Identifiable Markov Switching Models with Instantaneous Effects and Exponential Families

Roel Hulsman[1,2] Carles Balsells-Rodas[3] Sara Magliacane[1,4]

[1]University of Amsterdam [2]Adyen [3]Imperial College London [4]Saarland University

Contact: r.p.hulsman@uva.nl

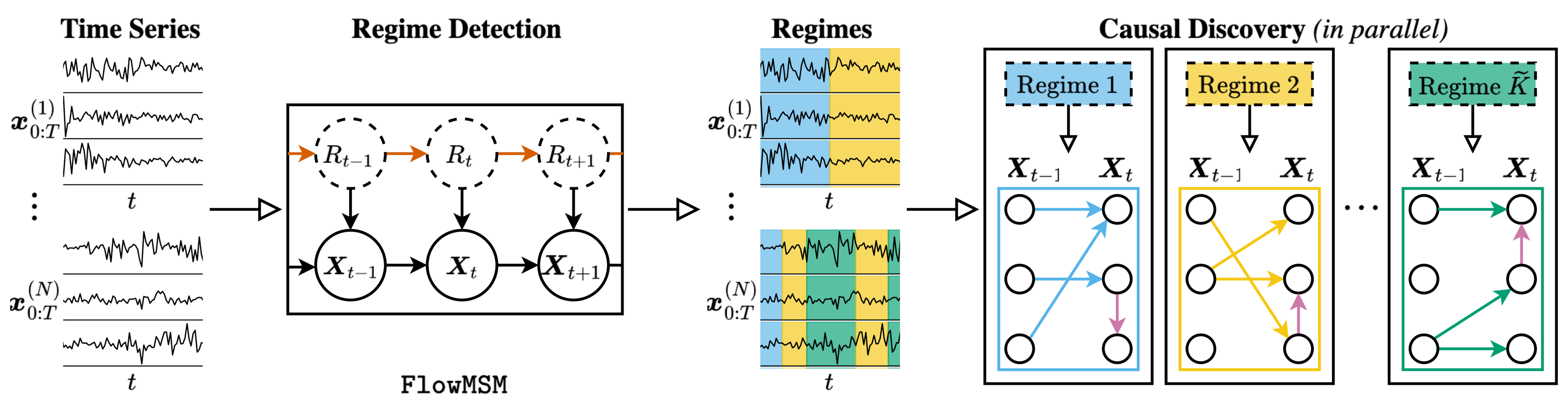
Our Work in Short

Motivation. Many systems move between hidden phases of stable behaviour called **regimes**, *e.g.*, ENSO weather patterns alternate between El Niño, La Niña and a neutral period.

Problem. *Detecting regimes* is challenging when *(i)* switches are frequent and *(ii)* the dynamics are nonlinear and non-Gaussian, which particularly occurs when causal effects appear faster than the rate of measurement, called **instantaneous effects**. For example for diabetes patients, the effect of insulin delivery is typically faster than the sampling frequency of glucose monitors.

This work.

- We establish **identifiability** for a broad class of **regime-switching structural causal models** (SCMs) with regime-dependencies (orange edges below), (nonlinear) instantaneous effects (pink edges) and independent exponential family noise.
- We introduce **FlowMSM**, a regime detection method based on normalising flows, which extends to discover regime-dependent temporal causal structures called **window graphs**.

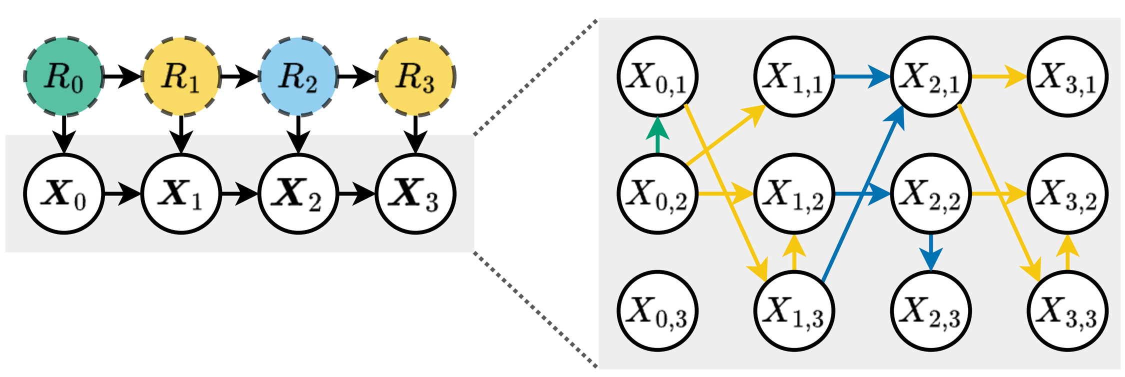


Regime-Switching Structural Causal Models

Consider *continuous* time series $X_{0:T} \in \mathcal{X}^{\times(T+1)}$ and discrete latent **regimes** $R_t \in \mathcal{A}_K$, $K < \infty$. Leaving out the initial equations for brevity, the transition structural equations are

