A State-Dependent Model for Identification of Time-varying Directed Graphs

Yuzhe Li, Hangjing Zhang and H. Vicky Zhao
Department of Automation, Tsinghua University, Beijing, China
E-mail: {yz-li24, hangjing23}@mails.tsinghua.edu.cn, vzhao@tsinghua.edu.cn

***Abstract*—Identifying dynamic directed graphs from graph signals has many applications in neuroscience and sociology. To identify graphs with time-varying weights, we propose a novel state-dependent model (SDM), integrating the linear time-varying sparse vector autoregressive model (SVAR) and nonlinear SVAR model. SDM models the time-varying property through feedback mechanisms that incorporate past graph signals, and extends the time-varying smoothness regularizer for model identification. We use the kernel method based on random Fourier features to formulate a convex optimization problem for identification, and propose an effective ADMM-based algorithm. Numerical results on synthetic datasets show that the proposed algorithm outperforms existing SVAR-based models, and demonstrate the importance of time-varying smoothness regularizer. Furthermore, experiments on US Senator ideology data showcase the model's interpretability, revealing state associations consistent with known political alignments.**

I. Introduction

Complex systems can be modeled as graphs and graph signals. The process of identifying graphs, which involves determining the structure of a graph from observed graph signals, has become a central focus in both graph signal processing [1], [2] and machine learning [3]. It reveals the internal structure of the complex systems and has many applications in neuroscience[4], sociology[5] and finance[6]. Also, graph identification helps improve the performance of downstream tasks such as signal prediction[7], signal sampling[8] and signal control[9].

Many models have been proposed for graph identification, e.g., smoothness-based models (SBM)[10], [11], structural equation models (SEM) [5], [12], [13] and sparse vector autoregressive model (SVAR) [14], [15]. SBM assume that the graph signal varies smoothly across the graph and identify the graph by minimizing the total variation of the graph signals. SEM focuses on modeling instantaneous directed interactions on graphs. SBM and SEM capture only the spatial relationships between nodes, making them unsuitable for modeling spatial-temporal interactions. SVAR models formulate the graph signal as a linear combination of historical graph signals, thereby encoding spatiotemporal dependencies. So far, SVAR models have been widely used in identifying graphs from the time series of the graph signals [16]. In various scenarios like weather networks, graphs are dynamic with time-varying weights and topology. These changes can be attributed to various factors, such as time [17], random noises [18], and dynamically changing graph signals [19]. Many works have extended the aforementioned models to a time-varying framework [20]–[22]. In these works, the time-varying graphical Lasso has gained popularity for learning dynamic graphs, and it models the time-varying graphs by regularizes the temporal variation of graph weights.

All the above traditional SVAR models model the interactions in a linear form, failing to discover nonlinear dependencies among nodes. Nonlinear SVAR models [4], [23], [24] were proposed to tackle this issue, where the time-varying graph signal is modeled as a nonlinear function of past graph signals. A generalized additive model based nonlinear SVAR model was proposed in [4], [24]. In this model, the nonlinear function of the N-dimensional vector graph signal is simplified to a sum of N nonlinear functions of the scalar node signals. Given sufficient graph signal samples, the identification of nonlinear model can be solved by the kernel method. However, existing nonlinear SVAR models do not consider edge weights, which is useful in graph signal processing applications.

In this paper, we study the time-varying directed graph identification problem. We focus on the scenario that the topology is fixed and the weights are time-varying due to feedbacks in graph signals. We propose a novel **State-Dependent Model** (SDM) to model the directed graphs with time-varying weights based on observed time series of graph signals. The feedback effects make the edge weights become the nonlinear functions of graph signals, hence the corresponding identification is beyond existing linear SVAR framework. To address the nonlinear function identification problem, our approach uses random Fourier features (RFF)-based kernel learning method to formulate a convex optimization problem for identification, which we solve via an ADMM-based algorithm (RF-SDM-Id). We use simulations and real data to validate the proposed state-dependent model and the effectiveness of the proposed RF-SDM-Id algorithm.

II. Problem Formulation

A. Sparse Vector Autoregressive Framework

Consider a directed graph denoted by $\mathcal{G}(\mathcal{V},\mathcal{E})$, where $\mathcal{V}$ is the set of N nodes, $\mathcal{E}$ is the set of edges. Graph signals observed on nodes are time-varying, denoted as $\mathbf{x}(t) := [x_1(t),\ldots,x_N(t)]^\top \in \mathbb{R}^N$ at each time stamp t. Sparse vector autoregressive (SVAR) model [15] is a useful tool to model time-varying graph signals by capturing the temporal

979-8-3315-7206-8/25/$31.00 ©2025 IEEE

dependencies of signals in linear form:

$$\mathbf{x}(t)=\sum_{p=1}^{P}\mathbf{W}^{(p)}\mathbf{x}(t-p)+\mathbf{e}(t). \quad (1)$$

