Constrained information flows in temporal networks reveal intermittent communities

Ulf Aslak, University of Copenhagen, SODAS

Martin Rosvall, Umeå University, Department of Physics

Sune Lehmann, Technical University of Denmark, DTU Compute

NetSci, June 2018

A **method** for finding intermittent communities in temporal networks

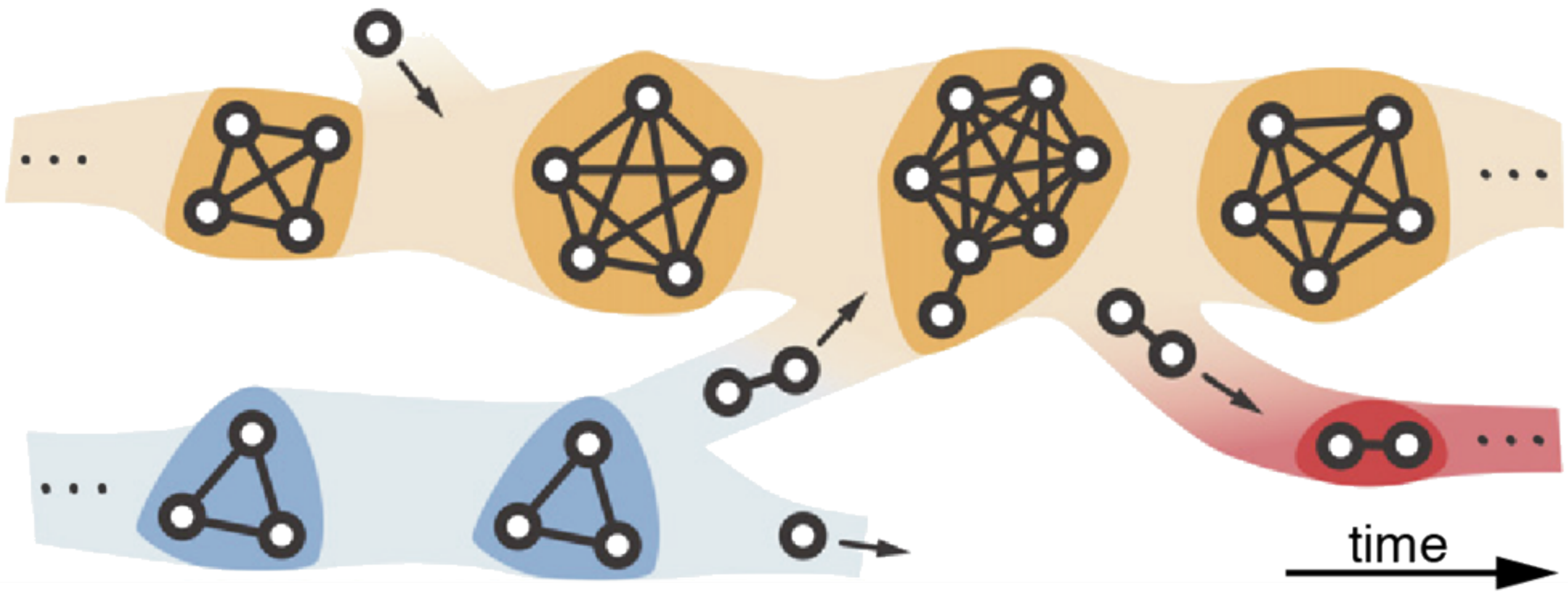
Ulf Aslak, University of Copenhagen, SODAS

Martin Rosvall, Umeå University, Department of Physics

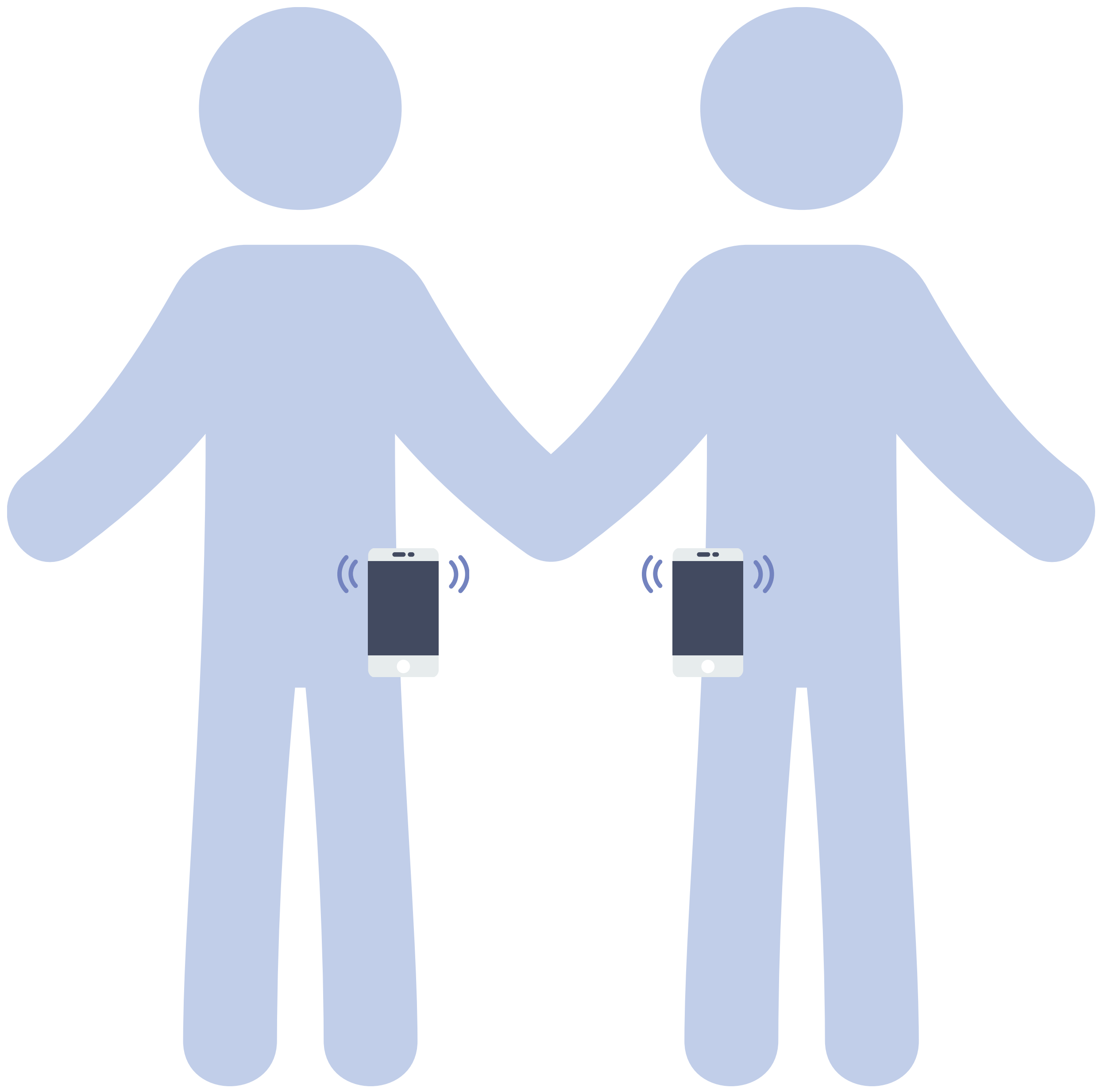
Sune Lehmann, Technical University of Denmark, DTU Compute

NetSci, June 2018

We want to understand social communities and how they change in time



To measure temporal proximity networks we can use smartphones



2 years

1000 individuals

bluetooth (5 min)

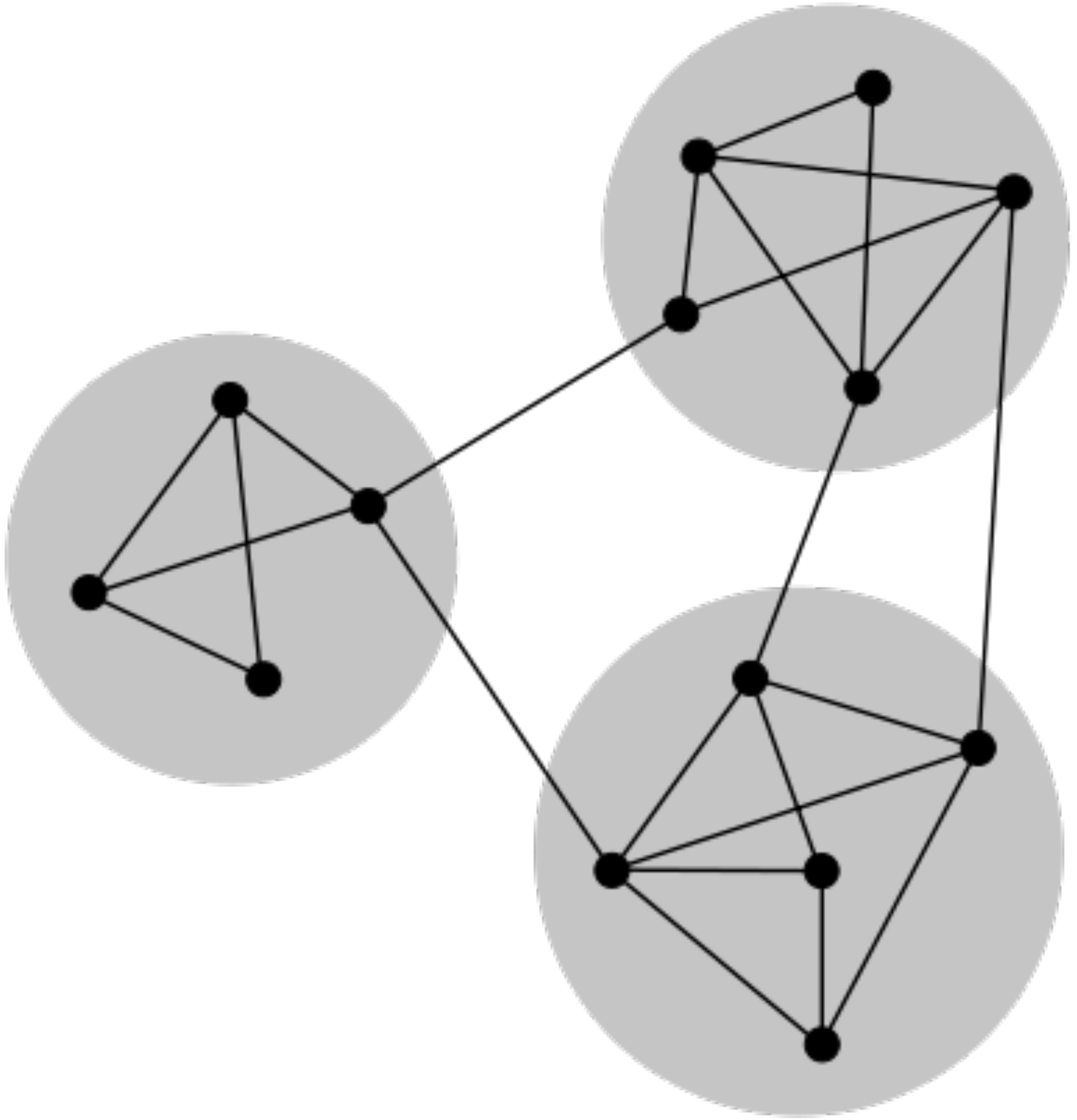
High accuracy
Frequent sampling

“Not too bad” population
Dense connections

Stopczynski, A., Sekara, V., Sapiezynski, P., Cuttone, A., Madsen, M. M., Larsen, J. E., & Lehmann, S. (2014). Measuring large-scale social networks with high resolution. PloS one, 9(4), e95978.


Pizza Pizza Pizza

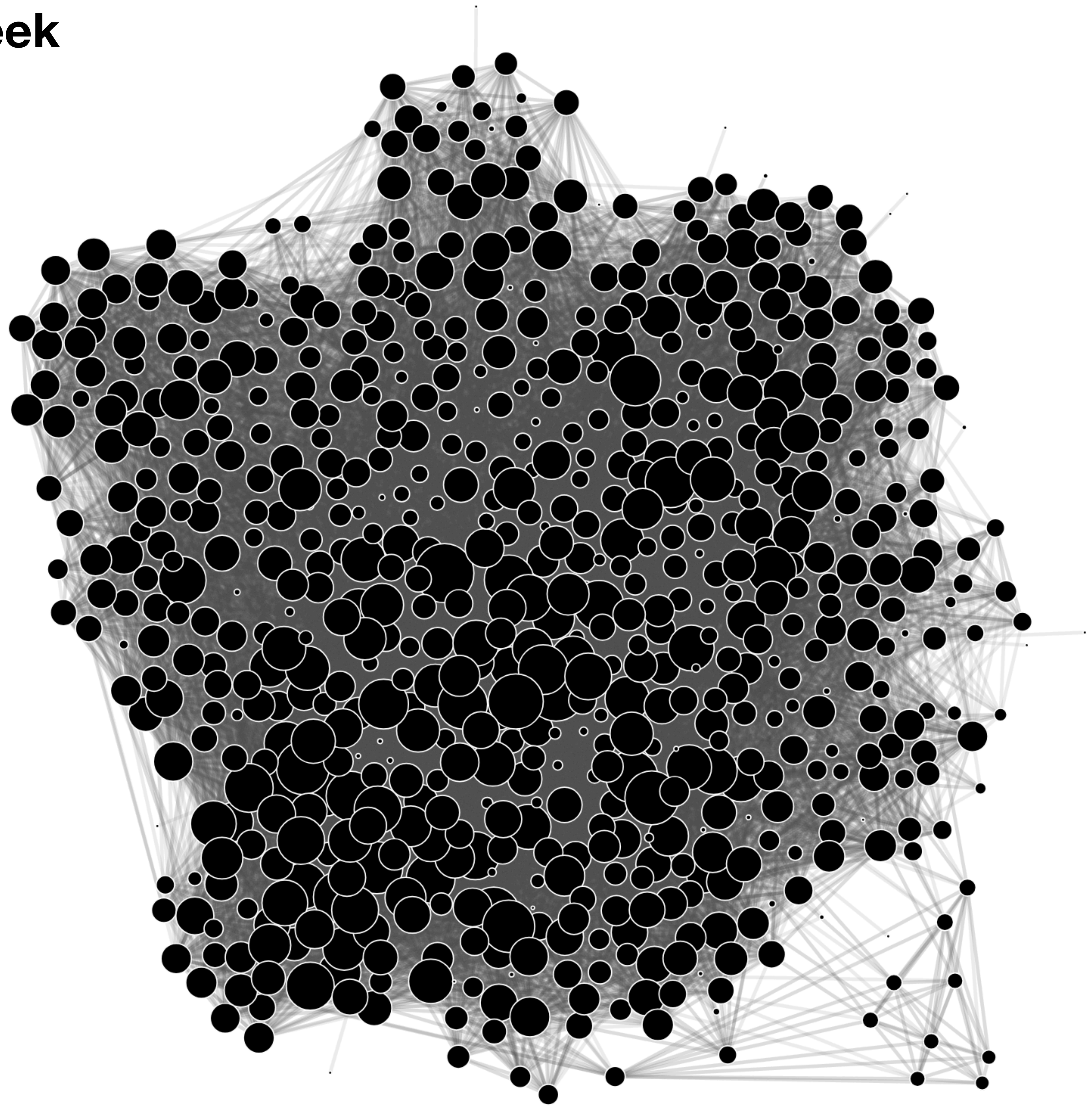
Can we expect beautiful community structure to emerge?



Wikipedia (https://commons.wikimedia.org/w/index.php?curid=17125894)

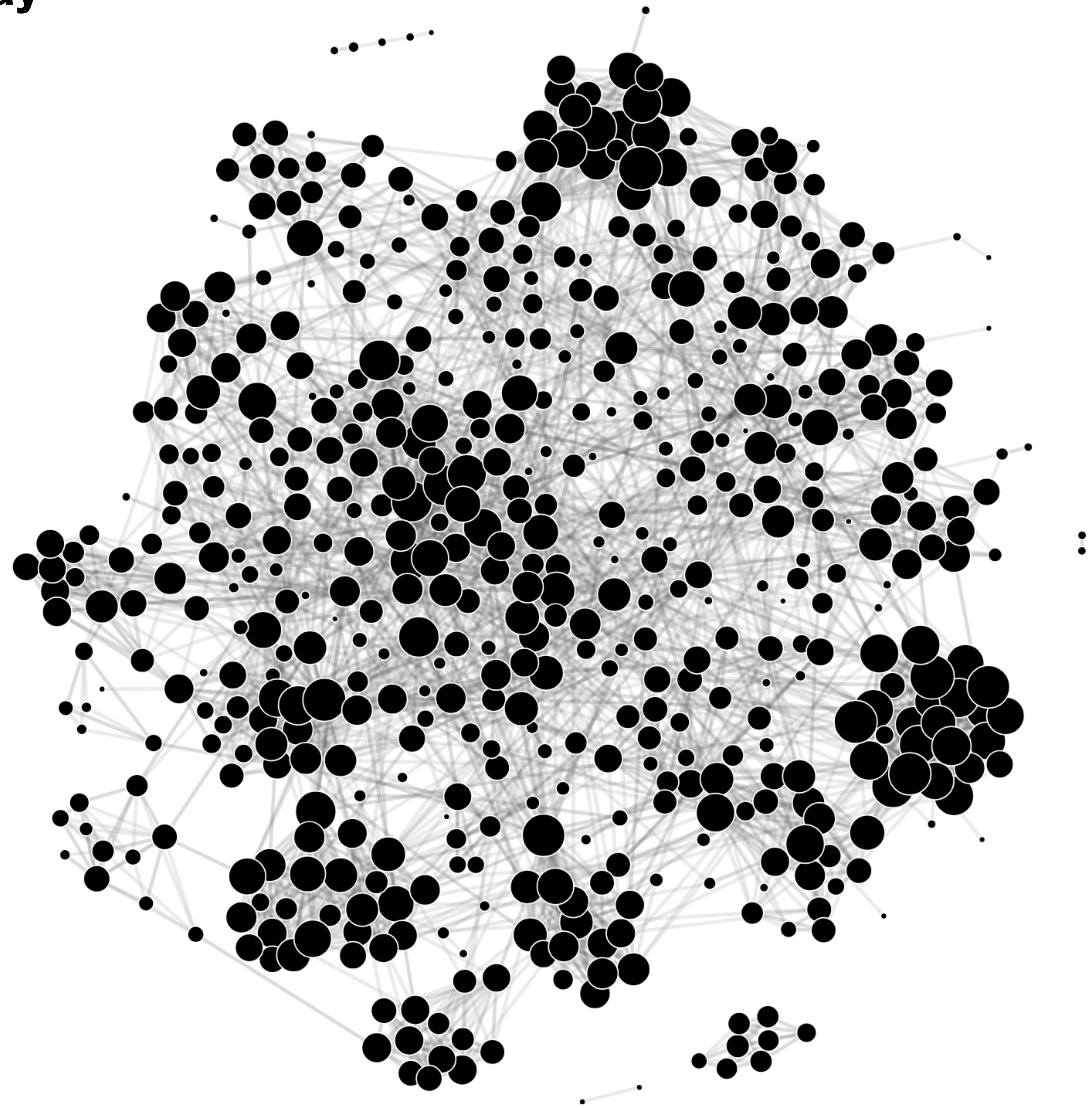
Not immediately: aggregating over time, we just get a big hairball