$$X_t \leftarrow f\big(\mathsf{Pa}(R_t), R_t, \epsilon_t\big), \qquad t \in \{1, \ldots, T\},$$

for (instantaneous) causal parents $\mathsf{Pa}(\cdot) \subset X_{t-1:t}$ and *i.i.d.* exogenous noises $\epsilon_t \in \mathcal{X}_\epsilon$.



The *dynamic process* induced by a regime-switching SCM is a **Markov Switching Model** (MSM). The joint distribution of $X_{0:T}$ can be written as a **finite mixture** over K^{T+1} regime sequences.

Overview of our assumptions

- *Conditional* causal stationarity: Regimes fully govern structural changes over time;
- *Conditional* causal sufficiency: No latent confounders *within each regime*;
- Acyclicity: Structural equations are recursive;
- Causal Markov & faithfulness: Conditional independency in data $\Longleftrightarrow$ d-separation in DAG.

For estimation, $R_{0:T}$ follows a *first-order stationary Markov chain* and f_a^0, f_a are *contractive*.

Identifiability Theory

Problem. Classic identifiability results for finite-state HMMs do not trivially extend to MSMs due to the autoregressive connections between observed variables, studied by [1].

Only for affine instantaneous effects and $\epsilon_t \sim \mathcal{N}(0, \sigma^2 I)$, we obtain a **classic Gaussian mixture model**. Distinct Gaussian parameters guarantees identifiability of the regimes [2].

If instantaneous effects are not affine, or the noise is non-Gaussian, we lose these guarantees...

Key question. When are the latent regimes identifiable in such **non-Gaussian mixtures**?

Upon regime identification, *known causal theory applies* to identify stationary causal graphs.

Definition. MSMs are **identifiable** *(up to permutation)* if the likelihood $p_\theta(x_{0:T})$ *uniquely determines* the regime prior $p_\theta(r_{0:T})$ and mixture components $p_\theta(x_{0:T} \mid r_{0:T})$, *up to a regime relabelling*.

Theorem 3.5 – Identifiable Regime-Switching SCMs

Consider an acyclic regime-switching SCM that satisfies conditional causal stationarity and sufficiency, and Ass. 3.1 to 3.4. Then the induced MSM is **identifiable** *(up to permutation)*.

Technical contributions. A necessary and sufficient condition is **linear independence** of the family of mixture components [2]. Inspired by [3], we contribute:

- Q1. Can we ensure linear independence *beyond Gaussian families*? ✓
- Q2. Can we translate the abstract notion of linear independence to *more fine-grained conditions* on regime-switching SCMs? ✓

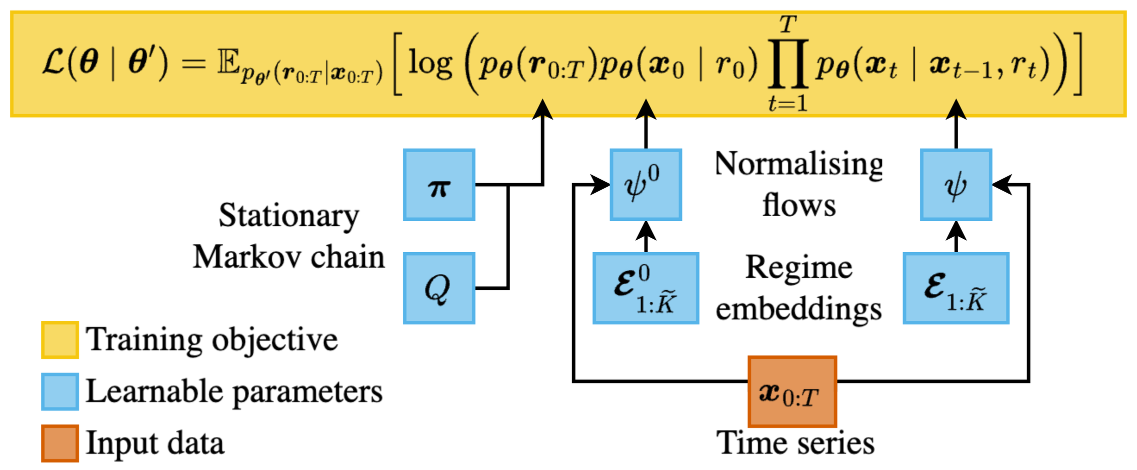
Method – FlowMSM

Input. Data $\{x_{0:T}^{(n)}\}_{n=1}^N$ of *i.i.d.* realisations of time series $X_{0:T}$. In practice, $N = 1$ often suffices.

