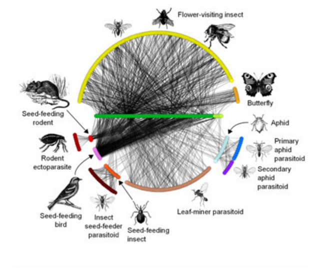
DYNAMIC NETWORK IDENTIFICATION FROM NON-STATIONARY VECTOR AUTOREGRESSIVE TIME SERIES

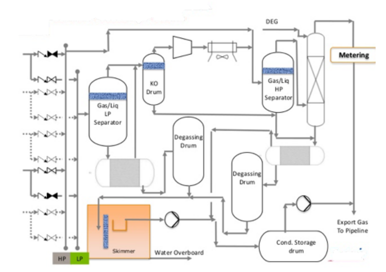
LUIS M. LOPEZ-RAMOS, DANIEL ROMERO, BAKHT ZAMAN, AND BALTASAR BEFERULL-LOZANO
University of Agder, Grimstad, Norway. Contact: luismiguel.lopez@uia.no

OVERVIEW

Graph-based modeling: learn complex dynamics
Exploiting intrinsic sparsity of direct interactions


Social networks


Biology - Ecology


Industrial processes

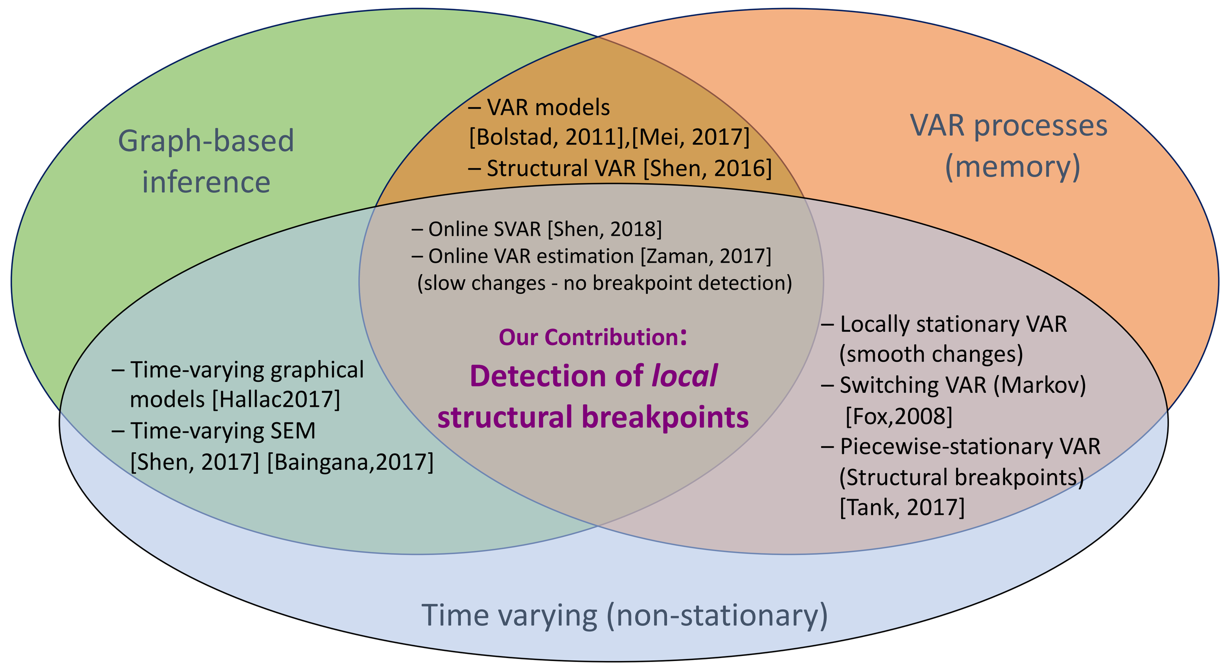
Interactions in many real-world applications are not stationary

Learning a time-varying VAR [1] model from data **enables**:

- Time-series forecasting, denoising, data completion, compression
- Parsimonious, human-readable models
- Unveiling behavioral patterns and changes
- Direct (unmediated) $\neq$ indirect (mediated) influence

Inferring nonstationary interaction model $\leftrightarrow$ unveiling time-varying causality net (graph)