One week



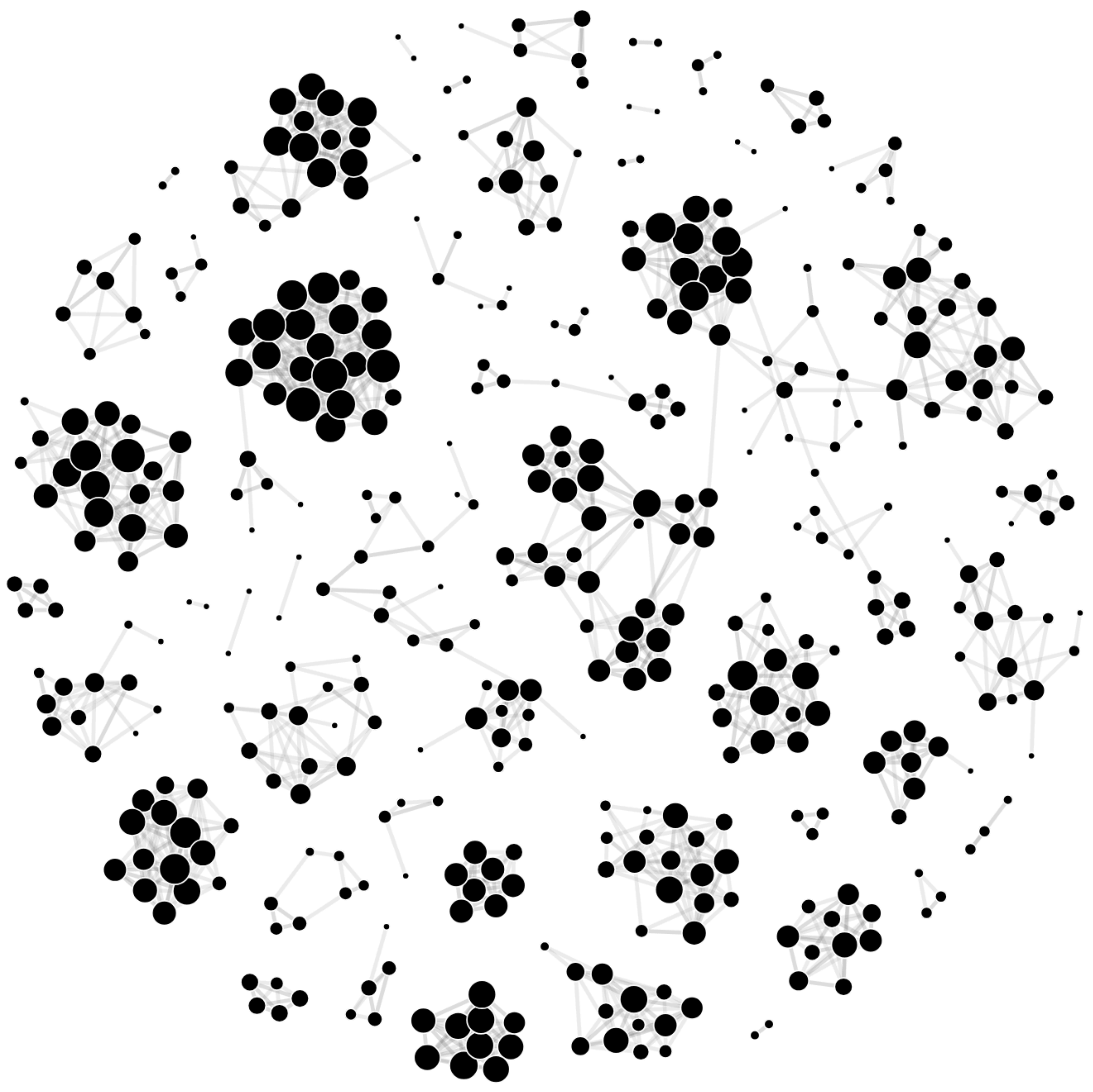
Shorter aggregates reveal more modularity

One day



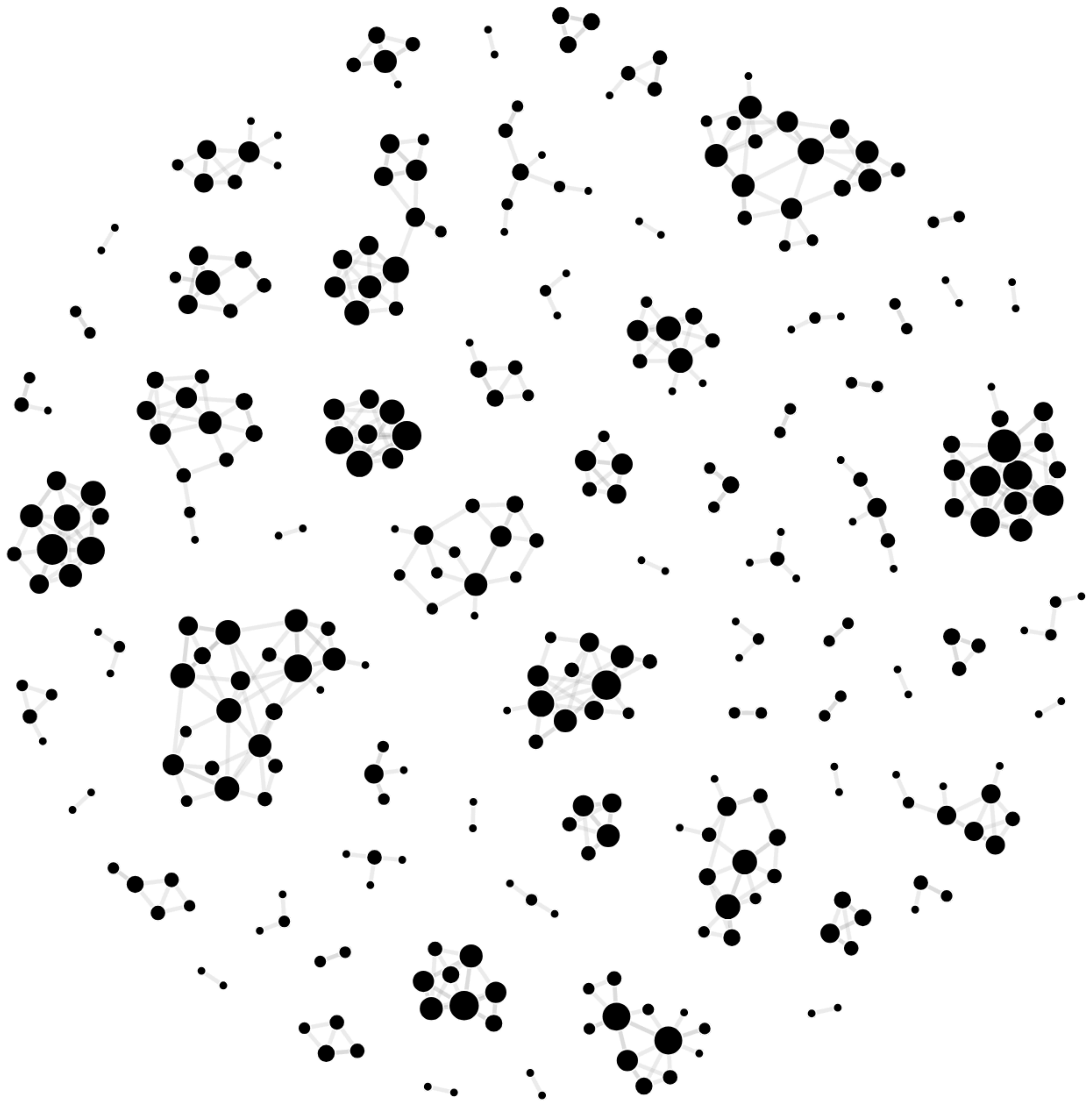
Shorter aggregates reveal more modularity

One hour

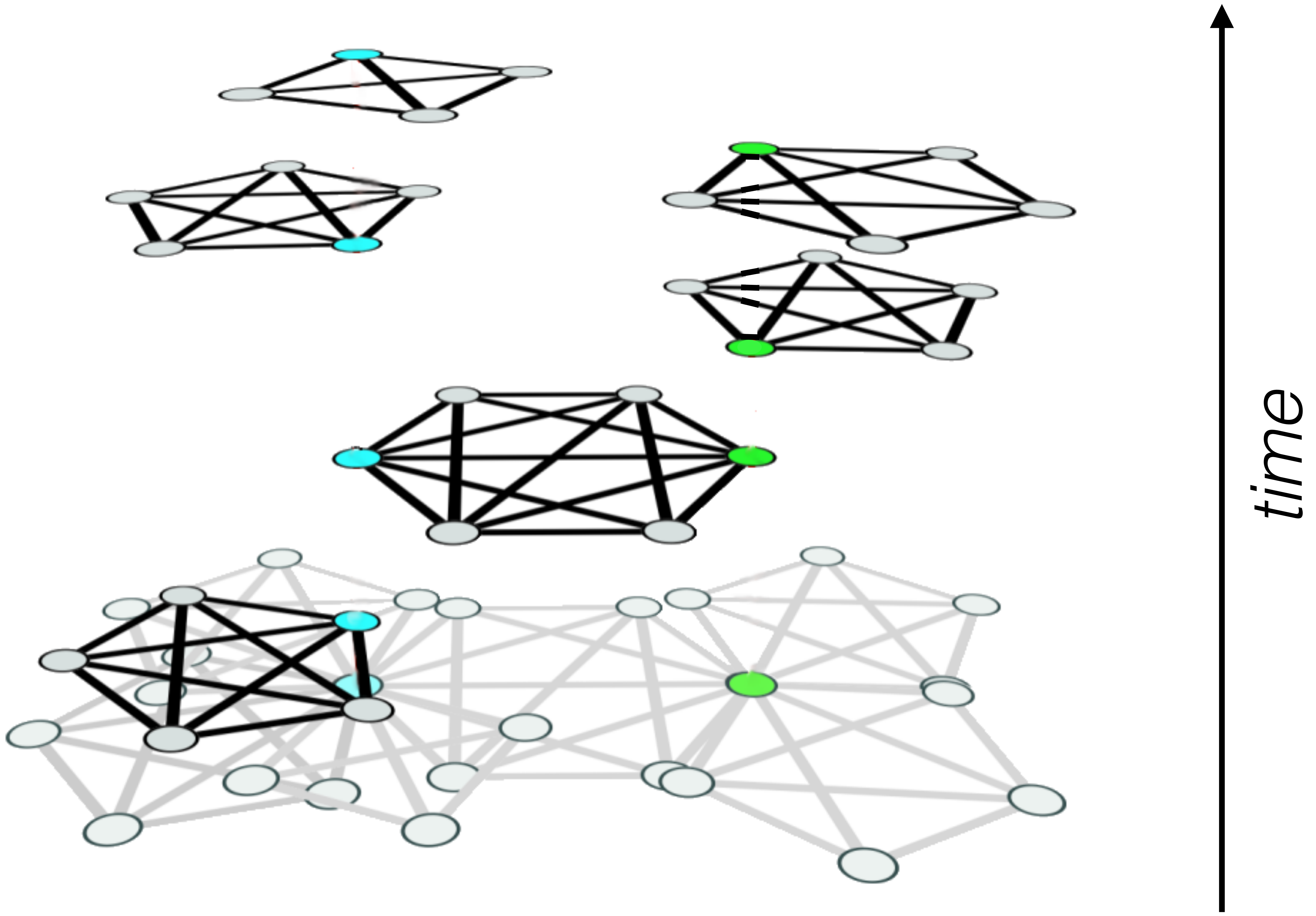


Shorter aggregates reveal more modularity

5 minutes

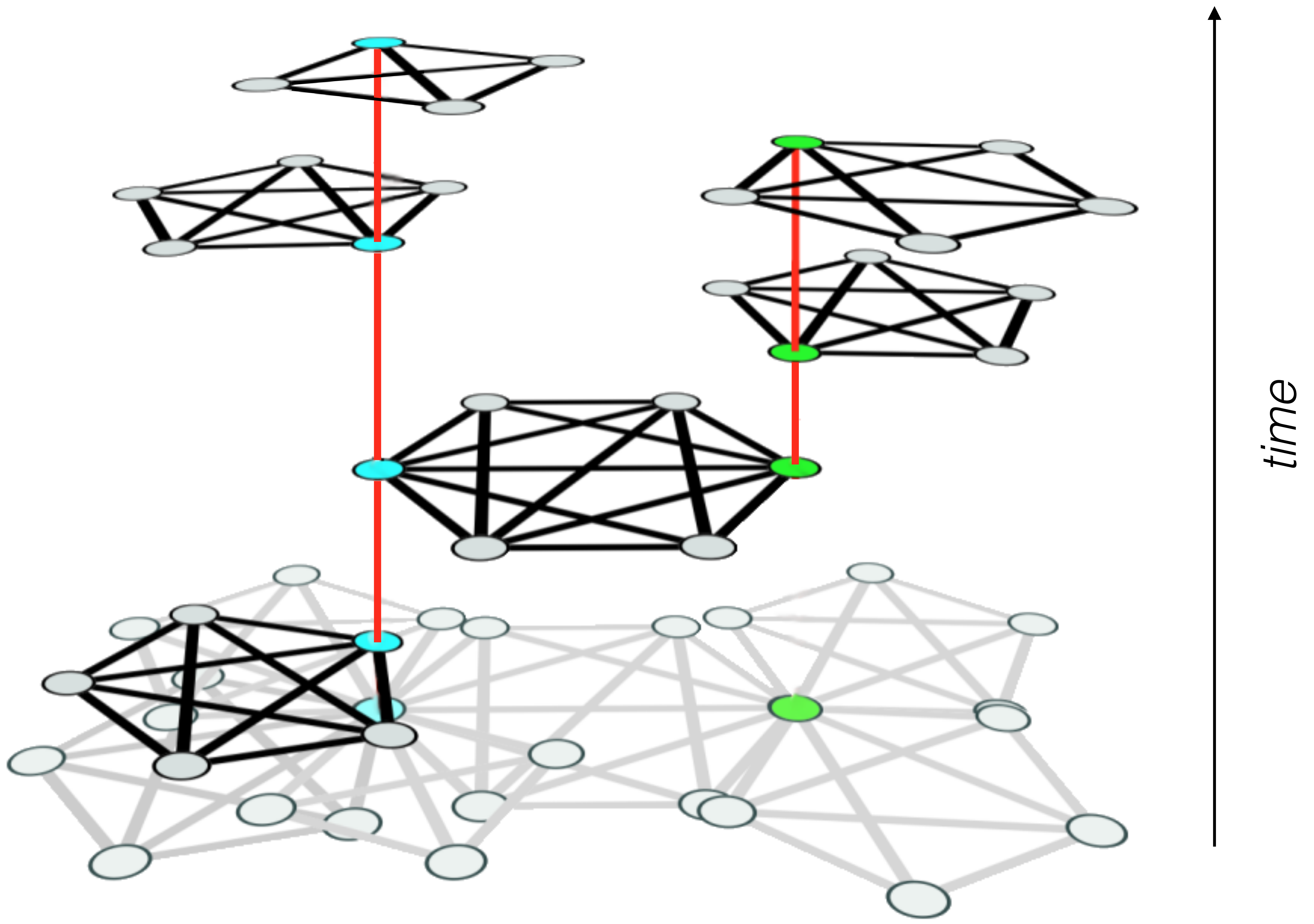


There are many overlapping communities which are intermittent in time

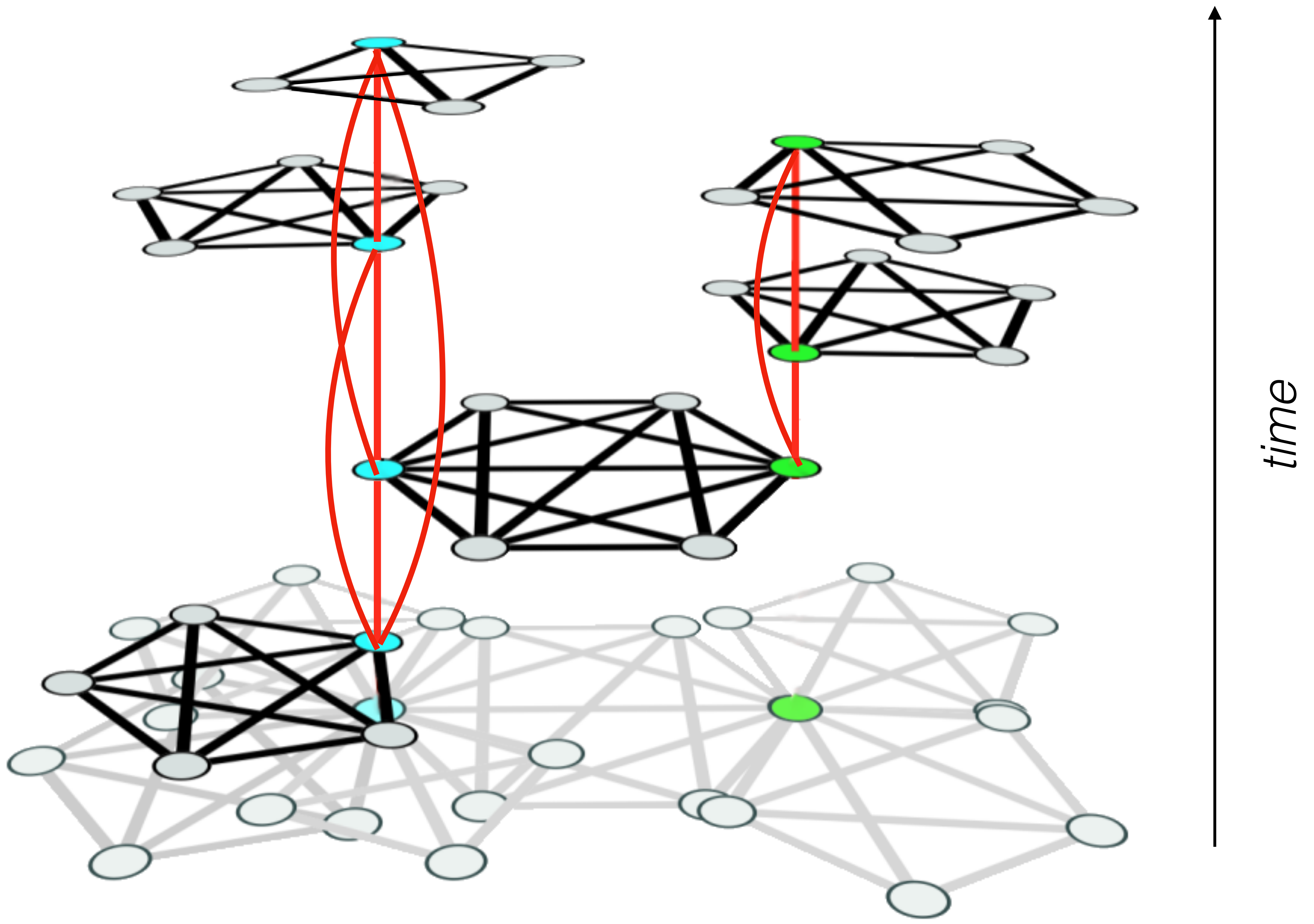


We need a way to find **overlapping intermittent** communities

Approach: Self-connect nodes through time



There are many ways to create temporal dependencies



Holme, P. (2015). Modern temporal network theory: a colloquium. The European Physical Journal B, 88(9), 234.

