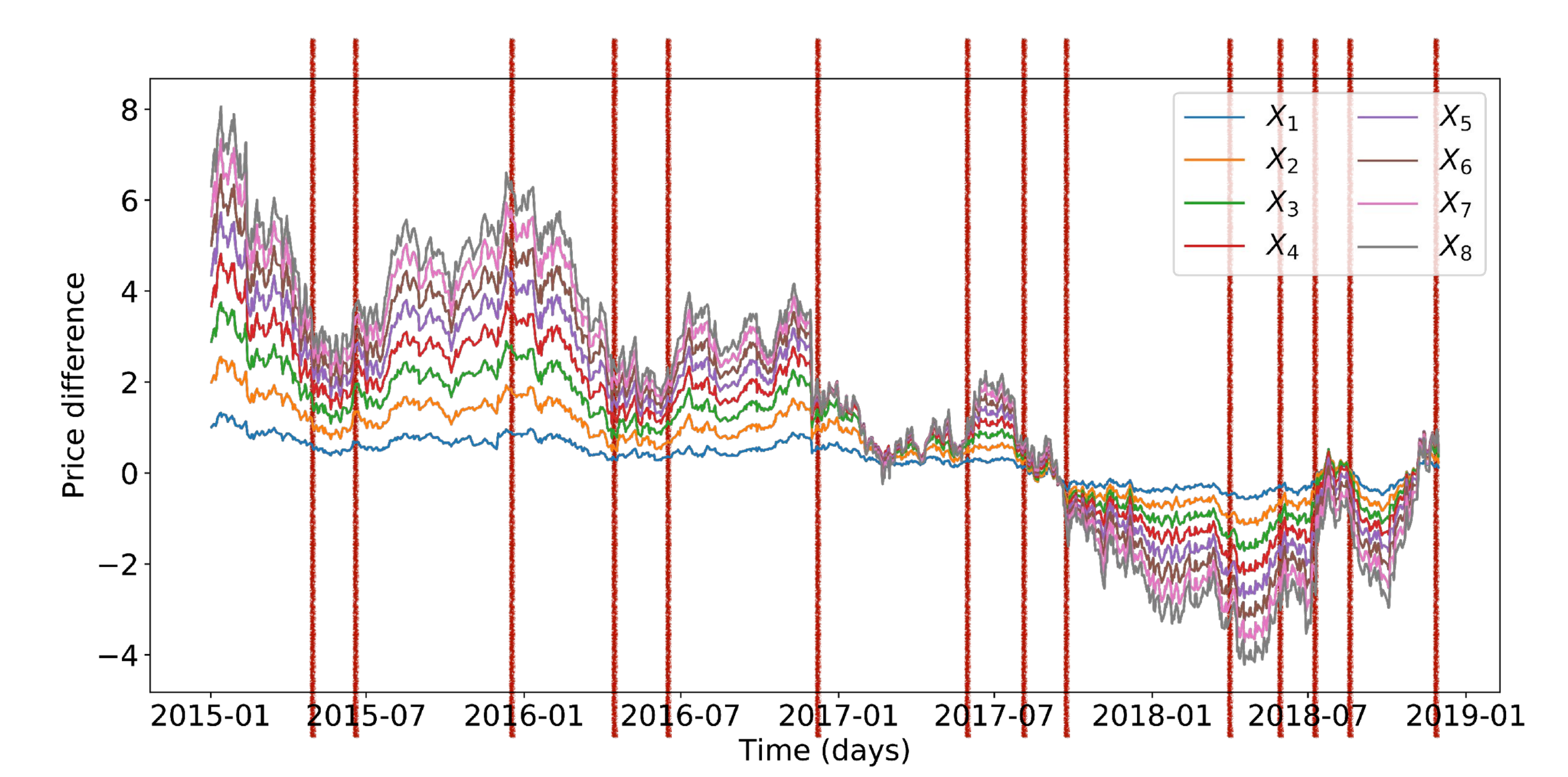
Temporal Pattern Detection in Time-Varying Graphical Models

Federico Tomasi, Spotify
Veronica Tozzo, Massachusetts General Hospital, Harvard Medical School
Annalisa Barla, Università di Genova

Introduction

Many real-world phenomena include variables which change their behaviour over time, leading to a change in their interactions.

