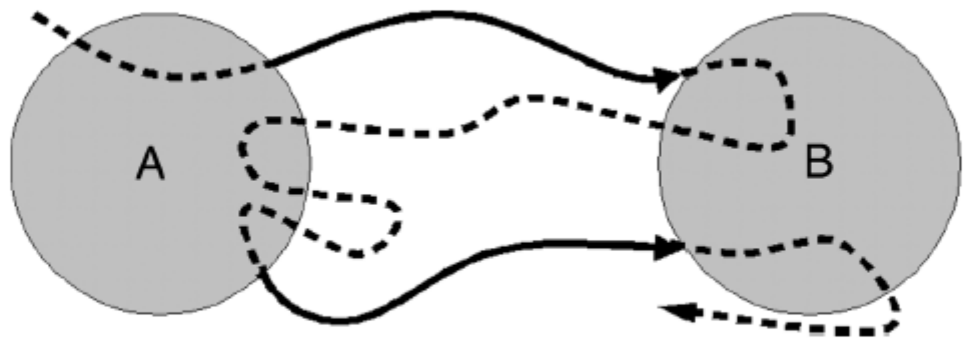
Transition Path Theory

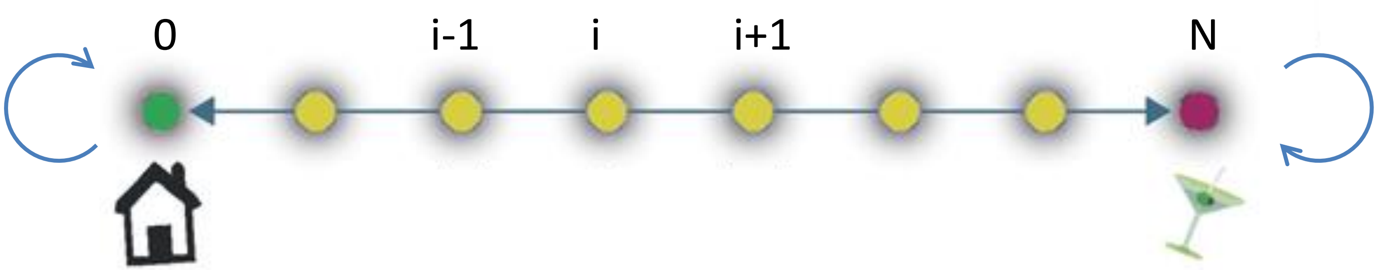
- TPT: method to study the ensemble of reactive trajectories.



- **reactive trajectory : came from A and goes next to B**
  - rate at which they occur
  - mechanisms (parallel pathways, traps, sequence of events, …)
  - committor: trajectory start starts in *i*, goes it next to A or to B?
    also known as “P_{fold}”
    transition states have $P_{\text{fold}} = 1/2$

The committor

- A drunk man walks one block left with P=1/2 and one block right with P=1/2.
- **Probability that the man starting in block *i* will reach home before the bar?**

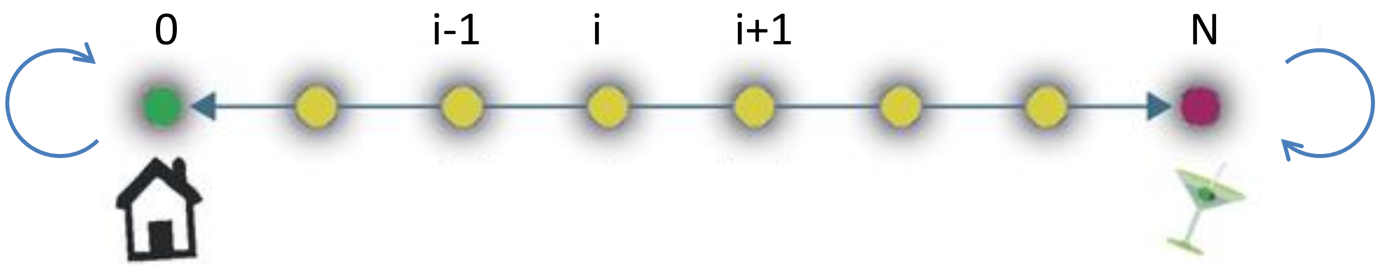


- $P(N) = 0$ N=bar. No chance to reach home if already at bar.
- $P(0) = 1$ 0=home. If at home, 100% chance to reach home.
- $P(i)$ for $\notin \{0, N\}$?

Doyle, Snell, *Random Walks and Electric Networks*, Carus (1984)
Image: Valleriani, *Nat. Scientific Reports* 5, 17986 (2015).

The committor q_i^+

$P(0) = 1,\ \ P(N) = 0,\ \ P(i)$ for $\notin \{0, N\}$?



Start with general statement from probability theory:

E event, F and G event s. t. only one of G or F will occur

$$P(E) \ = \ P(E|F)\,P(F) \ + \ P(E|G)\,P(G)$$

E = the man reaches home first

F = the first step is to the right

G = the first step is to the left

P(home)= P(home| went right) P(went right) + P(home | went left) P(went left)

P(home from *i*)= P(home from *i*+1) P(went right) + P(home from *i*-1) P(went left)

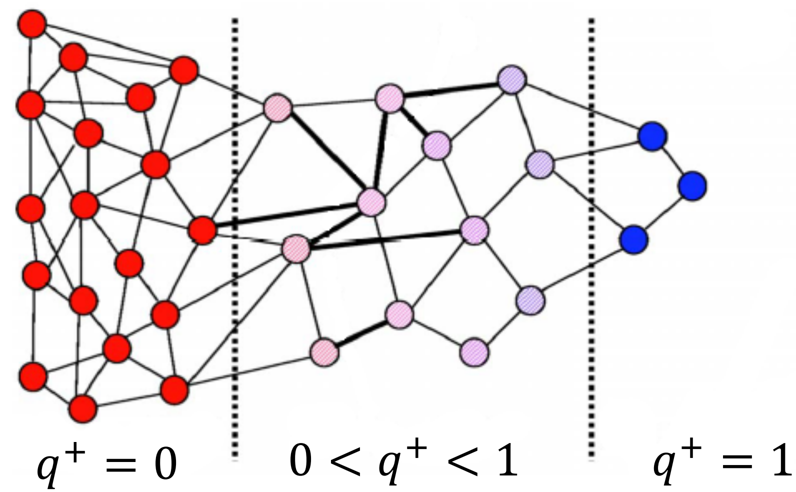
$$q_i^+ \ = q_{i+1}^+ p_{i,i+1} + q_{i-1}^+ p_{i,i-1}$$

The committor q_i^+

The same argument that we made for a 1-D random walk can be make for a general kinetic network (MSM).