In (1), $\mathbf{e}(t)$ represents the innovation process, while $\mathbf{W}^{(p)}$ denotes a set of weighted adjacency matrices at their respective time stamps, capturing spatiotemporal dependencies. Each entry $w_{n,n'}^{(p)}$ in $\mathbf{W}^{(p)}$ corresponds to the influence strength from node n' to node n at lag p. Crucially, a directed edge from n' to n exists if and only if there exists at least one p such that $w_{n,n'}^{(p)}>0$. The matrices $\{\mathbf{W}^{(p)}\}_p$ are sparse due to the sparsity of the topology. We assume that there are self-loops in the graph, indicating node n's graph signal $x_n(t)$ is influenced by past graph signal $x_n(t-p)$. With aforementioned concepts, the graph is denoted by $\mathcal{G}(\mathcal{V},\mathcal{E},\{\mathbf{W}^{(p)}\}_p)$.

To describe dynamic graphs, SVAR is extended to the time-varying form [21]:

$$\mathbf{x}(t)=\sum_{p=1}^{P}\mathbf{W}^{(p)}(t)\mathbf{x}(t-p)+\mathbf{e}(t). \quad (2)$$

Time-varying graphical smoothness regularizer $\|\mathbf{W}(t)-\mathbf{W}(t-1)\|_F^2$ was proposed for the above time-varying model, which allows a smooth transition between adjacent timestamps [20].

In this paper, we assume that topology is fixed in given time period, while the weights are time-varying. The topology may change in a long term, and it can be solved in a online fashion, which is beyond the scope of this paper. Our problem can be expressed as follows: given the observed graph signals at every time stamps $\mathcal{X}:=\{\mathbf{x}(t)\}_{t=0}^{T}$, we need to estimate a dynamic directed graph with static $\mathcal{E}$ and time-varying weights$\{\mathbf{W}^{(p)}(t)\}_{t=1}^{T}$.

B. The Proposed State-Dependent Model (SDM)

Changes in weights can be attributed to factors such as noise, random variables or time-varying deterministic variables. In this paper, we assume that the weights are influenced by past graph signals and formulate graph signals as:

$$\mathbf{x}(t)=\sum_{p=1}^{P}\mathbf{W}^{(p)}(\mathbf{x}(t-1),\cdots,\mathbf{x}(t-p))\mathbf{x}(t-p)+\mathbf{e}(t). \quad (3)$$

Note that (3) is actually a nonlinear model expressed in linear time-varying form, we adopt a simplifying method similar to those in [24], [25]. We assume that the weight $w_{n,n'}^{(p)}(\cdot)$ is a function of node n''s signals $\{x_{n'}(t-1),\dots,x_{n'}(t-p)\}$, implying that the current weights $w_{n,n'}^{(p)}$ are affected by past signals of node n'. Because of the weights' dependencies on past signals (system states), we name this model "**State-Dependent Model**" (SDM). Then, we propose two special SDMs for graph identification.

Type I. In SDM-I, the weights $\{w_{n,n'}^{(p)}(\cdot)\}_{p=1,\dots,P}$ are only dependent on the graph signals at the previous time stamp, i.e.,

$$x_n(t)=\sum_{n'=1}^{N}\sum_{p=1}^{P}w_{n,n'}^{(p)}(x_{n'}(t-1))x_{n'}(t-p)+e_n(t). \quad (4)$$

It can be viewed as an extension of the exponential AR model in [26].

Type II. In SDM-II, the weights $\{w_{n,n'}^{(p)}(\cdot)\}_{p=1,\dots,P}$ are dependent on the graph signals at multiple time stamps $\{x_{n'}(t-p)\}_p$, i.e.,

$$x_n(t)=\sum_{n=1}^{N}\sum_{p=1}^{P}w_{n,n'}^{(p)}(x_{n'}(t-p))x_{n'}(t-p)+e_n(t). \quad (5)$$

It can be viewed as a variant of the nonlinear SVAR in [24].

Note that SDM can express a wide range of nonlinear SVAR models in a pseudo-linear form. Compared with nonlinear model in [24], SDM has an explicit concept of weights, which benefit other signal processing tasks like influential node detection and portfolio selection [27]. If the functions of weights are constants, SDM is simplified to the standard linear SVAR model (1). We see that SDM unifies linear and nonlinear models within a single framework.

