Dynamics, Modularity and Robustness of Complex Networks

R. Lambiotte
Institute for Mathematical Sciences
Imperial College London

with M. Barahona and J.-C. Delvenne

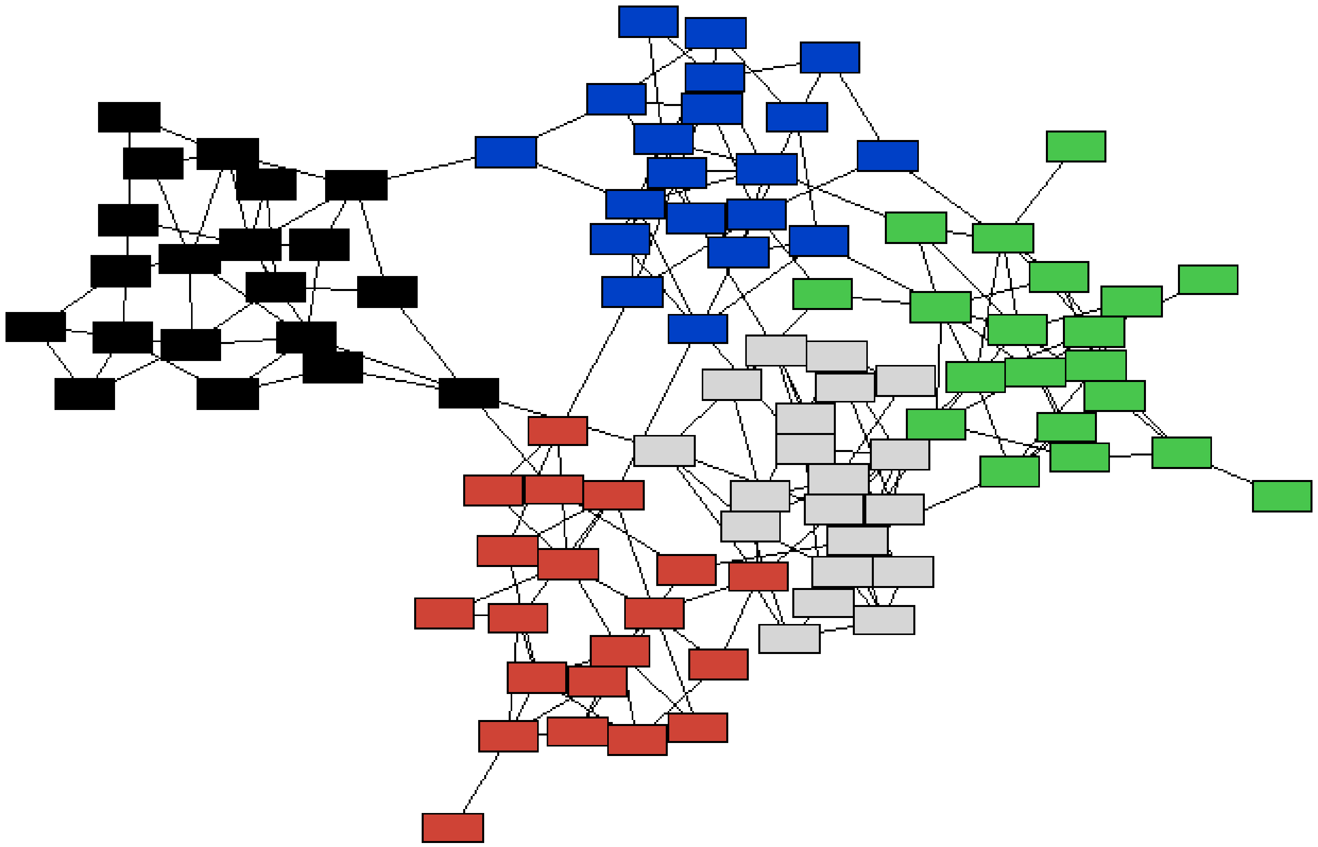
Imperial College London
100 years of living science
100

R. Lambiotte, J.-C. Delvenne and M. Barahona, *arXiv:0812:1770*

1. Modules and Hierarchies
2. Stability of a partition
    a. Stability vs Modularity
    b. Time as a resolution parameter
    c. Different processes lead to different stabilities
3. Optimisation and selection of the most relevant time scales/robustness

Modular Networks

Most networks are very inhomogeneous and are made of modules: many links within modules and a few links between different modules



Internet

Power grids

Food webs

Metabolic networks

Social networks

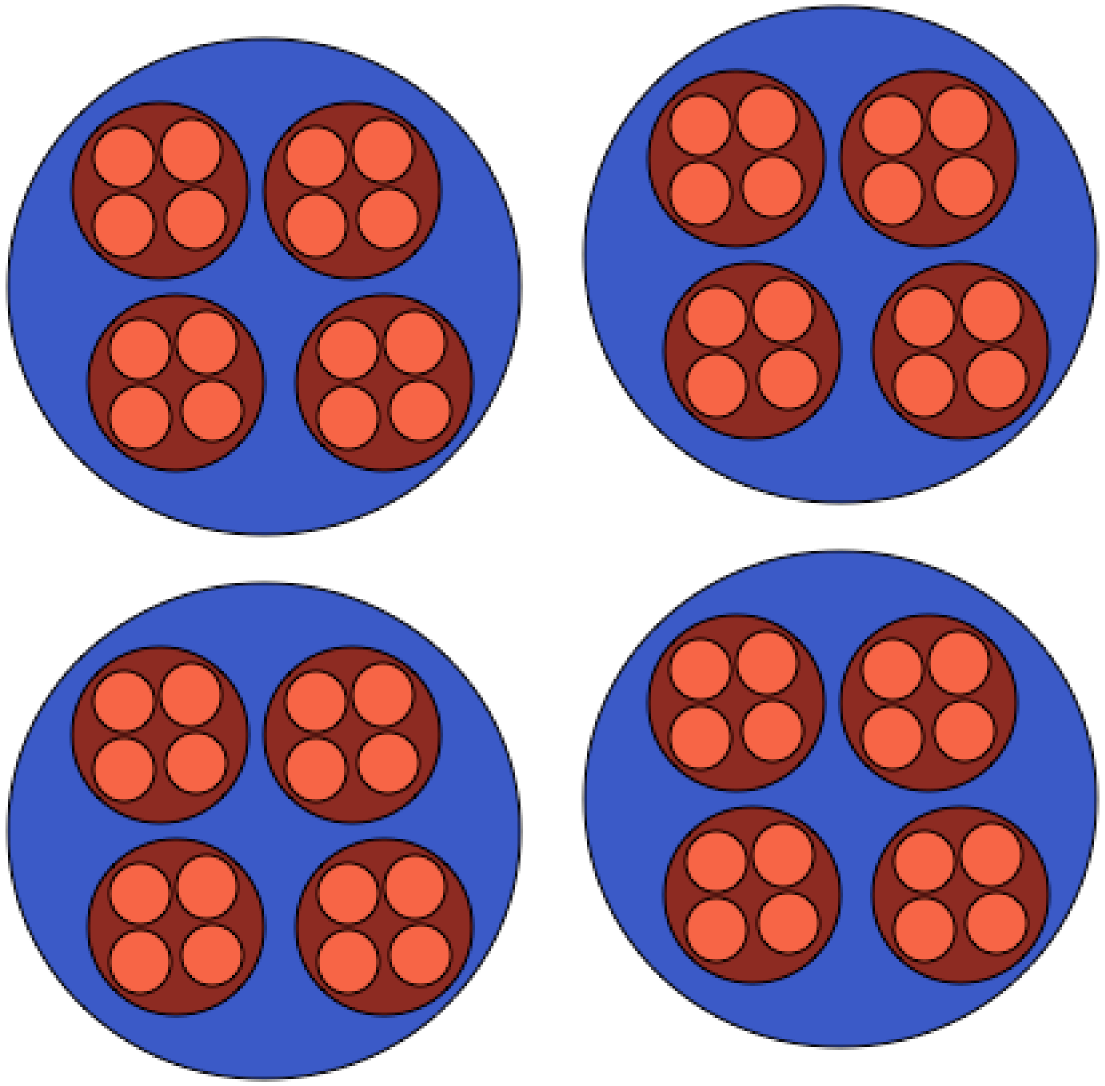
The brain

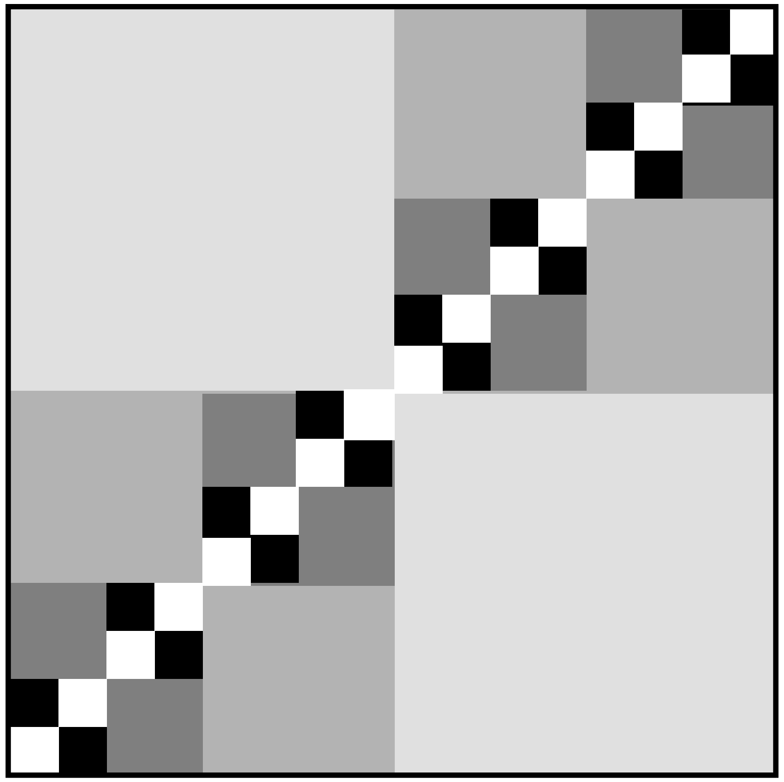
Etc.