The equations from the last slides are generalized to:

- $q_i^+ = 0$ for $i \in A$
- $q_i^+ = 1$ for $i \in B$
- $q_i^+ = \sum_{j \in I} p_{ij} q_j^+$ for $i \notin \{A, B\}$



There exists also a committor q_i^- that gives the probability of (immediately) **coming from** a set *A* without having visited *B* in between.

For reversible systems

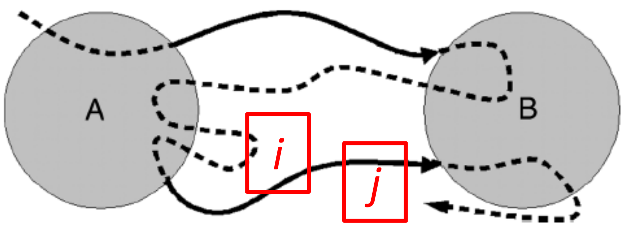
$$q_i^- = 1 - q_i^+$$

Reactive probability flux $\mathbf{F}^{AB}$

Reactive flux : average number of reactive trajectories per time unit making a transition from *i* to *j* on their way from *A* to *B*.

Misnomer, really should be "current", unit = 1/time unit

$$F_{ij}^{AB} := \begin{cases} q_i^- \pi_i P_{ij} q_j^+ & \text{if } i \neq j \\ 0 & \text{if } i = j \end{cases}$$



Properties (essentially the properties of electric current):

- Flux conservation within the intermediate states (Kirchhoff's law)

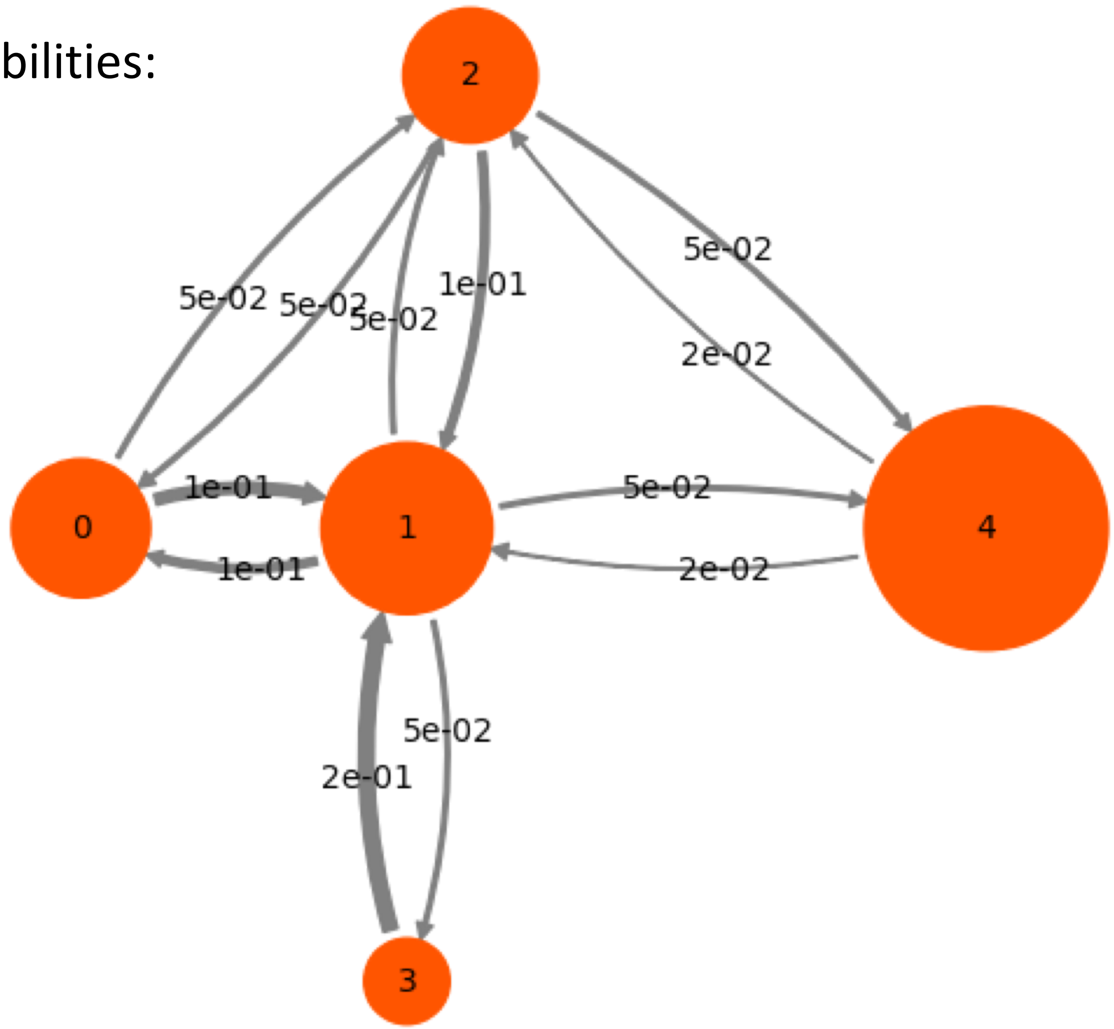
$$\sum_j (F_{ij}^{AB} - F_{ji}^{AB}) = 0 \text{ for all } i \notin \{\mathrm{A}, \mathrm{B}\}$$

- What goes into the network in A comes out at B:

$$\sum_{i \in A, j \notin A} F_{ij}^{AB} = \sum_{i \notin B, j \in B} F_{ij}^{AB}$$

