Cellular Autoregressive Higher-Order Models

Abdullah Canbolat
SIGIPRO Department
SimulaMet
Oslo, Norway
abdullahc@simula.no

Rohan Money
SIGIPRO Department
SimulaMet
Oslo, Norway
rohan@simula.no

Baltasar Beferull-Lozano
SIGIPRO Department
SimulaMet
Oslo, Norway
baltasar@simula.no

***Abstract*—Vector autoregressive (VAR) models have long been valued for their comprehensive representation capabilities in signal processing. However, this requires learning a matrix of parameters that grows with the number of individual time series for various time lags, making the overall estimation of a VAR model from data prohibitively complex. To address this challenge, graph, and topological signal processing approaches have introduced more structured models, such as graph VAR (G-VAR) and simplicial complex VAR (SC-VAR), which reduce the complexity by assuming that VAR models are derived from the polynomial graph and simplicial filters, respectively. In this work, we generalize structured VAR models to operate over abstract cellular complexes, allowing more flexible structural representation, which we refer to as cellular complex VAR (CC-VAR). Moreover, we propose an online learning algorithm to learn the parameters of CC-VAR. Our theoretical and experimental results demonstrate that the CC-VAR model effectively represents multivariate time series data while significantly reducing the number of parameters compared to the current state-of-the-art approaches.**

***Index Terms*—Vector autoregressive model, topological signal processing, abstract cellular complex, online learning.**

I. Introduction

Multivariate time series data generated from real-world networks are amenable to being represented using Vector Autoregressive (VAR) models [1]–[4] due to their ability to capture complex time-lagged interactions in a tractable manner [5], [6]. However, VAR models typically ignore the inherent higher-order structure present in many real-world networks, such as those found in water distributions, road transportation, and power distribution systems, including multiway dependencies. In standard VAR models, the parameter count is N^2P (where N is the number of time series and P is the time lag), which can become prohibitive as N increases. Incorporating network-like structures into VAR models introduces valuable inductive biases, leading to more accurate modeling and significantly reducing the number of parameters. In this work, we incorporate a cellular complex-based network structure into the model.

Although methods such as factor modeling [7], shrinkage estimators [8], and low-rank transformation matrices [9] have been proposed to address over-parameterization in VAR models, they often overlook the physical structure and higher-order dependencies inherent in many real-world networks. The rise of graph and topological signal processing has enabled G-VAR and SC-VAR models [10], [11], which incorporate network structures into the VAR framework using convolution-based filters over graphs and simplicial complexes, reducing the parameter count independent of vector dimension. However, the G-VAR model [10] is limited to node-level signals and overlooks higher-order relationships (e.g., edge flows) present in real datasets, while the SC-VAR model [11] may be unsuitable for networks not conforming to more complex structures, such as rectangular water pipelines and online learning of those models in the literature includes redundant parameters. Moreover, real-world datasets are often non-stationary, and sufficient batch data may not always be available (e.g., fluctuating users in wireless networks). To address this, the TopoLMS algorithm [12] operates over a single dimension of a cellular complex to minimize prediction error at each point. In contrast, this work introduces an online algorithm with fewer parameters that minimizes error over a time window, leveraging data across multiple dimensions to improve stability in multi-step-ahead predictions.

In this work, we introduce an extension of existing models to cellular complexes, called the cellular complex VAR (CC-VAR) model. The CC-VAR model applies to a wider range of network structures than simplicial models, as cellular complexes, encompass simplicial complexes while enabling online parameter learning. Additionally, the CC-VAR model reduces the computational complexity encountered in the SC-VAR model by lowering the number of operations.

This paper is organized as follows: Section II introduces preliminary notation and mathematics, Section III introduces the CC-VAR model, Section IV describes the online learning algorithm, and Section V examines the model's performance on real-world data.

II. Preliminaries

A. Vector Autoregressive Models

A P-order VAR model represents the evolution of a multivariate time-varying process $\mathbf{x}^t \in \mathbb{R}^N$ as a linear combination of its P past realizations $\mathbf{x}^{t-1}, \ldots, \mathbf{x}^{t-P}$:

$$\mathbf{x}^t = \sum_{p=1}^{P} \mathbf{W}^p \mathbf{x}^{t-p} + \varepsilon^t, \tag{1}$$

where $\mathbf{W}^p \in \mathbb{R}^{N \times N}$ represents the parameters for the temporal lag p, with $[\mathbf{W}^p]_{ij}$ capturing the spatio-temporal

The study was supported by the IKTPLUSS DISCO grant 338740/AAU.