III. Graph Identification Problem and Algorithm

A. Identification Problem Formulation

Given the above SDM, the identification of dynamic graphs is transformed into learning the functions $\{w_{n,n'}^{(p)}(\cdot)\}_{n,n',p}$. To solve this problem, we use kernels to represent these functions. We take SDM-I as an example to show how to formulate the identification optimization problem, and the analysis of SDM-II is similar and omitted. We assume that the function $w_{n,n'}^{(p)}(\cdot)$ belongs to a reproducing kernel Hilbert space (RKHS) $\mathcal{H}_n^{(p)}$. Then the function can be expressed as the linear combination of kernel functions, i.e., $w_{n,n'}^{(p)}(\cdot)=\sum_{t=p}^{\infty}a_{n,n'}^{(p)}\kappa_{n'}^{(p)}(\cdot,x_{n'}(t-p))$. Hence, we can formulate the identification problem as:

$$\{w_{n,n'}^{(p)}\}_{n',p}=\arg\min_{w_{n,n'}^{(p)}\in\mathcal{H}_n^{(p)}} l_{LMS}+\beta\, l_{WTV}+\lambda\, l_{spars}, \quad (6)$$

where

$$l_{LMS}=\frac{1}{2(T-P)}\sum_{t=P}^{T}\left[x_n(t)-\sum_{n'=1}^{N}\sum_{p=1}^{P}w_{n,n'}^{(p)}(x_{n'}(t-1))x_{n'}(t-p)\right]^2, \quad (7)$$

$$l_{WTV}=\frac{1}{2(T-P)}\sum_{t=P}^{T}\sum_{p=1}^{P}\|\mathbf{W}_n^{(p)}(\mathbf{x}(t))-\mathbf{W}_n^{(p)}(\mathbf{x}(t-1))\|_2^2, \quad (8)$$

$$l_{spars}=\sum_{n'=1}^{N}\|[w_{n,n'}^{(1)},\cdots,w_{n,n'}^{(P)}]\|_{\mathcal{H}_n^{(p)}}. \quad (9)$$

In the objective function in (6), l_{LMS} is the least mean square loss of predicted graph signals. $\mathbf{W}_n^{(p)}$ in the l_{WTV} is defined as $\mathbf{W}_n^{(p)}=[w_{n,1}^{(p)},\cdots,w_{n,N}^{(p)}]^\top$, and l_{WTV} is a time-varying graphical smoothness regularizer[20] that reduces the weights' variation on time. The third term is a group-lasso penalty to ensure edge sparsity.

According to the Representer Theorem[28], the solution can be obtained through a finite number of kernel evaluations, i.e.:

$$w_{n,n'}^{(p)}(\cdot)=\sum_{t=p}^{p+T}a_{n,n'}^{(p)}(t-p)\kappa^{(p)}(\cdot,x_{n'}(t-p)). \quad (10)$$

As the number of kernel evaluations increases with time T, computational complexity becomes a significant concern. To address this challenge, we employ the random Fourier features (RFF) in [29] to approximate the RKHS function, fixing the optimization variables in a low-dimensional space while accommodating growing T kernel evaluations. A Gaussian kernel with variance σ^2 can be approximated by:

$$\hat{\kappa}^{(p)}(\cdot, x_{n'}(t-p)) = \mathbf{z}^{(p)}(\cdot)^\top \mathbf{z}^{(p)}(x_{n'}(t-p)), \tag{11}$$

where

$$\mathbf{z}^{(p)}(\cdot) = \frac{1}{\sqrt{D}}[\sin(\langle v_1, \cdot\rangle), \ldots, \sin(\langle v_D, \cdot\rangle), \\ \cos(\langle v_1, \cdot\rangle), \ldots, \cos(\langle v_D, \cdot\rangle)]^\top. \tag{12}$$

In (12), $\langle \cdot, \cdot \rangle$ is the inner product. Scalars $\{v_1, \ldots, v_D\}$ are sampled from a zero-mean Gaussian distribution with variance σ^{-2}. For simplicity, we denote $\mathbf{z}_{n'}^{(p)}(t) := \mathbf{z}^{(p)}(x_{n'}(t-1))$. Substituting (11) into (10), we can derive the function $w_{n,n'}^{(p)}(x_{n'}(t-1))$ as:

$$w_{n,n'}^{(p)}(x_{n'}(t-1)) = \boldsymbol{\alpha}_{n,n'}^{(p)\top} \mathbf{z}_{n'}^{(p)}(t), \tag{13}$$

where $\boldsymbol{\alpha}_{n,n'}^{(p)} = \sum_{t=p}^{p+T} a_{n,n'}^{(p)}(t-p)\mathbf{z}^{(p)}(x_{n'}(t-p)) \in \mathbb{R}^{2D}$. Given the observed graph signals $x_{n'}(t-p)$, $\mathbf{z}_{n'}^{(p)}(t)$ can be directly computed using (12). Then the function learning problem is transformed into learning the vectors $\{\boldsymbol{\alpha}_{n,n'}^{(p)}\}_{n,n',p}$.