CONTRIBUTION IN CONTEXT



- Inference of **time-varying** models accounting for
  - memory (VAR coefficients)
  - network structure (few active edges)
- Detection of **local structural breakpoints**
  - changes at individual edges
  - changes in coefficients are not synchronized across edges
  - Previous works assume that changes at all nodes are aligned in time
- Low-complexity solver
- Data windowing (optional) to reduce computation
- Scheme to choose regularization parameters

NOTATION

$y_{i,t}$	Value of time series in i-th node at time t
$\mathbf{y}_t$	$:= [y_{1,t}, y_{2,t}, \ldots, y_{P,t}]^\top$ Vector at time t
$\varepsilon_{i,t}$	Innovation (process noise)
P	Number of nodes
L	Order of the VAR process
T	Time series duration
$\mathcal{E}_t$	Edge set at time t
$a_{ij,t}^{(\ell)}$	VAR coefficient for edge $j \to i$, lag ℓ, at time t
$\mathbf{A}_t^{(\ell)}$	Matrix of VAR coefficients for lag ℓ at time t
$\mathbf{a}_{ij,t}$	LTV filter coefficients at edge $j \to i$ at time t

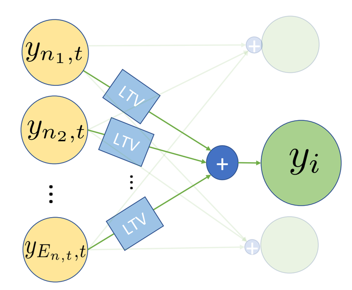
TIME-VARYING VAR $\leftrightarrow$ TIME-VARYING GRAPH

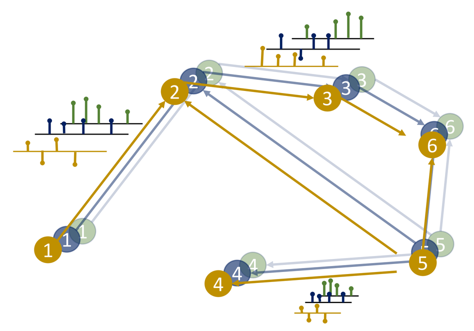
L-th order TVAR model [1, Ch. 1]:

$$\mathbf{y}_t = \sum_{\ell=1}^{L} \mathbf{A}_t^{(\ell)} \mathbf{y}_{t-\ell} + \boldsymbol{\varepsilon}_t \quad (1)$$

$$y_{i,t} = \sum_{j=1}^{P} [y_{j,t-1}, y_{j,t-2}, \ldots, y_{j,t-L}]\, \mathbf{a}_{ij,t} + \varepsilon_{i,t} \quad (2)$$

$\mathbf{a}_{ij,t} := [a_{ij,t}^{(1)}, a_{ij,t}^{(2)}, \ldots, a_{ij,t}^{(L)}]^\top$ coefficients of a *linear time-varying* (LTV) filter





Graph associated with a time-invariant (stationary) VAR process [2]
Generalization to *time-varying* VAR models

- One time series per node i
- Time-varying edge set $\mathcal{E}_t := \{(i,j) \in \mathcal{V} \times \mathcal{V} : \mathbf{a}_{ij,t} \neq \mathbf{0}\}$

PROBLEM DEFINITION

Goal: to estimate coefficients $\{\{\mathbf{A}_t^{(\ell)}\}_{\ell=1}^{L}\}_{t=L+1}^{T}$ given $\{\mathbf{y}_t\}_{t=1}^{T}$

- $(T-L)P^2L$ unknowns, PL samples
- Exploit sparse spatial structure (few active edges)
- Exploit spatio-temporal locality of changes (local breakpoints)

ONLINE METHODS

- Approximate estimation of VAR parameters
- Diminishing stepsize -> Asymptotic convergence for stationary VAR
- Constant stepsize -> Lightweight heuristic for tracking TVAR
- Two novel algorithms [12] (submitted to TSP)
- TISO $\approx$ group-sparse LMS
- TIRSO $\approx$ group-sparse RLS

GLOBAL AND LOCAL BREAKPOINTS

Previous works [9, 10, 11] enforce group-sparse first difference
Structural breakpoint set $\mathcal{T} := \{t : \mathbf{A}_t^{(\ell)} \neq \mathbf{A}_{t-1}^{(\ell)} \text{for some } \ell\}$
For instance, [10] computes

$$\min_{\{\mathbf{A}_t^{(\ell)}\}} \sum_{t=L+1}^{T} \left\| \mathbf{y}_t - \sum_{\ell=1}^{L} \mathbf{A}_t^{(\ell)} \mathbf{y}_{t-\ell} \right\|_2^2 + \gamma \sum_{t=L+2}^{T} \sqrt{\sum_{(i,j)} \|\mathbf{a}_{ij,t} - \mathbf{a}_{ij,t-1}\|_2} \quad (3)$$