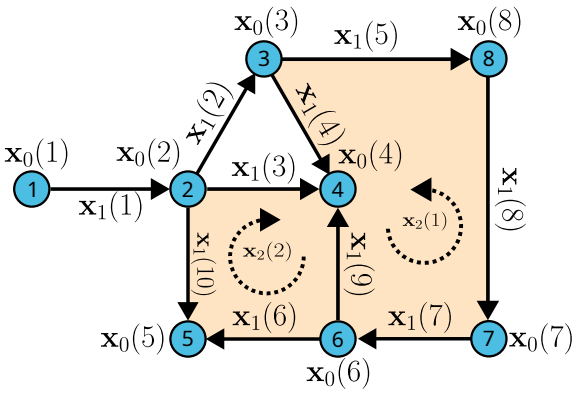


Fig. 1: An example cellular complex with 8 0-cells (nodes), 10 1-cells (edges), 2 2-cells (polygons) and signals defined over them

dependency between $[\mathbf{x}^{t-p}]_j$ and $[\mathbf{x}^t]_i$. The variable ε^t denotes model uncertainty, typically assumed to be zero-mean Gaussian. The VAR model in (1) faces limitations in real-world networks due to a high number of parameters, especially when there are many time series, limited data samples, or rapidly changing environments. To address this, we use the cellular complex structure as an inductive bias, capturing more dependencies while maintaining a similar parameter count.

B. Cellular Signal Processing

Without loss of generality, for the sake of simplicity in the explanation, in this work, we will mainly consider signals whose support is the underlying 2-dimensional cell complex, which is constructed with nodes, two tuples of nodes called edges, and three or more tuples of nodes that are connected with edges called polygons (see Fig. 1). These objects can be formally described in the framework of abstract cell complexes, which are general enough to incorporate graphs, simplicial complexes, and several geometric structures [13].

Definition 1 (Abstract Cell Complex (ACC) [14], [15])**.** *An abstract cell complex is a tuple $\mathcal{C} = \{\mathcal{S}, \prec_b, dim\}$, which is composed of the set $\mathcal{S}$ together with a strict partial order $\prec_b$[1] called boundary relation and a non-negative dimension function $dim : \mathcal{S} \to \mathbb{Z}$, which is monotone with respect to this boundary relation and the usual order on $\mathbb{Z}$.*

An ACC involves a set $\mathcal{S}$ of cells. A k-cell of $\mathcal{S}$, is an element $x \in \mathcal{S}$ with $\dim(x) = k$, where 0-cells are termed as vertices, 1-cells as edges and 2-cells as polygons. The set of all k-cells is denoted as $\mathcal{C}_k$ and the number of k-cells is denoted by N_k. We call K as the dimension of the cellular complex $\mathcal{C}$ if there exists $K = \max_{\sigma \in \mathcal{S}} \dim(\sigma)$. Throughout this text, we assume that K exists and is finite. An example of an ACC with $K = 2$ is illustrated in Fig. 1, which could represent for instance a water distribution network, where the junctions are 0-cells, the pipes connecting them are 1-cells and the closed surface between nodes are 2-cells.

[1]Subscript b is placed to distinguish the relation from partial order over the matrices.

For each $\mathcal{C}_k$, we can associate k-cellular signal $f_k : \mathcal{C}_k \to \mathbb{R}$. The set of signals over $\mathcal{C}_k$ is denoted by $\Omega^k(\mathcal{C})$ for all $k = 0, 1, \ldots, K$. One can define a vector $\mathbf{f}_k \in \mathbb{R}^{N_k}$, whose entries are the signal values over $\mathcal{C}_k$, where $\mathbf{f}_k(\sigma)$ corresponds to the value of the signal on $\sigma \in \mathcal{C}_k$. In Fig. 1, the signals $\mathbf{x}_i(\sigma)$ are named with this convention. In a water network for example, $\mathbf{x}_0(v)$ can be the pressure at a junction v, $\mathbf{x}_1(e)$ can be the flow rate through pipe e and $\mathbf{x}_2(\tau)$ can be total circulating flow on the surface τ. We also define $\Omega^{-1}(\mathcal{C}) = \Omega^{K+1}(\mathcal{C}) = \{0\}$ for simplicity and boundary maps $\mathbf{B}_k : \mathbb{R}^{N_k} \to \mathbb{R}^{N_{k-1}}$, satisfying $\mathbf{B}_k\mathbf{B}_{k+1} = 0$. A canonical way that accounts for the topology is to define boundary maps according to the boundary relation and the orientation given to the cells [14], [16], where the entries of boundary maps are typically from the set $\{-1, 0, 1\}$. Fig. 1 shows the established orientations. If the signal is time-varying, we will denote the signal with matrix notation $\mathbf{F}_k \in \mathbb{R}^{N_k \times T}$ where T is the total number of time instants and $\mathbf{f}_k^t$ denotes a $k-$cellular signal at time instant t.