We first define

$$\mathbf{Z}(t) = \mathrm{blkdiag}(\mathbf{z}_1^{(1)}(\mathrm{t})^\top, \ldots, \mathbf{z}_\mathrm{N}^{(1)}(\mathrm{t})^\top, \mathbf{z}_1^{(2)}(\mathrm{t})^\top, \ldots, \mathbf{z}_\mathrm{N}^{(\mathrm{P})}(\mathrm{t})^\top),$$
$$\boldsymbol{W}_n(t) = \left[w_{n,1}^{(1)}(t), \cdots, w_{n,N}^{(1)}(t), \cdots, w_{n,N}^{(P)}(t)\right]^\top,$$
$$\boldsymbol{\alpha}_n = \left[\boldsymbol{\alpha}_{n,1}^{(1)\top}, \cdots, \boldsymbol{\alpha}_{n,N}^{(1)\top}, \boldsymbol{\alpha}_{n,1}^{(2)\top}, \cdots, \boldsymbol{\alpha}_{n,N}^{(P)\top}\right]^\top, \tag{14}$$

where $\mathrm{blkdiag}(\cdot)$ denotes the block-diagonal operator, and we introduce the shorthand notation $\boldsymbol{W}n(t) := \boldsymbol{W}n(\mathbf{x}(t-1), \cdots, \mathbf{x}(t-p))$ for compactness. Therefore, we have $\boldsymbol{W}_n(t) = \mathbf{Z}(t)\boldsymbol{\alpha}_n$. Define $\boldsymbol{\alpha}_{n,n'} = [\boldsymbol{\alpha}_{n,n'}^{(1)}, \cdots, \boldsymbol{\alpha}_{n,n'}^{(P)}]^\top$ and $\tilde{\mathbf{x}}(t-1) = [\mathbf{x}(t-1)^\top, \cdots, \mathbf{x}(t-P)^\top]^\top$. The identification problem of SDM-I can be rewritten as:

$$\hat{\boldsymbol{\alpha}}_n = \arg\min_{\boldsymbol{\alpha}_n} l_{LMS}^\alpha + \beta\, l_{WTV}^\alpha + \lambda\, l_{spars}^\alpha, \tag{15}$$

where

$$l_{LMS}^\alpha = \sum_{t=P}^{T}\left[x_n(t) - \tilde{\mathbf{x}}^\top(t-1)\mathbf{Z}(t)\boldsymbol{\alpha}_n\right]^2, \tag{16}$$

$$l_{WTV}^\alpha = \frac{1}{2(T-P)}\sum_{t=P+1}^{T}\|\mathbf{Z}(t)\boldsymbol{\alpha}_n - \mathbf{Z}(t-1)\boldsymbol{\alpha}_n\|_2^2, \tag{17}$$

$$l_{spars}^\alpha = \sum_{n'=1}^{N}\|\boldsymbol{\alpha}_{n,n'}\|_2. \tag{18}$$

The identification optimization formulation for SDM-II shares the same form as (15), with a slight difference in the calculation of $\mathbf{Z}(t)$, i.e. $\mathbf{z}_{n'}^{(p)}(t) = \mathbf{z}^{(p)}(x_{n'}(t-p))$. It arises from the different independent variables of the function $w_{n,n'}^{(p)}(\cdot)$,

The optimization problem in (15) for both SDM-I and SDM-II can be further expressed as an unconstrained convex quadratic optimization problem $QP(\boldsymbol{\Phi}, \mathbf{g}_n)$ with group Lasso regularizer, which can be expressed as:

$$\min_{\boldsymbol{\alpha}_n} \frac{1}{2}\boldsymbol{\alpha}_n^\top \boldsymbol{\Phi}\boldsymbol{\alpha}_n + \mathbf{g}_n^\top \boldsymbol{\alpha}_n + \lambda \sum_{n'=1}^{N}\sum_{p=1}^{P}\|\boldsymbol{\alpha}_{n,n'}\|_2, \tag{19}$$

where

$$\boldsymbol{\Phi} = \frac{\sum_{t=P}^{T}\left[\mathbf{Z}(t)^\top \tilde{\mathbf{x}}(t-1)\tilde{\mathbf{x}}^\top(t-1)\mathbf{Z}(t) + \beta\Big(\mathbf{Z}(t+1) - \mathbf{Z}(t)\Big)\right]}{2(T-P)}, \tag{20}$$

and

$$\mathbf{g}_n = -\frac{1}{T-P}\sum_{t=P}^{T} x_n(t)\mathbf{Z}(t)^\top \tilde{\mathbf{x}}(t-1). \tag{21}$$