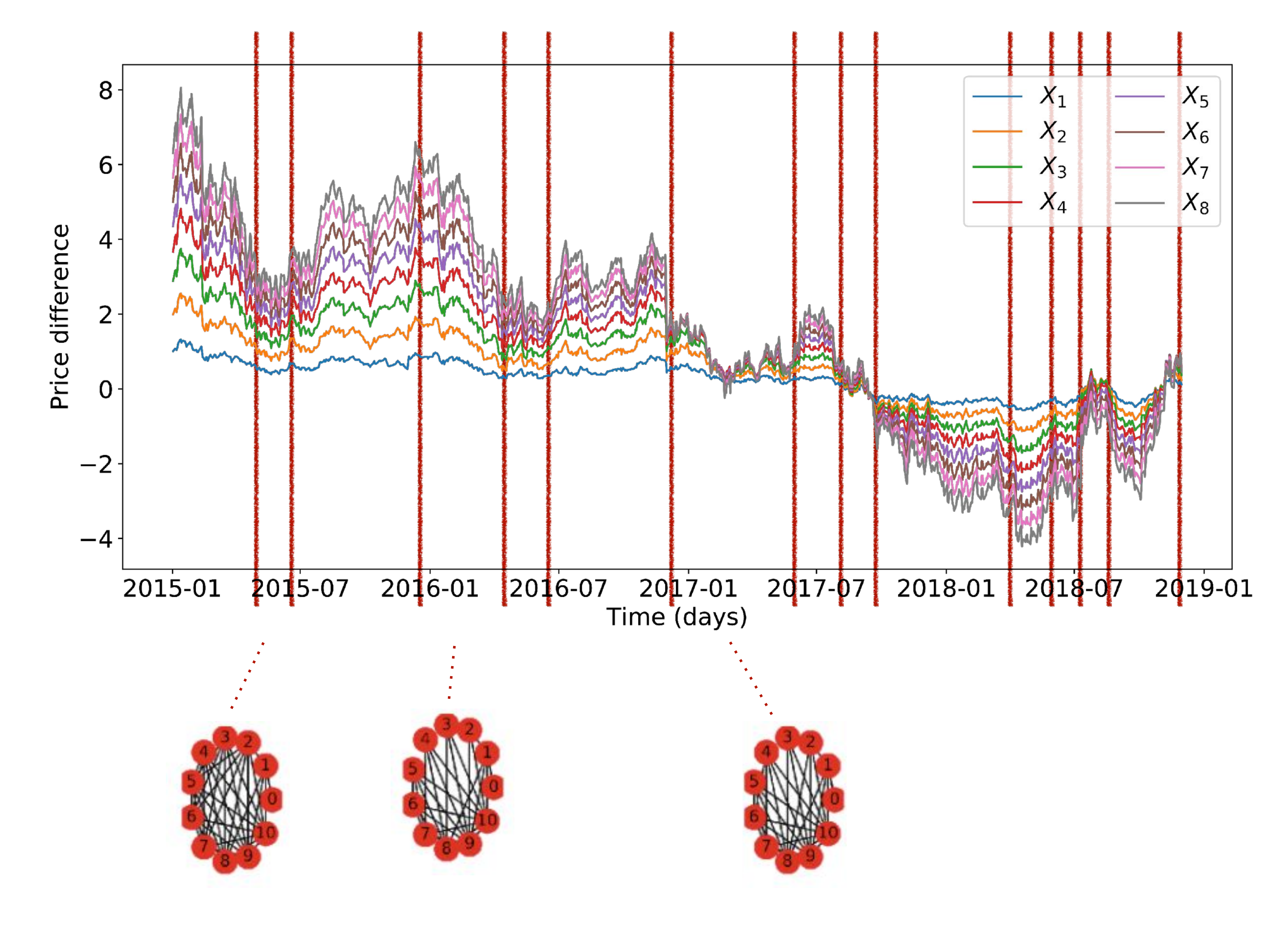


Introduction

Many real-world phenomena include variables which change their behaviour over time, leading to a change in their interactions.