Let $r_k = \text{rank}(\mathbf{B}_k)$, then each boundary map admits a Singular Value Decomposition (SVD) $\mathbf{B}_k = \mathbf{U}_k\mathbf{\Lambda}_k\mathbf{V}_k^T$, where $\mathbf{U}_k \in \mathbb{R}^{N_{k-1} \times r_k}, \mathbf{V}_k \in \mathbb{R}^{N_k \times r_k}$ and $\Lambda_k \in \mathbb{R}^{r_k \times r_k}$. We will denote the set of nonzero singular values of $\mathbf{B}_k$ as $\mathcal{B}_k$. Moreover, upper, lower, and Hodge $k-$Laplacian [14] are defined as:

$$\mathbf{L}_k = \mathbf{L}_{k,l} + \mathbf{L}_{k,u} = \mathbf{B}_k^T\mathbf{B}_k + \mathbf{B}_{k+1}\mathbf{B}_{k+1}^T = \mathbf{Y}_k\mathbf{\Phi}_k{\mathbf{Y}_k}^T, \quad (2)$$

where $\mathbf{Y}_k \in \mathbb{R}^{N_k \times N_k}$ are the eigenvectors of the Laplacian matrix and $\mathbf{\Phi}_k$ is a diagonal matrix whose elements are coming from the set $\mathcal{F}_k = \left\{\lambda^2 : \lambda \in \{0\} \cup \mathcal{B}_k \cup \mathcal{B}_{k+1}\right\}$ [17]. The frequency response [13] of $\mathbf{f}_k \in \Omega^k(\mathcal{S})$ is given by

$$\hat{\mathbf{f}}_k = {\mathbf{Y}_k}^T\mathbf{f}_k.$$

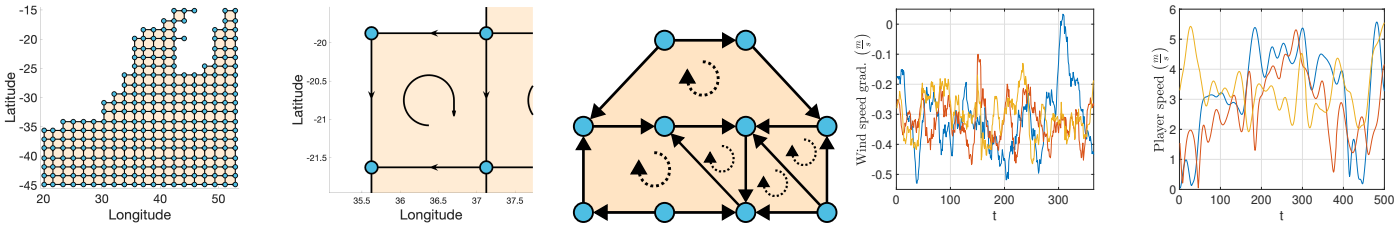
We can associate each element of $\hat{\mathbf{f}}_k$ with singular values of $\mathbf{B}_k$, $\mathbf{B}_{k+1}$ and the value 0, since the eigenvalues of the k-Laplacian are the squared singular values of $\mathbf{B}_k$, $\mathbf{B}_{k+1}$ and value 0. Therefore, for a frequency response $\hat{\mathbf{f}}_k$, $\hat{\mathbf{f}}_k\left(\lambda^2\right)$ means the frequency response at singular value λ. An input k-cell signal $\mathbf{f}_k$ can be filtered with a kernel $\hat{h}_k : \mathbb{R} \to \mathbb{R}$ as:

$$\mathbf{y}_k = \mathbf{Y}_k\hat{h}_k\left(\mathbf{\Phi}_k\right){\mathbf{Y}_k}^T\mathbf{f}_k = \mathbf{H}_k\left(\mathbf{L}_k\right)\mathbf{f}_k$$

where $\hat{h}_k\left(\mathbf{\Phi}_k\right)$ is a diagonal matrix whose diagonal entries are the element-wise evaluations of $\hat{h}_k$ at $\mathcal{F}_k$. $\mathbf{H}_k\left(\mathbf{L}_k\right)$ is called as a k-cellular filter which captures the spatial relationships between k-cell signals.