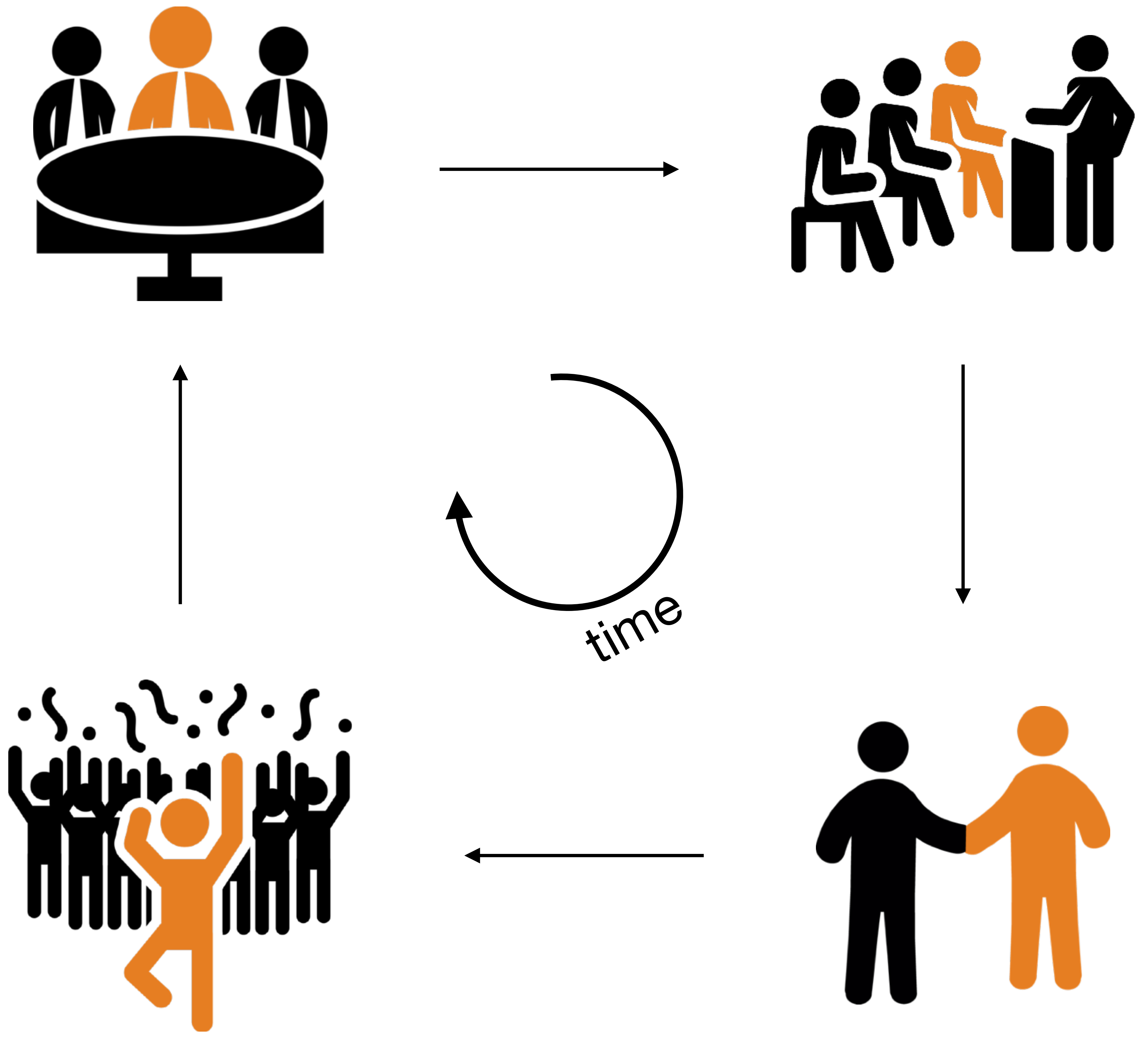
Mucha, P. J., Richardson, T., Macon, K., Porter, M. A., & Onnela, J. P. (2010). Community structure in time-dependent, multiscale, and multiplex networks. science, 328(5980), 876-878.

Petri, G., & Expert, P. (2014). Temporal stability of network partitions. Physical Review E, 90(2), 022813.

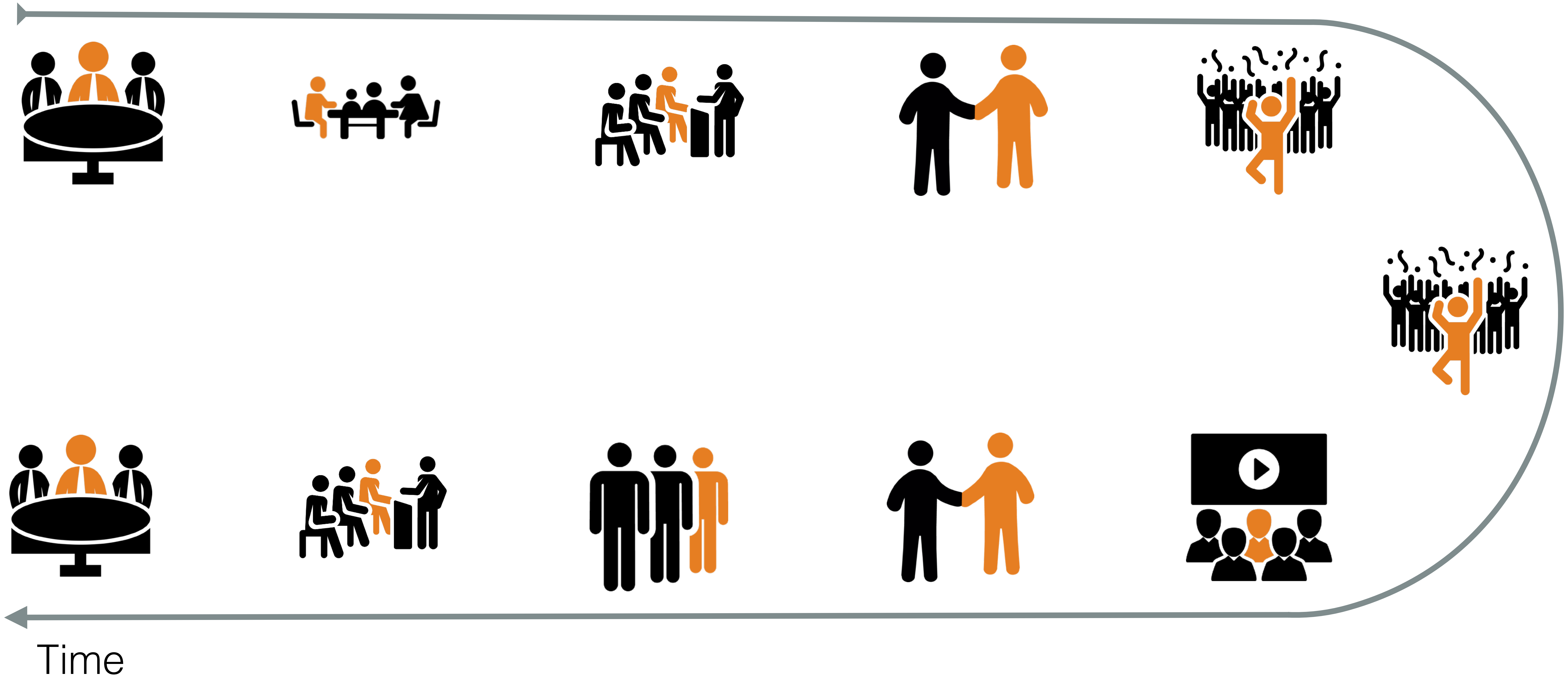
De Bacco, C., Power, E. A., Larremore, D. B., & Moore, C. (2017). Community detection, link prediction, and layer interdependence in multilayer networks. Physical Review E, 95(4), 042317.

We want to create inter-temporal dependencies that are **meaningful** and **interpretable**

We suggest a principle for defining meaningful and interpretable temporal dependencies

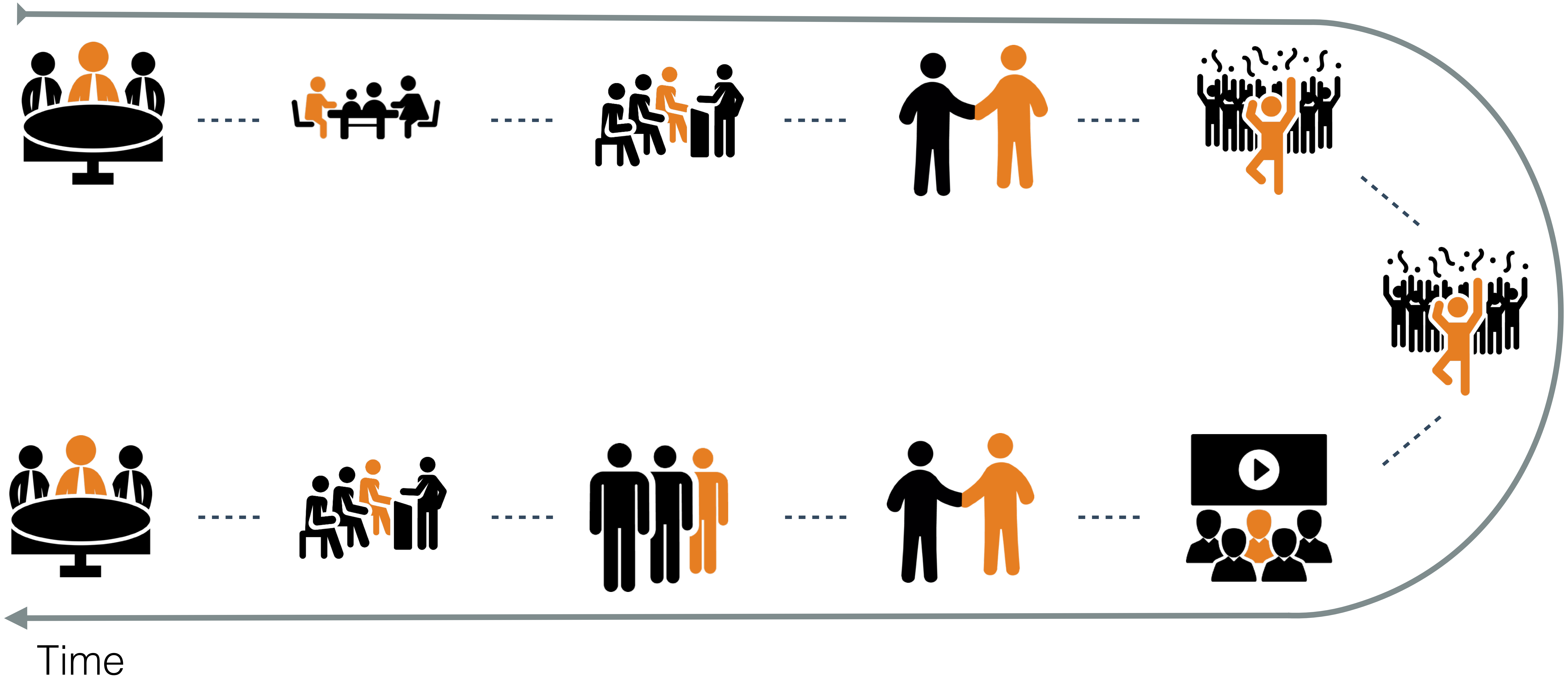


We suggest a principle for defining meaningful and interpretable temporal dependencies



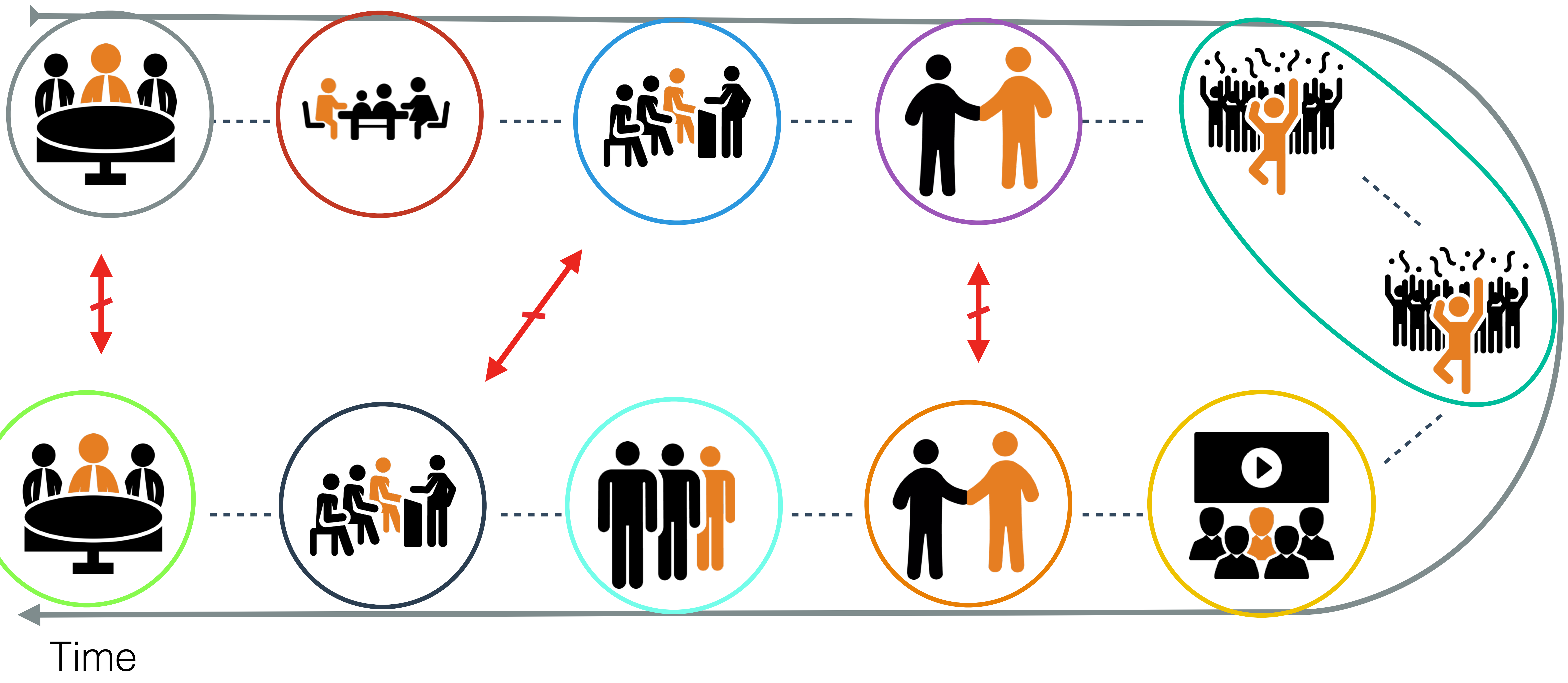
We suggest a principle for defining meaningful and interpretable temporal dependencies

- **Adjacent coupling** results in temporally continuous communities



We suggest a principle for defining meaningful and interpretable temporal dependencies

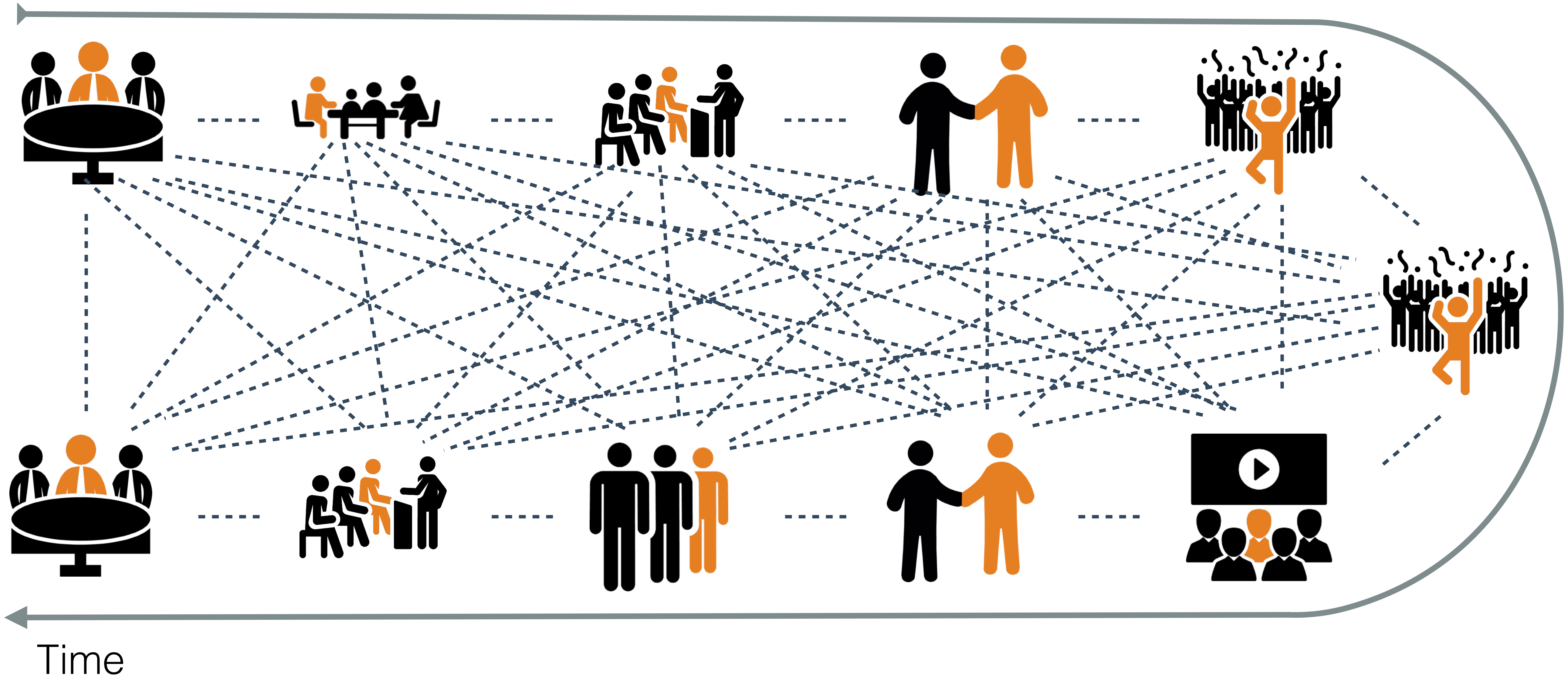
- **Adjacent coupling** results in temporally continuous communities