B. Identification Algorithm

We propose an ADMM-based algorithm [30] to solve (15). Using the ρ-augmented Lagrangian method, we introduce auxiliary variable $\boldsymbol{\omega}_n$ and rewrite the problem as:

$$\min_{\boldsymbol{\alpha}_n, \boldsymbol{\omega}_n} \frac{1}{2}\boldsymbol{\alpha}_n^\top \boldsymbol{\Phi}\boldsymbol{\alpha}_n + \mathbf{g}_n^\top \boldsymbol{\alpha}_n + \lambda \sum_{n'=1}^{N}\sum_{p=1}^{P}\|\boldsymbol{\omega}_{n,n'}\|_2 \\ + \mathbf{y}^\top(\boldsymbol{\alpha}_n - \boldsymbol{\omega}_n) + \frac{\rho}{2}\|\boldsymbol{\alpha}_n - \boldsymbol{\omega}_n\|_2^2, \tag{22}$$

where $\mathbf{y} \in \mathbb{R}^N$ is the dual variable. Using this split, the update of $\boldsymbol{\omega}_n$ can be easily solved by MSTO [31]. The variables iterates according to the following formulas:

$$\boldsymbol{\alpha}_n^{k+1} = -(\boldsymbol{\Phi} + \rho\mathbf{I})^{-1}(\mathbf{g}_n + \mathbf{y}^k - \rho\boldsymbol{\omega}_n^k)^\top, \tag{23}$$

$$\boldsymbol{\omega}_{n,n'}^{k+1} = \frac{1}{\rho}(\mathbf{y}^k + \rho\boldsymbol{\alpha}_{n,n'}^{k+1})\left(1 - \frac{\lambda}{\|\mathbf{y}^k + \rho\boldsymbol{\alpha}_{n,n'}^{k+1}\|_2}\right)_+, \tag{24}$$

$$\mathbf{y}^{k+1} = \mathbf{y}^k + \rho(\boldsymbol{\alpha}_n^{k+1} - \boldsymbol{\omega}_n^{k+1}), \tag{25}$$

where $(x)_+ = \max\{0, x\}$.

After deriving $\hat{\boldsymbol{\alpha}}_n$, we use a threshold δ to classify edges. Define normalized strength of presence $\hat{e}_{n,n'}$ as:

$$\hat{e}_{n,n'} = \|\hat{\boldsymbol{\alpha}}_{n,n'}\|_2 / (\max_{n'}\|\hat{\boldsymbol{\alpha}}_{n,n'}\|_2). \tag{26}$$

If $\hat{e}_{n,n'} >= \delta$, the edge exists.

We summarize the process of identifying dynamic graphs in Algorithm 1 named "Random-Feature-based SDM Identification" (RF-SDM-Id).

Algorithm 1: RF-SDM-Id

Input: Observed graph signals $\{\mathbf{x}_t\}_{t=0}^{T}$, λ, β
Output: $\{\hat{\boldsymbol{\alpha}}_n\}_{n=1}^{N}$, $\hat{\mathcal{E}}$, $\{\widehat{\mathbf{W}}(t)\}_{t=P}^{N}$

```
Compute {Z(t)}_{t=P}^{T} via (12),(14)
for n = 1, ..., N do
    while not convergence do
        Update α_n via (18)
        for n' = 1, ..., N do
            Update ω_{n,n'} via (19)
        end for
        Update y via (20)
    end while
    Compute {Ŵ(t)}_{t=P}^{N} via (12)
    Compute Ê via (21)
end for
```

IV. Numerical Results

A. Simulations

We use simulations to verify the effectiveness of the proposed identification criterion and RF-SDM-Id algorithm. The simulation dataset is constructed as follows. We first construct an Erdős-Rényi random graph with N=20 nodes, where each possible edge is included independently with probability p=0.2, and augment the graph by adding self-loops to all nodes. Then, the weight functions are set as $w_{n,n'}^{(p)}(\cdot) = c_{n,n'}^{(p)}[1 + k\sin(\cdot)]$ with order $P = 2$, where $\{c_{n,n'}^{(p)}\}_{n,n',p}$ is a set of stable time-invariant VAR coefficients based on the topology and k controls the nonlinearity. Finally, we generate $T = 400$ intervals of time-varying graph signals using the aforementioned SDM. In our experiment, we prepare 20 graphs for each kind of SDM setting: SDM-I and SDM-II, with $k = 0.15$ and $k = 0.3$.