III. CC-VAR Model

In this section, we propose the CC-VAR model, which combines the VAR model's ability to learn time-lagged dependencies together with the cellular filter's ability to capture spatial dependencies. We first propose the CC-VAR model which extends the previous models over the simplicial complexes and graphs, [10], [11] model by using cellular filters, leading to a better alignment with the cell structures found in most real-world networks. Notably, the previous models, G-VAR and SC-VAR can be seen as a special case of the CC-VAR model. Additionally, through a frequency analysis, we demonstrate that the convolve-transform-convolve approach in



(a) Ocean Dataset oriented cubical complex (b) Football Dataset oriented cellular complex (c) Signal values at selected edges on Ocean Dataset (wind gradient) and selected nodes on Football Dataset (player speed)

Fig. 2: Cellular complexes and signal values at selected cells used in experiments

the previous models [11], [18] is substituted by a transform-convolve approach, allowing us to propose a compact CC-VAR model, with better generalizability.

In [11], the frequency response of the SC-VAR model is given in terms of the transformed cell functions. Here, we show that there is no need to transform the cell functions for the frequency analysis, instead their frequency characteristics are transferred through the different dimensional cells of the complex, which allows simplifying the model expression. This idea is formalized in **Theorem 1**.

Theorem 1. *Let* $k \in \{0, 1, \ldots, K\}$, $\mathbf{H}_k\left(\mathbf{L}_k\right)$, $\mathbf{H}_{k+1}\left(\mathbf{L}_{k+1}\right)$ *be cellular filters and* $\mathbf{f}_{k+1}$ *be a signal on* $k+1$*-dimensional cells. Then, the signal:*

$$\mathbf{f}_{k,1} = \mathbf{H}_k\left(\mathbf{L}_k\right)\mathbf{B}_{k+1}\mathbf{H}_{k+1}\left(\mathbf{L}_{k+1}\right)\mathbf{f}_{k+1}$$

has frequency response

$$\hat{\mathbf{f}}_{k,1}(\lambda^2) = \hat{h}_k\left(\lambda^2\right)\hat{h}_{k+1}\left(\lambda^2\right)\lambda\hat{\mathbf{f}}_{k+1}\left(\lambda^2\right), \quad \textit{if } \lambda \in \mathcal{B}_{k+1}$$

and 0 *otherwise.*

Proof. See Appendix A. □

The result is similar for upward transformation from lower dimensional cells. **Theorem 1** states that the resulting signal is invariant to the order of the convolution, i.e. the resulting filtering output is consistent whether it is applied directly across different cells, or applied to a single cell and transformed across different dimensional cells. In other words, convolve-transform-convolve operations can be achieved by performing single convolve-transform or transform-convolve operations.

We present the CC-VAR model based on the transform-and-then-convolve principle

$$\begin{aligned}\mathbf{x}_k^t = &\sum_{p=1}^{P}\mathbf{H}_{k,-1}^p\left(\mathbf{L}_{k,u}\right)\mathbf{B}_k^T\mathbf{x}_{k-1}^{t-p}\\ &+\mathbf{H}_{k,1}^p\left(\mathbf{L}_{k,l}\right)\mathbf{B}_{k+1}\mathbf{x}_{k+1}^{t-p}+\mathbf{H}_k^p\left(\mathbf{L}_k\right)\mathbf{x}_k^{t-p}+\frac{1}{P}\epsilon_k^t\end{aligned} \tag{3}$$

where the time series vector of a k-cell at time t, denoted as $\mathbf{x}_k^t$, is expressed as a filtered versions of P time-lagged values of the signal defined over the k-cell, $k-1$-cell, and $k+1$-cell. Before applying the convolution filters $\{\mathbf{H}_{k,-1}^p(\mathbf{L}_{k,u})\}_{p=1}^P$ to the signals defined over the $k-1$-cell, the signals are transformed into a k-cell signals using a transformation via the boundary map $\mathbf{B}_k^T$. Similarly, the signal defined over the $k+1$-cell is transformed into a k-cell signal via $\mathbf{B}_{k+1}$, before applying the filters $\{\mathbf{H}_{k,1}^p(\mathbf{L}_{k,l})\}_{p=1}^P$. In cases where the signal has a smooth variation over the structure across different cells, these transformations are natural, intuitive, and informative about the signal on the k-cell. For example, in a water network, transforming a 1-cell signal (flow rate in pipes) to a 0-cell signal via $\mathbf{B}_1$ results in a transformed node signal that represents the sum of directed flows at a junction. Similarly, transforming a 0-cell signal (node pressure) to an edge signal via $\mathbf{B}_1^T$ produces an edge signal representing the potential difference between nodes.