Gaussian graphical models (GGMs) can be used to infer a dynamical network on discrete chunks of time-stamps by assuming a temporal consistency between consecutive time points (Markovianity assumption).

What about complex, periodic patterns?



Contribution

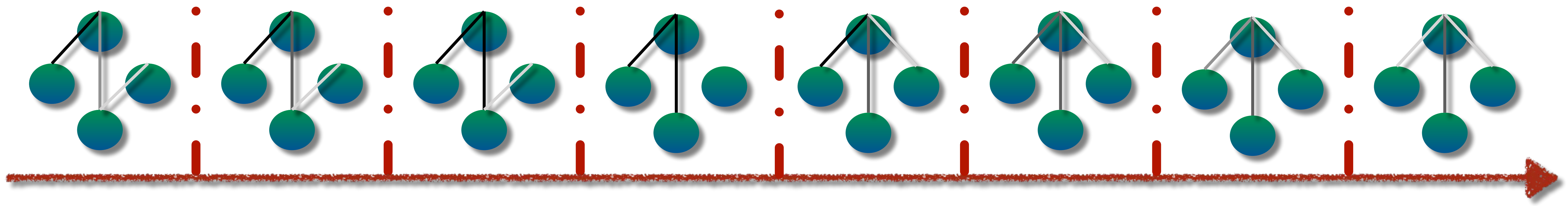
Here we address the problem of possible non-Markovian temporal relationships by exploiting similarity kernels to infer a dynamical network.

We consider two cases:

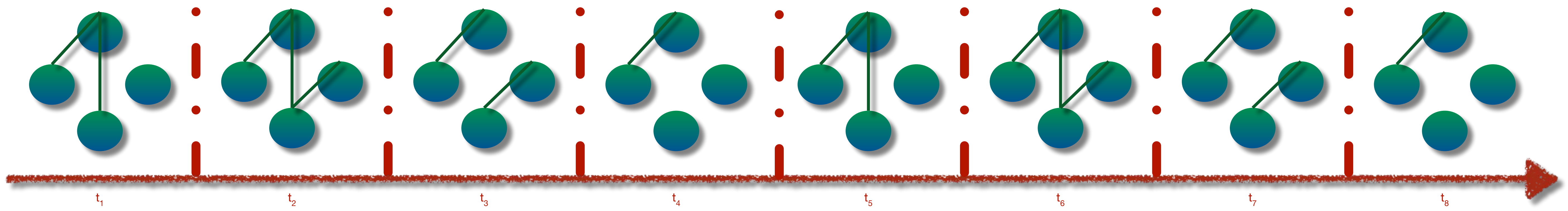
- when the similarity pattern is known a priori (e.g., smooth, periodic) we use such knowledge by incorporating a suitable kernel in the model;
- (when the similarity is unknown or cannot easily be modelled by kernel functions we resort to the simultaneous inference of the kernel and the network structure.