We evaluate the performance from the following two aspects: weights estimation error (normalized mean square error $\mathrm{NMSE_w}$) and edge classification AUC (AUC). $\mathrm{NMSE_w}$ is defined as:

$$\mathrm{NMSE_w} = \frac{\sum_{\mathrm{t=P}}^{\mathrm{T}} \|\widehat{\mathbf{W}}(\mathrm{t}) - \mathbf{W}(\mathrm{t})\|_2^2}{\sum_{\mathrm{t=P}}^{\mathrm{T}} \|\mathbf{W}(\mathrm{t})\|_{\mathrm{F}}^2} \tag{27}$$

We compare our model and identification algorithm with time-varying SVAR model in (2) (TVSVAR) [21], time-invariant SVAR model in (1) (TISVAR) and GAM-based nonlinear SVAR model (Nonliear) [24]. We construct identification algorithms for TVSVAR, TISVAR and Nonlinear using similar ADMM-based method respectively, since their identification is also a convex optimization problem. It is crucial to highlight that the proposed SDM framework fundamentally combines time-varying SVAR model and nonlinear SVAR model. Comparing SDM with TVSVAR and nonlinear models highlights the necessity of our proposed model in capturing state-dependent properties. The settings of some important parameters are as follows: $D = 15, \beta = 10^{-3}, \lambda = 5\times10^{-5}$ for $k = 0.3$, and $D = 15, \beta = 5\times10^{-3}, \lambda = 5\times10^{-5}$ for $k = 0.15$. Following the approach proposed in [21], we use a data window of 10 intervals for TVSVAR identification.

TABLE I
Identification Performance Comparison.

Metrics	Data Type	Proposed	TVSVAR	TISVAR	Nonlinear
$\mathrm{NMSE_w}$	I, $k = 0.3$	**0.083**	0.809	1.322	-
	I, $k = 0.15$	**0.100**	0.691	0.747	-
	II, $k = 0.3$	**0.177**	0.995	2.881	-
	II, $k = 0.15$	**0.184**	0.712	0.691	-
AUC	I, $k = 0.3$	**0.954**	0.721	0.813	-
	I, $k = 0.15$	**0.944**	0.730	0.834	-
	II, $k = 0.3$	**0.920**	0.679	0.738	0.656
	II, $k = 0.15$	**0.900**	0.688	0.787	0.646

Table I shows the simulation results, and it can be seen that our proposed model and algorithm achieve the best identification performance on the synthetic dataset we generate. We can also see that the weight estimation error $\mathrm{NMSE_w}$ of the identification algorithm based on TISVAR is significantly larger than k, in spite of the synthetic weight functions being around constant values. This discrepancy stems from the fundamental mismatch between the nonlinear dynamics of the graph signals and the linear assumptions inherent in the TISVAR framework.

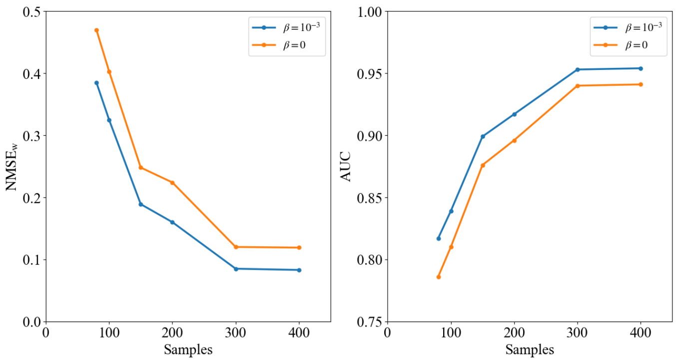

Fig. 1. Time-varying Graphical Smoothness Regularizer Ablation Experiment Results. Left: weights estimation error; Right: edges classification AUC. Parameter settings: $\beta = 10^{-3}$ and $\lambda = 5\times10^{-5}$ in control group, $\beta = 0$ and $\lambda = 8\times10^{-5}$ in experimental group.

In addition, we conduct ablation experiment to examine the effectiveness of time-varying graphical smoothness regularizer. Fig. 1 shows that with different observation samples, the identification performance with time-varying graphical smoothness regularizer is always better. Moreover, we calculate the numerical difference between control group $\beta = 10^{-3}$ and experimental group $\beta = 0$. As sampling number grows, the difference of $\mathrm{NMSE_w}$ decreases from 0.085 to 0.036, and the difference of AUC decreases from 0.031 to 0.013. This suggests that the time-varying graphical smoothness regularizer provides more pronounced performance.

B. Real-Data Tests

To validate the proposed model, this experiment utilizes US Senator ideology data in [32]. Each state is represented by two Senators in the 100-member Senate, making their policy positions and ideology reflective of their state's interests. The dataset employs the first dimension of the "nominate" score to measure each legislator's position on the liberal-conservative economic policy spectrum, ranging from -1 (liberal) to 1 (conservative). Scores are available for every Senator per Congress.

We analyze ideology data spanning 63 Congresses, from the 51st (1889-1891) to the 113th (2013-2015). The ideological association network comprises $N = 48$ states excluding two overseas states that joined the US relatively late. The graph signal of each state is computed as the mean value of two Senators' scores. With a time length $T = 63$, we set the model order $P = 2$. Random feature parameters are: dimension $D = 15$, variable $v \sim \mathcal{N}(0, 4)$, smoothness parameter $\beta = 5 \times 10^{-4}$, and sparsity parameter $\gamma = 5 \times 10^{-4}$.