Simon, H. (1962). The architecture of complexity. Proceedings of the American Philosophical Society, 106, 467–482.

Multi-scale Modular Networks

Networks have a hierarchical structure: modules within modules





Simon, H. (1962). The architecture of complexity. Proceedings of the American Philosophical Society, 106, 467–482.

Modular Networks and dynamics

Is it possible to uncover those modules/hierarchies in large networks?

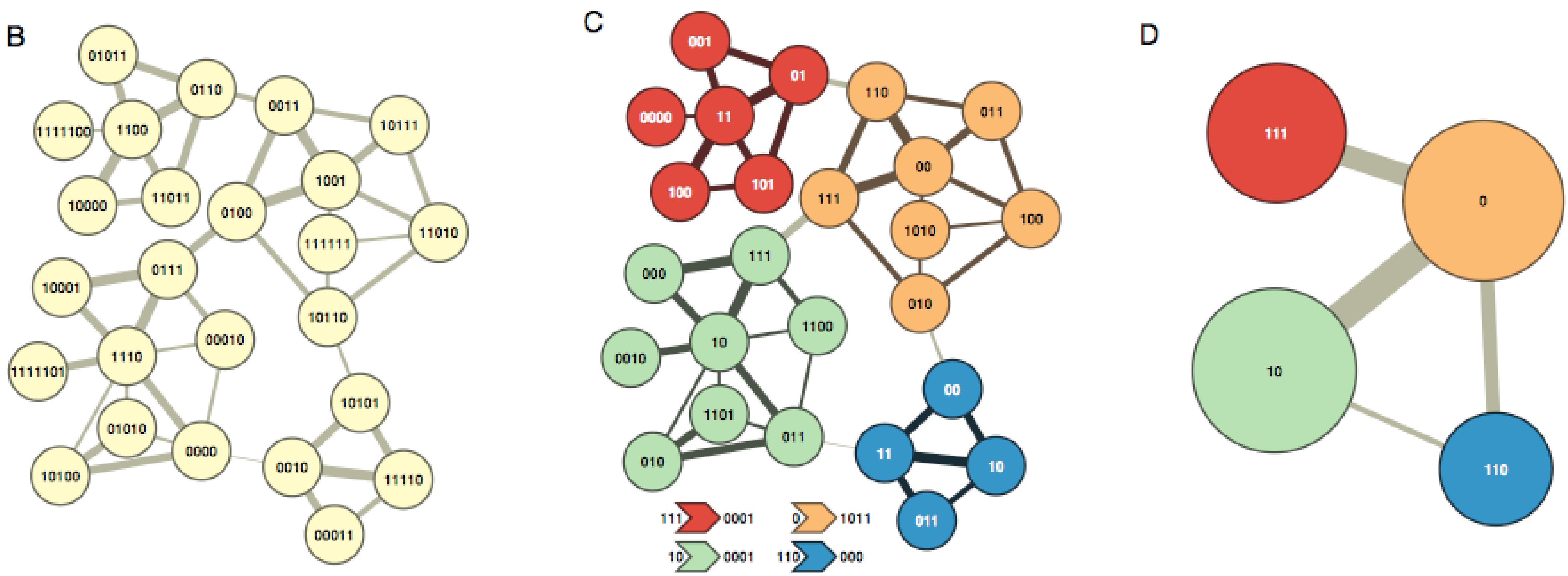
Hundreds of heuristics to optimise modularity.

How does such modularity affect dynamics?

A. Arenas, A. Diaz-Guilera and C.J. Pérez-Vicente (*Phys. Rev. Lett.*, 2006).
R. Lambiotte, M. Ausloos and J.A. Holyst, *Phys. Rev. E* **75**, 030101(R) (2007).

Modular Networks

Uncovering communities/modules helps to understand the structure of the network, to uncover similar nodes and to draw a readable map of the network (when N is large).



Find a partition of the network into communities

Coarse-grained description

Martin Rosvall and Carl T. Bergstrom, PNAS 105, 1118 –1123 (2008)

Modular Networks and dynamics

Is it possible to uncover those modules/hierarchies in large networks?

Hundreds of heuristics to optimise modularity.

How does such modularity affect dynamics?

A. Arenas, A. Diaz-Guilera and C.J. Pérez-Vicente (*Phys. Rev. Lett.*, 2006).
R. Lambiotte, M. Ausloos and J.A. Holyst, *Phys. Rev. E* **75**, 030101(R) (2007).

Modular Networks and dynamics

Is it possible to uncover those modules/hierarchies in large networks?

Hundreds of heuristics to optimise modularity.

Is it possible to use dynamics to characterize (and uncover?) the modular structure of a network?

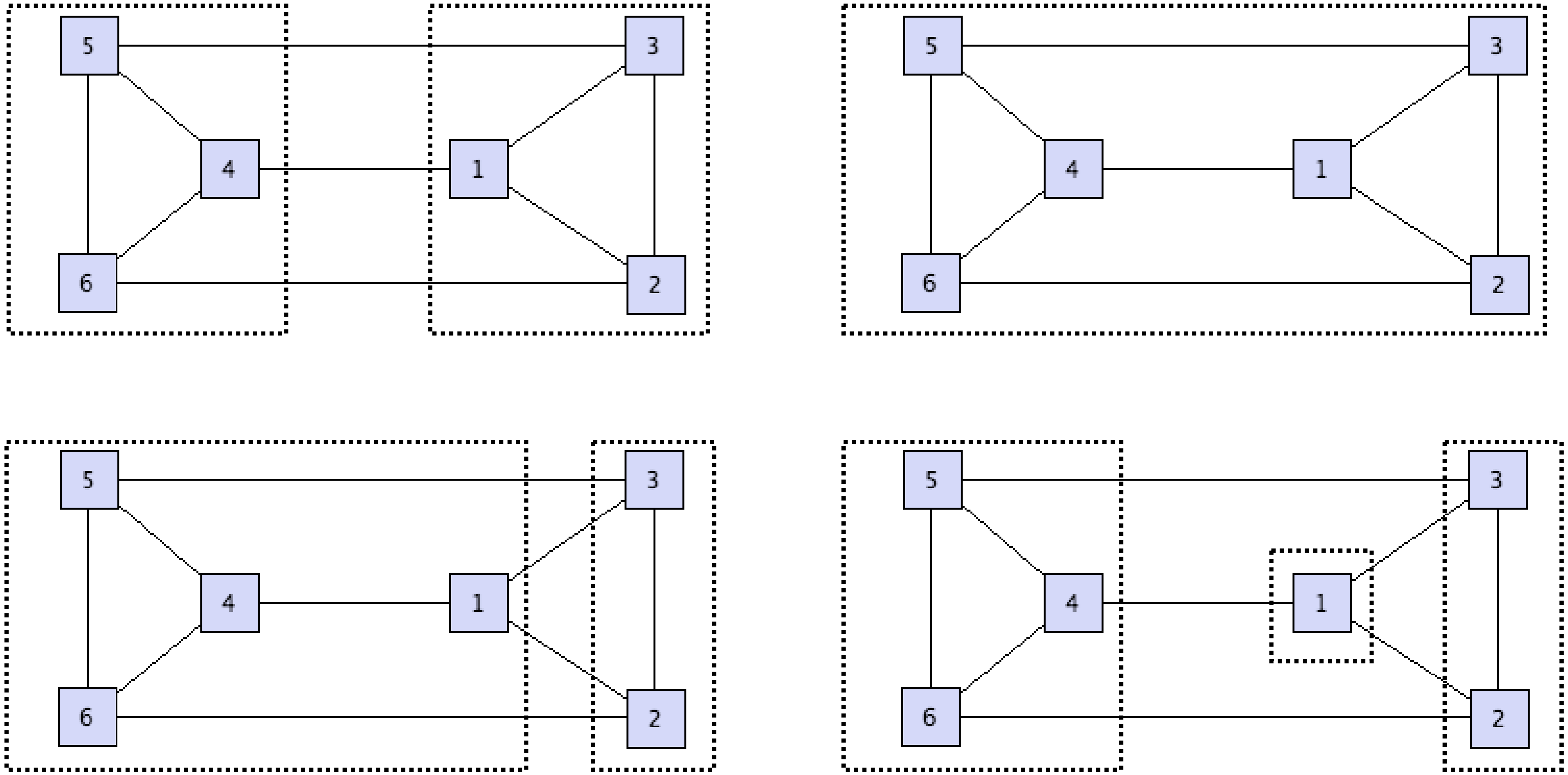
e.g. Walktrap (RW exploration), Rosvall and Bergstrom (PNAS, 2008)

How does such modularity affect dynamics?

A. Arenas, A. Diaz-Guilera and C.J. Pérez-Vicente (*Phys. Rev. Lett.*, 2006).
R. Lambiotte, M. Ausloos and J.A. Holyst, *Phys. Rev. E* **75**, 030101(R) (2007).

Quality of a partition

What is the best partition of a network into modules?



....

Modularity

Q = fraction of edges within communities - expected fraction of such edges

Let us attribute each node *i* to a community c_i

$$Q = \frac{1}{2m} \sum_{i,j} \Big[A_{ij} - P_{ij} \Big] \delta(c_i, c_j)$$

$$Q \in [-1, 1]$$

$$P_{ij} = \frac{k_i k_j}{2m}$$

expected number of links between i and j

→

$$Q_C = \frac{1}{2m} \sum_{i,j} \Big[A_{ij} - k_i k_j / 2m \Big] \delta(c_i, c_j)$$

Modularity optimisation

Different types of algorithm for different applications:

Small networks ($<10^3$): Simulated Annealing

Intermediate size (10^3 - 10^4): Spectral methods, PL, etc.