However, in some real applications (e.g. industrial processes):

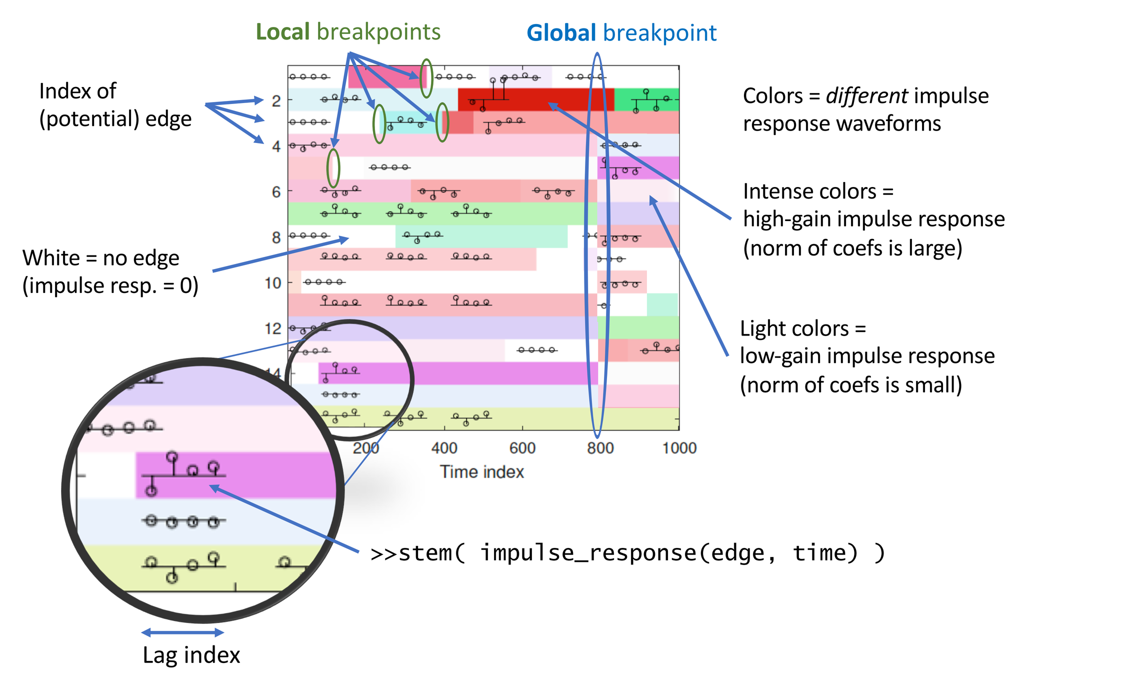
- Links (local interaction patterns) change, but only a few at a time
- ***Local*** breakpoint set: $\mathcal{T}_{i,j} := \{t : \mathbf{a}_{ij,t} \neq \mathbf{a}_{ij,t-1}\}$

Proposed estimation criterion:

$$\min_{\{\mathbf{A}_t^{(\ell)}\}} \sum_{t=L+1}^{T} \left\| \mathbf{y}_t - \sum_{\ell=1}^{L} \mathbf{A}_t^{(\ell)} \mathbf{y}_{t-\ell} \right\|_2^2 \quad (4)$$
$$+ \sum_{(i,j)} \left(\lambda \sum_{t=L+1}^{T} \|\mathbf{a}_{ij,t}\|_2 + \gamma \sum_{t=L+2}^{T} \|\mathbf{a}_{ij,t} - \mathbf{a}_{ij,t-1}\|_2 \right)$$

Enables spatial location of change events

VISUALIZING TVAR PROCESSES



SOLUTION VIA ADMM

$$\arg\min_{\mathbf{B},\boldsymbol{\Theta},\mathbf{C}} \frac{1}{2} \|\mathbf{Y} - \mathbf{Z}\mathbf{B}\|_F^2 + \lambda \Omega_{GL}(\boldsymbol{\Theta}) + \gamma \Omega_{GL}(\mathbf{C}),$$
$$\text{s.to } \mathbf{D}\mathbf{B} = \boldsymbol{\Theta}, \ \mathbf{B} = \mathbf{C} \quad (5)$$

where: $\mathbf{B}$ stacks all $\mathbf{A}_\ell^{(t)}$,
$\mathbf{Z}$ and $\mathbf{Y}$ stacks all $\mathbf{y}$ (regressors and targets, resp.)