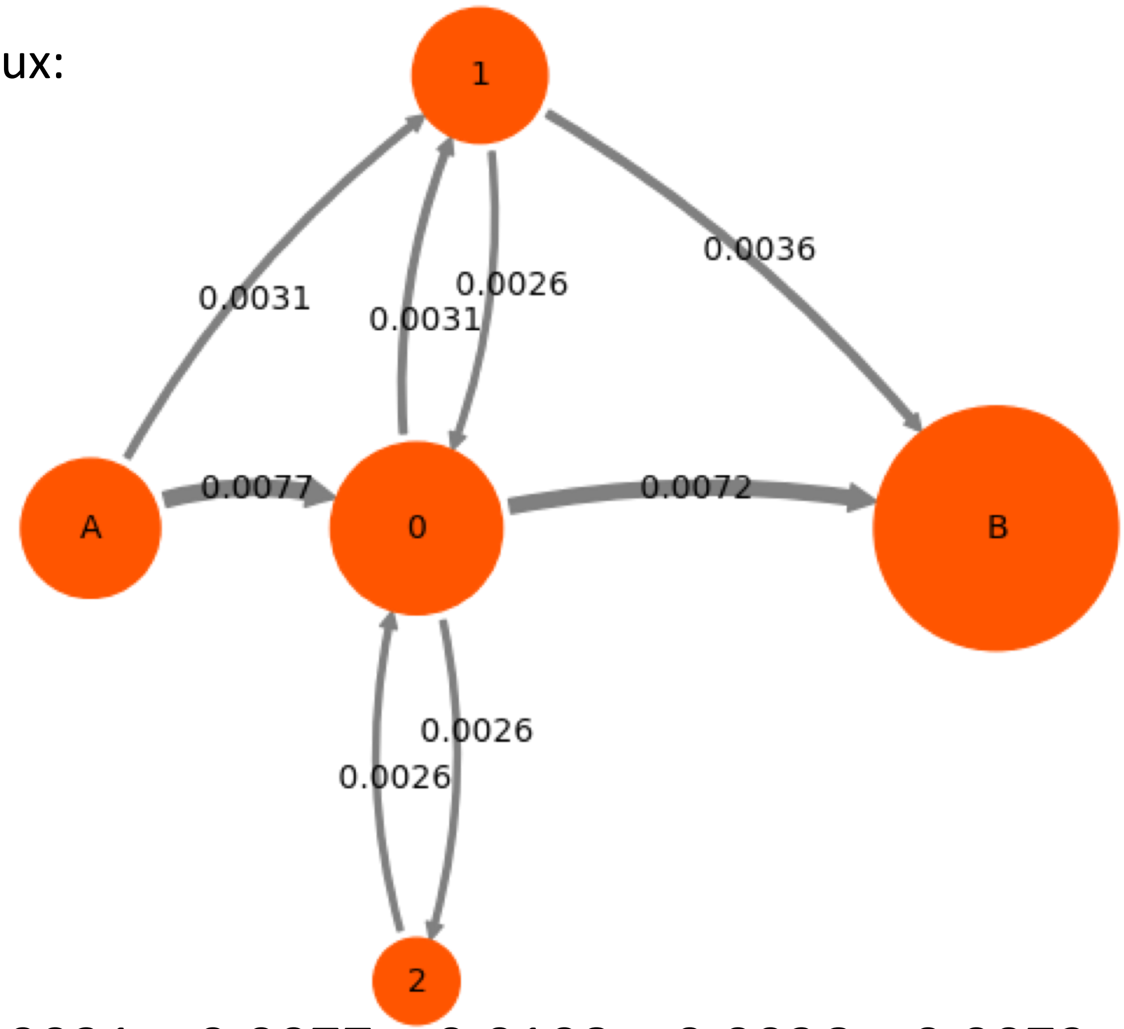
Example

transition probabilities:



Example

reactive gross flux:



0.0031 + 0.0077 = 0.0108 = 0.0036 + 0.0072

Gross flux vs. Net flux

On their way from *A* to *B*, trajectories might take forward and backward steps.

What if we were only interested in the productive flux, that is in steps that take us closer to *B*?

Define the **net flux** :

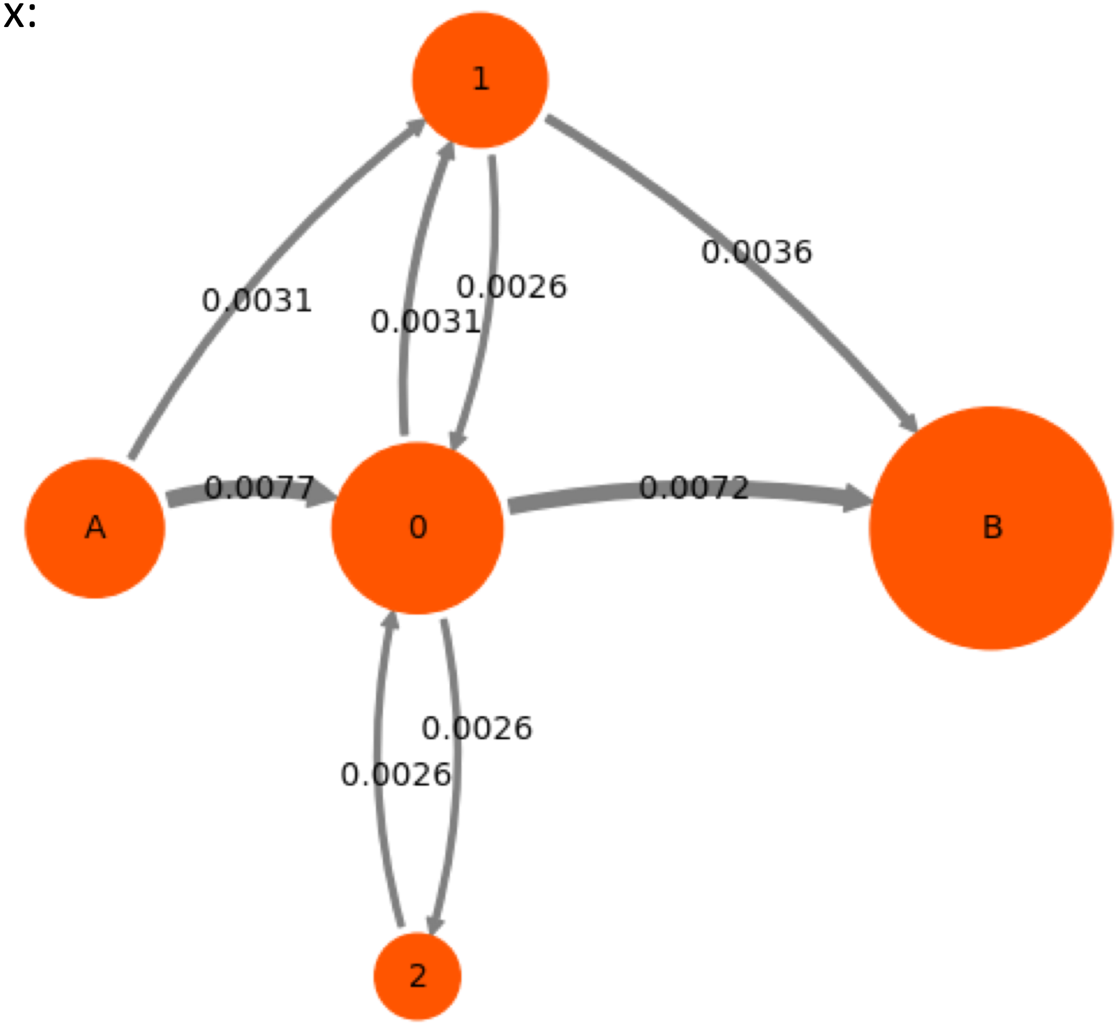
$$F_{ij}^{+} = \max\left(F_{ij}^{AB} - F_{ji}^{AB}, 0\right)$$