IV. Online Learning of CC-VAR Models

In this section, first, we formulate the optimization problem that can extract the parameters of the CC-VAR model in (3) online when the filters are parametrized by polynomials. Then, we will discuss complexity and dynamic regret properties of the proposed problem.

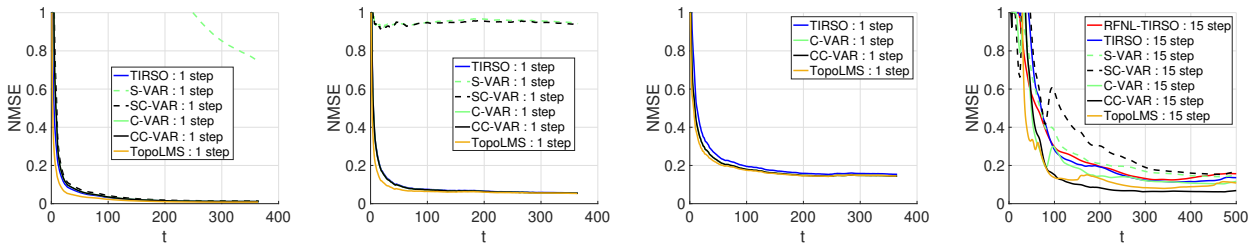
We assume that the cellular filters admit a polynomial representation with polynomial orders $L_{k,1}, L_{k,-1}, L_k$ for their respective filters, i.e., cellular filter $\mathbf{H}_{k,1}^p$ is generated from the parametric kernel as follows:

$$\hat{h}_{k,1}^p(\lambda) = \sum_{l=0}^{L_{k,1}-1} a_{k,1}^{l,p}\lambda^l \tag{4}$$

and similarly for the other filters. Let us denote the vector of all polynomial coefficients over k-dimensional cells by $\mathbf{a}_k$. We consider an online optimization problem of the form for time instant t:

$$\begin{aligned}\mathbf{a}_k^*(t) = \arg\min_{a_k}\ &\frac{1}{2}(1-\gamma)\sum_{t'=P+1}^{t}\gamma^{t-t'}\|\mathbf{x}_k^t-\tilde{\mathbf{x}}_k^t\left(\mathbf{a}_k\right)\|_2^2\\ &+\gamma^{t-t'}\mu_{k,-1}\|\mathbf{B}_k\tilde{\mathbf{x}}_k^t\left(\mathbf{a}_k\right)\|_2^2+\gamma^{t-t'}\mu_{k,1}\|\mathbf{B}_{k+1}^T\tilde{\mathbf{x}}_k^t\left(\mathbf{a}_k\right)\|_2^2\\ &+\lambda_k\|\mathbf{a}_k\|_2^2, \quad \forall k\in\{0,1,\ldots,K\}\end{aligned} \tag{5}$$

where $\tilde{\mathbf{x}}_k^t\left(\mathbf{a}_k\right)$ denotes the estimate of $\mathbf{x}_k^t$ using (3) and (4) at time t with respect to parameter vector $\mathbf{a}_k$. In (5), the first term is the data fitting term where $\gamma \in (0,1)$ is the forgetting factor. The second and third terms control the frequency content of the model over the squared values of $\mathcal{B}_k$ and $\mathcal{B}_{k+1}$, respectively, where $\mu_{k,-1}$ and $\mu_{k,1}$ are regularization parameters controlling the smoothness of the estimated signal over different frequencies. The last term represents the regularization for the parameter vector $\mathbf{a}_k$, reducing the effect of random noise, where λ_k is the regularization parameter. To solve the optimization problem (5), we use Stochastic Gradient Descent (SGD). The transform-and-then-convolve approach reduces