Large size ($>10^5$): greedy algorithms

Modularity

Optimising modularity uncovers one partition

What about sub (or hyper)-communities in a hierarchical network?

Reichardt & Bornholdt

$$Q_\gamma = \frac{1}{2m}\sum_{i,j}\Big[A_{ij} - \gamma P_{ij}\Big]\delta(c_i, c_j)$$

Arenas et al.

$$Q(A_{ij} + rI_{ij})$$

Tuning parameters allow to uncover communities of different sizes

Reichardt & Bornholdt different of Arenas, except in the case of a regular graph where

$$\gamma = 1 + r/\langle k \rangle$$

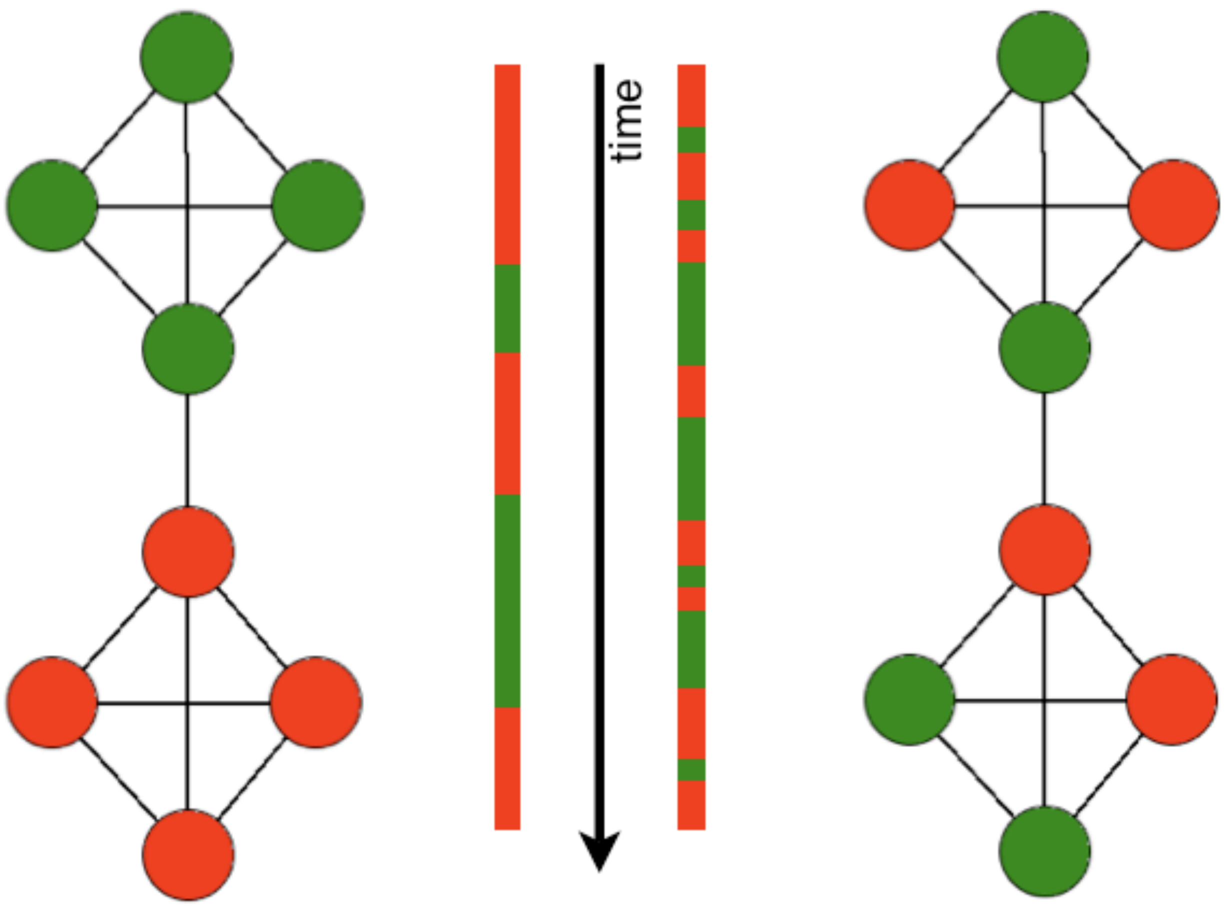
*J. Reichardt and S. Bornholdt, Phys. Rev. E **74**, 016110 (2006). Statistical mechanics of community detection*

*A Arenas, A Fernandez, S Gomez, New J. Phys. **10**, 053039 (2008). Analysis of the structure of complex networks at different resolution levels*

Stability

The quality of a partition is determined by the patterns of a flow within the network: a flow should be trapped for long time periods within a community before escaping it.

The stability of a partition is defined by the statistical properties of a random walker moving on the graph:



M. Rosvall and C. T. Bergstrom, PNAS 105, 1118 –1123 (2008)

J.-C. Delvenne, S. Yaliraki & M. Barahona, Stability of graph communities across time scales. arXiv:0812.1811.

Stability

The quality of a partition is determined by the patterns of a flow within the network: a flow should be trapped for long time periods within a community before escaping it.

The stability of a partition is defined by the statistical properties of a random walker moving on the graph:

$$R(t) = \sum_{C \in \mathcal{P}} P(C, t_0, t_0 + t) - P(C, t_0, \infty)$$

$P(C, t_0, t_0 + t)$ probability for a walker to be in the same community at times t_0 and $t_0 + t$ when the system is at equilibrium

$P(C, t_0, \infty)$ probability for two independent walkers to be in C (ergodicity)

J.-C. Delvenne, S. Yaliraki & M. Barahona, Stability of graph communities across time scales. arXiv:0812.1811.

a. Modularity vs Stability

Let us consider a random walk on an undirected network:

$$p_{i;n+1} = \sum_j \frac{A_{ij}}{k_j} p_{j;n} \quad \xrightarrow{\text{equilibrium}} \quad p_i^* = k_i/2m$$

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

Probability that a walker is in the same community initially and at time t=1

Same probability for independent walkers

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

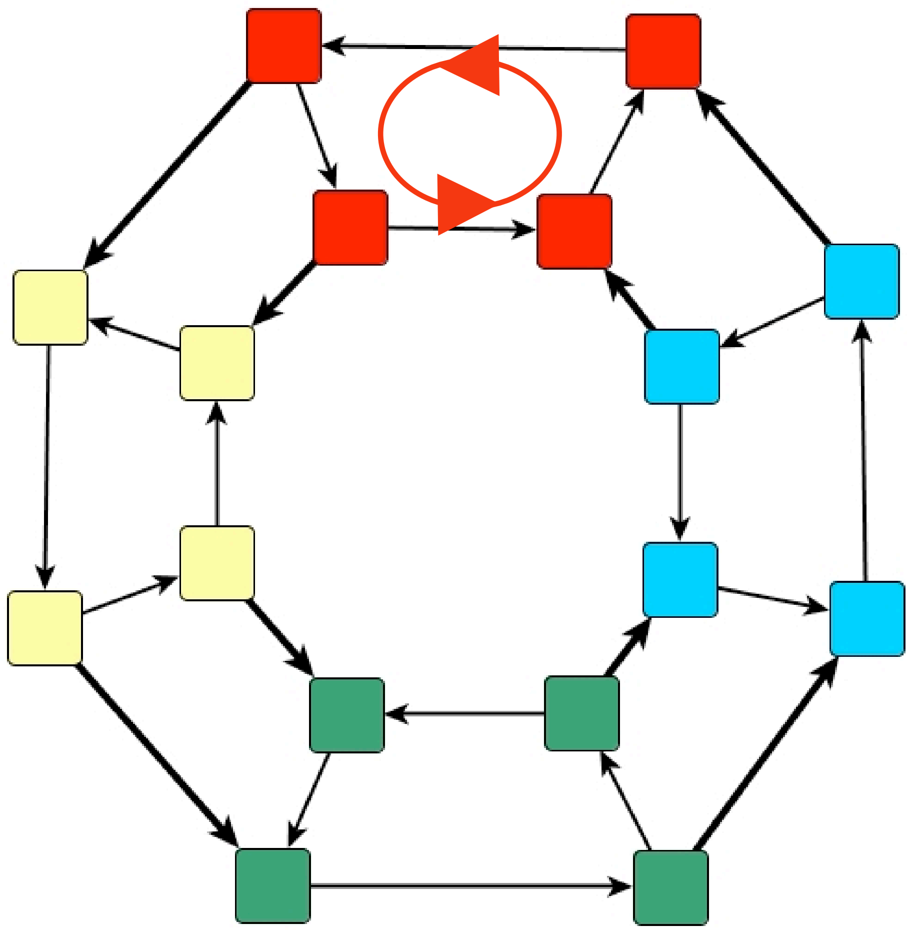
a. Modularity vs Stability

Let us consider a random walk on a directed network:

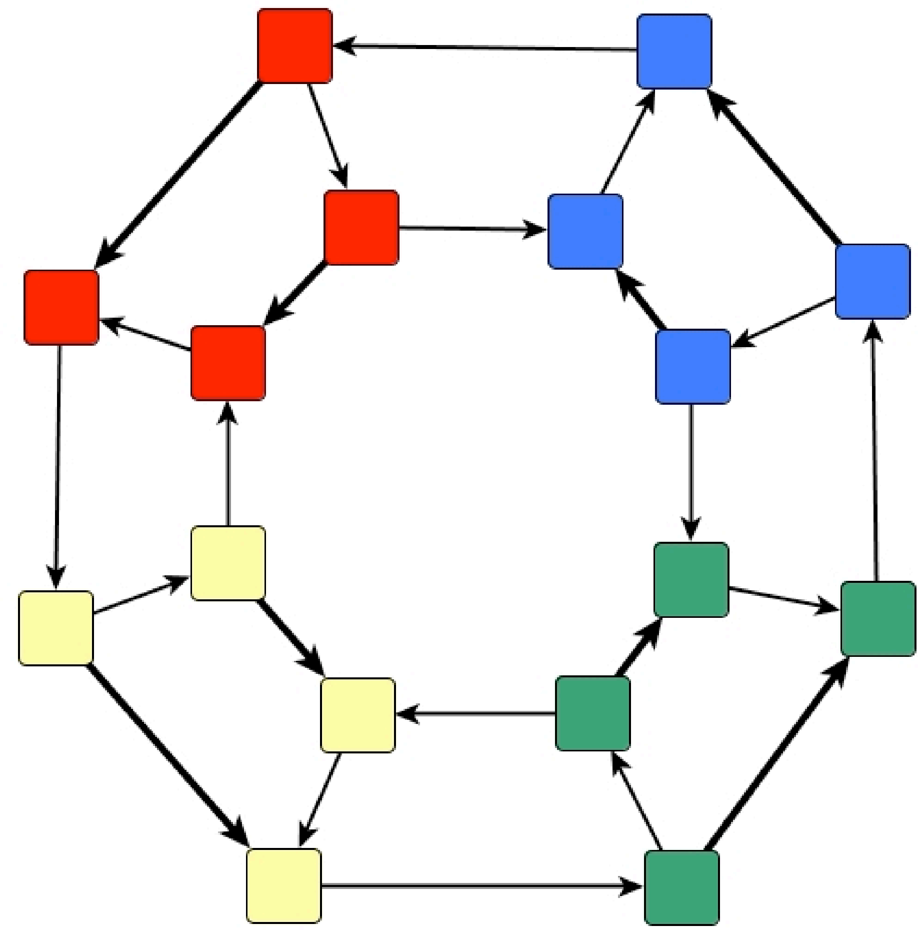
$$p_{i;n+1} = \sum_j \frac{A_{ij}}{k_j^{out}} p_{j;n} \quad \xrightarrow{\text{equilibrium}} \quad p_i^* = \pi_i$$

$$R(1) = \sum_{i,j} \left[\frac{A_{ij}}{k_j^{out}} \pi_j - \pi_i \pi_j \right] \delta(c_i, c_j) \neq Q$$

a. Modularity vs Stability



Flow-based modules



Combinatorial modules

M. Rosvall and C. T. Bergstrom, PNAS 105, 1118 –1123 (2008)

a. Modularity vs Stability

Let us consider a random walk on a directed network:

$$p_{i;n+1} = \sum_j \frac{A_{ij}}{k_j^{out}} p_{j;n} \quad \xrightarrow{\text{equilibrium}} \quad p_i^* = \pi_i$$

$$R(1) = \sum_{i,j} \left[\frac{A_{ij}}{k_j^{out}} \pi_j - \pi_i \pi_j \right] \delta(c_i, c_j) \neq Q$$

$R(1) \neq Q(A)$ but $R(1) = Q(Y)$

$$Y = \frac{X + X^T}{2} \qquad X_{ij} = \frac{A_{ij}}{k_j^{out}} \pi_j$$

b. Stability: time as a resolution parameter

Let us consider a continuous-time random walk with Poisson waiting times

$$\dot{p}_i = \sum_j \frac{A_{ij}}{k_j} p_j - p_i \quad \longrightarrow \quad p_i^* = k_i/2m$$

equilibrium

$$R(t) = \sum_{i,j} \left[\left(e^{t(B-I)} \right)_{ij} \frac{k_j}{2m} - \frac{k_i k_j}{(2m)^2} \right] \delta(c_i, c_j)$$

$$B_{ij} = A_{ij}/k_j$$

Probability that a walker is in the same community initially and at time t

Same probability for independent walkers

b. Stability: time as a resolution parameter

What are the optimal partitions of R_t?

t=0 $$R(0) = 1 - \sum_{i,j} \frac{k_i k_j}{(2m)^2} \delta(c_i, c_j)$$ → Communities=single nodes

t small $$R(t) \approx (1-t)R(0) + tQ_C \equiv Q(t)$$

favours single nodes — modularity

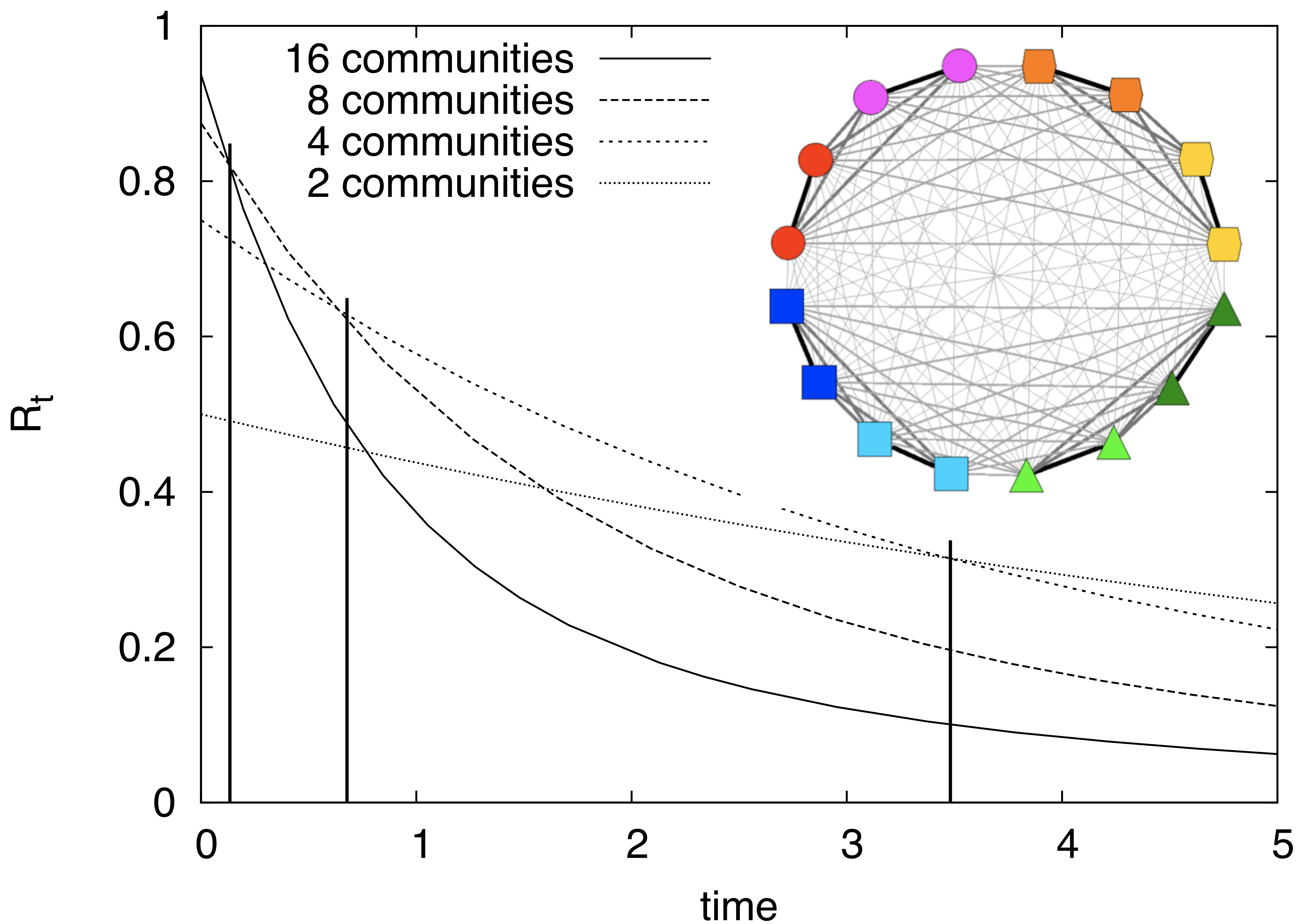
!! Q_t equivalent to the Hamiltonian formulation of Reichardt and Bornholdt (t=1/gamma)

....

When t goes to infinity, the optimal partition is made of 2 communities (by spectral decomposition)