for reversible systems one can show that

$$F_{ij}^{+} = \max\left(\pi_i T_{ij}\left(q_j^{+} - q_i^{+}\right), 0\right)$$

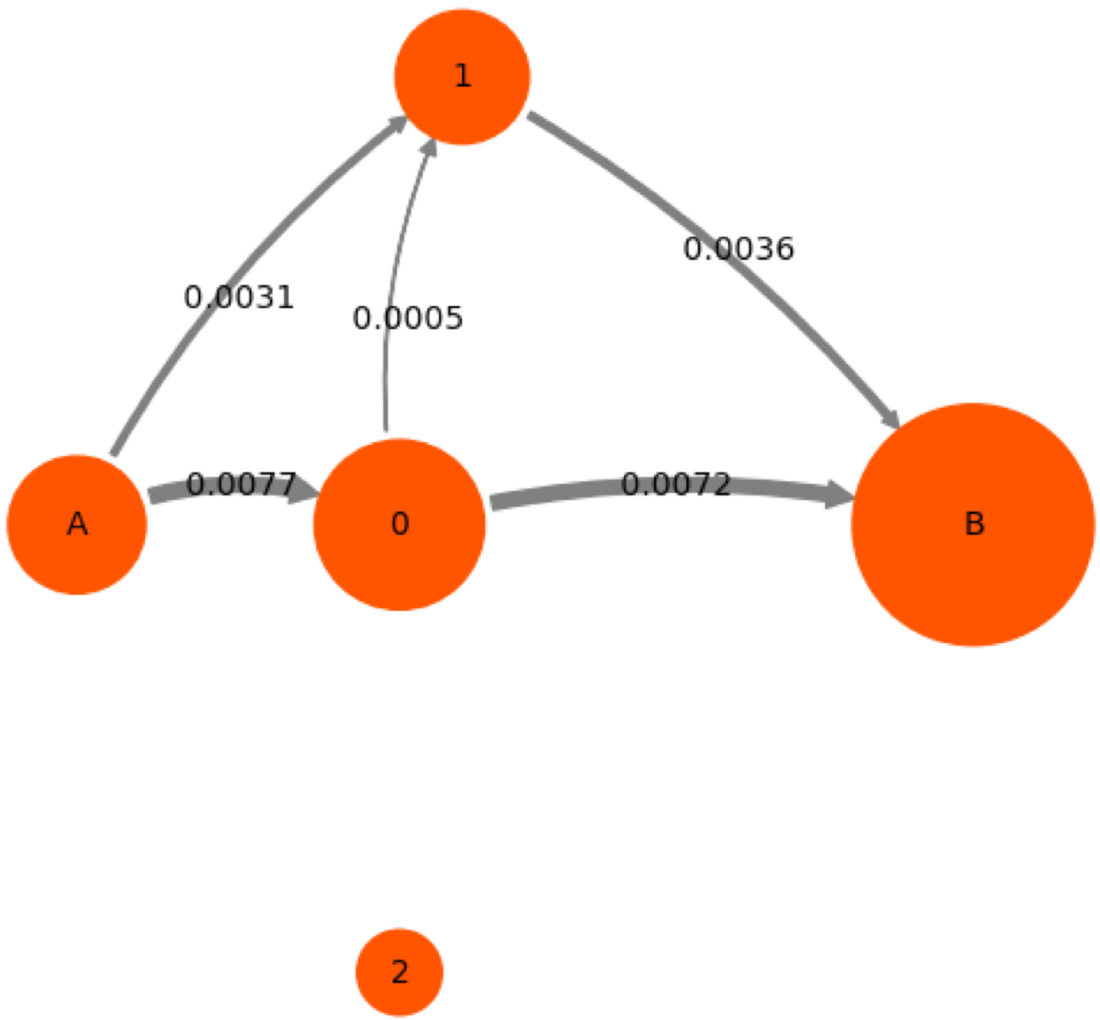
Example

reactive **gross** flux:



Example

reactive **net** flux:



Pathway decomposition

- A pathway is a sequence of states that starts with a state in *A* and ends with a state in *B*

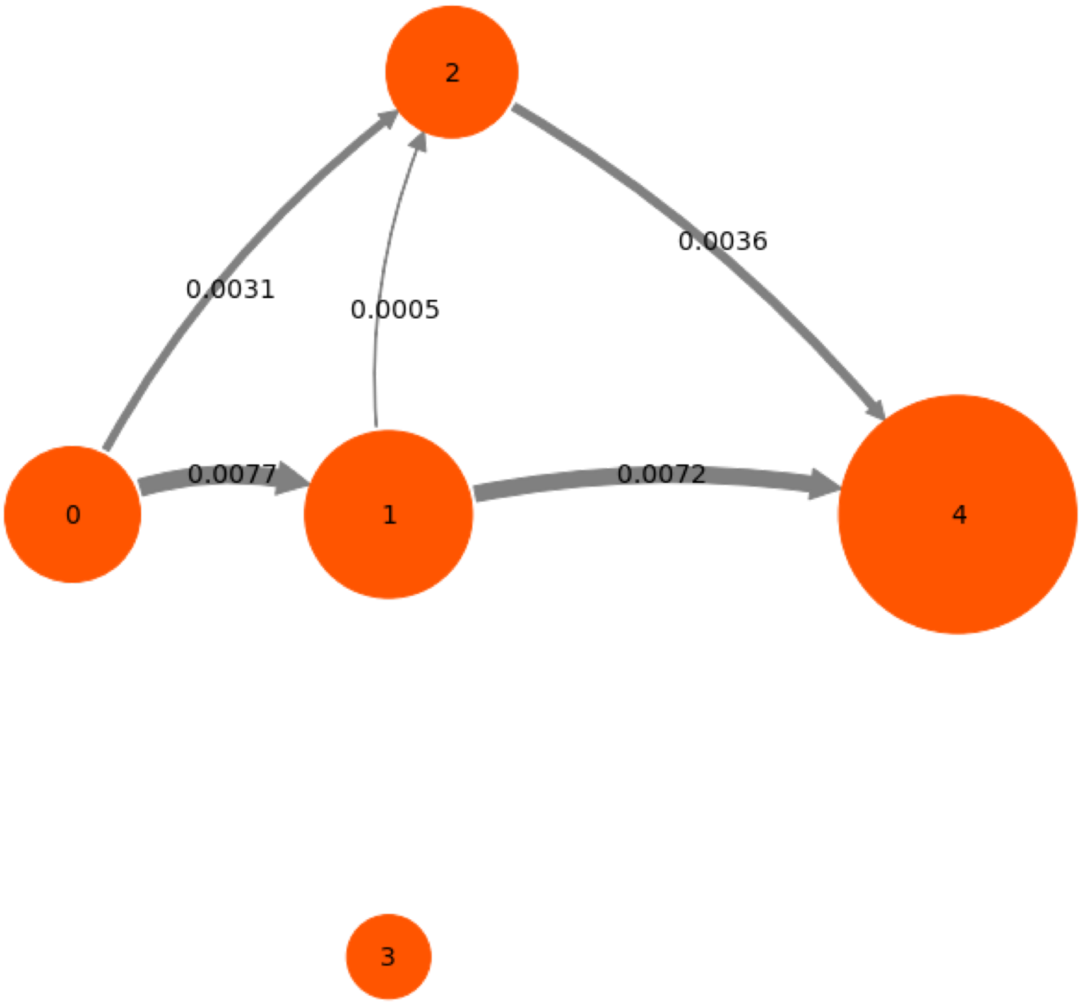
$$P = (i_1, i_2, \dots, i_k) \text{ such that } i_1 \in A, i_k \in B$$