Intermittence not captured

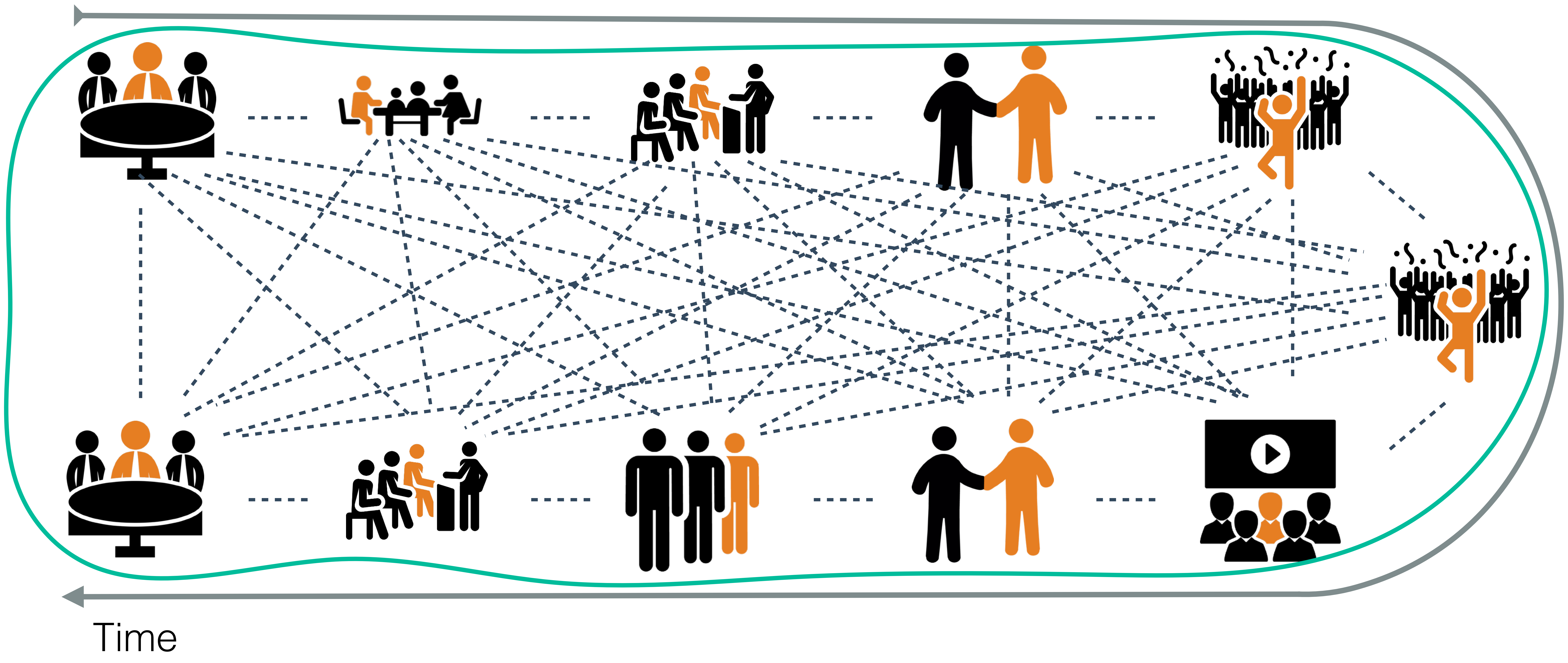
We suggest a principle for defining meaningful and interpretable temporal dependencies

- **Full coupling** merges overlapping communities



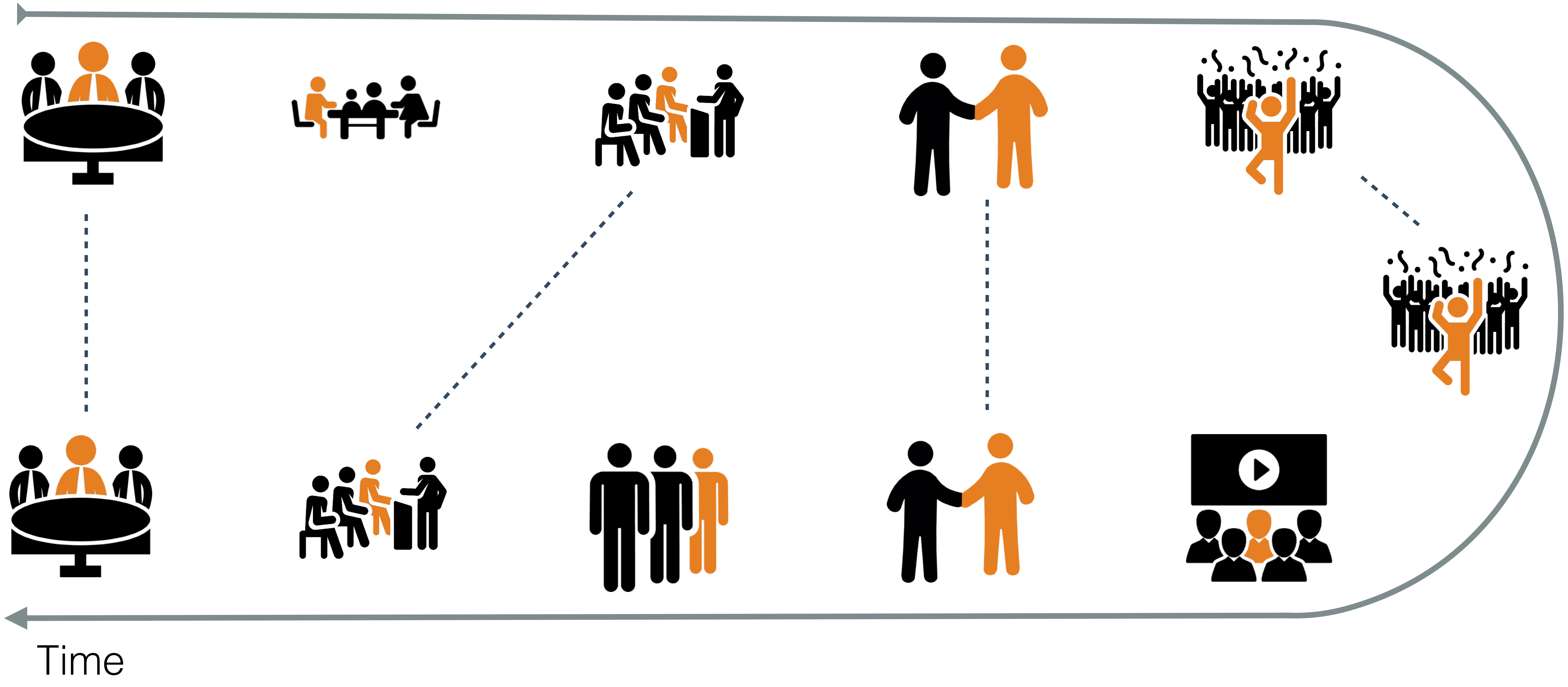
We suggest a principle for defining meaningful and interpretable temporal dependencies

- **Full coupling** merges overlapping communities



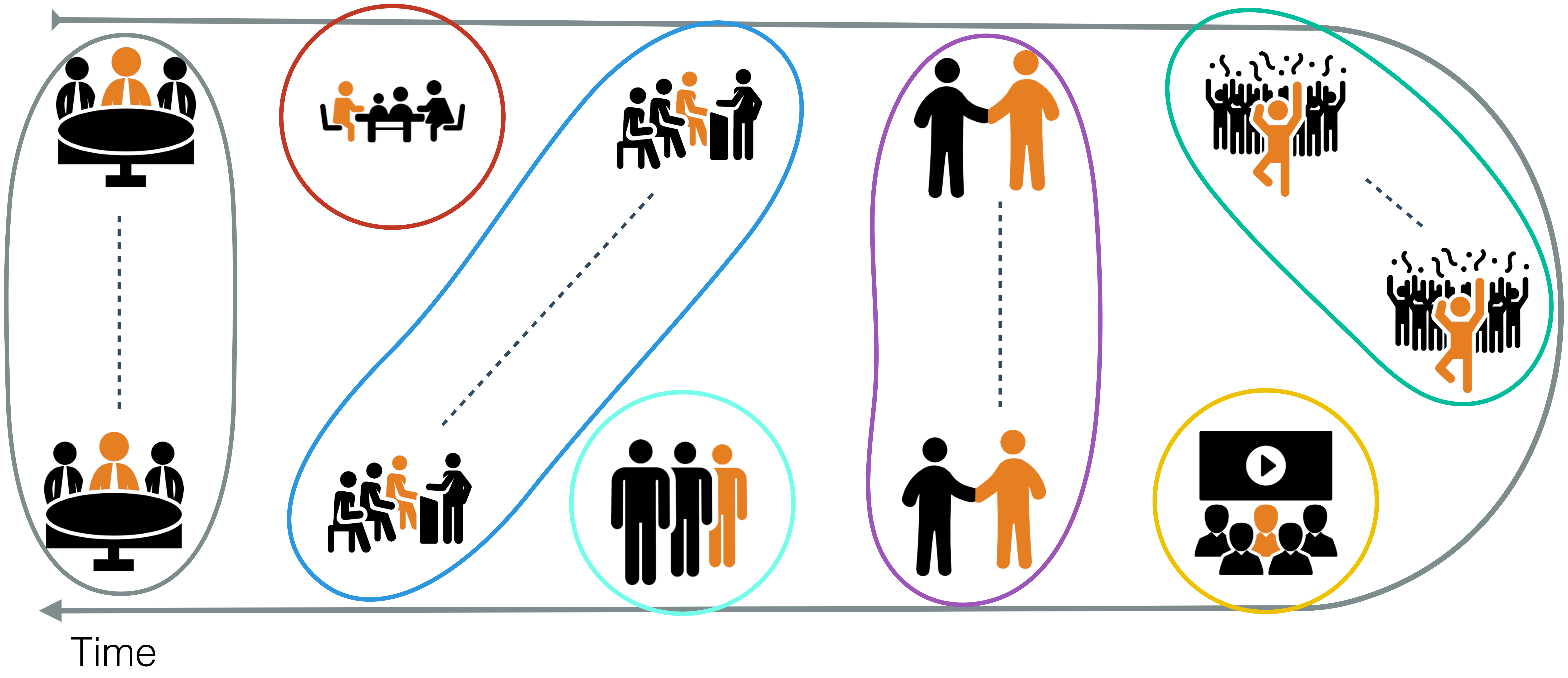
We suggest a principle for defining meaningful and interpretable temporal dependencies

- We suggest: **neighborhood flow coupling.** Finds intermittent and overlapping communities

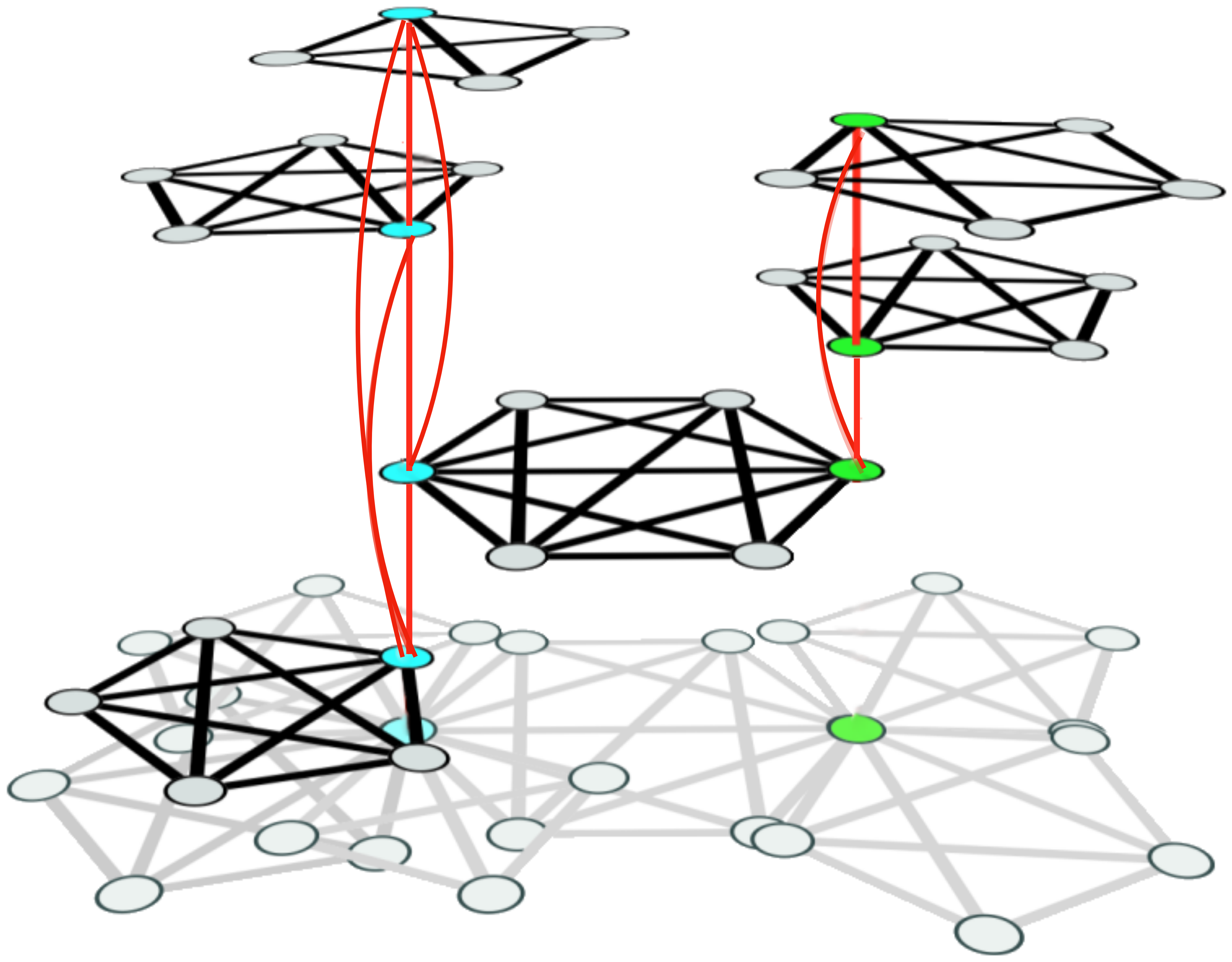


We suggest a principle for defining meaningful and interpretable temporal dependencies

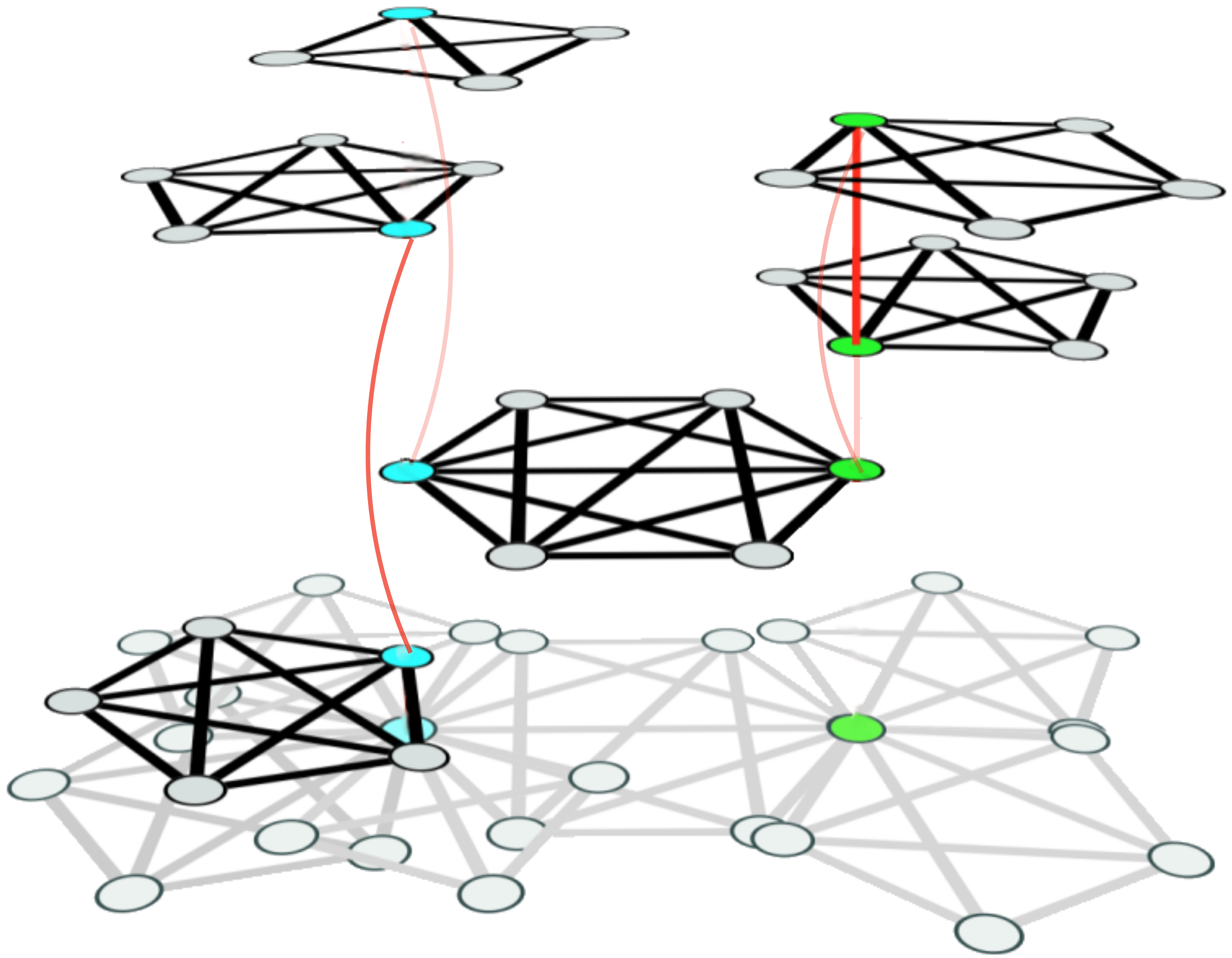
- We suggest: **neighborhood flow coupling.** Finds intermittent and overlapping communities