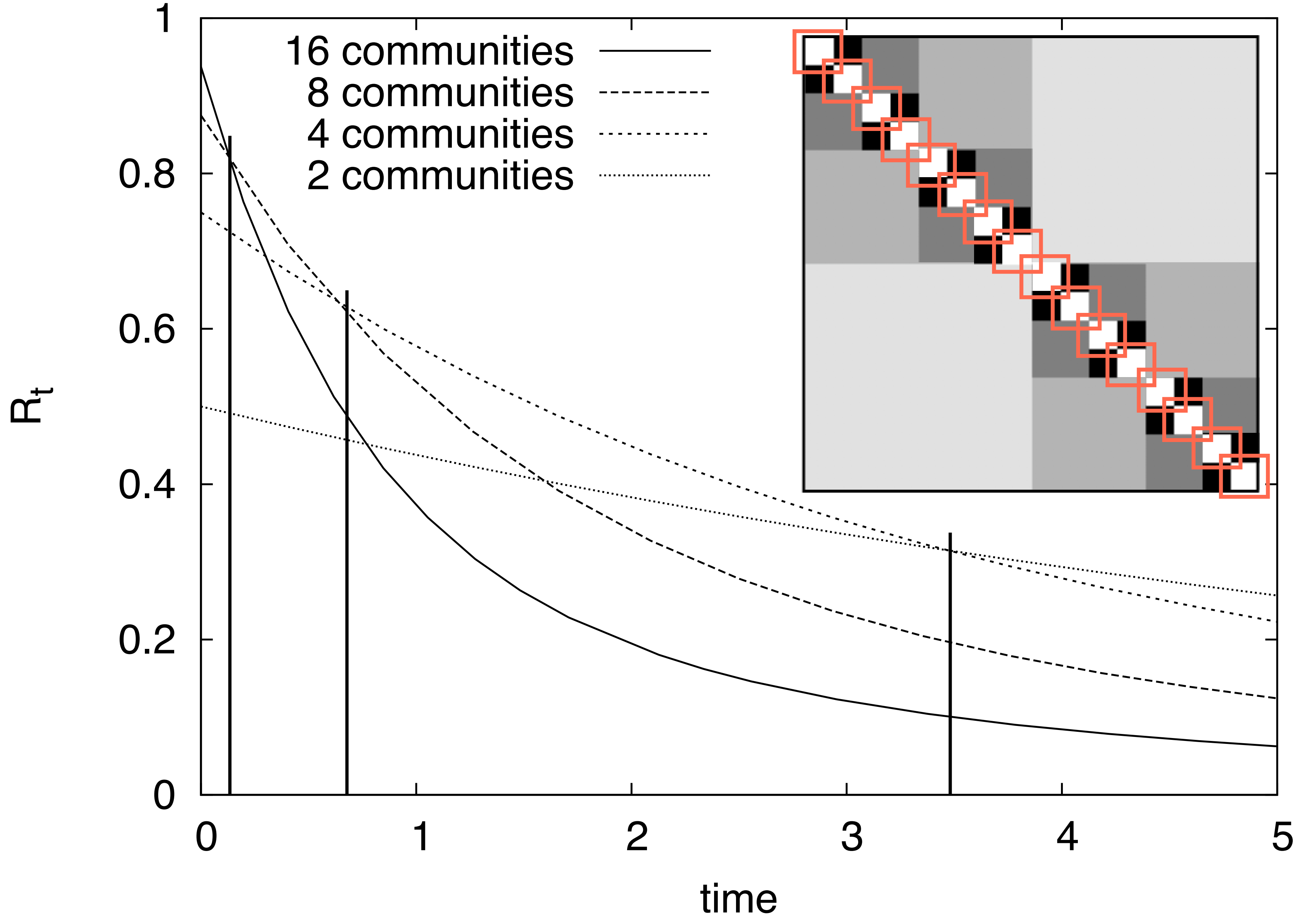
b. Stability: time as a resolution parameter

Time is a “resolution parameter”: larger and larger communities when time is increased



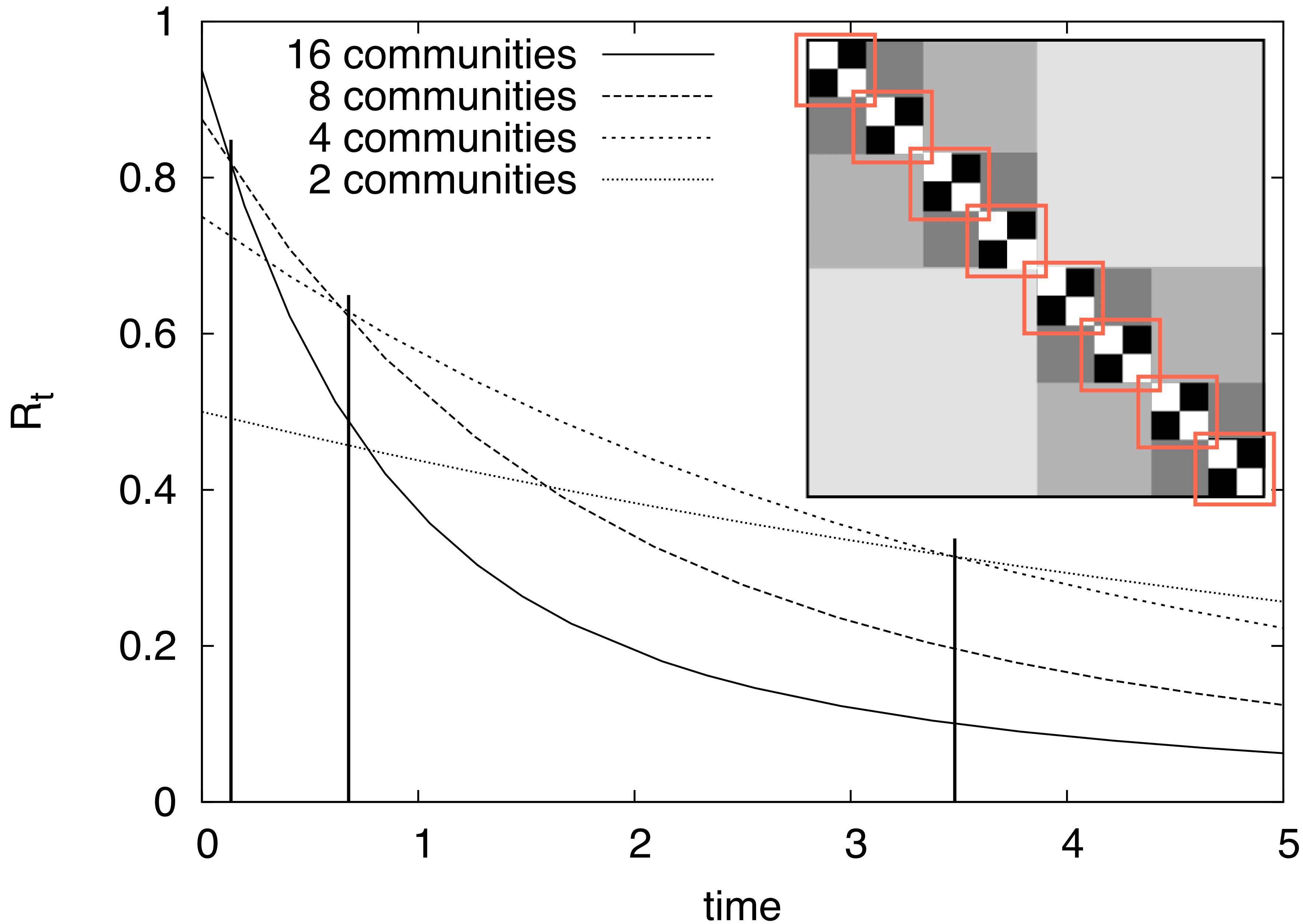
b. Stability: time as a resolution parameter

Time is a “resolution parameter”: larger and larger communities when time is increased



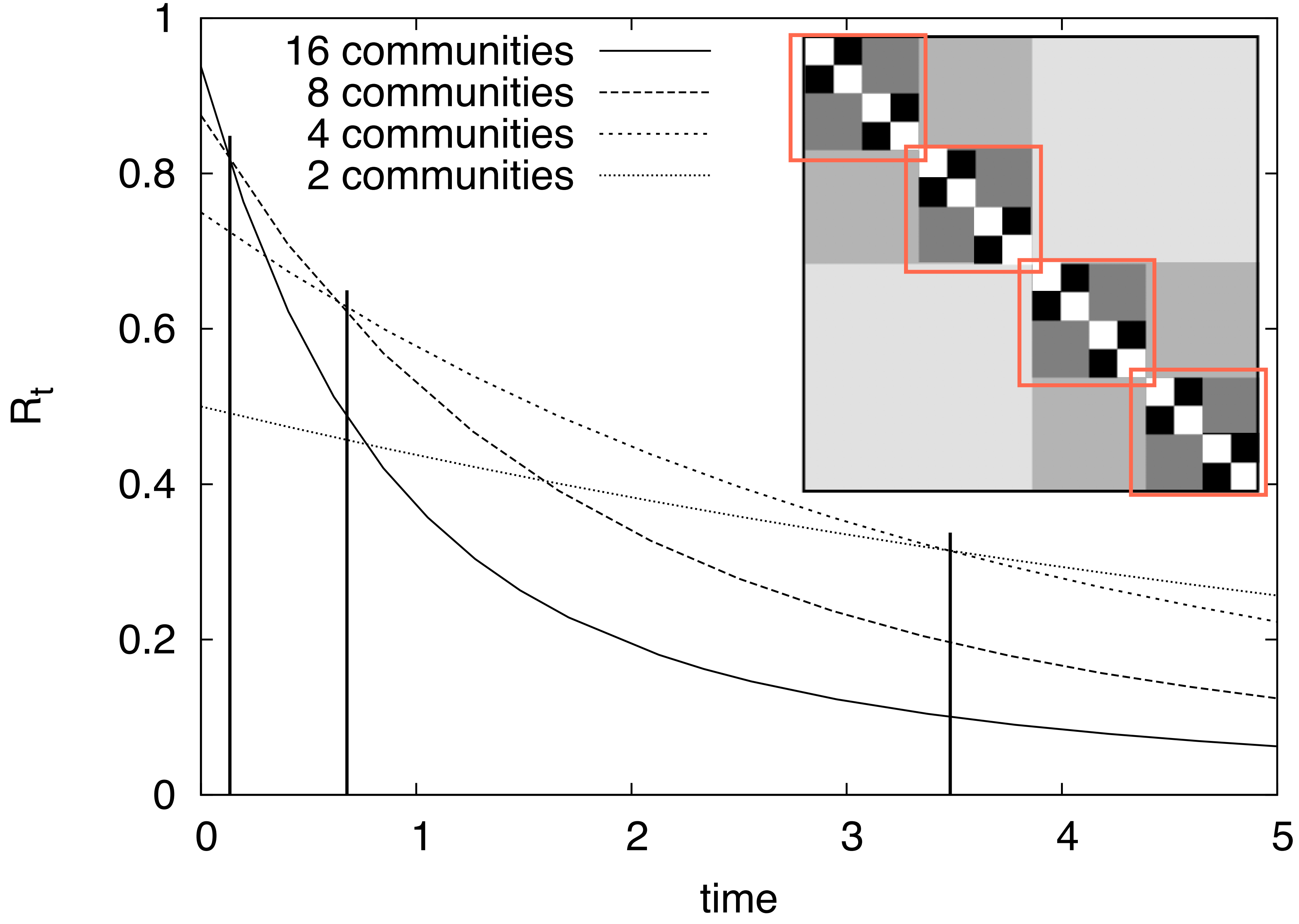
b. Stability: time as a resolution parameter