- You can think of a pathway a special network without meshes (loops). There is also a flux matrix $\mathbf{F}_p$ for the pathway.
- We aim to decompose the network into a number of pathways.
- The original network should be the "sum" of all the pathways.

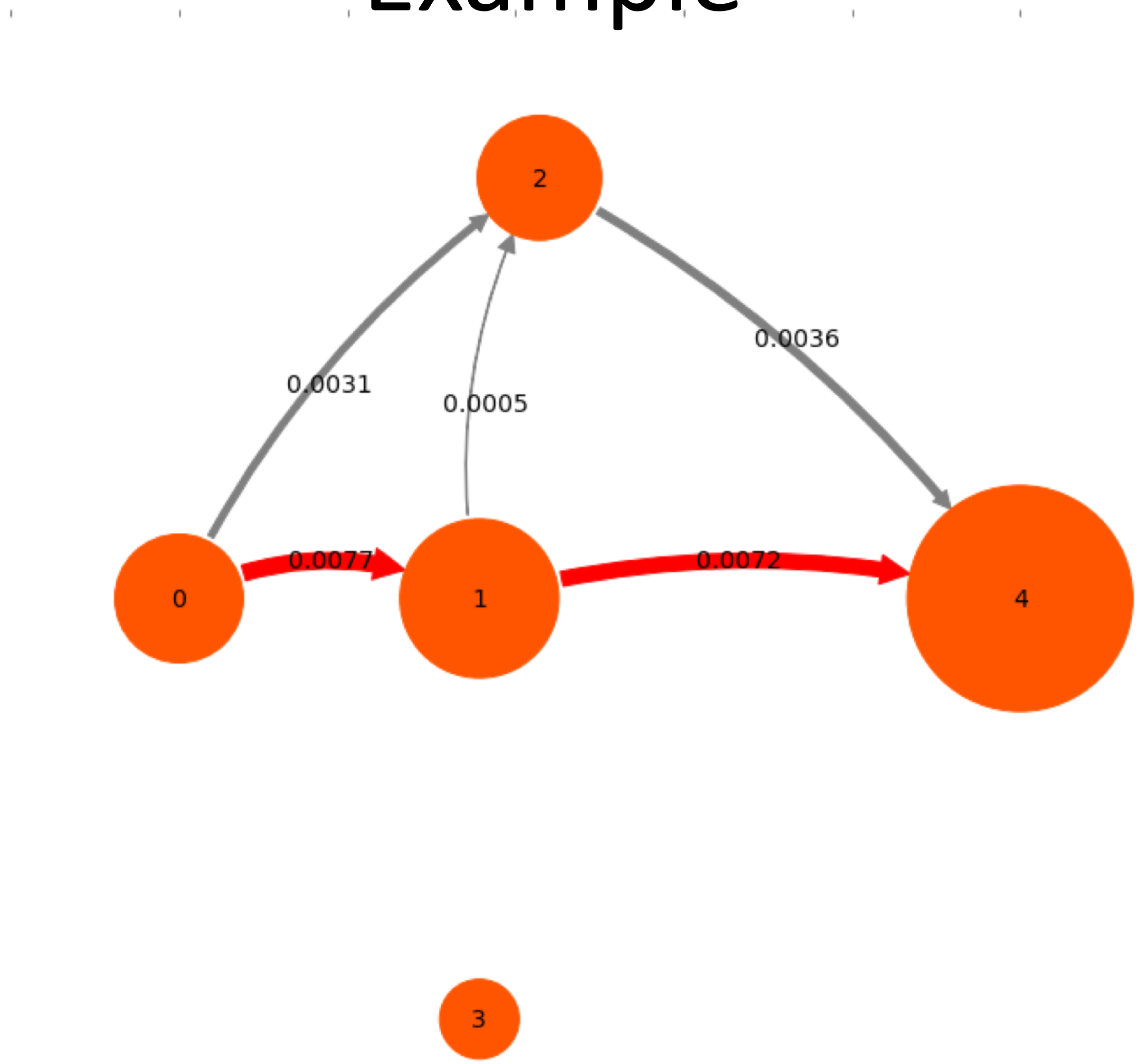
$$\mathbf{F}^+ = \sum_k \mathbf{F}_{\text{path } k}$$

- The pathway decomposition of the network will use an algorithm that "subtracts" pathways from the original network.