Regime detection. **FlowMSM** outputs estimates $\widehat{p}_\theta(r_t \mid x_{0:T}^{(n)})$ of the regime posterior likelihood.

We use **Generalised Expectation-Maximisation** (GEM) for mixture model estimation, where a *conditional normalising flow* models the transition distributions, using shared parameters ψ and regime-embeddings $\mathcal{E}_{1:\widetilde{K}}$. Likelihood updates are guaranteed to converge to a *local optimum*.

The hyperparameter $\widetilde{K}$ models the number of regimes K, chosen as a large value exceeding the oracle value. Redundant regimes are effectively ignored in practice.



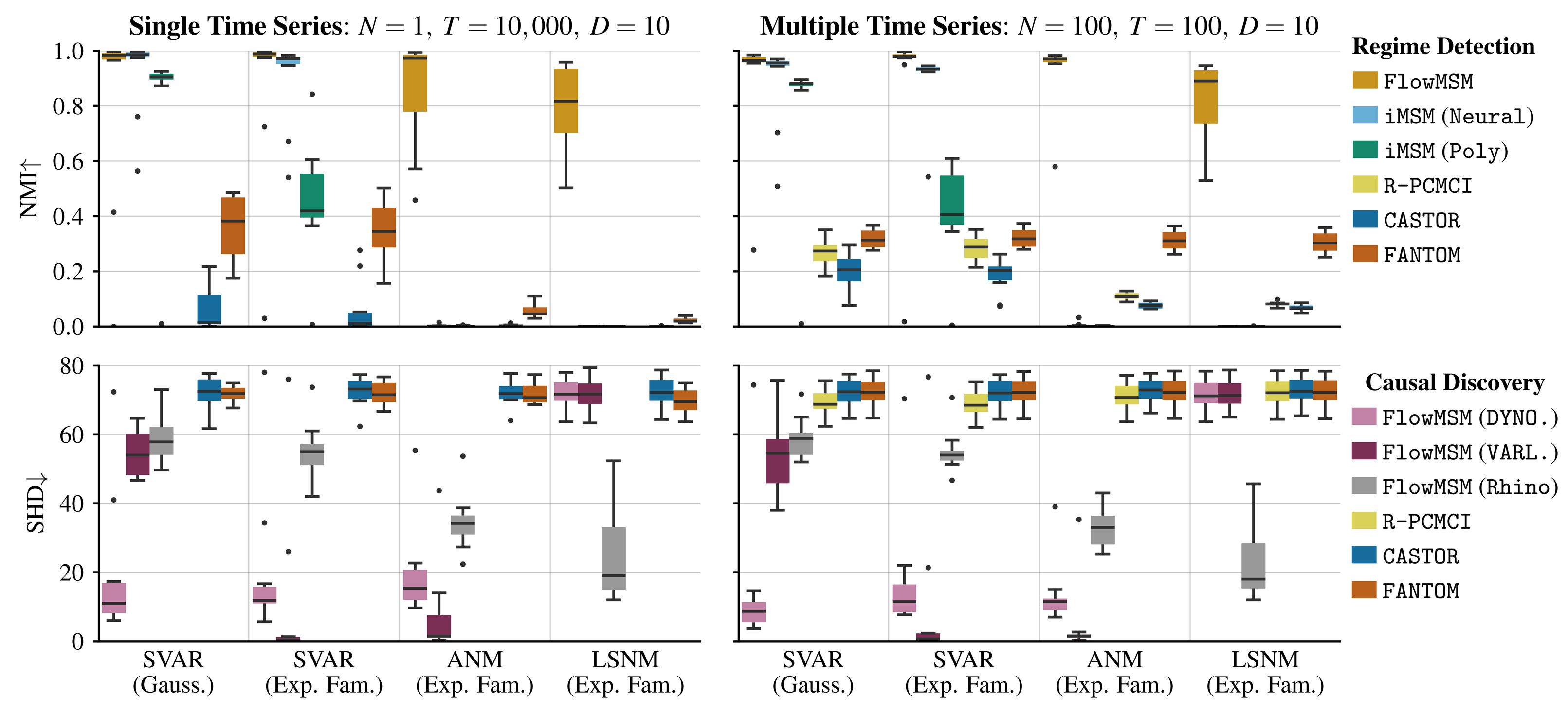
Sample splitting. We assign the sequences $x_{t-1:t}^{(n)}$ to the MAP regime. This creates $\widetilde{K}$ clusters with partially overlapping *causally stationary* **sliding windows**, contingent on regime accuracy.

