples are then concatenated, i.e., $\sigma_t = 1$ for $t = 1, \dots, 250$ and

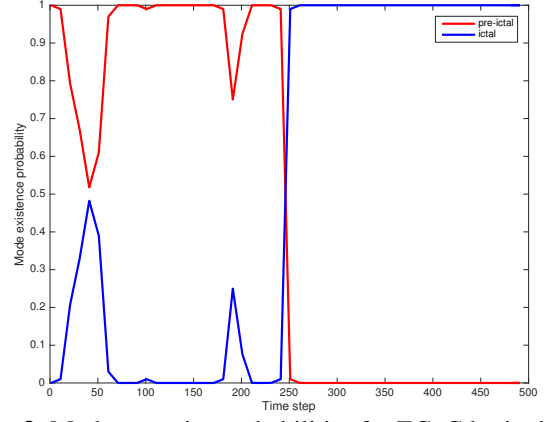


Fig. 3. Mode posterior probabilities for ECoG brain data.

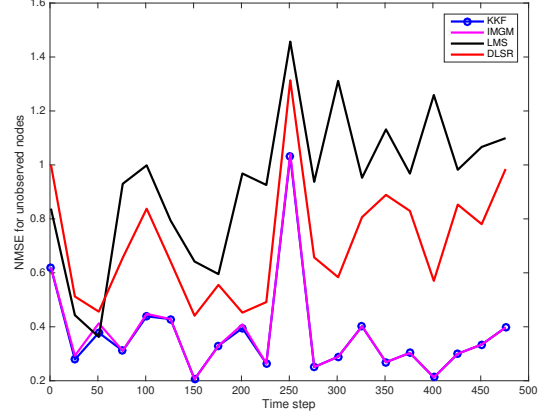


Fig. 4. NMSE for ECoG brain data ($\mu_{\text{LMS}} = 0.6$, $B_{\text{LMS}} = 2$, $\mu_{\text{DLSR}} = 1.2$, $B_{\text{DLSR}} = 6$, $\beta_{\text{DLSR}} = 0.5$).

$\sigma_t = 2$ for $t = 251, \dots, 500$. A time-invariant symmetric correlation graph is generated for each of the two modes. The ECoG signals are modeled to evolve based on (2), where the transition function $\mathbf{F}_t^{\sigma_t} = 0.15(\mathbf{A}^{\sigma_t} + \mathbf{I}_N)$, and process noise covariance $\mathbf{K}_t^{\sigma_t}$ is a diffusion kernel with parameter $\sigma = 2$ (see Table I in [6]). The observations are generated as in (1) with $M = 53$ and $\mathbf{R} = 10^{-2} \mathbf{I}_M$. Matrix $\mathbf{H}_t$ is invariant over $T = 500$ slots. The performance is averaged over 100 random sample realizations. Fig. 3 shows the IMGM probabilities $\{w_t^s\}_{s=1}^2$. Here, IMGM acts as a “neurophysiologist” that detects the onset of an epileptic seizure. In addition, the NMSE of IMGM is comparable to that of the mode-clairvoyant KKF, while markedly outperforming the other two alternatives.

5. CONCLUSIONS

This paper dealt with tracking dynamic graph processes that evolve over switching graph topologies. Given observations at a subset of nodes and candidate mode-conditioned topologies, a scalable Bayesian algorithm, termed IMGM, was introduced to learn the dynamic graph processes and discrete network modes online. Numerical tests on synthetic and real data corroborated the performance of the IMGM algorithm.

Acknowledgement. This work was supported by NSF grants 1442686, 1508993, and 1711471.

6. REFERENCES

- [1] B. Baingana and G. B. Giannakis, "Tracking switched dynamic network topologies from information cascades," *IEEE Transactions on Signal Processing*, vol. 65, no. 4, pp. 985–997, 2017.
- [2] Y. Bar-Shalom, X. R. Li, and T. Kirubarajan, *Estimation with Applications to Tracking and Navigation: Theory Algorithms and software*. John Wiley & Sons, 2004.
- [3] H. A. Blom and Y. Bar-Shalom, "The interacting multiple model algorithm for systems with Markovian switching coefficients," *IEEE Transactions on Automatic Control*, vol. 33, no. 8, pp. 780–783, 1988.
- [4] P. Di Lorenzo, S. Barbarossa, P. Banelli, and S. Sardellitti, "Adaptive least mean-squares estimation of graph signals," *IEEE Transactions on Signal and Information Processing over Networks*, vol. 2, no. 4, pp. 555–568, 2016.
- [5] G. B. Giannakis, Y. Shen, and G. V. Karanikolas, "Topology identification and learning over graphs: Accounting for nonlinearities and dynamics," *Proceedings of the IEEE*, vol. 106, no. 5, pp. 787–807, 2018.
- [6] V. N. Ioannidis, D. Romero, and G. B. Giannakis, "Inference of spatio-temporal functions over graphs via multikernel kriged Kalman filtering," *IEEE Transactions on Signal Processing*, vol. 66, no. 12, pp. 3228–3239, 2018.
- [7] E. D. Kolaczyk, *Statistical Analysis of Network Data: Methods and Models*. Springer, 2009.
- [8] R. I. Kondor and J. Lafferty, "Diffusion kernels on graphs and other discrete structures," in *Proc. Intl. Conf. on Machine Learning*, 2002, pp. 315–322.
- [9] M. A. Kramer, E. D. Kolaczyk, and H. E. Kirsch, "Emergent network topology at seizure onset in humans," *Epilepsy Research*, vol. 79, no. 2-3, pp. 173–186, 2008.
- [10] X.-R. Li and Y. Bar-Shalom, "Design of an interacting multiple model algorithm for air traffic control tracking," *IEEE Transactions on Control Systems Technology*, vol. 1, no. 3, pp. 186–194, 1993.
- [11] E. Mazor, A. Averbuch, Y. Bar-Shalom, and J. Dayan, "Interacting multiple model methods in target tracking: A survey," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 34, no. 1, pp. 103–123, 1998.
- [12] K. P. Murphy, *Machine Learning: A Probabilistic Perspective*. MIT press, 2012.
- [13] D. Romero, V. N. Ioannidis, and G. B. Giannakis, "Kernel-based reconstruction of space-time functions on dynamic graphs," *IEEE Journal of Selected Topics in Signal Processing*, vol. 11, no. 6, pp. 856–869, 2017.
- [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 Processing Magazine*, vol. 30, no. 3, pp. 83–98, 2013.
- [15] A. J. Smola and R. Kondor, "Kernels and regularization on graphs," in *Learning Theory and Kernel Machines*. Springer, 2003, pp. 144–158.
- [16] P. A. Traganitis and G. B. Giannakis, "Sketched subspace clustering," *IEEE Transactions on Signal Processing*, vol. 66, no. 7, pp. 1663–1675, 2018.
- [17] X. Wang, M. Wang, and Y. Gu, "A distributed tracking algorithm for reconstruction of graph signals," *IEEE Journal of Selected Topics in Signal Processing*, vol. 9, no. 4, pp. 728–740, 2015.
- [18] Q. Yu, Y. Du, J. Chen, J. Sui, T. Adali, G. D. Pearlson, and V. D. Calhoun, "Application of graph theory to assess static and dynamic brain connectivity: Approaches for building brain graphs," *Proceedings of the IEEE*, vol. 106, no. 5, pp. 886–906, 2018.