$$\mathbf{D} := \begin{bmatrix} -\mathbf{I} & \mathbf{I} & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & -\mathbf{I} & \mathbf{I} & & \vdots \\ \vdots & \ddots & \ddots & \ddots & \\ \mathbf{0} & & \cdots & -\mathbf{I} & \mathbf{I} \end{bmatrix},$$

$\Omega_{GL}(\mathbf{B}) = \sum_{(i,j)} \sum_{t=L+1}^{T} \|\mathbf{a}_{ij,t}\|_2$
$\Omega_{GTV}(\mathbf{B}) = \sum_{(i,j)} \sum_{t=L+1}^{T} \|\mathbf{a}_{ij,t+1} - \mathbf{a}_{ij,t}\|_2 = \Omega_{GL}(\mathbf{D}\mathbf{B})$

For each iteration k:

$$\mathbf{B}^{[k+1]} := \left(\mathbf{Z}^\top\mathbf{Z}/\rho + \mathbf{I} + \mathbf{D}^\top\mathbf{D}\right)^\dagger \left(\mathbf{Z}^\top\mathbf{Y}/\rho + \mathbf{C}^{[k]} - \mathbf{V}^{[k]} + \mathbf{D}^\top(\boldsymbol{\Theta}^{[k]} - \mathbf{U}^{[k]})\right) \quad (6)$$
$$\boldsymbol{\theta}_{ij,t}^{[k+1]} := \text{prox}_{\lambda/\rho\|\cdot\|_2}(\mathbf{b}_{ij,t}^{[k+1]} - \mathbf{b}_{ij,t-1}^{[k+1]} + \mathbf{u}_{ij,t-1}^{[k+1]}) \quad (7)$$
$$\mathbf{c}_{ij,t}^{[k+1]} := \text{prox}_{\lambda/\rho\|\cdot\|_2}(\mathbf{b}_{ij,t}^{[k+1]} + \mathbf{v}_{ij,t}^{[k+1]}) \quad (8)$$
$$\mathbf{U}^{[k+1]} := \mathbf{U}^{[k]} + (\mathbf{D}\mathbf{B}^{[k+1]} - \boldsymbol{\Theta}^{[k+1]}) \quad (9)$$
$$\mathbf{V}^{[k+1]} := \mathbf{V}^{[k]} + (\mathbf{B}^{[k+1]} - \mathbf{C}^{[k+1]}) \quad (10)$$

REFERENCES

[1] H. Lütkepohl, *New Introduction to Multiple Time Series Analysis*. Springer, 2005.

[2] A. Bolstad, B. D. V. Veen, and R. Nowak, "Causal netwk inference group sparse regularization," *IEEE Trans. Sig. Process.*, vol. 59, no. 6, pp. 2628–2641, 2011.

[3] J. Mei and J. M. Moura, "Signal processing on graphs: Causal modeling of unstructured data." *IEEE Trans. Sig. Process.*, vol. 65, no. 8, pp. 2077–2092, 2017.

[4] Y. Shen, B. Baingana, and G. B. Giannakis, "Nonlinear structural vector autoregressive models for inferring effective brain network connectivity," *arXiv preprint arXiv:1610.06551*, 2016.

[5] D. Hallac, Y. Park, S. Boyd, and J. Leskovec, "Network inference via the time-varying graphical lasso," in *Proc. 23rd Intl. Conf. Knowledge Discovery, Data Mining*. ACM, 2017, pp. 205–213.

[6] Y. Shen, B. Baingana, and G. B. Giannakis, "Tensor decompositions for identifying directed graph topologies and tracking dynamic networks," *IEEE Trans. Sig. Process.*, vol. 65, no. 14, pp. 3675–3687, 2017.

[7] B. Baingana and G. B. Giannakis, "Tracking switched dynamic network topologies from information cascades," *IEEE Trans. Sig. Process.*, vol. 65, no. 4, pp. 985–997, 2017.