Time is a “resolution parameter”: larger and larger communities when time is increased



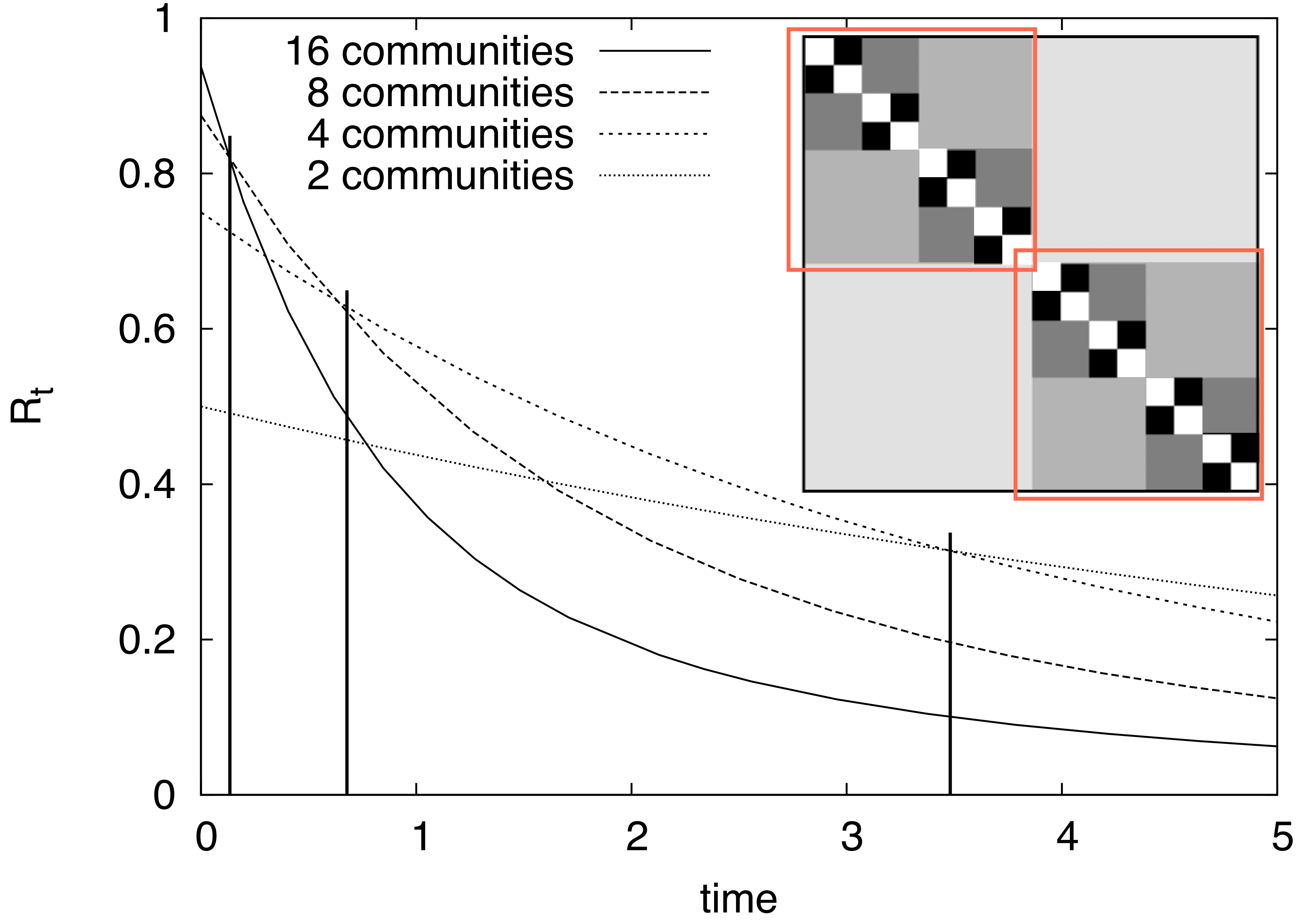
b. Stability: time as a resolution parameter

Time is a “resolution parameter”: larger and larger communities when time is increased



b. Stability: time as a resolution parameter

Time is a “resolution parameter”: larger and larger communities when time is increased



c. Other processes

Modularity is therefore based on the probability to be in the same module over two subsequent time steps, which suggests to generalize the definition over longer time intervals, but also to look at different dynamical processes.

$$\dot{p}_i = \sum_j \frac{A_{ij}}{\langle k \rangle} p_j - \frac{k_i}{\langle k \rangle} p_i$$

→ the probability to leave a node is proportional to the degree of the node

$$p_i^* = 1/N \equiv \langle k \rangle / 2m$$

which is different from:

$$\dot{p}_i = \sum_j \frac{A_{ij}}{k_j} p_j - p_i$$

→ the probability to leave a node does not depend on the degree of the node

$$p_i^* = k_i / 2m$$

$$R'(t) = \sum_{i,j} \left[\left(e^{t/\langle k \rangle (A-K)} \right)_{ij} \frac{1}{N} - \frac{1}{N^2} \right] \delta(c_i, c_j)$$

In practice: Optimisation

The stability *R(t)* of the partition of a graph with adjacency matrix *A* is equivalent to the modularity *Q* of a time-dependent graph with adjacency matrix *X(t)*

$$X_{ij}(t) = \left(e^{t(B-I)}\right)_{ij} k_j \qquad X_{ij}(t) = X_{ji}(t)$$

which is the flux of probability between 2 nodes at equilibrium and whose generalised degree is

$$\sum_j X_{ij}(t) = k_i$$

$$R(t) = \sum_{i,j} X_{ij}(t)/2m - k_i k_j/(2m)^2 \;\; \delta(c_i, c_j) = Q(X(t))$$

For very large networks: $R(t) \approx (1-t)R(0) + tQ_C \equiv Q(t)$

In practice: Selection of the most relevant scales

The optimization of *R(t)* over a period of time leads to a sequence of partitions that are optimal at different time scales.
How to select the most relevant scales of description?

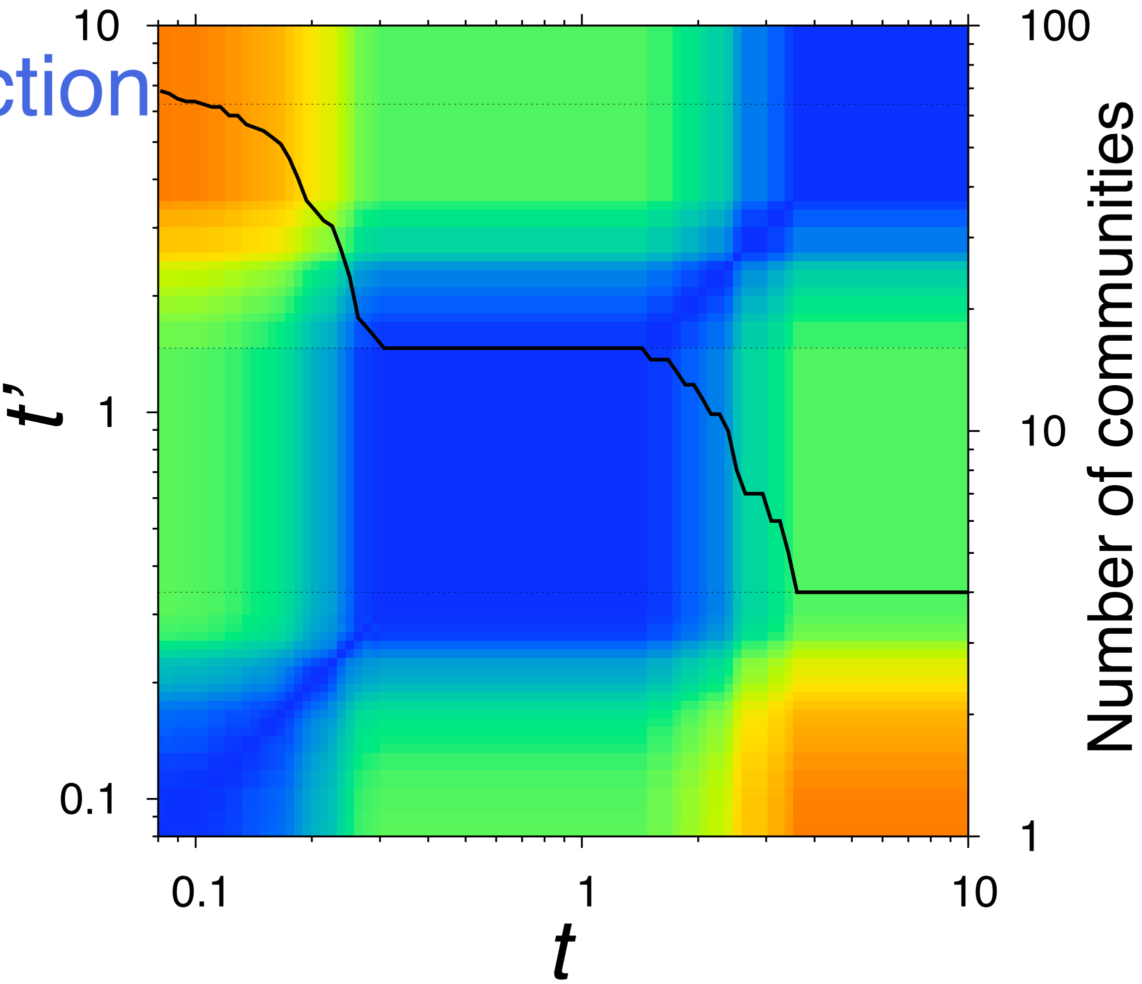
The significance of a particular scale is usually associated to a certain notion of the robustness of the optimal partition. Here, robustness indicates that a small modification of the optimization algorithm, of the network, or of the quality function does not alter this partition.