As there is no absolute objective measure for interstate ideological association weights, we illustrate the model's plausibility by examining the connections of one state as the example. For clarity, we aggregate weights across different time lags for each node pair. Figure 2 shows Alabama's (AL) connection weights to other states at $t = 62$ (approx. 2011-2013).

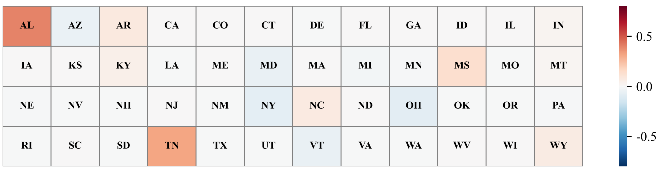

Fig. 2. Alabama's Connection Weights to Other States

Figure 2 shows that besides self-connection, AL exhibits strong positive weights with Missouri (MO), Tennessee (TN), and South Carolina (SC), indicating AL positively considers these states' economic policies. Conversely, AL shows weak negative weights with Ohio (OH), New York (NY), and Maryland (MD), suggesting policy divergence. According to Gallup polling during $t = 61,\ 62$ (approx. 2009-2013) [33], MO, TN, and SC were strongly conservative states, aligning with AL's stance. In contrast, OH was politically moderate, while NY and MD were strongly liberal states – opposing AL's orientation. This alignment supports the interpretability of our model regarding AL's economic policy considerations.

V. Conclusion

In this paper, we propose a novel state-dependent model (SDM) and corresponding identification algorithm for learning directed graphs with time-varying weights from graph signals. The SDM combines the time-varying linear SVAR models and the nonlinear SVAR models. To solve the identification problem, we utilize kernel method based on random Fourier features and then propose an ADMM-based algorithm. Comprehensive experiments on synthetic datasets, including comparative and ablation studies, demonstrate the accuracy of our identification algorithm and the effectiveness of the proposed regularizer. Our results indicate that the SDM and its identification method effectively combine the advantages of linear time-varying and nonlinear modeling approaches, enabling the accurate learning of time-varying edge weights. Furthermore, an experiment on US Senator ideology data showcases the model's interpretability on real-world datasets.

References

[1] G. Leus, A. G. Marques, J. M. Moura, A. Ortega, and D. I. Shuman, "Graph signal processing: History, development, impact, and outlook," *IEEE Signal Processing Magazine*, vol. 40, no. 4, pp. 49–60, 2023.

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

[3] M. M. Bronstein, J. Bruna, Y. LeCun, A. Szlam, and P. Vandergheynst, "Geometric deep learning: Going beyond euclidean data," *IEEE Signal Processing Magazine*, vol. 34, no. 4, pp. 18–42, 2017.

[4] Y. Shen, G. B. Giannakis, and B. Baingana, "Nonlinear structural vector autoregressive models with application to directed brain networks," *IEEE Transactions on Signal Processing*, vol. 67, no. 20, pp. 5325–5339, 2019.

[5] B. Baingana, G. Mateos, and G. B. Giannakis, "Proximal-gradient algorithms for tracking cascades over social networks," *IEEE Journal of Selected Topics in Signal Processing*, vol. 8, no. 4, pp. 563–575, 2014.

[6] D. Y. Kenett, T. Preis, G. Gur-Gershgoren, and E. Ben-Jacob, "Dependency network and node influence: Application to the study of financial markets," *International Journal of Bifurcation and Chaos*, vol. 22, no. 07, p. 1 250 181, 2012.

[7] E. Isufi, A. Loukas, N. Perraudin, and G. Leus, "Forecasting time series with varma recursions on graphs," *IEEE Transactions on Signal Processing*, vol. 67, no. 18, pp. 4870–4885, 2019.

[8] Y. Li, H. Vicky Zhao, and G. Cheung, "Eigen-decomposition-free directed graph sampling via gershgorin disc alignment," in *ICASSP 2023 - 2023 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2023, pp. 1–5. DOI: 10.1109/ICASSP49357.2023.10096903.

[9] Y. Li, Z. Chen, and H. V. Zhao, "Robust opinion control under network perturbation," *IEEE Signal Processing Letters*, vol. 29, pp. 1649–1653, 2022. DOI: 10.1109/LSP.2022.3193039.

[10] X. Dong, D. Thanou, M. Rabbat, and P. Frossard, "Learning graphs from data: A signal representation perspective," *IEEE Signal Processing Magazine*, vol. 36, no. 3, pp. 44–63, 2019.

[11] V. Kalofolias, A. Loukas, D. Thanou, and P. Frossard, "Learning time varying graphs," in *2017 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, IEEE, 2017, pp. 2826–2830.

[12] W. T. Bielby and R. M. Hauser, "Structural equation models," *Annual review of sociology*, vol. 3, pp. 137–161, 1977.