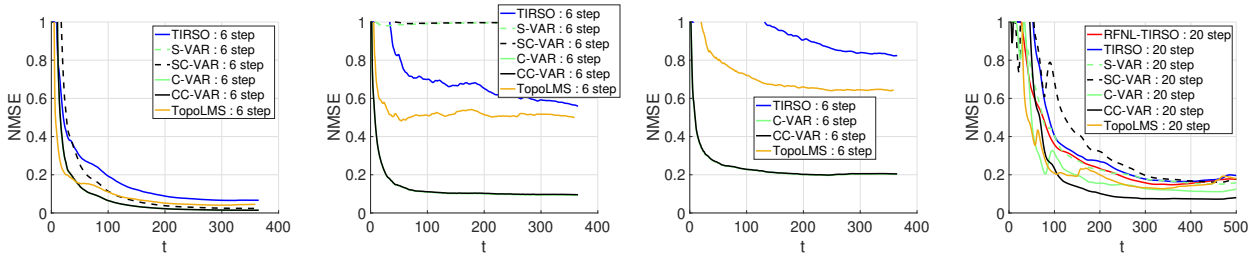


(a) 1-step ahead prediction of node signal (temperature)

(b) 1-step ahead prediction of edge signal (wind speed gradient)

(c) 1-step ahead prediction of polygon signal (geostrophic current curl)

(d) 15-step ahead prediction of node signal (player speed)



(e) 6-step ahead prediction of node signal (temperature)

(f) 6-step ahead prediction of edge signal (wind speed gradient)

(g) 6-step ahead prediction of polygon signal (geostrophic current curl)

(h) 20-step ahead prediction of node signal (player speed)

Fig. 3: Comparative evaluation of NMSE (a), (e) for temperature; (b), (f) gradient of wind speed, (c), (g) curl of geostrophic current on Ocean Dataset and (d), (h) player speed on Football Dataset

the complexity of gradient updates based on the polynomial order. Given that calculating the data fidelity term involves N_k elements and the regularization terms require N_k^2 operations for each polynomial term, the time complexity of the SGD term is $\mathcal{O}\left(N_k^2 L\right)$. Since cellular adjacency is expressed using sparse upper and lower Laplacians (e.g., in water networks), the time complexity of this algorithm is notably lower than the squared summation of polynomial orders L_k in the convolve-transform-convolve approach. Moreover, the transform-and-convolve principle promotes faster convergence as dynamic regret is linked to the path length of parameters [19], [20], which directly relates to the number of parameters. This allows faster convergence compared to [11].

V. Experiments

In this section, we test our algorithm's estimation and generalizability capabilities through T-step ahead predictions, where the algorithms recursively predict T-steps ahead from their 1-step ahead predictions. We evaluate our algorithm using two different datasets described below:

a) Ocean Dataset: The Ocean Dataset comprises an oriented cubical complex sampled from a 20×20 grid at the southern tip of Africa (latitudes -15 to -45, longitudes 20 to 53.1), as shown in Fig. 2a. Cellular signals include daily sea surface temperature over 225 nodes [21], sea surface wind speed gradient over 498 edges [22], and curl of the geostrophic current over 274 polygons [23] for 2023. Simplicial complex-based algorithms only consider node and edge data due to the absence of triangles. We assess performance for 1- and 6-step ahead predictions for node, edge, and polygon signals.

b) Football Dataset: The Football Dataset [24] includes 10 nodes (players), 14 edges, and 5 polygons representing closed loops, where the players' mean coordinates for the first minute encode structural information. Speed and distance between players, and the area of each polygon are used as node, edge, and polygon signals, respectively. We use 500 data samples collected at 0.05-second intervals within the first minute for parameter estimation and evaluate performance for 15- and 20-step ahead predictions of node signals.

We compare our online algorithm in (5) with state-of-the-art models including SC-VAR [11], S-VAR [18], TIRSO [1], RFNL-TIRSO [25], and TopoLMS [12]. We distinguish between CC-VAR and C-VAR, where C-VAR uses (3) assuming $\Omega^{k+1}(\mathcal{C}) = \Omega^{k-1}(\mathcal{C}) = 0$. G-VAR [10] is omitted, as no online algorithm exists for it; it is essentially C-VAR applied at the node level. Comparison is based on the NMSE metric defined for each time instant t over k-dimensional cells as $\text{NMSE} = \|\mathbf{z}_k^t - \tilde{\mathbf{z}}_k^t\|_2^2 / \|\mathbf{z}_k^t\|_2^2$, where $\mathbf{z}_k^t$ is the ground truth signal and $\tilde{\mathbf{z}}_k^t$ is the predicted signal. The results are shown in Fig. 3. RFNL-TIRSO is excluded from the Ocean Dataset due to instability issues.

