AutoInformation State aggregation

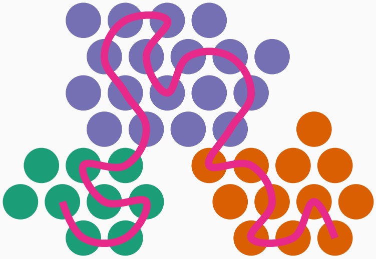
a dynamical point of view

Mauro Faccin
IRD/CEPED, Université de Paris
@ COMPLEX NETWORKS

Projected Markov Chain

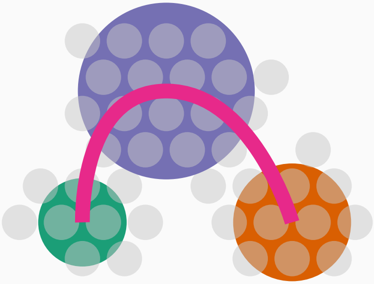
Markov Chain

$$\ldots, x_{\text{past}}, x_{\text{now}}, x_{\text{future}}, \ldots$$



Projection

$$\ldots, y_{\text{past}}, y_{\text{now}}, y_{\text{future}}, \ldots$$



Aggregation strategies

Non-linear correlations

AutoInformation

$$I(y_t; y_{t-\tau})$$

Non-linear *correlation* between successive time-steps

M.F. et al, Journal of Complex Networks, cnx055

AutoInformation

$$I(y_t; y_{t-\tau})$$

Non-linear *correlation* between successive time-steps

$$I(y_t; y_{t-\tau}) = \underbrace{I(y_t; y_{t-\tau}, \ldots)}_{{}^{1}\text{Predictability}} - \underbrace{I(y_t; y_{t-2\tau}, \ldots | y_{t-\tau})}_{{}^{2}\text{non-Markovianity}}$$

where τ represents a time-scale parameter.
Maximized by a Markov chain that:

1. Maximize predictability of the dynamics
2. Minimize non-Markovianity (effective memories from the projection)

M.F. et al, Journal of Complex Networks, cnx055

Random walker covariance

χ_c characteristic function of class c

$$Q = \sum_c \mathbf{Cov}\left(\chi_c(t), \chi_c(t+1)\right)$$

Auto-covariance of the dynamics on the partition space.

Modularity:

$$Q = \frac{1}{2m} \sum_{ij} \left[A_{ij} - \frac{k_i k_j}{2m} \right] \delta(c_i, c_j)$$

Linear correlation between consecutive time-steps.

Shen et al. (2010) PRE, 82, 016114

Generative models as particular case

Fitting a generative model (e.g. DC-SBM) to the data through log-likelihood maximization can be seen as maximizing the AutoInformation for paths of lenght $\tau = 1$ (e.g. links).

$$I(Y_t; Y_{t-1}) = H(Y_y) + H(Y_{t-1}) - H(Y_y, Y_{t-1})$$

$$H(Y_t) = -\sum_c \frac{e_c}{2m} \log \frac{e_c}{2m} \qquad e_c = \sum_{i \in c, j} A_{ij}$$

$$H(Y_t, Y_{y-1}) = -\sum_{cd} \frac{e_{cd}}{2m} \log \frac{e_{cd}}{2m} \qquad e_{cd} = \sum_{i \in c, j \in d} A_{ij}$$

DC-SBM

$$\mathcal{S} \propto \frac{1}{2} \sum_{cd} e_{cd} \log \frac{e_{cd}}{e_c e_d}$$

In binary symmetric networks

Karrer and Newman (2011), PRE 83, 016107.

AutoInformation

$$I(y_t; y_{t-\tau})$$

Non-linear *correlation* between successive time-steps

The parameter τ selects the *time-scale* of the aggregation.

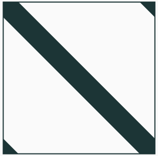
Maximizing in a naive way is not possible, one need to fix the number of classes or apply some **model selection**. E.g. a entropic Lagrange multiplier:

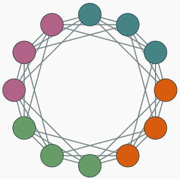
$$\mathcal{I}_{\alpha\tau} = I(y_t; y_{t-\tau}) - \alpha H(y_t)$$

Didactic Examples.

A regular ring lattice with N nodes, each connected with ***k*** neighbours.

How many classes?

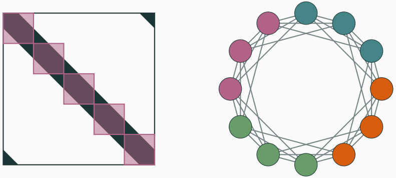
Adj: 



A regular ring lattice with N nodes, each connected with ***k*** neighbours.

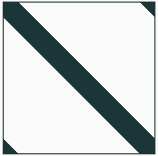
How many classes?