Two inferred networks are said to be exactly consistent if the distance between the related network structures is zero.
The kernel allows to specify different types of consistency:

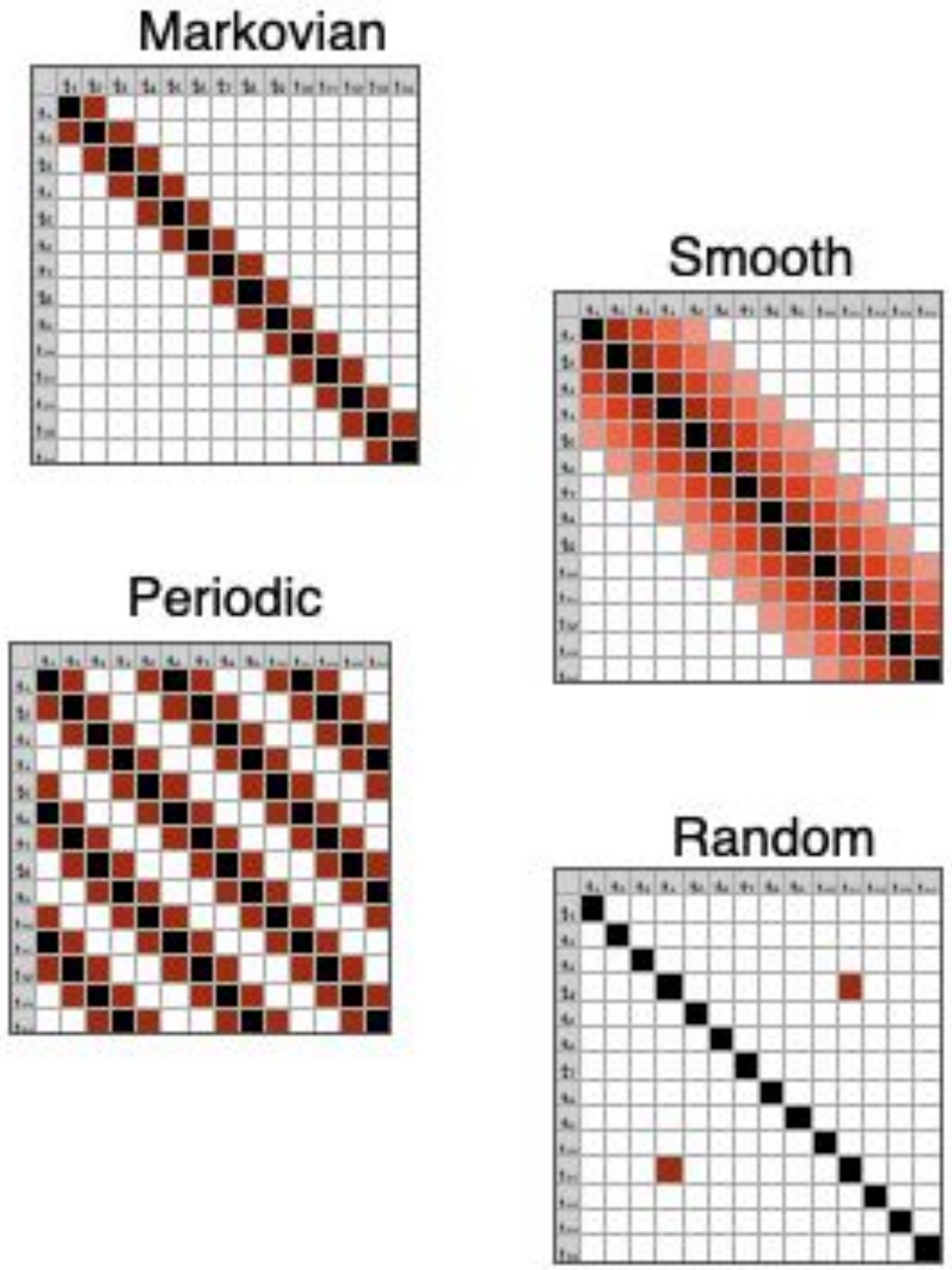
- promoting smooth changes