Neighborhood flow coupling: Couple two states with a strength that is proportional to how similar their surrounding network flows are



Neighborhood flow coupling: Couple two states with a strength that is proportional to how similar their surrounding network flows are

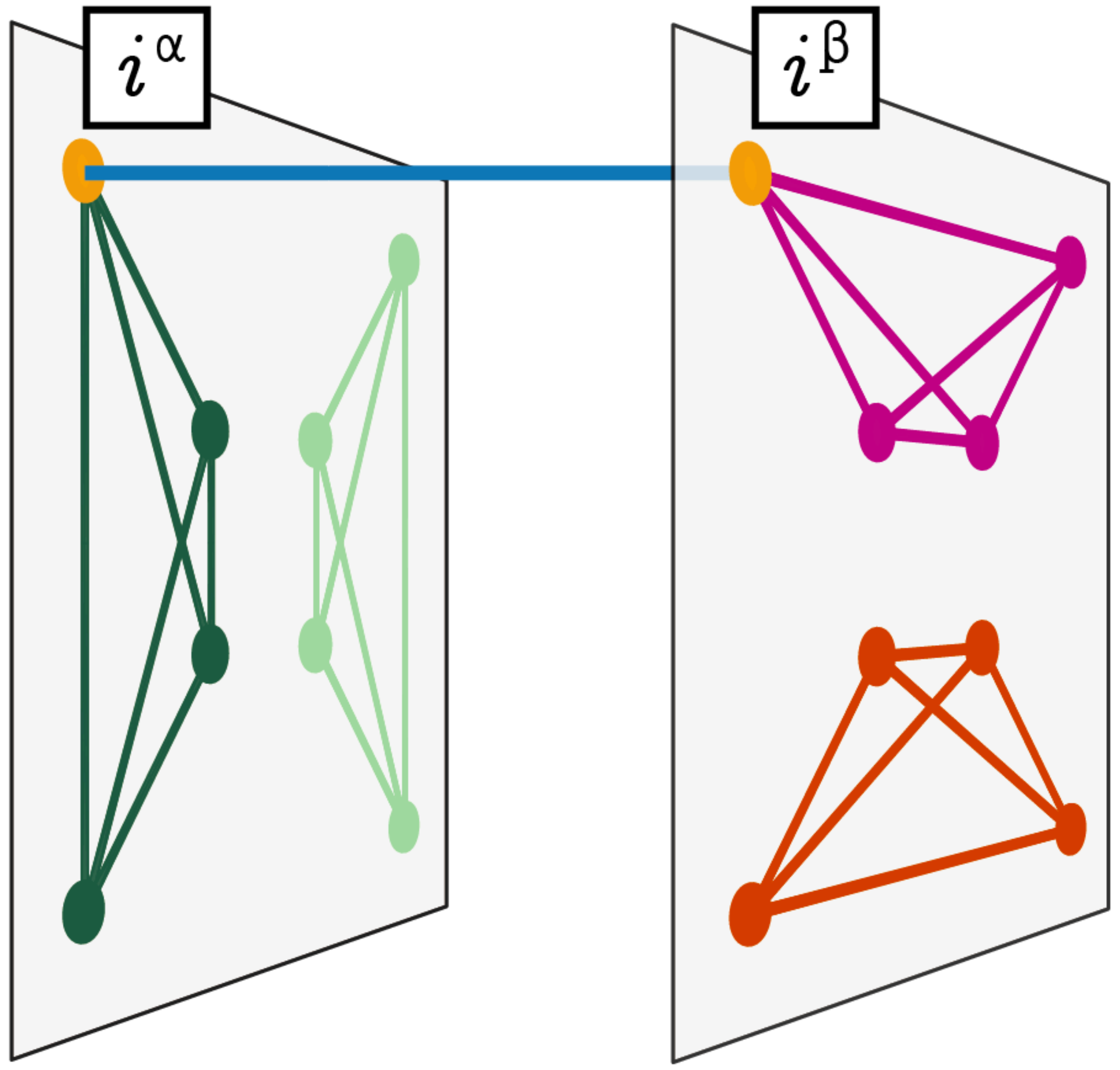


We use the Jensen-Shannon similarity to estimate temporal dependencies

$$D_i^{\alpha\beta} = 1 - \mathrm{JSD}\left(\boldsymbol{P}_i^{\alpha}, \boldsymbol{P}_i^{\beta}\right)$$

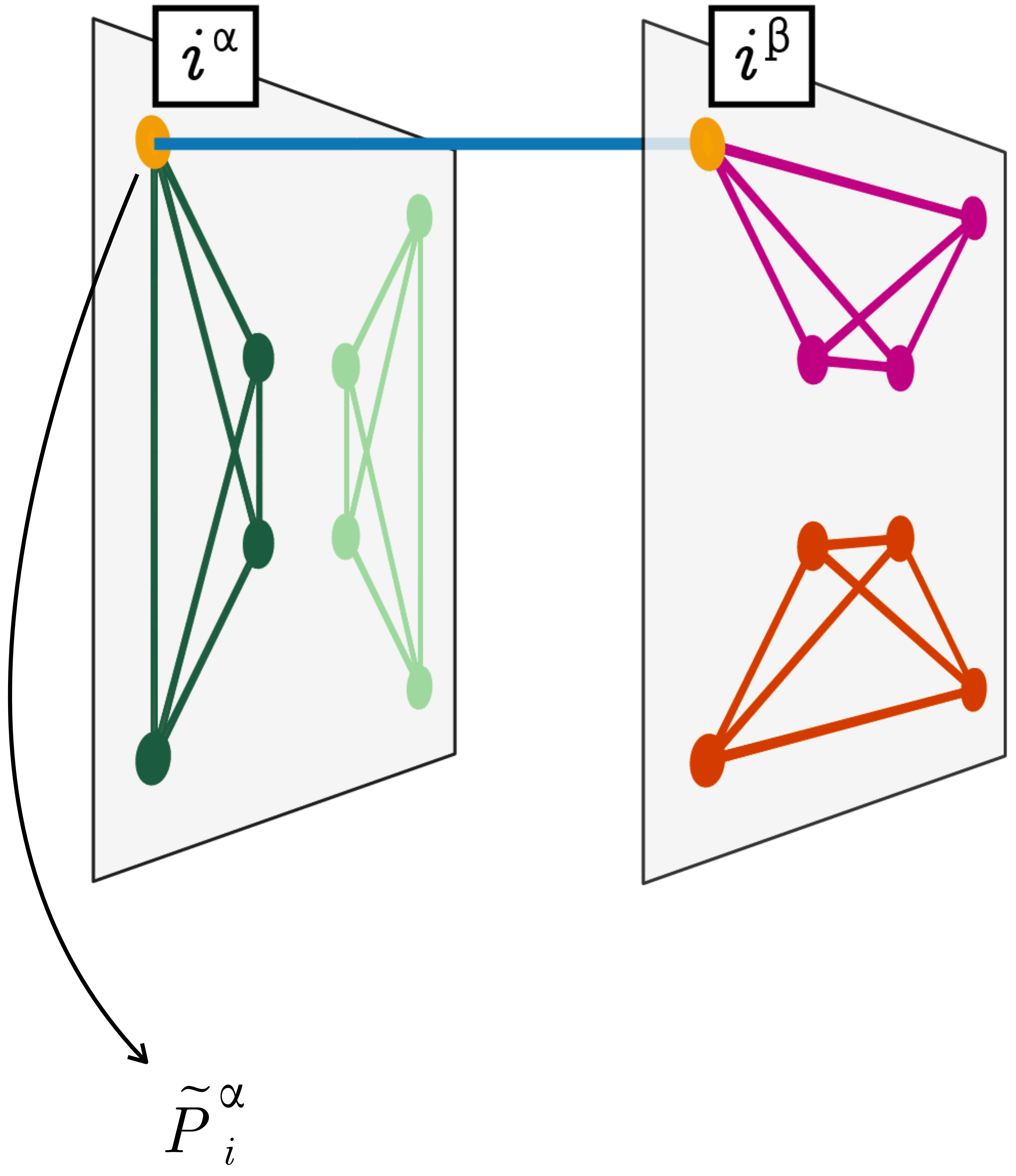
We use the Jensen-Shannon similarity to estimate temporal dependencies

$$D_i^{\alpha\beta} = 1 - \mathrm{JSD}\left(\boldsymbol{P}_i^{\alpha}, \boldsymbol{P}_i^{\beta}\right)$$



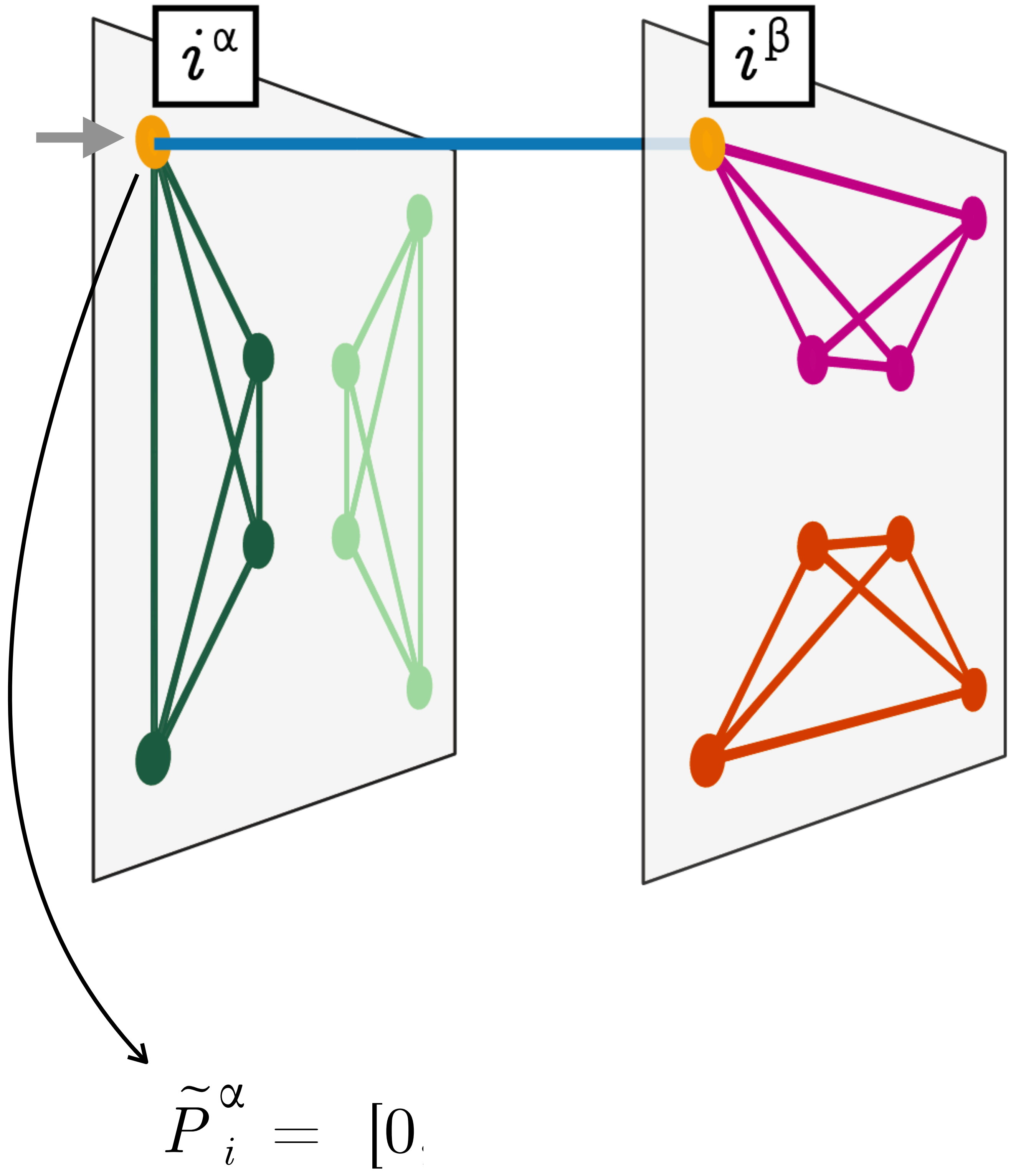
We use the Jensen-Shannon similarity to estimate temporal dependencies

$$D_i^{\alpha\beta} = 1 - \mathrm{JSD}\left(\boldsymbol{P}_i^{\alpha}, \boldsymbol{P}_i^{\beta}\right)$$



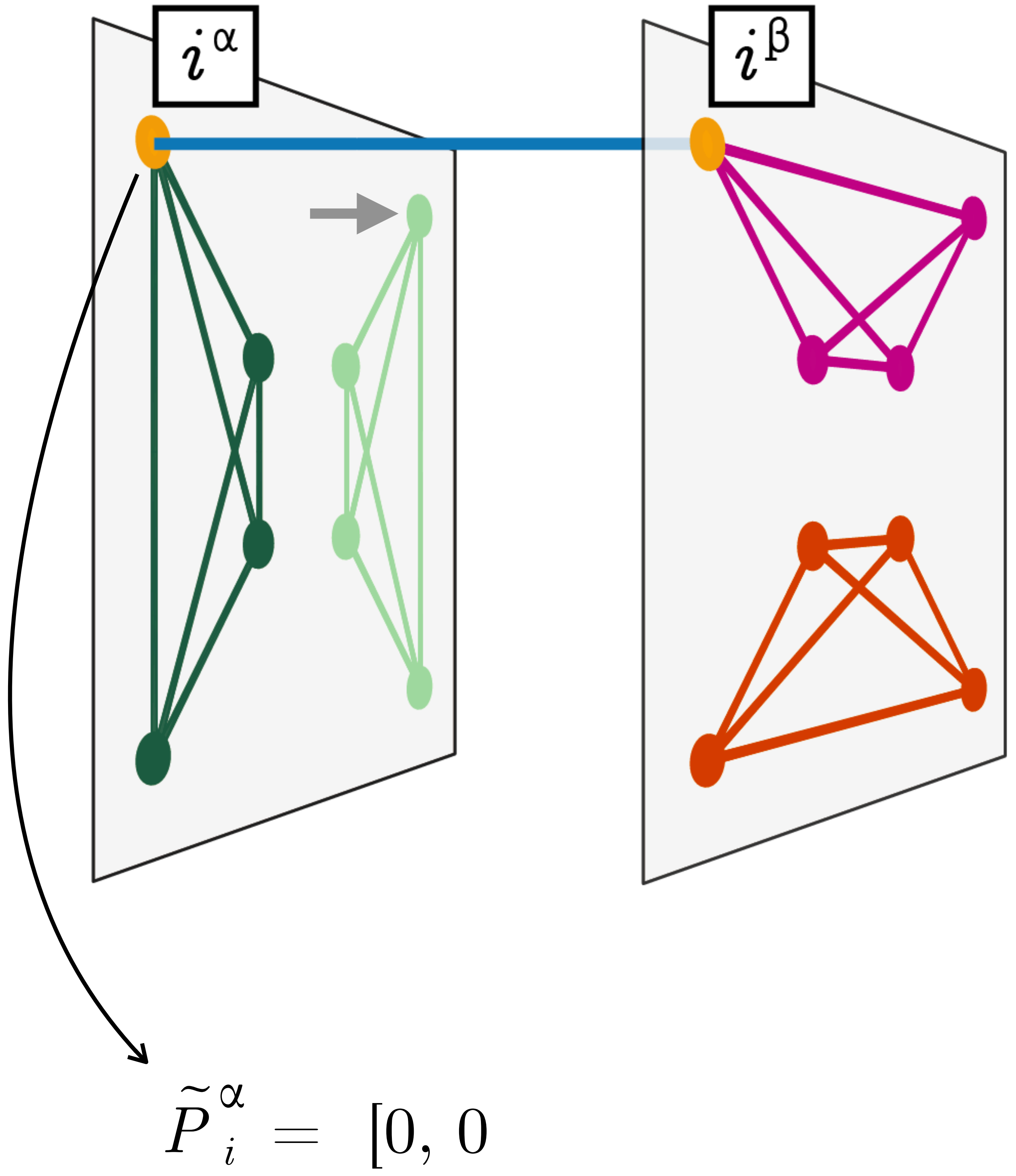
We use the Jensen-Shannon similarity to estimate temporal dependencies

$$D_i^{\alpha\beta} = 1 - \mathrm{JSD}\left(\boldsymbol{P}_i^{\alpha}, \boldsymbol{P}_i^{\beta}\right)$$