We look for regions of time where the optimal partitions are very similar. The similarity between two partition is measured by the *normalised variation of information*.

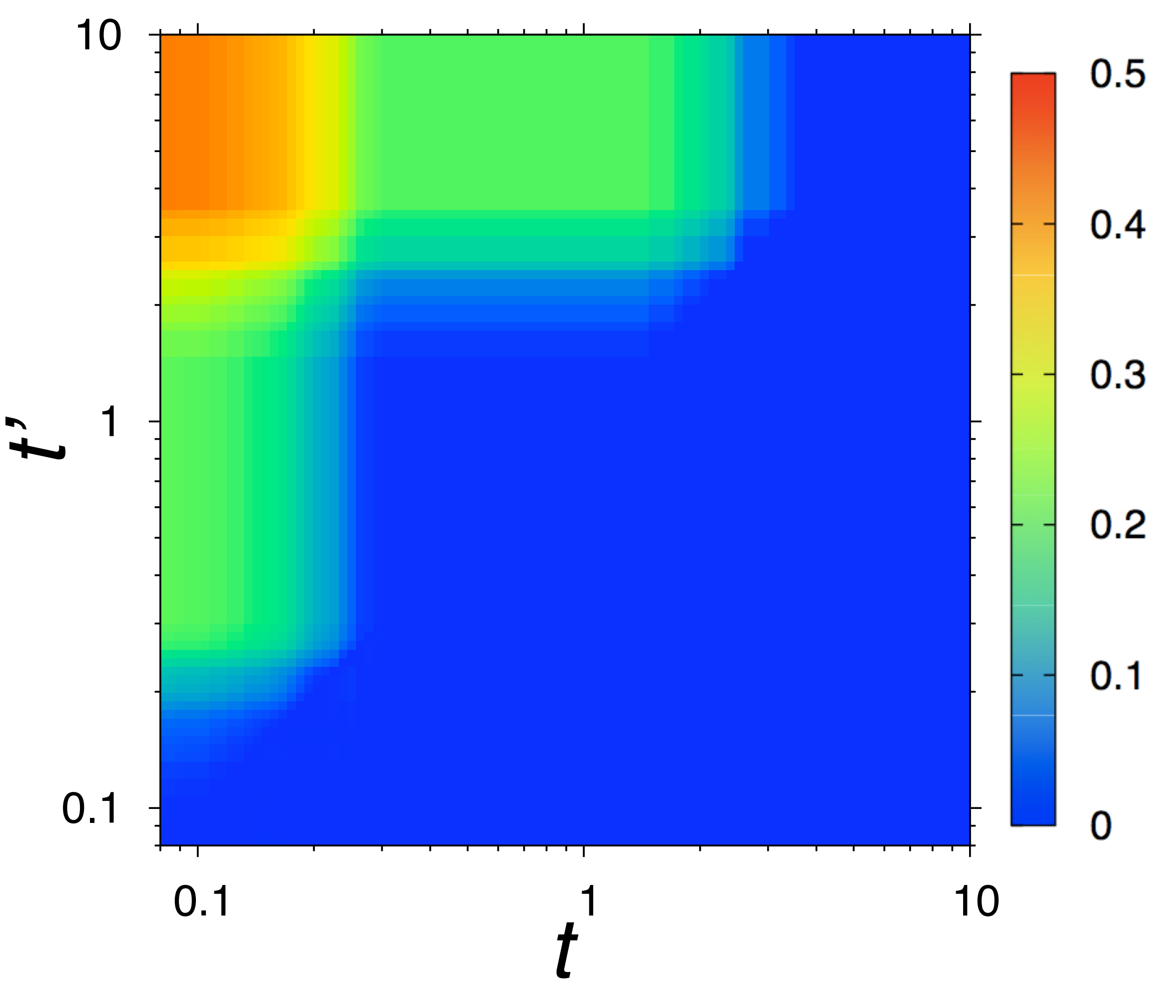
D.J. Fenn, M.A. Porter, M. McDonald, S. Williams, N.F. Johnson, N.S. Jones, *Chaos* 19, 033119 (2009).
M. Meilă, *J. Multivariate Anal.* 98, 873 (2007).

Modification of the quality function

Normalised variation of information vanishes only if partitions P_t and $P_{t'}$ are identical.



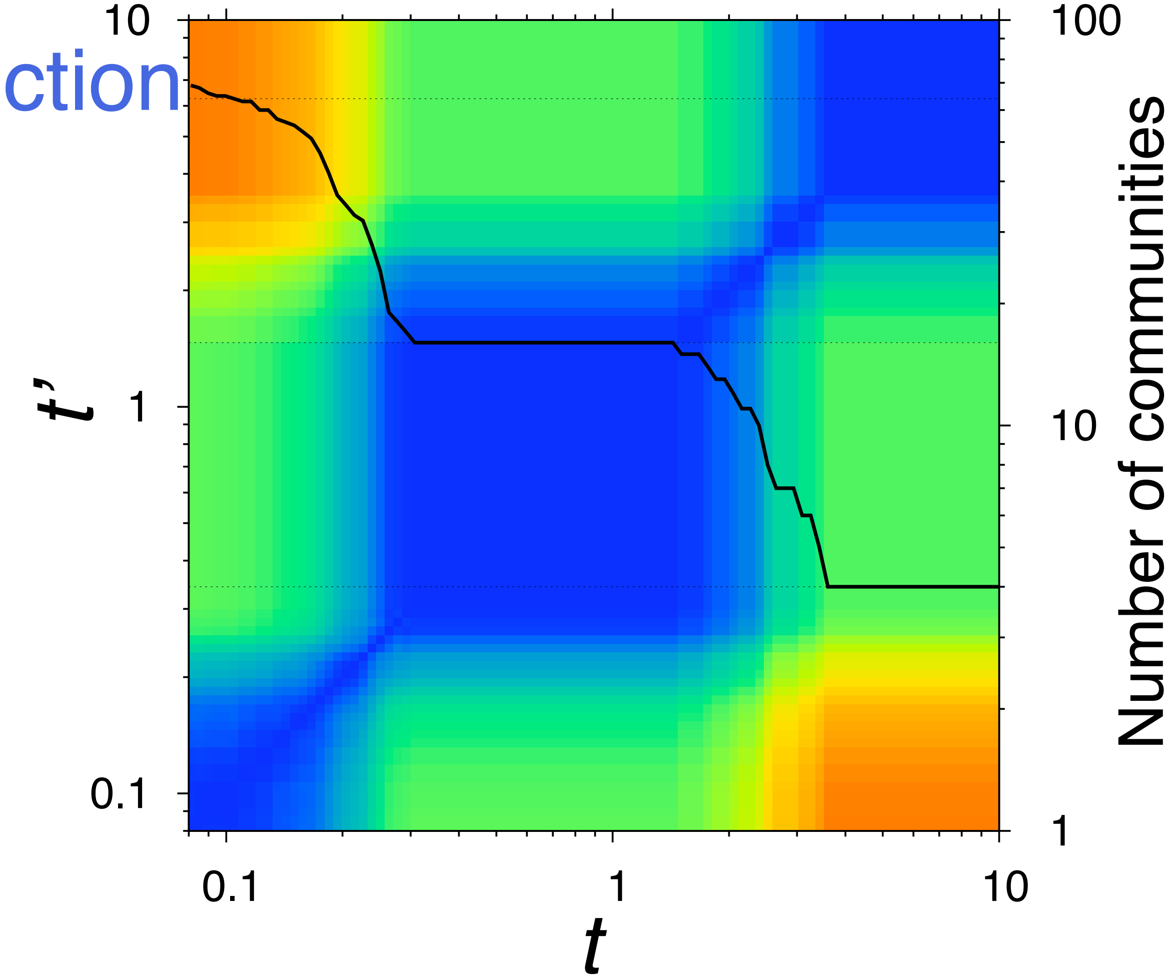
Normalized conditional entropy vanishes only if each community of P_t is the union of communities of $P_{t'}$



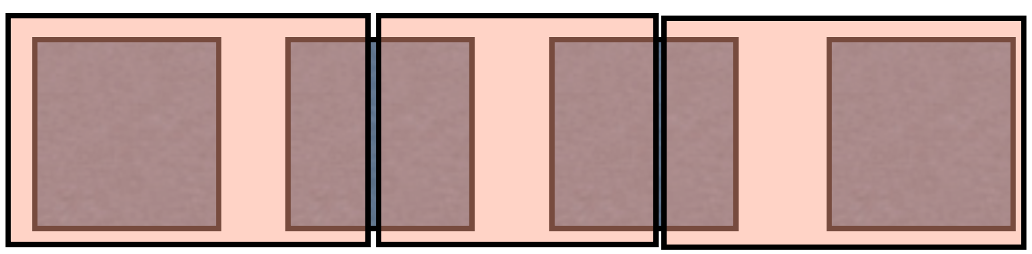
M. Meilă, *J. Multivariate Anal.* 98, 873 (2007).

Modification of the quality function

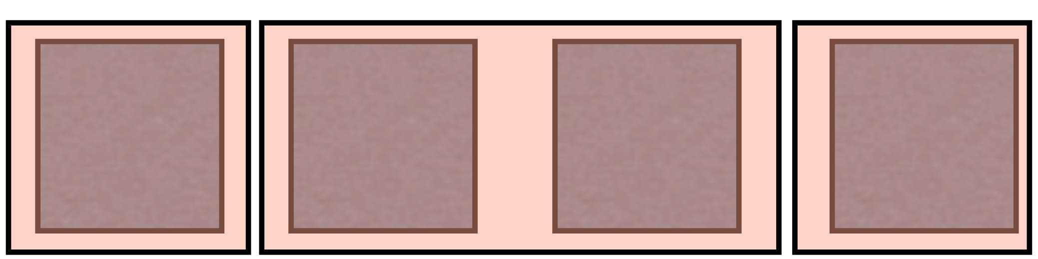
Normalised variation of information vanishes only if partitions P_t and $P_{t'}$ are identical.