For the Ocean Dataset, C-VAR (but not S-VAR) shows similar performance to SC-VAR and CC-VAR, while TopoLMS slightly outperforms CC-VAR for 1-step predictions over the nodes. SC-VAR and S-VAR struggle with edge data due to the absence of triangles and parameter redundancy from the convolve-transform-convolve principle. In the Football Dataset, CC-VAR outperforms all models for large step-ahead predictions (15-20 steps), reaching an NMSE near 0.1 and demonstrating faster convergence than SC-VAR, consistent with our discussion on dynamic regret in Section IV. Furthermore, CC-VAR surpasses C-VAR, highlighting the importance of utilizing node, edge, and polygon data for modeling. CC-VAR consistently outperforms TIRSO and TopoLMS for high

step-size predictions.

To explain the performance difference between TIRSO and TopoLMS across different step-ahead predictions, we plot selected edge values for the Ocean Dataset and node values for the Football Dataset in Fig. 2c. The Football dataset shows smooth variations, whereas the Ocean Dataset has rapid changes, making online estimation more challenging. As a result, we observe higher performance variability for different time-step predictions. The TIRSO model overfits due to a high number of parameters, while TopoLMS lacks a forgetting term γ as in (5) and considers node, edge, and polygon signals separately. Therefore, CC-VAR excels in adaptability with fewer parameters, integration of all cellular complex signals, and the use of a forgetting term.

VI. Conclusion

In this work, we introduce cellular complex VAR models (CC-VAR) to represent time series data defined over higher-order structures. The CC-VAR model leverages the cellular structure of real-world networks as an inductive bias, enabling effective modeling of time series generated by these networks. This approach proves particularly advantageous in scenarios with a large number of time series, limited available data, or when swift updates are required in an online setting with better adaptability guarantees.

Appendix A. Proof of Theorem 1

In this section, we will present a proof to **Theorem 1**.

*Proof to **Theorem 1**.* We prove the statement for $\mathbf{f}_{k,1}$. The basic idea is to use the fact that $\mathbf{B}_k\mathbf{B}_{k+1} = 0$. Considering the SVD of $\mathbf{B}_{k+1}$, we can write:

$$\mathbf{B}_{k+1}\mathbf{H}_{k+1}\left(\mathbf{L}_{k+1}\right)\mathbf{f}_{k+1} = \sum_{\lambda\in\mathcal{B}_{k+1},\phi\in\mathcal{F}_{k+1}} \lambda\mathbf{u}_{k+1}(\lambda)\mathbf{v}_{k+1}(\lambda)^\top\mathbf{u}^l_{k+1}(\phi)h_{k+1}(\phi)\hat{\mathbf{f}}_{k+1}(\phi) \quad (6)$$

Note that by **Proposition 1** in [17], $\mathbf{v}_{k+1}(\lambda)^\top\mathbf{u}^l_{k+1}(\phi) = 1$, if $\phi = \lambda^2$, $\mathbf{v}_{k+1}(\lambda)^\top\mathbf{u}^l_{k+1}(\phi) = 0$ otherwise. Then this expression becomes

$$\mathbf{B}_{k+1}\mathbf{H}_{k+1}\left(\mathbf{L}_{k+1}\right)\mathbf{f}_{k+1} = \sum_{\lambda\in\mathcal{B}_{k+1}} \lambda\mathbf{u}_{k+1}(\lambda)h_{k+1}(\lambda^2)\hat{\mathbf{f}}_{k+1}(\lambda^2) \quad (7)$$

Applying similar arguments yields the desired result:

$$\mathbf{H}_k\left(\mathbf{L}_k\right)\mathbf{B}_{k+1}\mathbf{H}_{k+1}\left(\mathbf{L}_{k+1}\right)\mathbf{f}_{k+1} = \sum_{\lambda\in\mathcal{B}_{k+1}} \mathbf{u}^l_{k+1}(\lambda^2)h_k(\lambda^2)h_{k+1}(\lambda^2)\lambda\hat{\mathbf{f}}_{k+1}(\lambda^2). \quad (8)$$

□

References

[1] 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.

[2] 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.

[3] M. S. Veedu, H. Doddi, and M. V. Salapaka, "Topology learning of linear dynamical systems with latent nodes using matrix decomposition," *IEEE Transactions on Automatic Control*, vol. 67, no. 11, pp. 5746–5761, 2021.