[13] A. Natali, E. Isufi, M. Coutino, and G. Leus, "Online graph learning from time-varying structural equation models," in *2021 55th Asilomar Conference on Signals, Systems, and Computers*, IEEE, 2021, pp. 1579–1585.

[14] J. Songsiri and L. Vandenberghe, "Topology selection in graphical models of autoregressive processes," *The Journal of Machine Learning Research*, vol. 11, pp. 2671–2705, 2010.

[15] A. Bolstad, B. D. Van Veen, and R. Nowak, "Causal network inference via group sparse regularization," *IEEE transactions on signal processing*, vol. 59, no. 6, pp. 2628–2641, 2011.

[16] B. Zaman, L. M. L. Ramos, D. Romero, and B. Beferull-Lozano, "Online topology identification from vector autoregressive time series," *IEEE Transactions on Signal Processing*, vol. 69, pp. 210–225, 2020.

[17] Y. Liu, C. Cui, M. Ajirak, and P. M. Djurić, "Estimation of time-varying graph topologies from graph signals," in *ICASSP 2023-2023 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, IEEE, 2023, pp. 1–5.

[18] H. Yu, S. Wu, and J. Dauwels, "Efficient variational bayes learning of graphical models with smooth structural changes," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 45, no. 1, pp. 475–488, 2022.

[19] J. Liu, M. Ye, B. D. Anderson, T. Basar, and A. Nedic, "Discrete-time polar opinion dynamics with heterogeneous individuals," in *2018 IEEE Conference on Decision and Control (CDC)*, IEEE, 2018, pp. 1694–1699.

[20] D. Hallac, Y. Park, S. Boyd, and J. Leskovec, "Network inference via the time-varying graphical lasso," in *Proceedings of the 23rd ACM SIGKDD international conference on knowledge discovery and data mining*, 2017, pp. 205–213.

[21] L. M. Lopez-Ramos, D. Romero, B. Zaman, and B. Beferull-Lozano, "Dynamic network identification from non-stationary vector autoregressive time series," in *2018 IEEE Global Conference on Signal and Information Processing (GlobalSIP)*, IEEE, 2018, pp. 773–777.

[22] K. Yamada, Y. Tanaka, and A. Ortega, "Time-varying graph learning based on sparseness of temporal variation," in *ICASSP 2019-2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, IEEE, 2019, pp. 5411–5415.

[23] N. Lim, F. d'Alché-Buc, C. Auliac, and G. Michailidis, "Operator-valued kernel-based vector autoregressive models for network inference," *Machine learning*, vol. 99, pp. 489–513, 2015.

[24] R. T. Money, J. P. Krishnan, and B. Beferull-Lozano, "Sparse online learning with kernels using random features for estimating nonlinear dynamic graphs," *IEEE Transactions on Signal Processing*, vol. 71, pp. 2027–2042, 2023.

[25] Y. Shen and G. B. Giannakis, "Online identification of directional graph topologies capturing dynamic and nonlinear dependencies," in *2018 IEEE Data Science Workshop (DSW)*, IEEE, 2018, pp. 195–199.

[26] M. Priestley, "State-dependent models: A general approach to non-linear time series analysis," *Journal of Time Series Analysis*, vol. 1, no. 1, pp. 47–71, 1980.

[27] A. Javaheri, J. Ying, D. P. Palomar, and F. Marvasti, *Time-varying graph learning for data with heavy-tailed distribution*, 2024. arXiv: 2501.00606 [cs.LG]. [Online]. Available: https://arxiv.org/abs/2501.00606.

[28] T. Hastie, *The elements of statistical learning: Data mining, inference, and prediction*, 2009.

[29] J. Lu, S. C. Hoi, J. Wang, P. Zhao, and Z.-Y. Liu, "Large scale online kernel learning," *Journal of Machine Learning Research*, vol. 17, no. 47, pp. 1–43, 2016.

[30] S. Boyd, N. Parikh, E. Chu, B. Peleato, J. Eckstein, *et al.*, "Distributed optimization and statistical learning via the alternating direction method of multipliers," *Foundations and Trends® in Machine learning*, vol. 3, no. 1, pp. 1–122, 2011.

[31] A. T. Puig, A. Wiesel, G. Fleury, and A. O. Hero, "Multidimensional shrinkage-thresholding operator and group lasso penalties," *IEEE Signal Processing Letters*, vol. 18, no. 6, pp. 363–366, 2011.

[32] J. B. Lewis, K. Poole, H. Rosenthal, A. Boche, A. Rudkin, and L. Sonnet, *Voteview: Congressional roll-call votes database*, https://voteview.com/, 2021.

[33] *Gallup polling*, https://news.gallup.com/poll/146348/mississippi-rates-conservative-state.aspx, 2024.