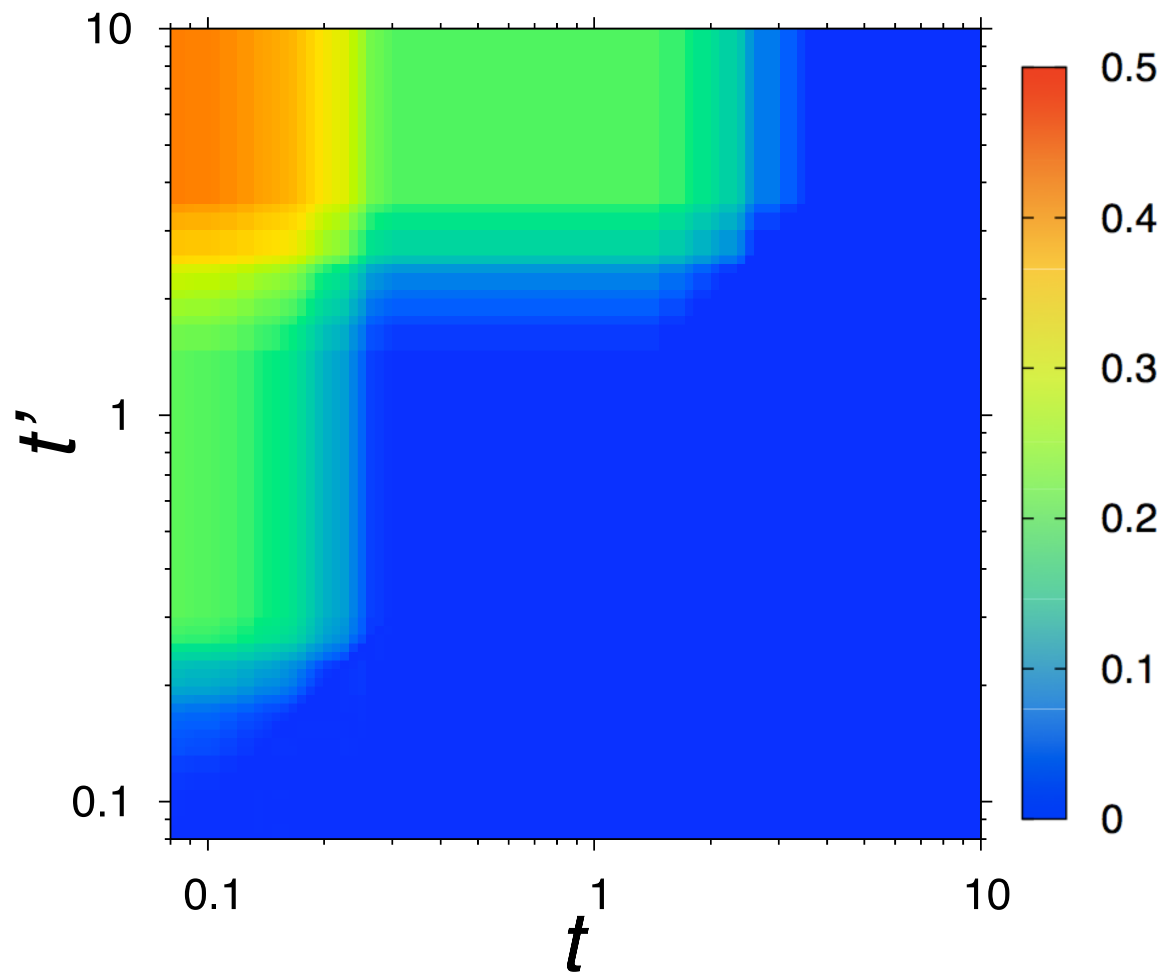


No hierarchy:



Hierarchy:





Conclusion

- Relation between dynamics and the hierarchical structure of networks
- Dynamical formulation for the quality of a partition
- Modularity and Stability are radically different in the case of directed networks
- Changing time allows to zoom in and out
- Different dynamics lead to different quality functions for the partition of a graph

Original Louvain method to optimise modularity available on *http://findcommunities.googlepages.com*

Generalized codes to optimise Q_t available on http://*www.lambiotte.be*

Thanks to J.-L. Guillaume (for providing his c++ code)

R. Lambiotte, J.-C. Delvenne and M. Barahona, *arXiv:0812:1770*

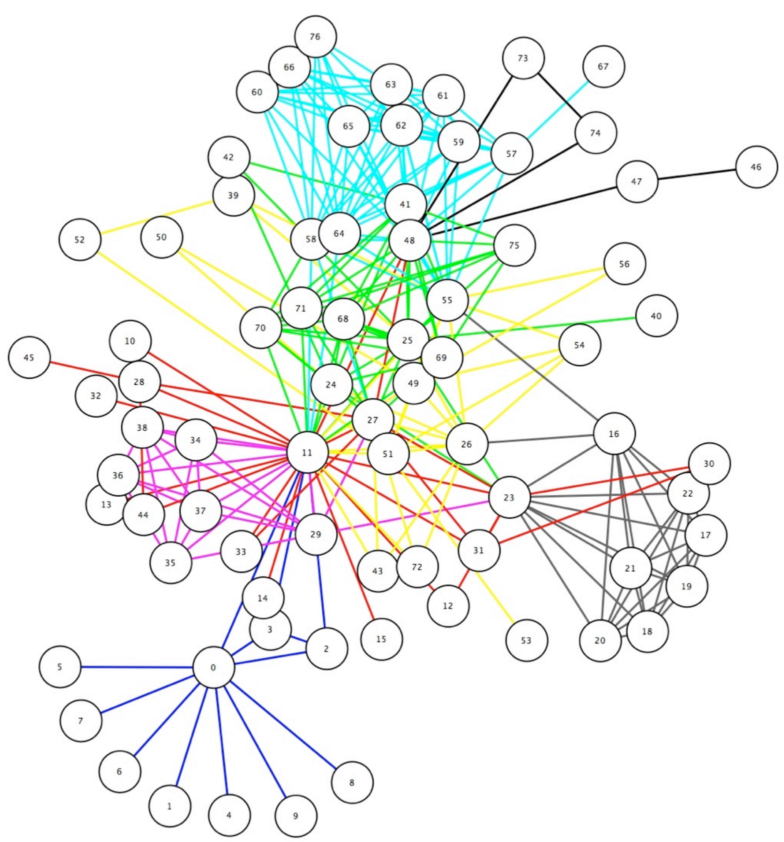
V.D. Blondel, J.-L. Guillaume, R. Lambiotte and E. Lefebvre, J. Stat. Mech., P10008 (2008).

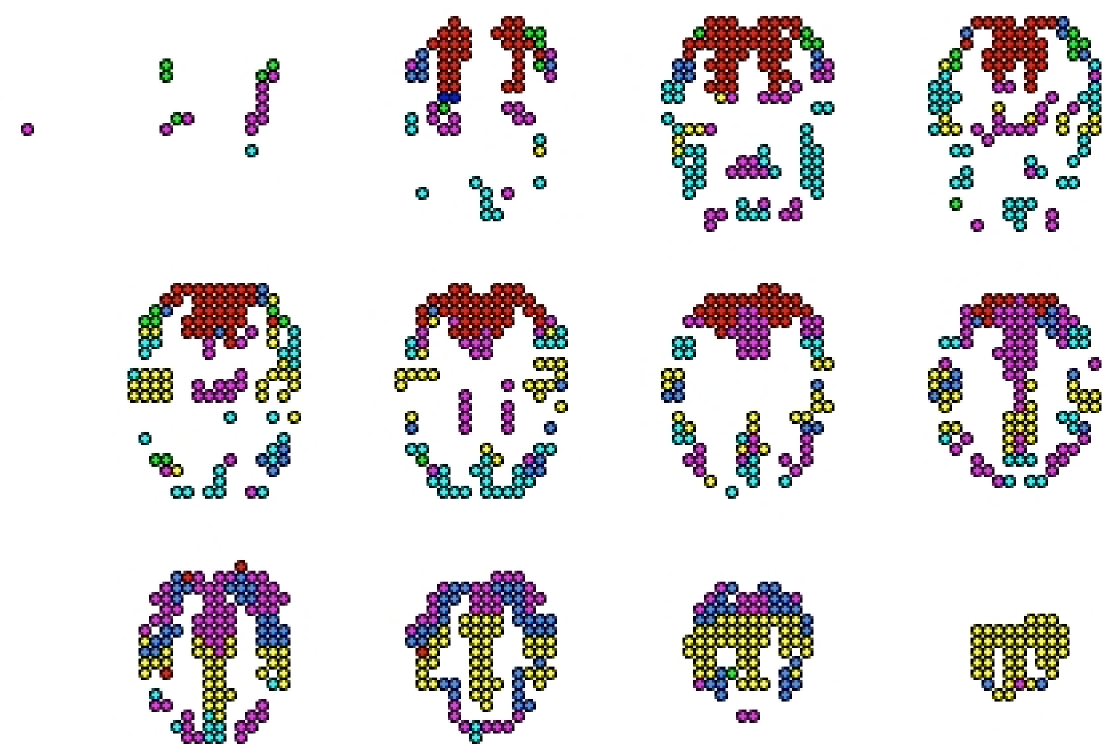
Thanks to:

S. Yaliraki, K. Christensen, P. Expert, T. Evans, H. Jensen (Imperial College)

D. Meunier (Cambridge)

V. Blondel (Louvain)





Line Graphs, Link Partitions and Overlapping Communities, T. Evans and R. Lambiotte, *Phys. Rev. E*, **80** (2009) 016105
Hierarchical modularity in human brain functional networks, D. Meunier, R. Lambiotte, A. Fornito, K.D. Ersche, E.T. Bullmore, submitted to *Frontiers in NeuroInformatics*