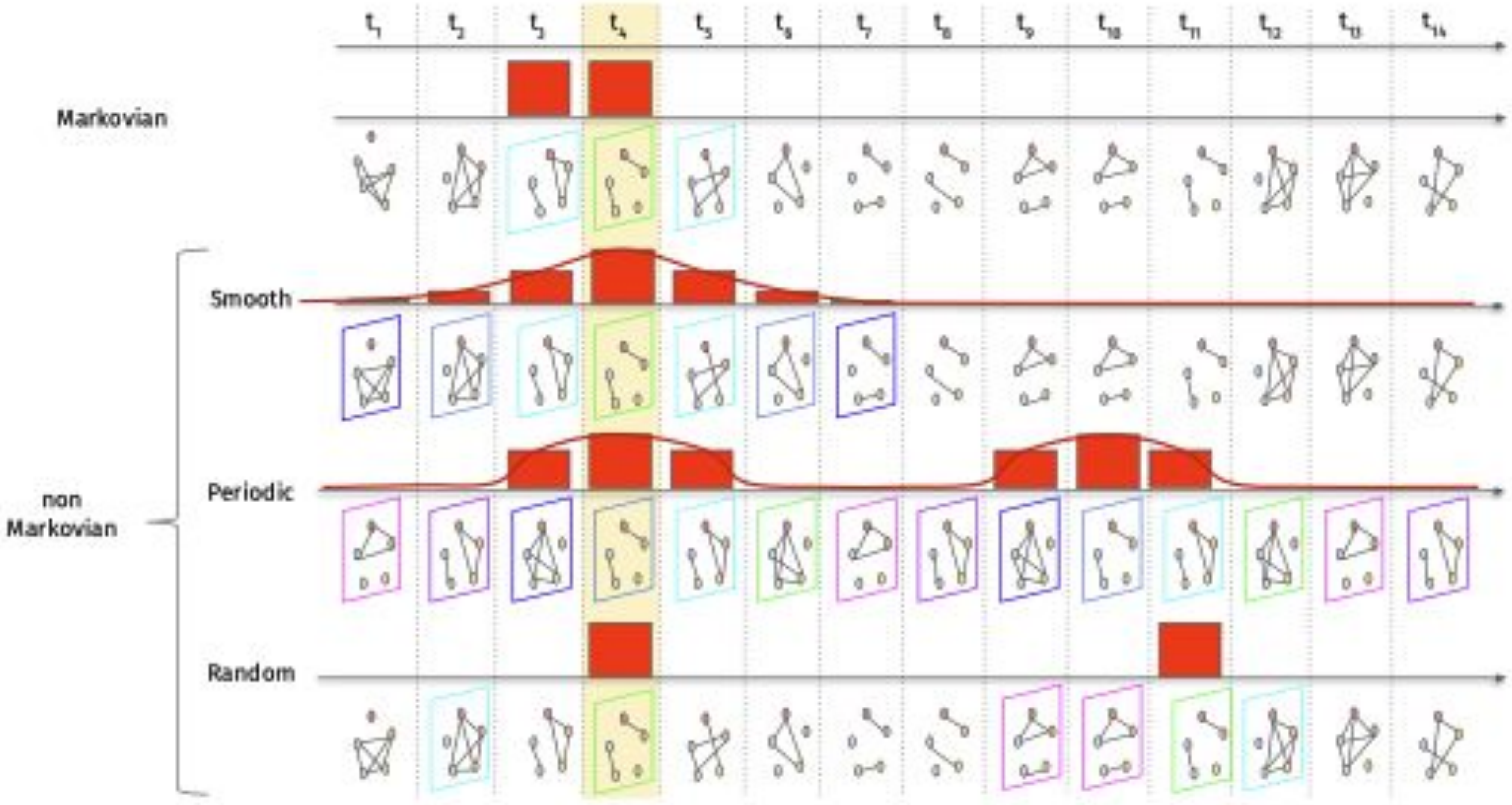


- promoting punctual changes



Two inferred networks are said to be **dependent** if they model the same or a similar state of the system. The dependency is modelled by a suitable kernel.