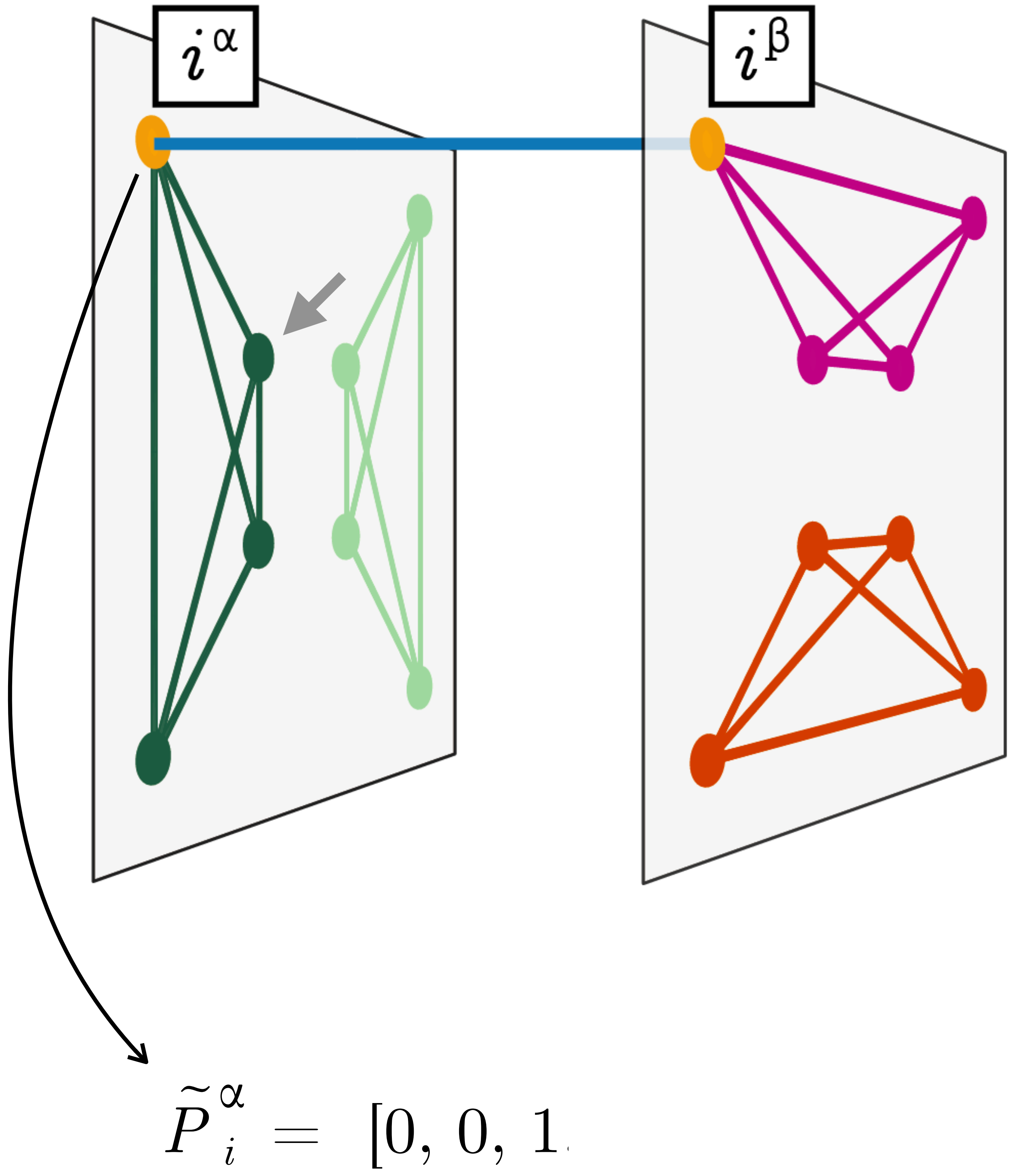
We use the Jensen-Shannon similarity to estimate temporal dependencies

$$D_i^{\alpha\beta} = 1 - \mathrm{JSD}\left(\boldsymbol{P}_i^{\alpha}, \boldsymbol{P}_i^{\beta}\right)$$



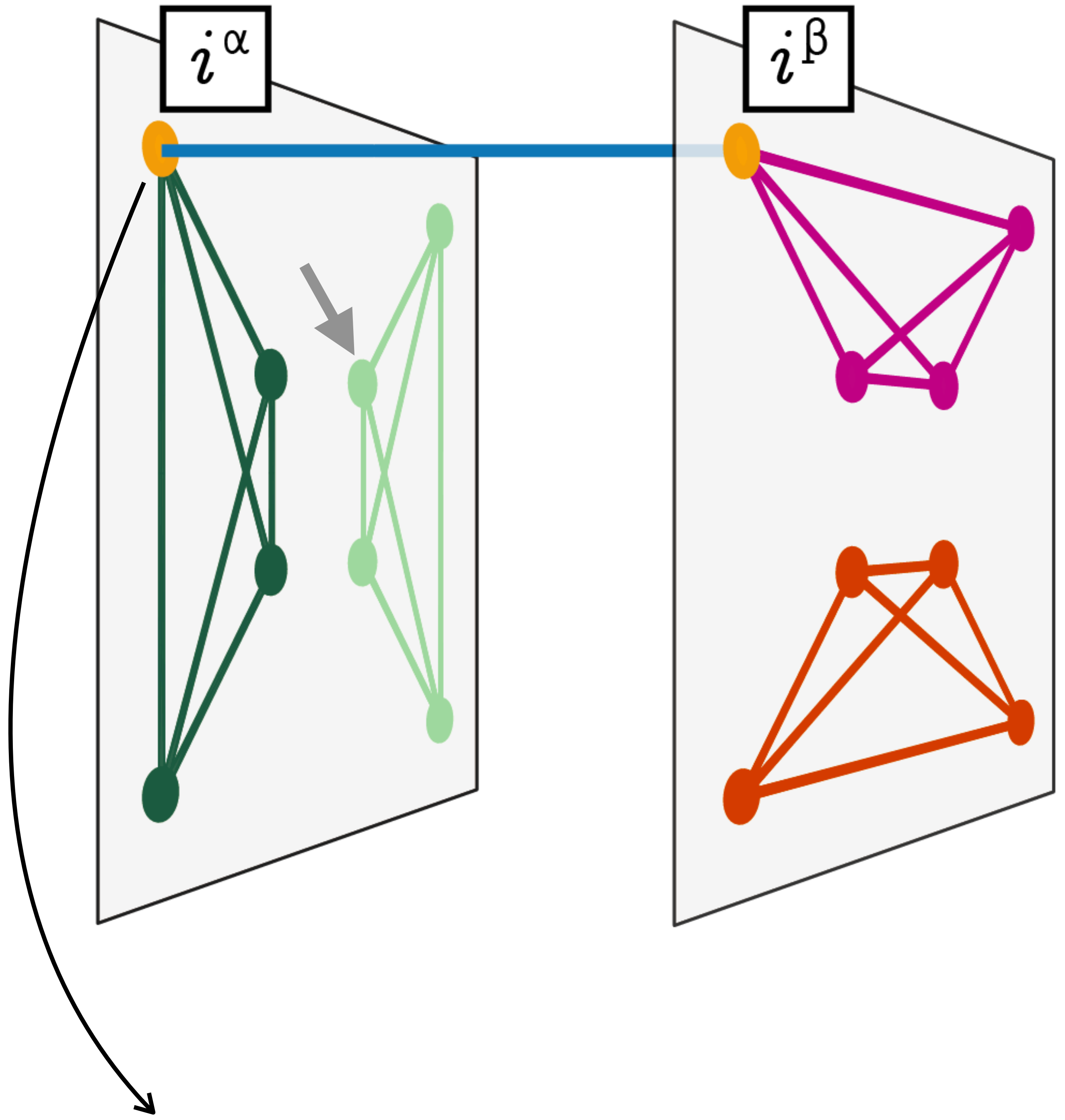
We use the Jensen-Shannon similarity to estimate temporal dependencies

$$D_i^{\alpha\beta} = 1 - \mathrm{JSD}\left(\boldsymbol{P}_i^{\alpha}, \boldsymbol{P}_i^{\beta}\right)$$



We use the Jensen-Shannon similarity to estimate temporal dependencies

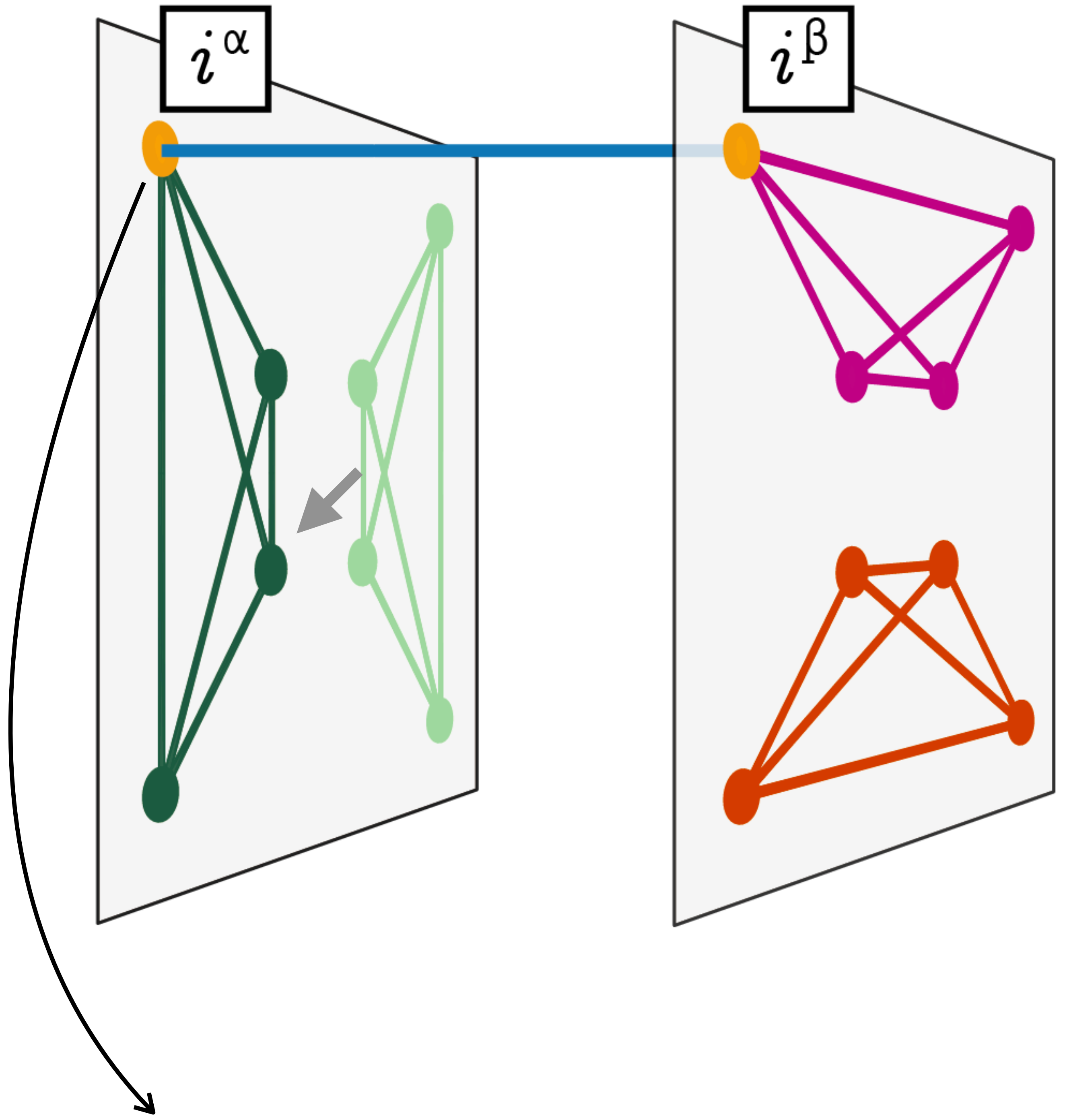
$$D_i^{\alpha\beta} = 1 - \mathrm{JSD}\left(\boldsymbol{P}_i^{\alpha}, \boldsymbol{P}_i^{\beta}\right)$$



$\widetilde{P}_i^{\alpha} = [0, 0, 1, 0$

We use the Jensen-Shannon similarity to estimate temporal dependencies

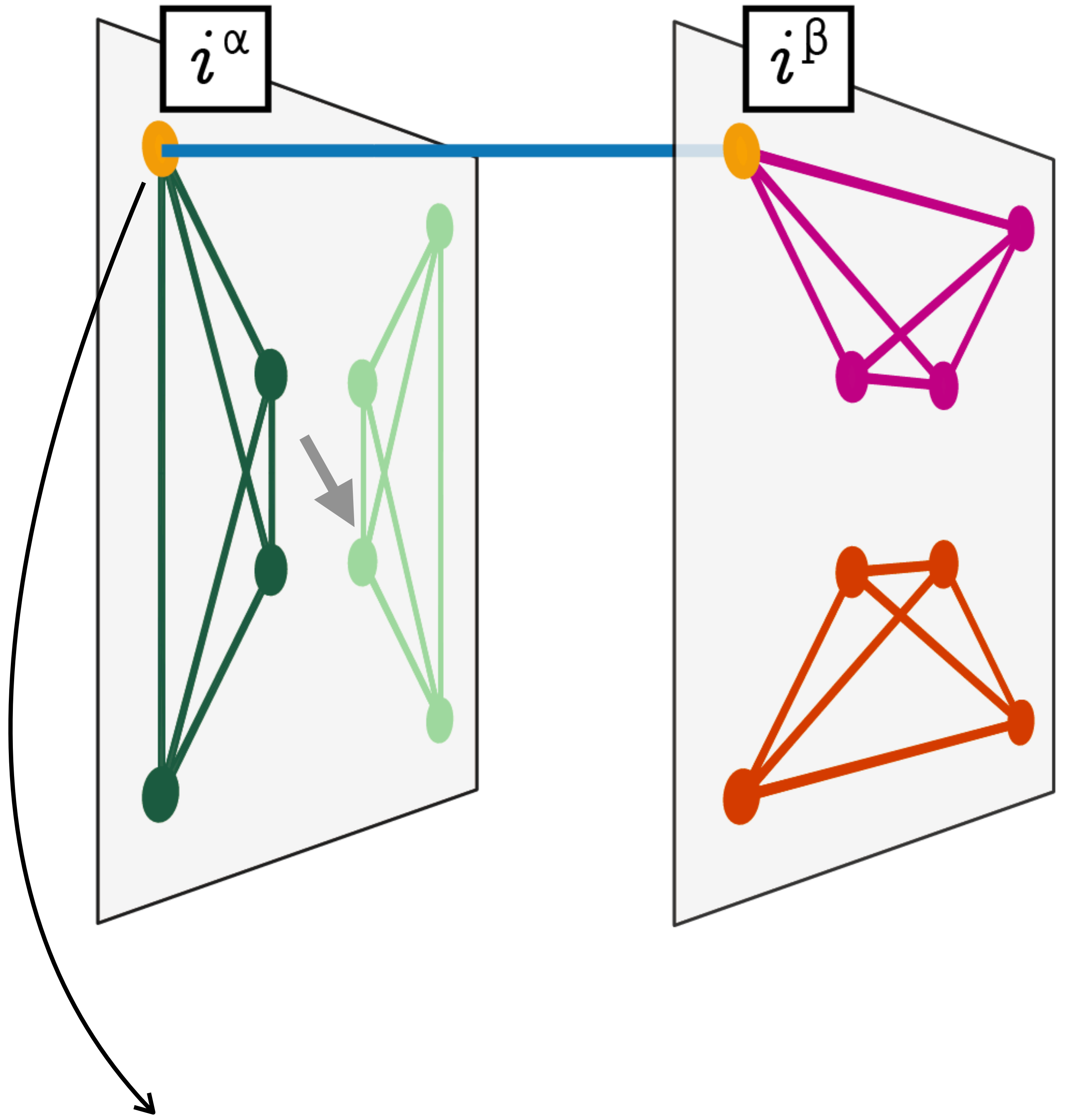
$$D_i^{\alpha\beta} = 1 - \mathrm{JSD}\left(\boldsymbol{P}_i^{\alpha}, \boldsymbol{P}_i^{\beta}\right)$$



$\widetilde{P}_i^{\alpha} = [0, 0, 1, 0, 1$

We use the Jensen-Shannon similarity to estimate temporal dependencies

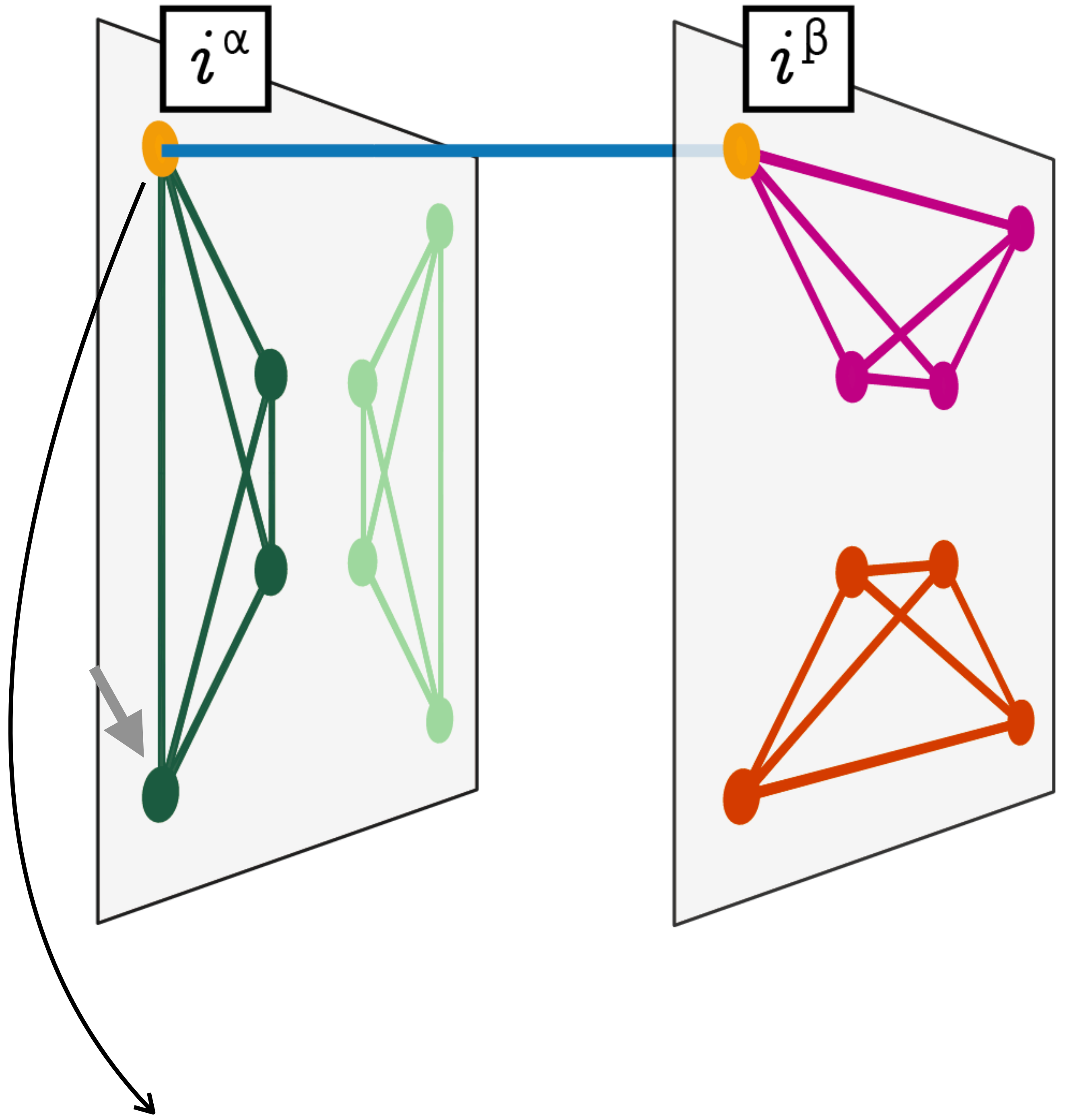
$$D_i^{\alpha\beta} = 1 - \mathrm{JSD}\left(\boldsymbol{P}_i^{\alpha}, \boldsymbol{P}_i^{\beta}\right)$$