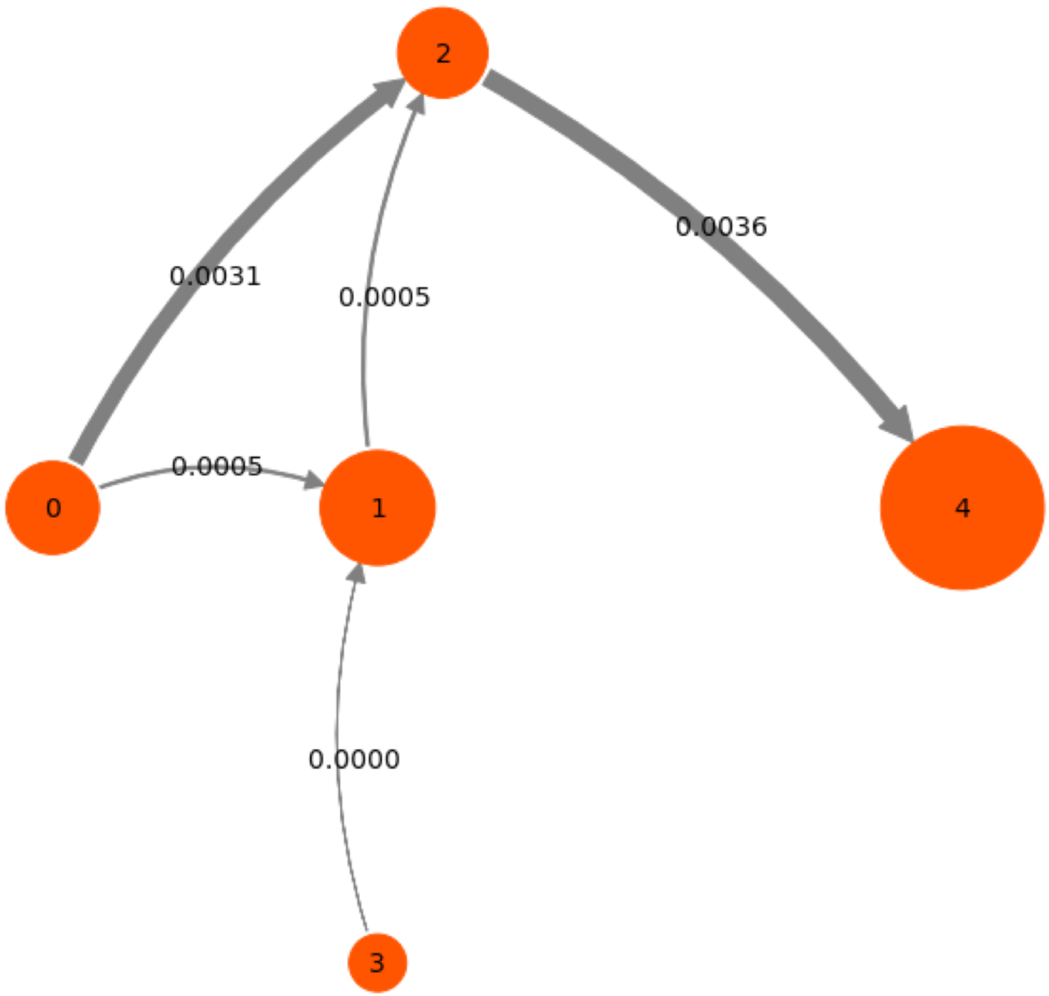
Example



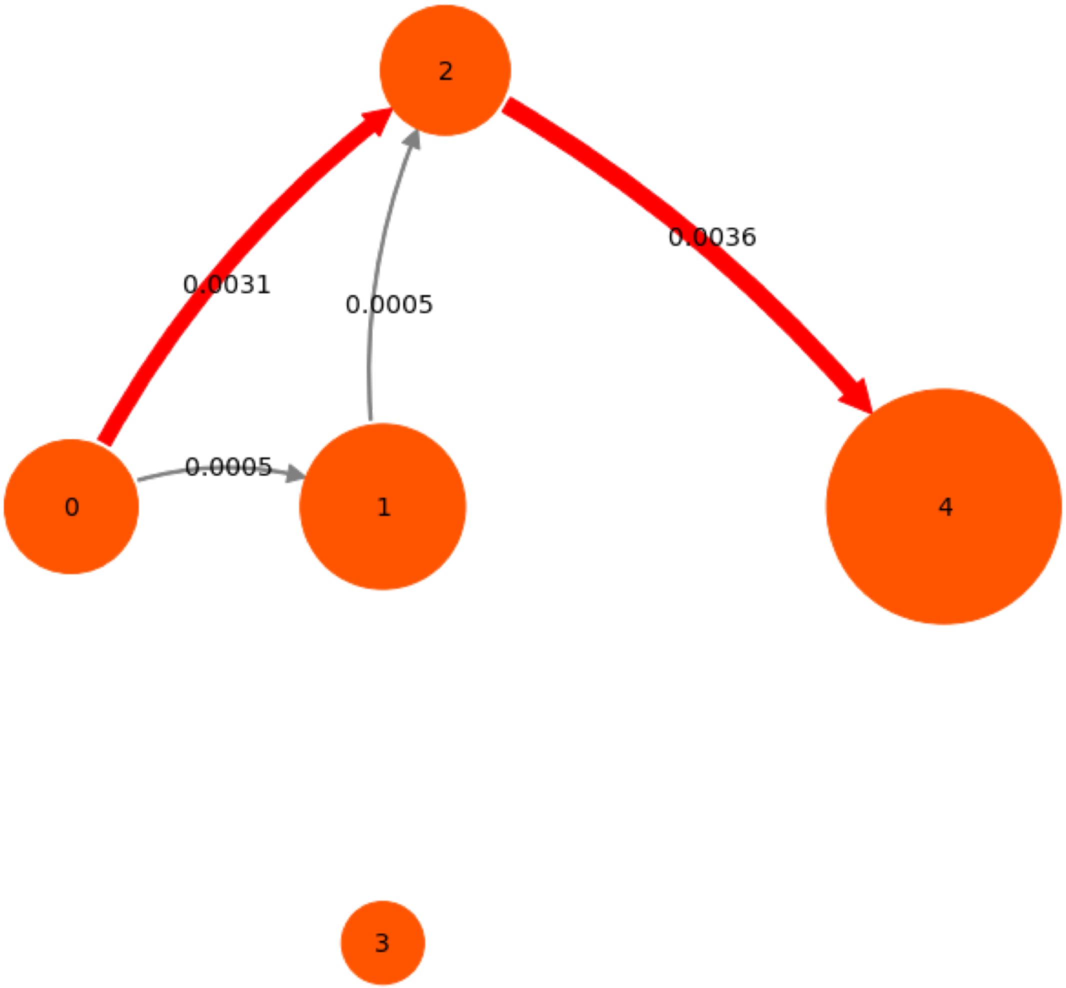
Example



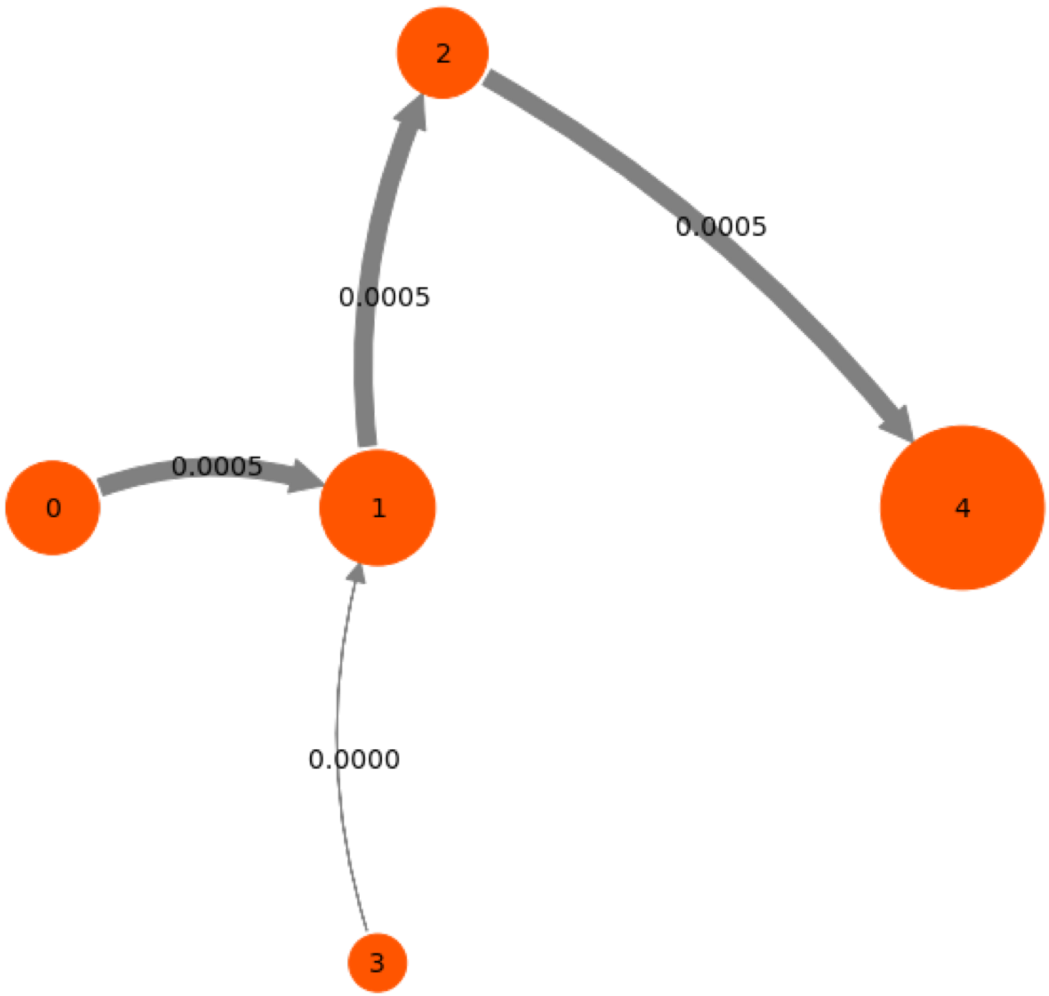
Example



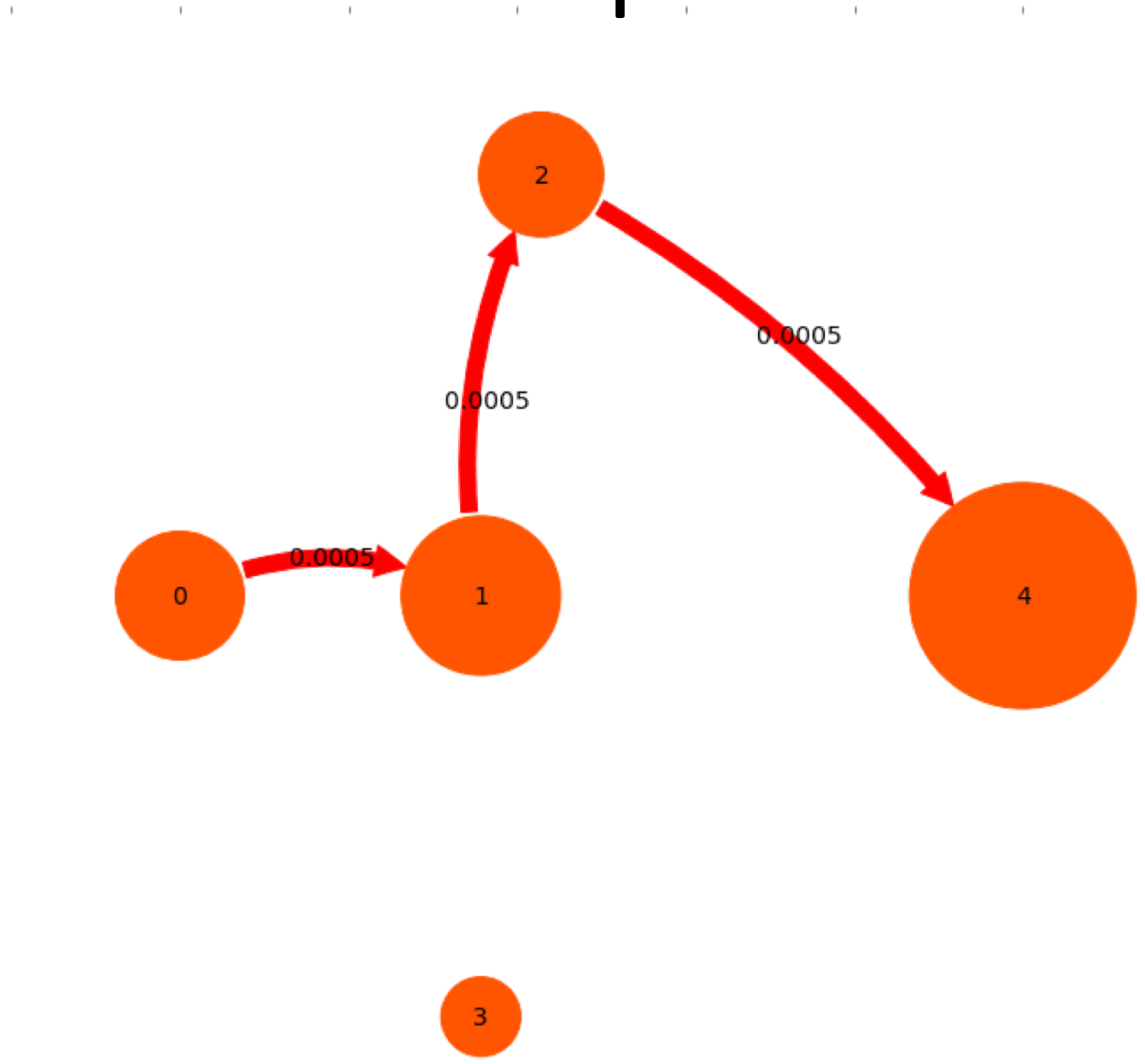
Example



Example



Example



Further reading

- F. Noé, C. Schütte, E. Vanden-Eijnden, L. Reich, T. Weikl: “Constructing the Full Ensemble of Folding Pathways from Short Off-Equilibrium Simulations”.
- P. Metzner, C. Schütte, and E. Vanden-Eijnden: “Transition Path Theory for Markov Jump Processes”. Mult. Mod. Sim. (2007)

- The capacity (or flux) of pathway is its weakest link

$$f(P) = \min\{f_{i_l i_{l+1}} \mid l = 1 \dots k-1\}$$

- Pathway decomposition: chose a pathway P_1 and remove its capacity from the flux along all edges of P_1. Repeat until no flux remains.

outline

- Committor
- Reactive flux
- Gross flux vs. Net flux
- Pathway decomposition
- A word of caution

Gross flux vs. Net flux

On their way from *A* to *B*, trajectories might take forward and backward steps.