Experiments

- On synthetic data, we compute the divergence from the ground truth in terms of structure of the network, edges weight and clustering score;
- Our method improves using both periodic, random or conditioned-random (covariances are dependent on consecutive states) patterns;
- TGL_κ has a high score when data exhibit a clear pattern;
  TGL_P (with automatic pattern discovery) has increasing accuracy in structure inference and reduced error in dynamical network estimation.

experiment		method	balanced accuracy	average precision	MCC	MSE	V-measure
(a) Periodic-pattern		GL	0.505 ± 0.002	0.029 ± 0.003	0.093 ± 0.017	0.190 ± 0.003	0.790 ± 0.096
		TGL	0.558 ± 0.005	0.301 ± 0.014	0.122 ± 0.003	1.470 ± 0.015	0.791 ± 0.089
		WP	0.498 ± 0.002	0.022 ± 0.001	0.002 ± 0.007	1.133 ± 0.003	0.730 ± 0.137
	Ours	TGL_κ	0.560 ± 0.008	$\mathbf{0.372 \pm 0.046}$	0.117 ± 0.005	$\mathbf{0.148 \pm 0.007}$	0.800 ± 0.090
	Ours	TGL_P	$\mathbf{0.577 \pm 0.003}$	0.341 ± 0.014	$\mathbf{0.130 \pm 0.002}$	0.707 ± 0.211	$\mathbf{0.822 \pm 0.113}$
(b) Random-pattern		GL	0.505 ± 0.001	0.029 ± 0.003	0.094 ± 0.014	0.191 ± 0.002	0.802 ± 0.088
		TGL	0.560 ± 0.004	0.299 ± 0.011	0.122 ± 0.004	1.475 ± 0.012	0.788 ± 0.093
		WP	0.497 ± 0.003	0.022 ± 0.001	0.007 ± 0.005	1.130 ± 0.006	0.766 ± 0.099
	Ours	TGL_κ	0.553 ± 0.007	0.284 ± 0.033	0.113 ± 0.005	$\mathbf{0.173 \pm 0.017}$	0.804 ± 0.088
	Ours	TGL_P	$\mathbf{0.574 \pm 0.004}$	$\mathbf{0.331 \pm 0.015}$	$\mathbf{0.130 \pm 0.003}$	0.758 ± 0.181	$\mathbf{0.811 \pm 0.081}$
(c) Conditioned-random pattern		LTGL	0.509 ± 0.002	0.278 ± 0.003	0.082 ± 0.010	12.229 ± 0.279	0.497 ± 0.037
	Ours	$LTGL_\kappa$	$\mathbf{0.521 \pm 0.016}$	$\mathbf{0.299 \pm 0.045}$	$\mathbf{0.107 \pm 0.117}$	$\mathbf{11.861 \pm 0.483}$	0.492 ± 0.058
	Ours	$LTGL_P$	0.500 ± 0.000	0.251 ± 0.002	-0.001 ± 0.002	13.711 ± 0.117	$\mathbf{0.580 \pm 0.047}$

Food-terms Search Trends

We used TGL_P for the inference of similarity patterns in the setting of web search terms, in particular on food-terms search in US from 2004 to 2016.