$\widetilde{P}_i^{\alpha} = [0, 0, 1, 0, 1, 0$

We use the Jensen-Shannon similarity to estimate temporal dependencies

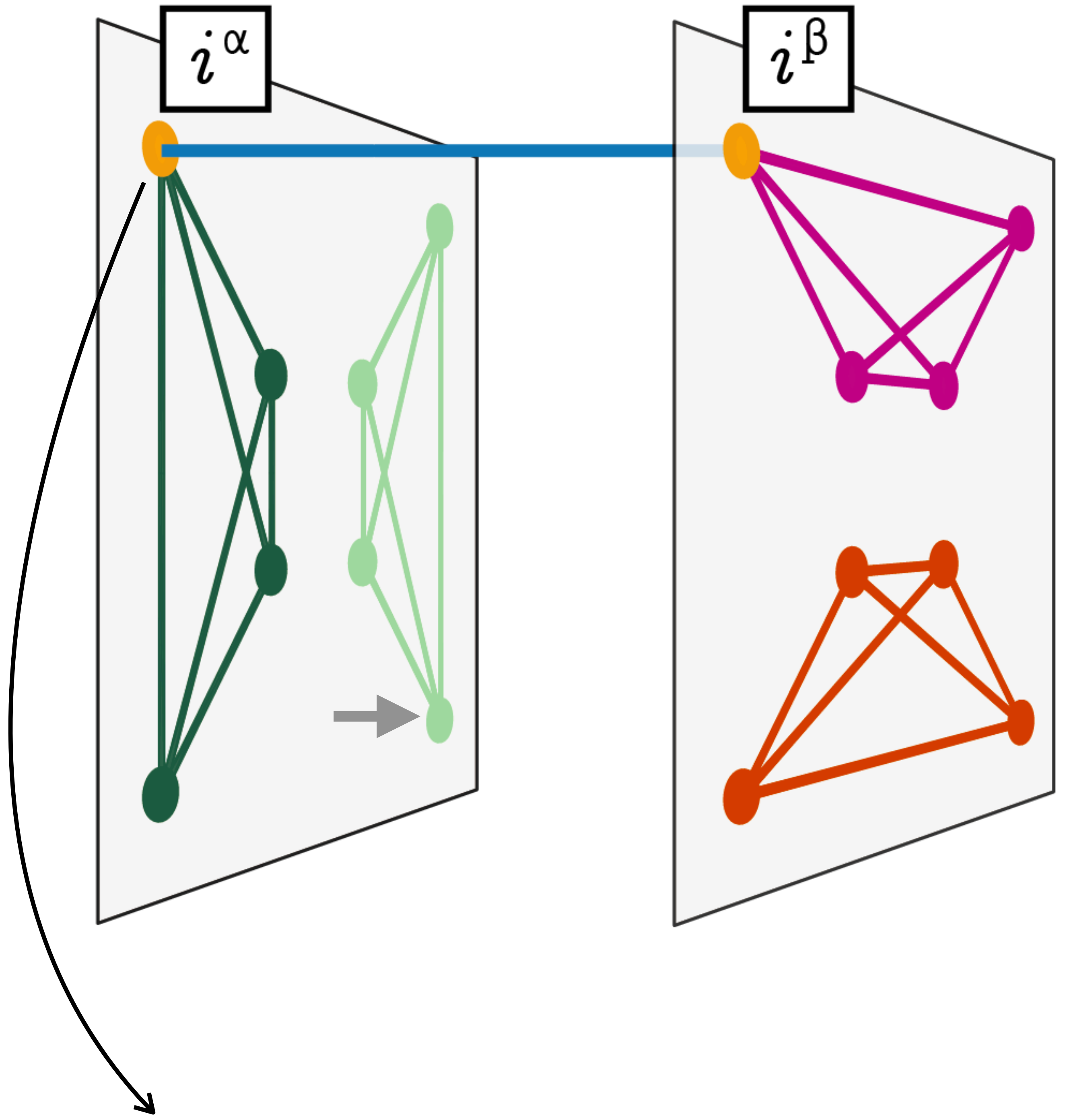
$$D_i^{\alpha\beta} = 1 - \mathrm{JSD}\left(\boldsymbol{P}_i^{\alpha}, \boldsymbol{P}_i^{\beta}\right)$$



$$\widetilde{P}_i^{\alpha} = [0,\ 0,\ 1,\ 0,\ 1,\ 0,\ 1$$

We use the Jensen-Shannon similarity to estimate temporal dependencies

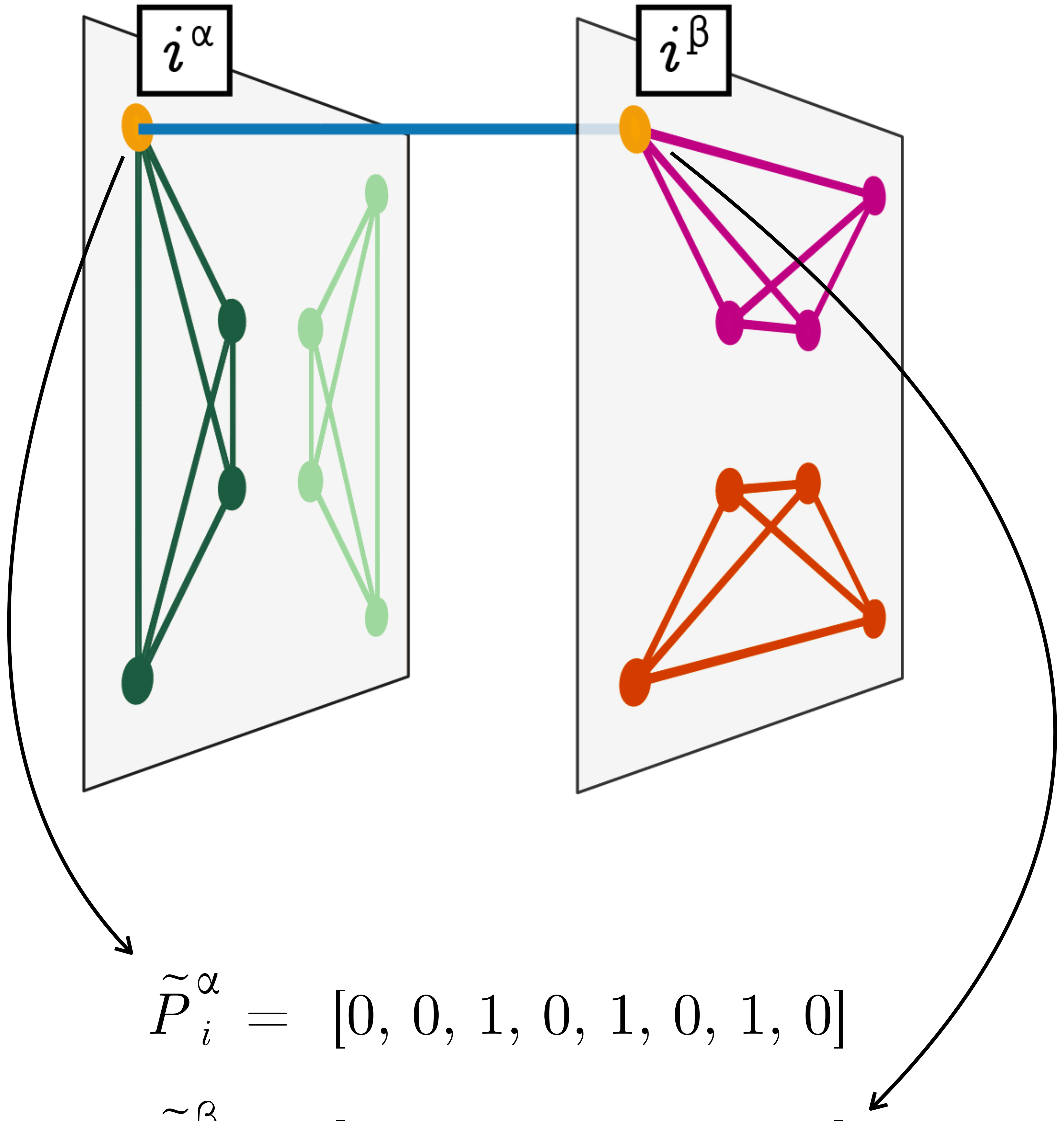
$$D_i^{\alpha\beta} = 1 - \mathrm{JSD}\left(\boldsymbol{P}_i^{\alpha}, \boldsymbol{P}_i^{\beta}\right)$$



$$\widetilde{P}_i^{\alpha} = [0,\ 0,\ 1,\ 0,\ 1,\ 0,\ 1,\ 0]$$

We use the Jensen-Shannon similarity to estimate temporal dependencies

$$D_i^{\alpha\beta} = 1 - \mathrm{JSD}\left(\boldsymbol{P}_i^{\alpha}, \boldsymbol{P}_i^{\beta}\right)$$



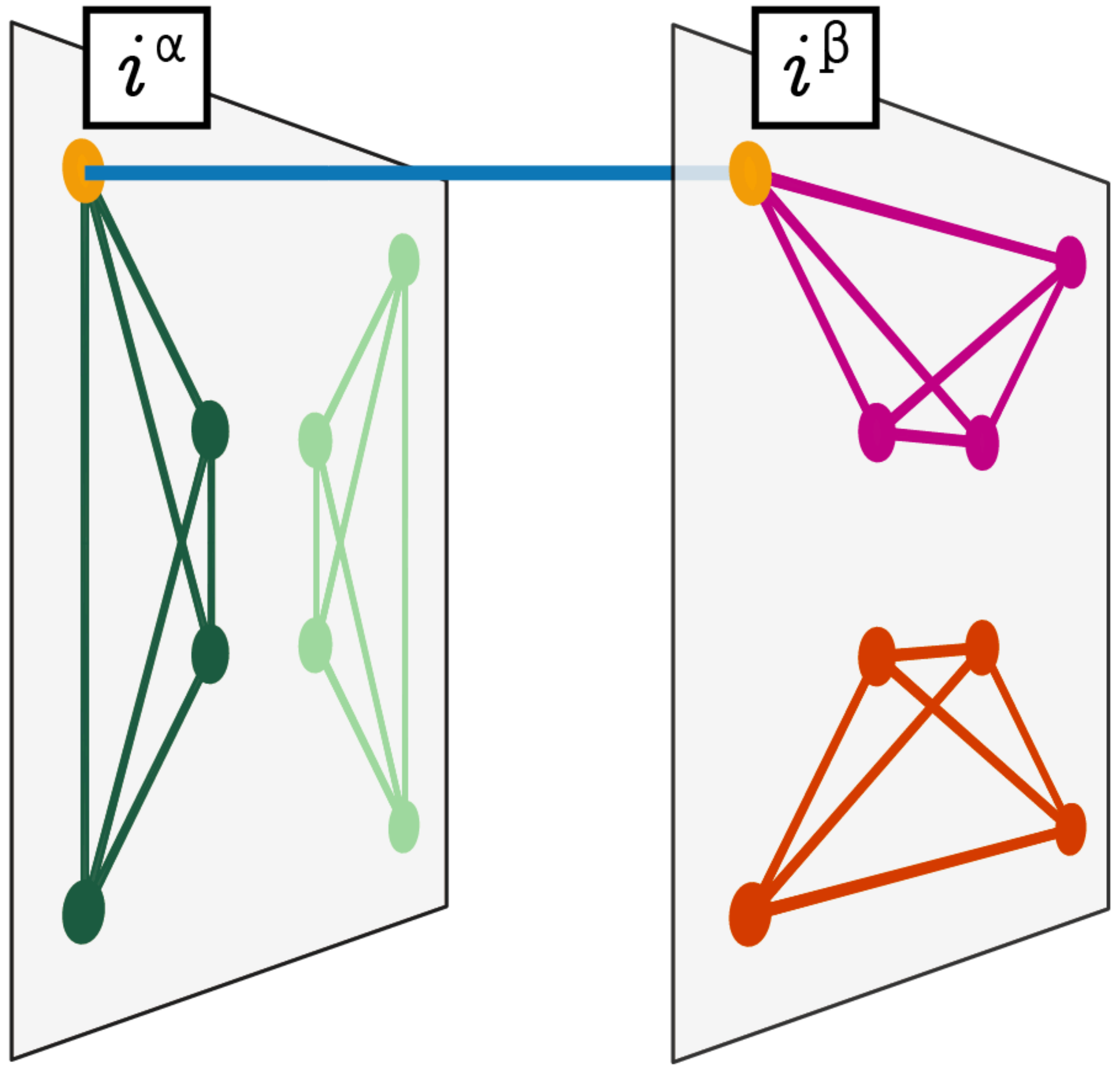
$$\widetilde{P}_i^{\alpha} = [0,\ 0,\ 1,\ 0,\ 1,\ 0,\ 1,\ 0]$$

$$\widetilde{P}_i^{\beta} = [0,\ 1,\ 1,\ 1,\ 0,\ 0,\ 0,\ 0]$$

We use the Jensen-Shannon similarity to estimate temporal dependencies

$$D_i^{\alpha\beta} = 1 - \mathrm{JSD}\left(\boldsymbol{P}_i^{\alpha}, \boldsymbol{P}_i^{\beta}\right)$$

$$= 1 - \mathrm{JSD}([0, 0, 1, 0, 1, 0, 1, 0] \;/\; 3, [0, 1, 1, 1, 0, 0, 0, 0] \;/\; 3)$$



$$\widetilde{P}_i^{\alpha} = [0, 0, 1, 0, 1, 0, 1, 0]$$

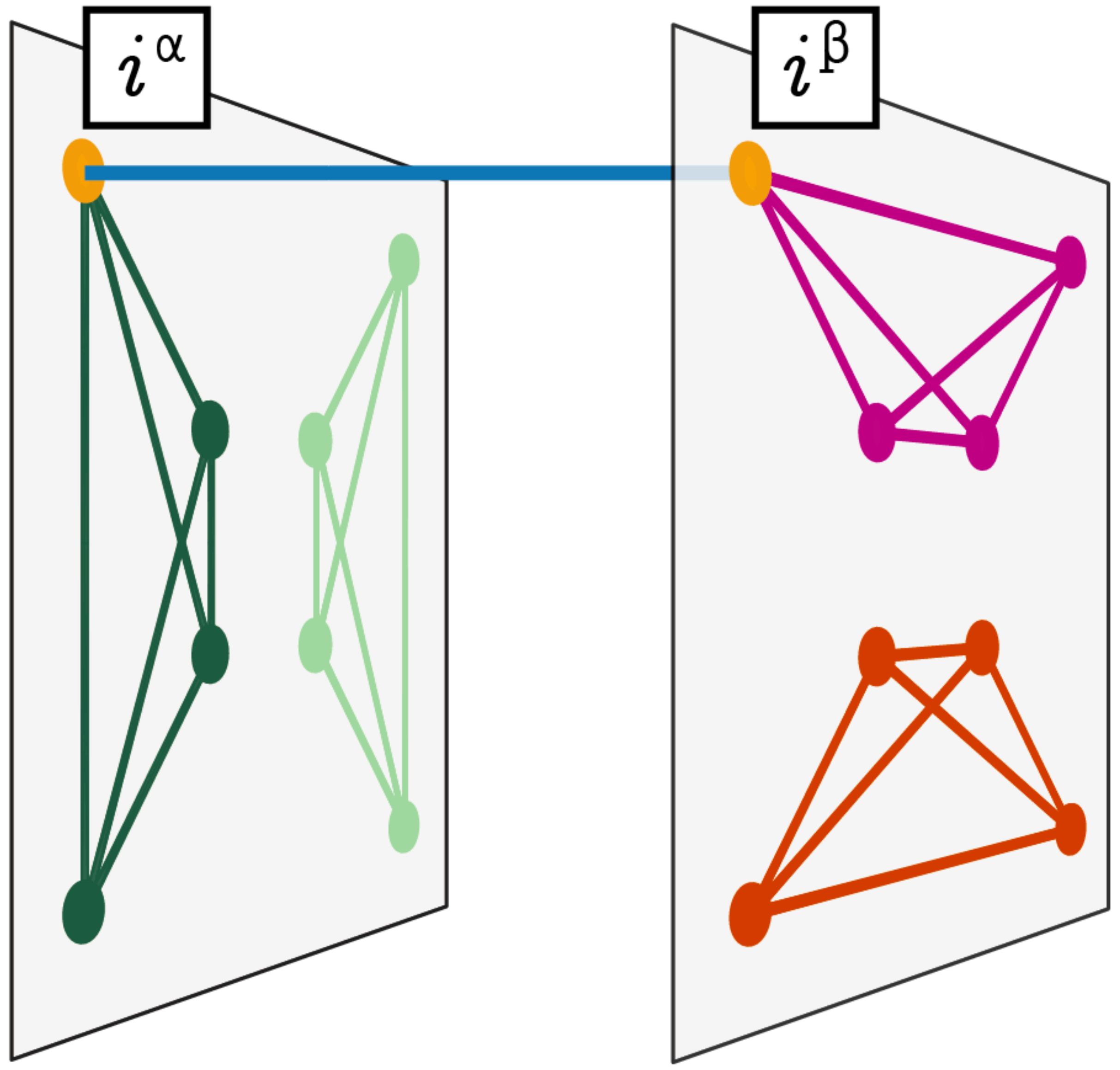
$$\widetilde{P}_i^{\beta} = [0, 1, 1, 1, 0, 0, 0, 0]$$