[8] E. Fox, E. B. Sudderth, M. I. Jordan, and A. S. Willsky, "Bayesian nonparametric inference of switching dynamic linear models," *IEEE Trans. Sig. Process.*, vol. 59, no. 4, pp. 1569–1585, Apr. 2011.

[9] A. Safikhani and A. Shojaie, "Structural break detection in high-dimensional non-stationary var models," *arXiv preprint arXiv:1708.02736*, 2017.

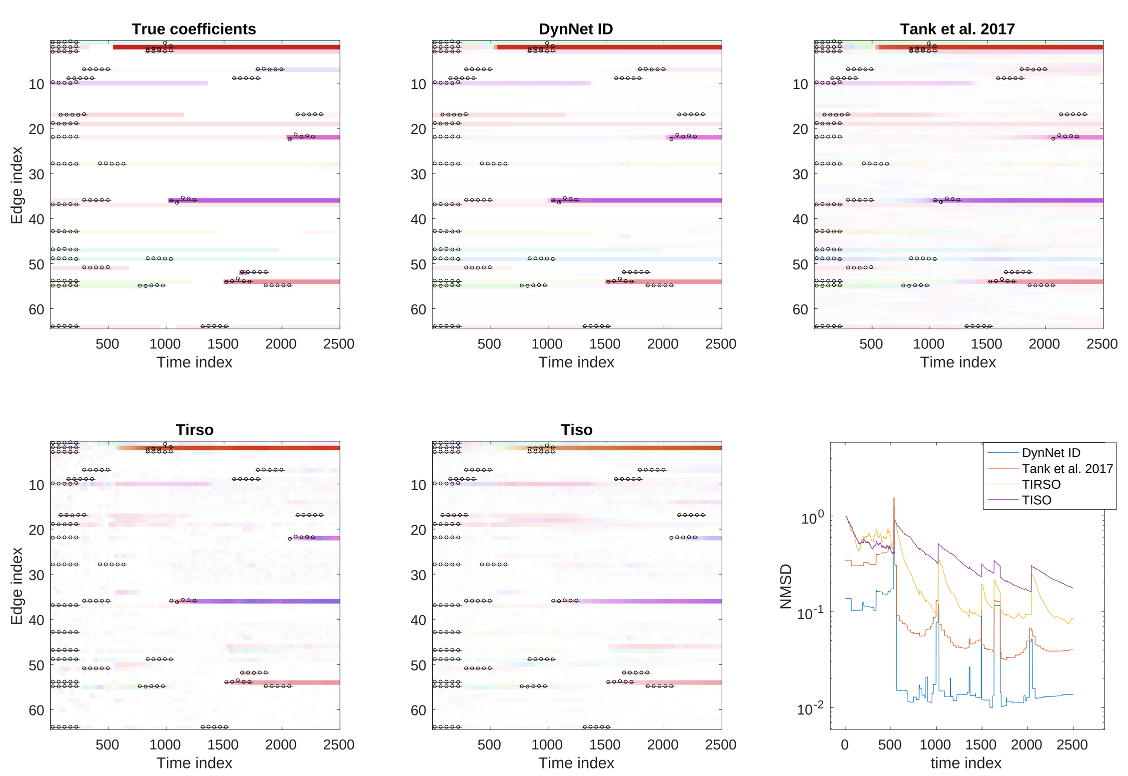
[10] A. Tank, E. B. Fox, and A. Shojaie, "An efficient admm algorithm for structural break detection in multivariate time series," *arXiv preprint arXiv:1711.08392*, 2017.

[11] H. Cho *et al.*, "Change-point detection in panel data via double cusum statistic," *Electronic J.of Statistics*, vol. 10, no. 2, pp. 2000–2038, 2016.

[12] B. Zaman, L. M. Lopez-Ramos, D. Romero, and B. Beferull-Lozano, "Online estimation of sparse topologies from vector autoregressive time-series," *Submitted to IEEE Trans. on Sig. Process.*

EMPIRICAL VALIDATION

- $P = 10$ nodes, 21 active edges (nonzero coefficients)
- Initial VAR coefficients $\sim \mathcal{N}(0,1)$, scaled to ensure stability
- Local breakpoints generated at uniformly spaced instants, edges chosen uniformly at random
- LTV coefficients redefined at each breakpoint:
  - with probability p=0.4, back to 0
  - otherwise, generated from $\sim \mathcal{N}(0,1)$, and scaled again to ensure stability



The work in this paper was supported by the SFI Offshore Mechatronics grant 237896/O30.