[4] X. Rong and V. Solo, "Symmetric var (1) modelling with guaranteed stability," in *ICASSP 2024-2024 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. IEEE, 2024, pp. 9786–9790.

[5] H. Lütkepohl, *New introduction to multiple time series analysis*. Springer Science & Business Media, 2005.

[6] E. J. Hannan, *Multiple time series*. John Wiley & Sons, 2009.

[7] C. Lam and Q. Yao, "Factor modeling for high-dimensional time series: inference for the number of factors," *The Annals of Statistics*, pp. 694–726, 2012.

[8] N. Lee, H. Choi, and S.-H. Kim, "Bayes shrinkage estimation for high-dimensional var models with scale mixture of normal distributions for noise," *Computational Statistics & Data Analysis*, vol. 101, pp. 250–276, 2016.

[9] J. Chang, B. Guo, and Q. Yao, "Principal component analysis for second-order stationary vector time series," *The Annals of Statistics*, vol. 46, no. 5, pp. 2094–2124, 2018.

[10] 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.

[11] J. Krishnan, R. Money, B. Beferull-Lozano, and E. Isufi, "Simplicial vector autoregressive models," *IEEE Transactions on Signal Processing*, vol. 72, pp. 5454–5469, 2024.

[12] L. Marinucci, C. Battiloro, and P. Di Lorenzo, "Topological adaptive learning over cell complexes," in *2024 32nd European Signal Processing Conference (EUSIPCO)*. IEEE, 2024, pp. 832–836.

[13] T. M. Roddenberry, M. T. Schaub, and M. Hajij, "Signal processing on cell complexes," in *ICASSP 2022-2022 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. IEEE, 2022, pp. 8852–8856.

[14] S. Sardellitti, S. Barbarossa, and L. Testa, "Topological signal processing over cell complexes," in *2021 55th Asilomar Conference on Signals, Systems, and Computers*. IEEE, 2021, pp. 1558–1562.

[15] R. Klette, "Cell complexes through time," in *Vision Geometry IX*, vol. 4117. SPIE, 2000, pp. 134–145.

[16] C. Battiloro, S. Sardellitti, S. Barbarossa, and P. Di Lorenzo, "Topological signal processing over weighted simplicial complexes," in *ICASSP 2023-2023 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. IEEE, 2023, pp. 1–5.

[17] S. Barbarossa and S. Sardellitti, "Topological signal processing over simplicial complexes," *IEEE Transactions on Signal Processing*, vol. 68, pp. 2992–3007, 2020.

[18] J. Krishnan, R. Money, B. Beferull-Lozano, and E. Isufi, "Simplicial vector autoregressive model for streaming edge flows," in *ICASSP 2023-2023 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. IEEE, 2023, pp. 1–5.

[19] S. M. Fosson, "Online optimization in dynamic environments: A regret analysis for sparse problems," in *2018 IEEE Conference on Decision and Control (CDC)*, 2018, pp. 7225–7230.

[20] P. Zhao, Y.-J. Zhang, L. Zhang, and Z.-H. Zhou, "Dynamic regret of convex and smooth functions," *Advances in Neural Information Processing Systems*, vol. 33, pp. 12 510–12 520, 2020.

[21] NOAA CoastWatch, "Noaa geo-polar blended global sea surface temperature analysis (level 4)," 2024, altimetry data are provided by the NOAA Laboratory for Satellite Altimetry. [Online]. Available: https://coastwatch.noaa.gov/cwn/products/noaa-geo-polar-blended-global-sea-surface-temperature-analysis-level-4.html

[22] Copernicus Marine Service, "Global ocean hourly sea surface wind and stress from scatterometer and model," 2024, this study has been conducted using E.U. Copernicus Marine Service Information; https://doi.org/10.48670/moi-00305.

[23] NOAA CoastWatch, "Sea level anomaly and geostrophic currents, multi-mission, global, optimal interpolation, gridded," 2024, altimetry data are provided by the NOAA Laboratory for Satellite Altimetry. [Online]. Available: https://coastwatch.noaa.gov/cwn/products/sea-level-anomaly-and-geostrophic-currents-multi-mission-global-optimal-interpolation.html

[24] S. A. Pettersen, D. Johansen, H. Johansen, V. Berg-Johansen, V. R. Gaddam, A. Mortensen, R. Langseth, C. Griwodz, H. K. Stensland, and P. Halvorsen, "Soccer video and player position dataset," in *Proceedings of the 5th ACM multimedia systems conference*, 2014, pp. 18–23.

[25] R. 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*, 2023.