What if we were only interested in the productive flux, that is in steps that take us closer to *B*?

Define the **net flux** : $f_{ij}^{+} = \max\left(f_{ij}^{AB} - f_{ji}^{AB}, 0\right)$

for the case of detailed balance one can show that

$$f_{ij}^{+} = \max\left(\pi_i T_{ij}\left(q_j^{+} - q_i^{+}\right), 0\right)$$

(For the general case, without detailed balance f_{ij}^{+} might still contain unproductive cycles/detours.)

Transition Path Theory

Computer tutorial in Markov modeling (PyEMMA)

20.2.2018

Fabian Paul

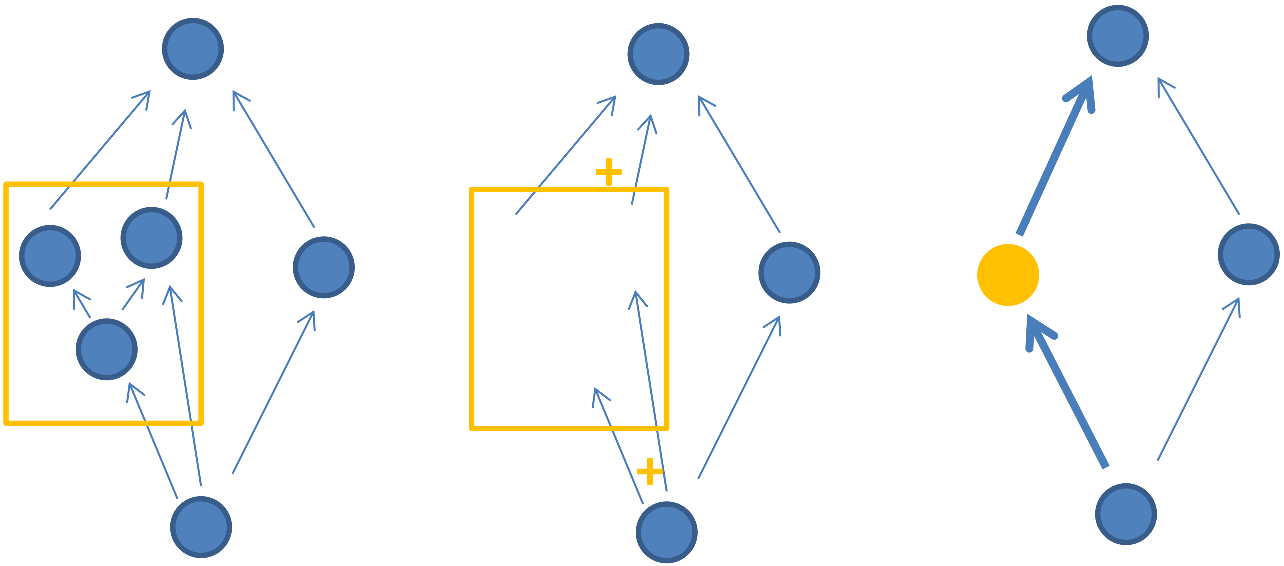
Coarse-graining of fluxes

Markov model construction is best done with many states.

For better interpretation you may be interested in a coarse-grained version of the state space (e.g. PCCA sets / metastable sets).

Reactive currents are a quantity that can be coarse-grained without systematic error. (In contrast a to coarse-grained transition matrix).

$$F_{IJ}^{AB} = \sum_{i \in I, j \in J} f_{ij}^{AB} \text{ where } I \cap (A \cup B) = \emptyset \text{ and } J \cap (A \cup B) = \emptyset$$



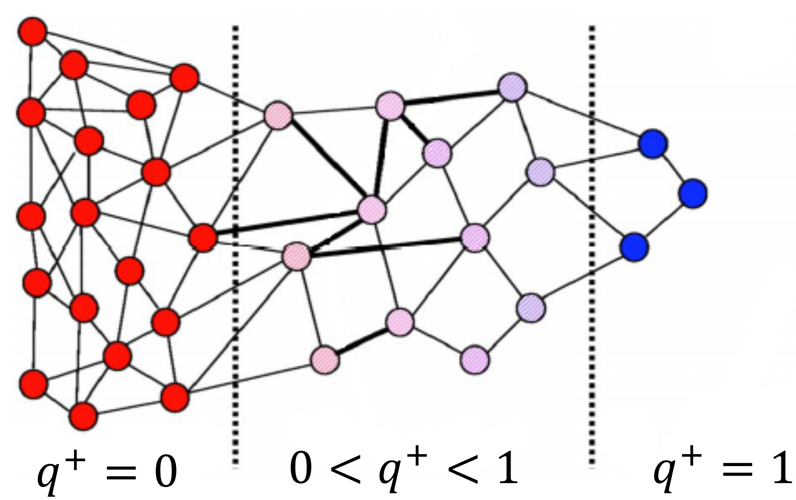
The committor

The discrete forward committor q_i^+ is defined as the probability that the process starting in *i* will reach first *B* (home) rather than *A* (bar).

$$q_i^+ = q_{i+1}^+ p_{i,i+1} + q_{i-1}^+ p_{i,i-1}$$

Same argument works for general MSM and a “home“ and “bar“ that consists of more than of one MSM state.

- $q_i^+ = 0$ for $i \in A$
- $q_i^+ = 1$ for $i \in B$
- $q_i^+ = \sum_{j \in I} p_{ij} q_j^+$ for $i \notin \{A, B\}$



For the reverse process

- $q_i^- = 1$ for $i \in A$
- $q_i^- = 0$ for $i \in B$
- $q_i^- = \sum_{j \in I} \frac{\pi_j}{\pi_i} p_{ji} q_j^-$ for $i \notin \{A, B\}$

With detailed balance $q_i^- = 1 - q_i^+$

Pathway decomposition

- pathway $P = (i_1, i_2, \dots, i_k)$ such that $i_1 \in A$, $i_k \in B$
- capacity (or flux) of pathway
$$f(P) = \min\{f_{i_l i_{l+1}} \mid l = 1 \dots k-1\}$$
- Pathway decomposition: chose a pathway P_1 and remove its capacity from the flux along all edges of P_1. Repeat until no flux remains.
- Decomposition is not unique, depends on the order in which $P_1, P_2, \dots$ are picked.
Reasonable choice: remove the strongest pathway (the one with largest capacity) first, then remove the strongest pathway of the remaining network.