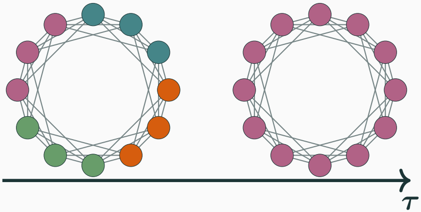
Adj:



A regular ring lattice with N nodes, each connected with ***k*** neighbours.

How many classes?

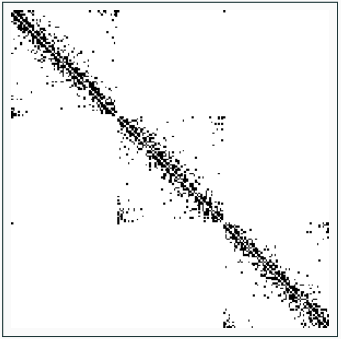
Adj: 



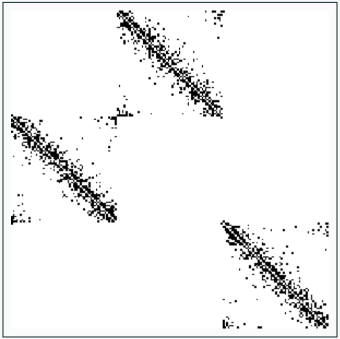
Example 1: Range dependant graphs

$$p_{ij} = \alpha_{c_i c_j} \cdot (\gamma_{c_i c_j})^{d_{ij}}$$
$$\alpha_{c_i c_j}, \gamma_{c_i c_j} \in [0, 1]$$

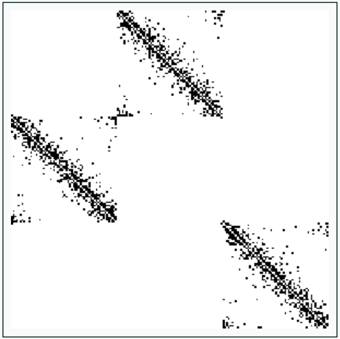
with d_{ij} a (normalized) distance between nodes aligned on a cycle.