Causal discovery. To estimate *window graphs* from the clustered samples, **FlowMSM** can be paired with any stationary causal discovery method that allows for instantaneous effects.

Experiments

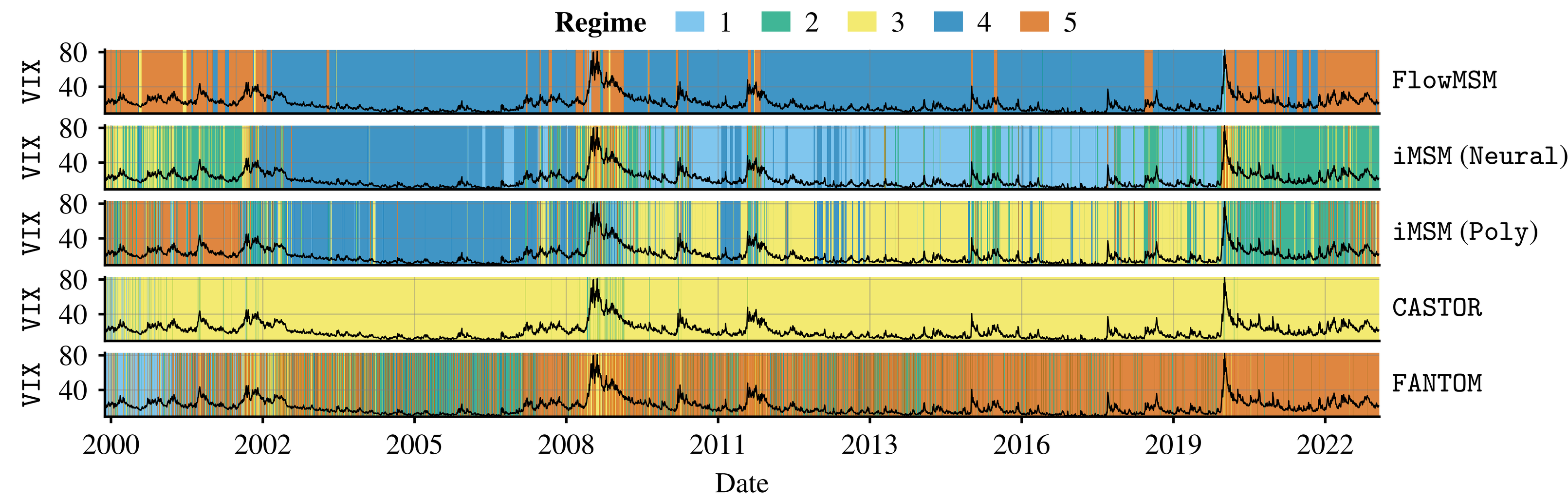
Synthetic data. One "easy" Gaussian and three "challenging" SCMs with skewed Gamma noise.

Results. Below we display the performance on (*top*) regime detection and (*bottom*) causal discovery on (*left*) a single long time series and (*right*) multiple smaller time series, with $K = 3$ regimes. Our framework achieves strong performance compared to baseline methods.



Real-world data. The Fama-French model captures systemic patterns in stock returns [4], which we supplement with **AAPL** stock. Without ground-truth regimes, we provide a qualitative study.

Results. We estimate regimes on daily data, with an overlay of the **VIX** volatility index *(not used in training)*. **FlowMSM** differentiates stable from volatile periods, such as the 2008 financial crisis. We also investigate causal interpretations of the Fama-French model, finding only partial support.



Interested to learn more? Scan the QR!

Future work. Non-stationary causal models with latent confounders other than regime variables, further relaxing *(conditional) causal sufficiency*.

References

[1] C. Balsells-Rodas et al. "On the Identifiability of Switching Dynamical Systems". In: *ICML*. 2024.
[2] S. J. Yakowitz et al. "On the identifiability of finite mixtures". In: *The Annals of Mathematical Statistics* (1968).
[3] O. Barndorff-Nielsen. "Identifiability of mixtures of exponential families". In: *Journal of Mathematical Analysis and Applications* (1965).
[4] E. F. Fama et al. "A five-factor asset pricing model". In: *Journal of Financial Economics* (2015).

University of Amsterdam AMLAB adyen Imperial College London Universität des Saarlandes