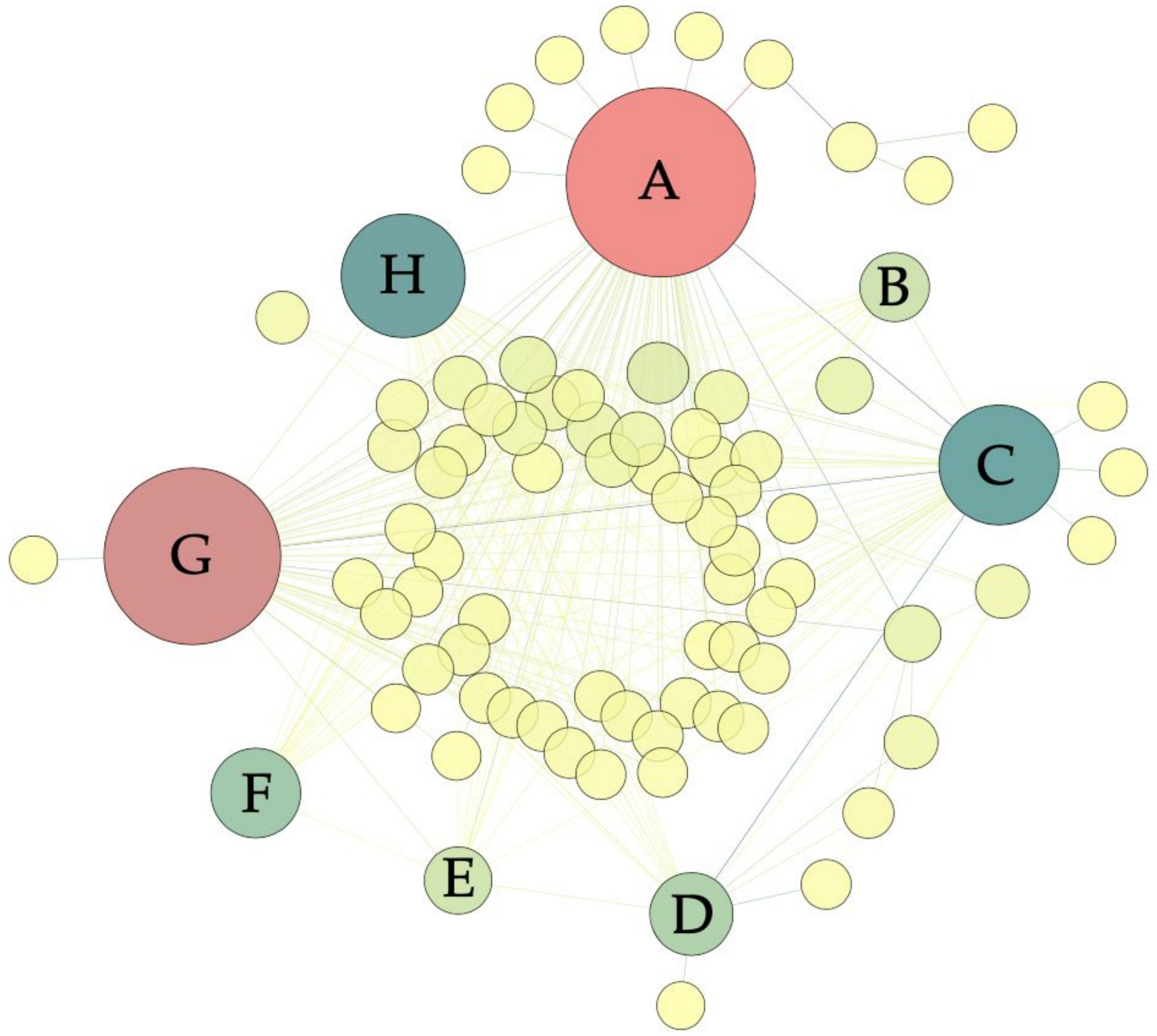


Fig. 5: Network related to holidays weeks (light green cluster) considering node degrees higher than 4 and their connected 1-degree neighbours. The hubs are the following terms: A-macaron, B-cauliflower, C-moscow-mule, D-quinoa, E-taco, F-brussel sprouts, G-kale, H-chia.

Conclusion

- Novel kernel-based extensions for graphical modelling. Dependency in time is driven by kernel functions, capturing evolving relationships between not necessarily contiguous time steps.

- The automatic discovery of the kernel can be used to cluster temporal series.

- Possible extensions of our models would incorporate additional prior knowledge on the problem, e.g., forcing subgroups of variables to behave consistently in time. Also, chunks length could potentially be non-homogeneous.