Example 1: Range dependent graphs

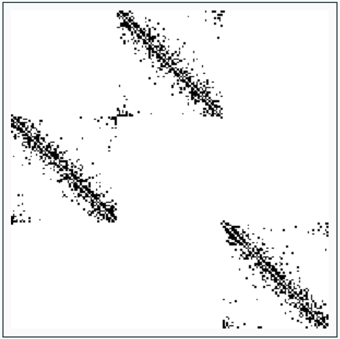


Example 1: Range dependent graphs

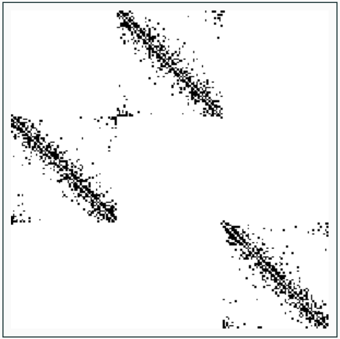


DC-SBM

Example 1: Range dependent graphs



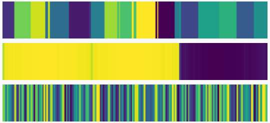
Example 1: Range dependent graphs



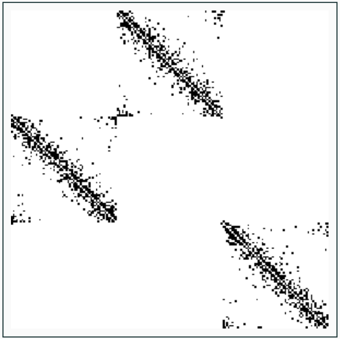
DC-SBM

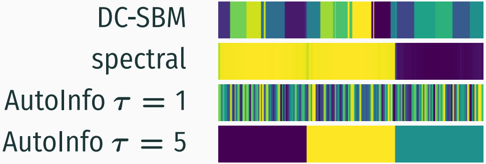
spectral

AutoInfo $\tau = 1$

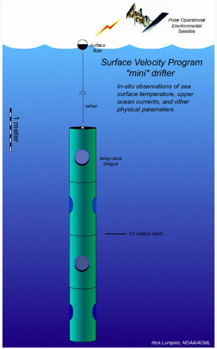


Example 1: Range dependent graphs





Example 2. Ocean buoys



Global Drifter Program

GDP Array
AOML Drifter Data Assembly Center
Mon, 04 Oct 2021

No. of Buoys = 1471

ID WMO

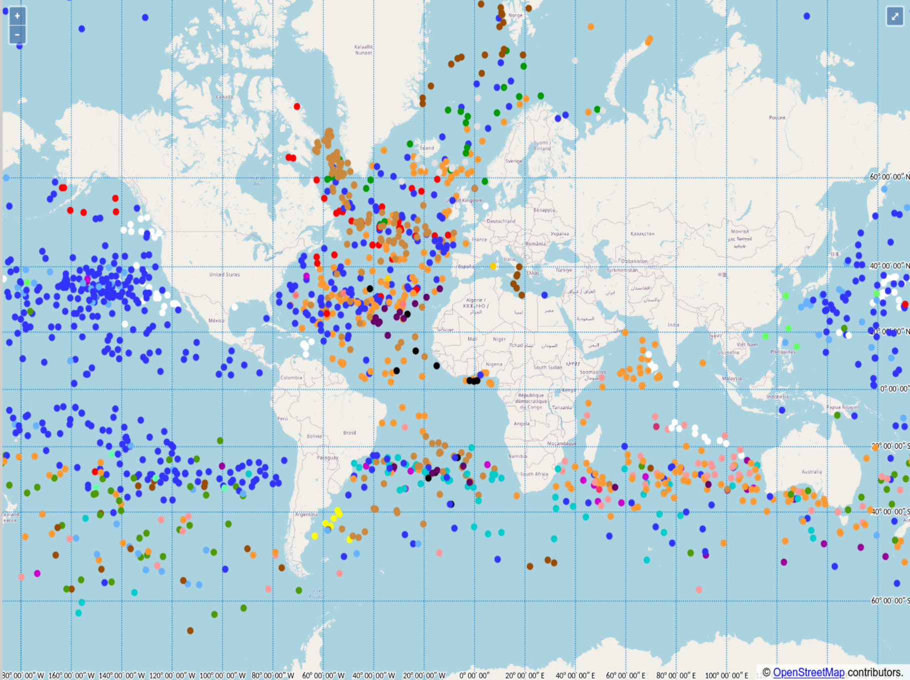
Search..

Map Viewing Options

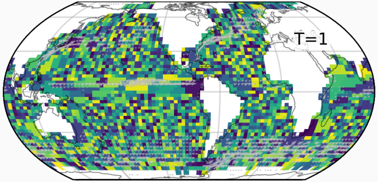
- Deploying Country
- Buoy Type
- Buoy Drogue Status

Deploying Country

Argentina (7)	Australia (48)
Barbados (3)	Brazil (12)
Canada (40)	Chile (4)
China (6)	Denmark (1)
France (272)	Germany (12)
Iceland (23)	India (3)
Indonesia (1)	Italy (51)
Japan (11)	Korea Rep. of (63)
New Zealand (52)	Netherlands (14)
Portugal (19)	Seychelles (1)
South Africa (59)	Spain (2)
Tonga (1)	UK (153)
USA (539)	Unknown (74)

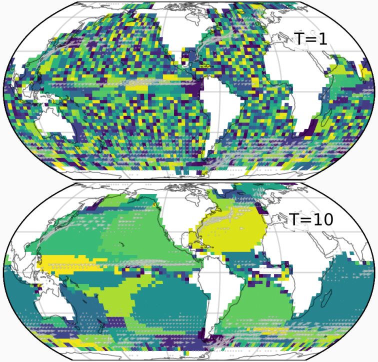


Example 3. Ocean buoys

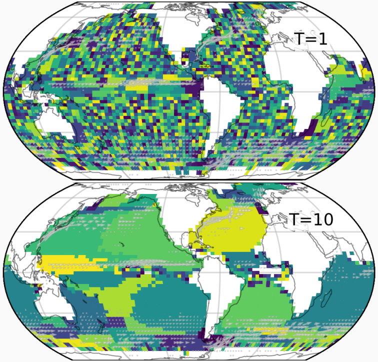


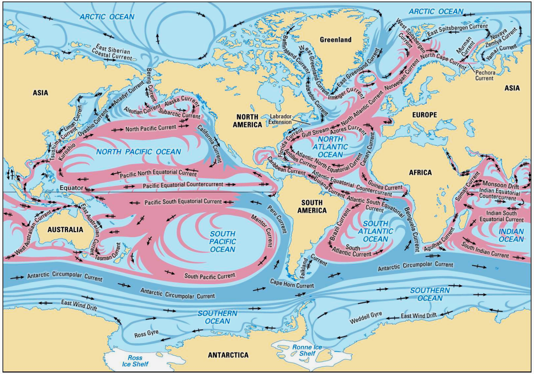
Each time step lasts 16 days.

Example 3. Ocean buoys



Each time step lasts 16 days.





Each time step lasts 16 days.

🏃 Finally...

Questions?

FIG. 173.—FRANKLIN'S CHART OF THE GULF STREAM.

Joint work with:

JC Delvenne
UCLouvain

M Schaub
RWTH AACHEN UNIVERSITY

https://maurofaccin.github.io
mauro.fccn@gmail.com

Code at:
https://maurofaccin.github.io/aisa