Meaning: Jensen-Shannon divergence measures lost information when two states are merged

$$D_i^{\alpha\beta} = 1 - \mathrm{JSD}\left(\boldsymbol{P}_i^{\alpha}, \boldsymbol{P}_i^{\beta}\right)$$

$$= 1 - \mathrm{JSD}(\\ [0, 0, 1, 0, 1, 0, 1, 0] \ / \ 3, \\ [0, 1, 1, 1, 0, 0, 0, 0] \ / \ 3 \\ )$$

$$= 0.333...$$



$$\widetilde{P}_i^{\alpha} = [0, 0, 1, 0, 1, 0, 1, 0]$$

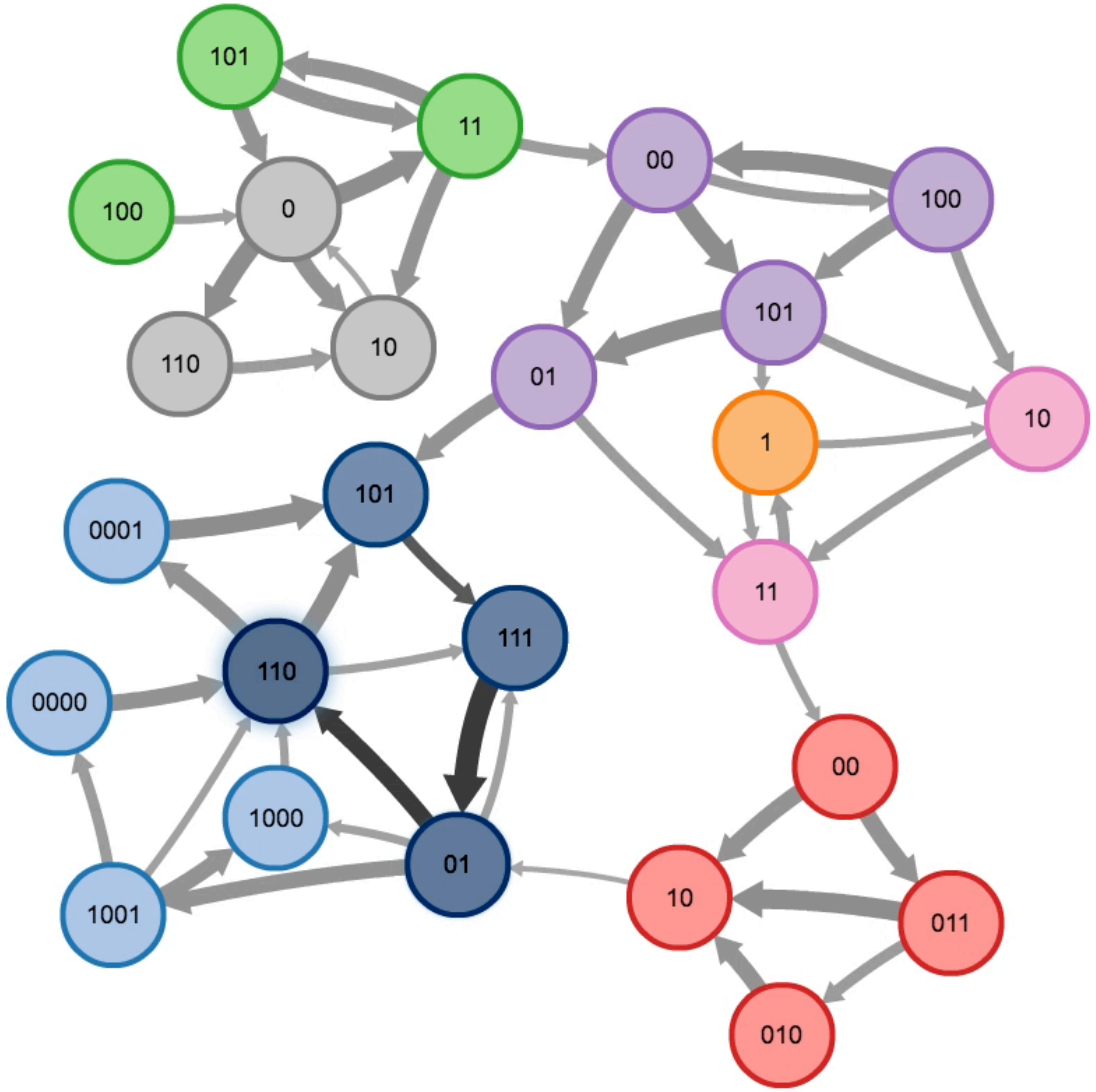
$$\widetilde{P}_i^{\beta} = [0, 1, 1, 1, 0, 0, 0, 0]$$

Putting this principle into a **flow-based** modeling framework (and Infomap)

Flow-based community detection minimizes description length of random walk on a network

your cost function

$L\,(\mathrm{M})$

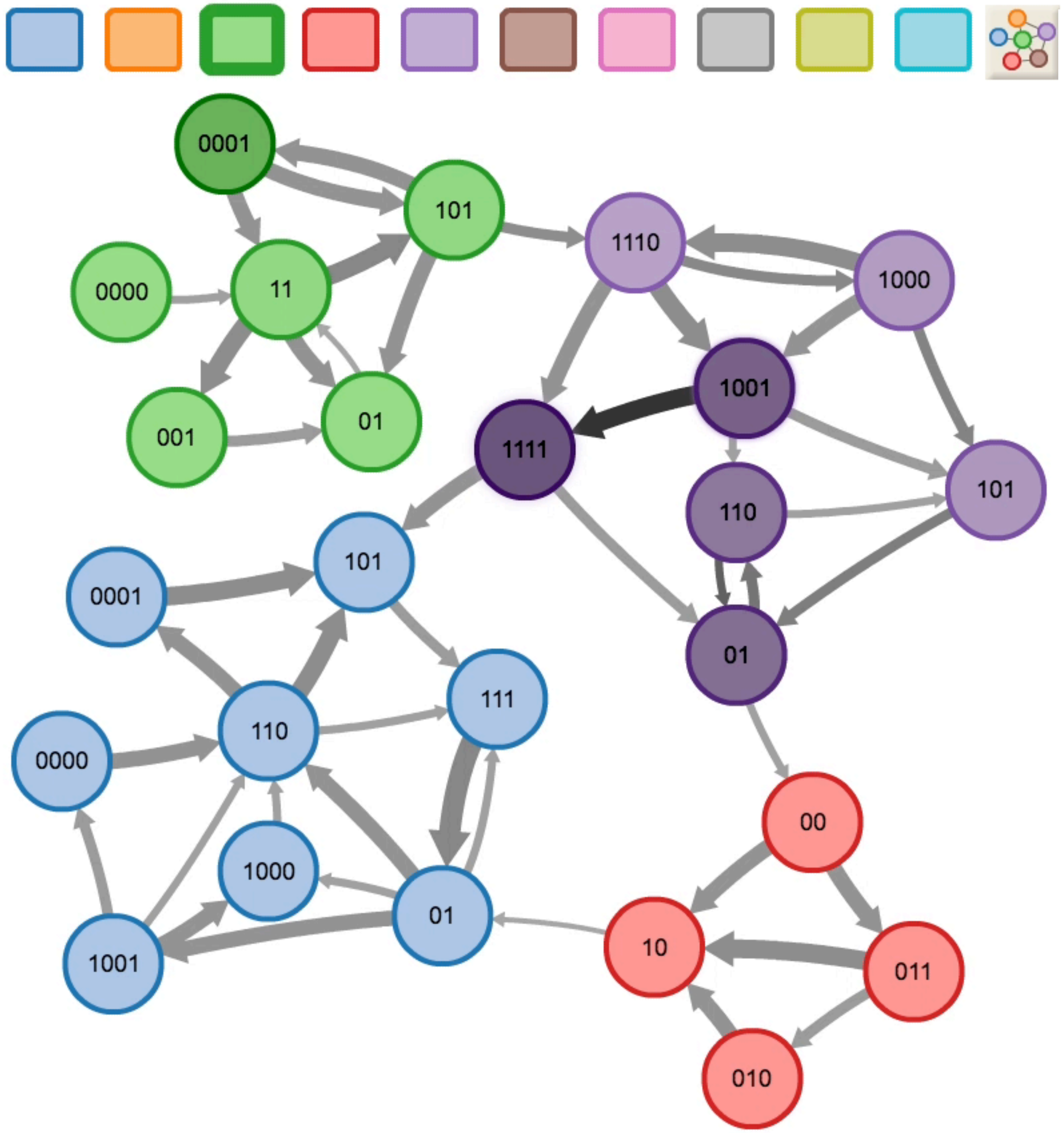


The Map Equation website (http://www.mapequation.org/)

Rosvall, M., Axelsson, D., & Bergstrom, C. T. (2009). The map equation. The European Physical Journal Special Topics, 178(1), 13-23.

Flow-based community detection minimizes description length of random walk on a network

your cost function

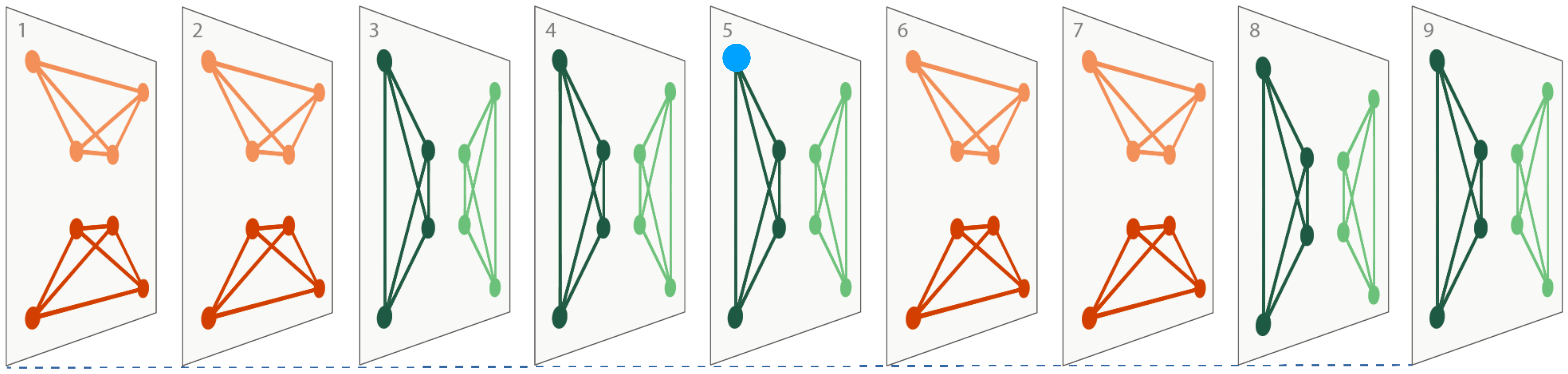


$$L(\mathsf{M}) = q_{\curvearrowleft} H(\mathcal{Q}) + \sum_{j=1}^{m} p_{j\circlearrowright} H(\mathcal{P}_j)$$

The Map Equation website (http://www.mapequation.org/)

Rosvall, M., Axelsson, D., & Bergstrom, C. T. (2009). The map equation. The European Physical Journal Special Topics, 178(1), 13-23.

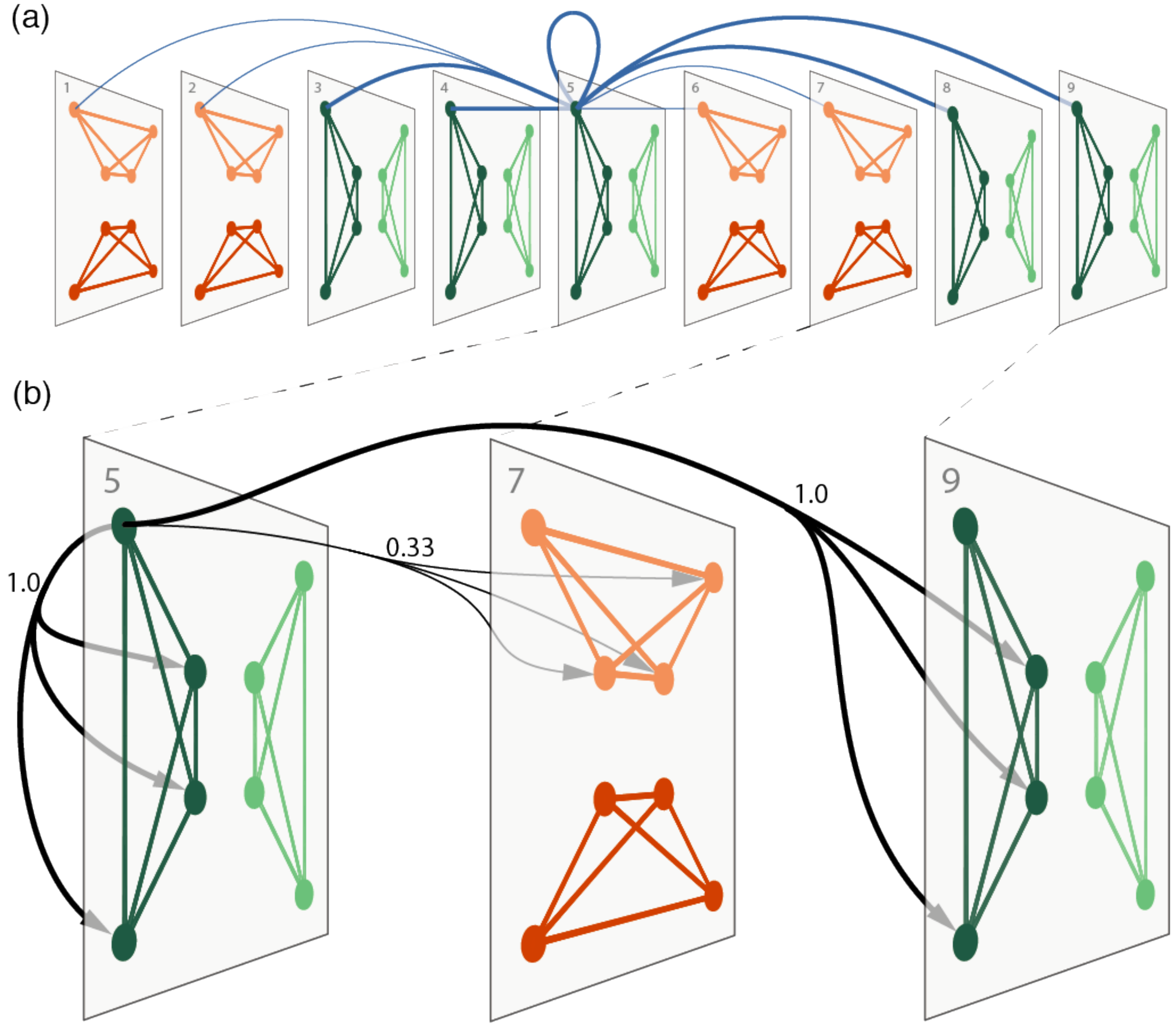
A random walker on a multilayer network changes layer with probability r



Create “new network” where weights represent random walker transition probabilities

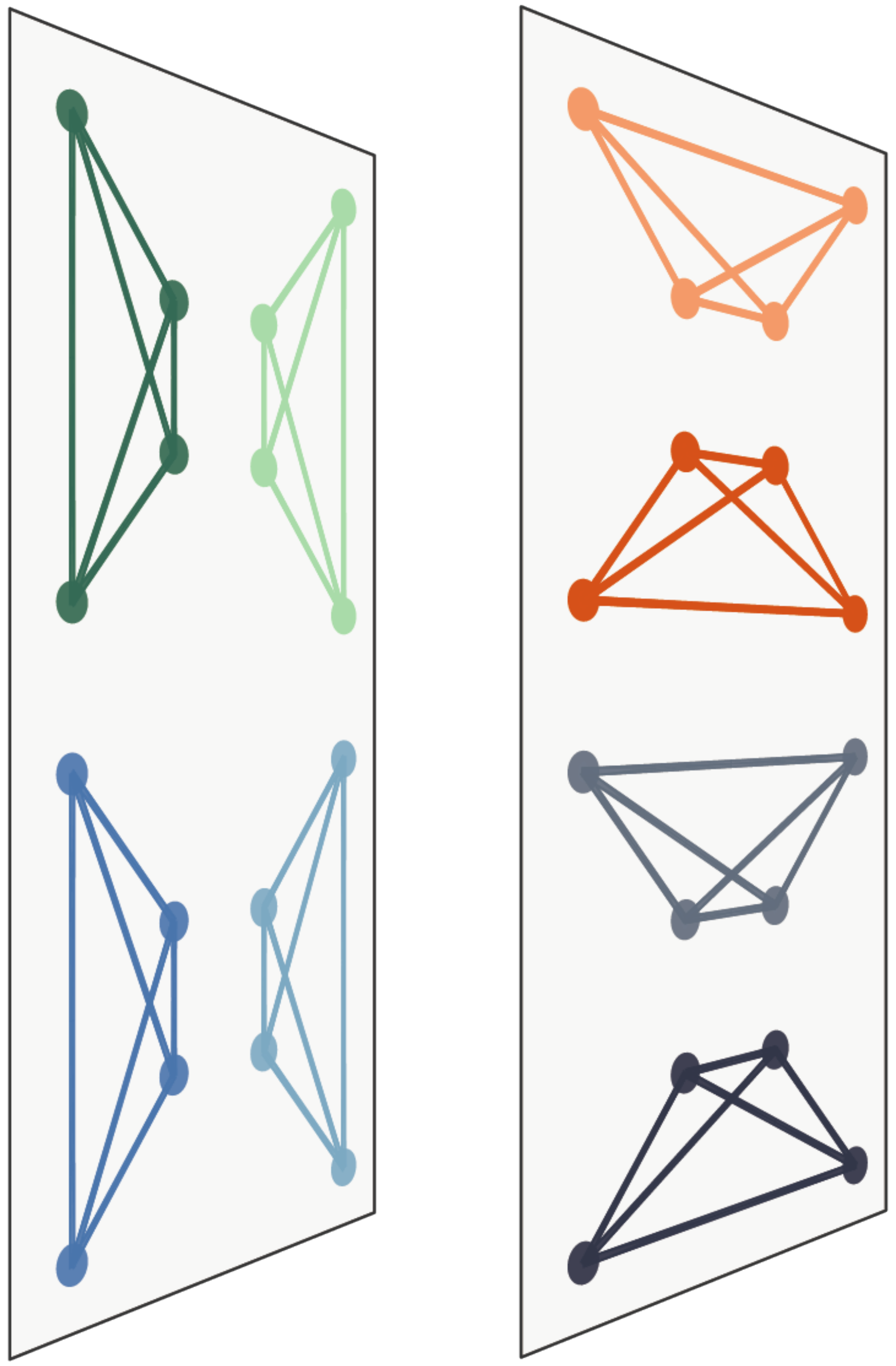
$$P_{ij}^{\alpha\beta}(r) = (1-r)\frac{W_{ij}^{\beta}}{s_i^{\beta}}\delta_{\alpha\beta} + r\frac{D_i^{\alpha\beta}}{S_i^{\alpha}}\frac{W_{ij}^{\beta}}{s_i^{\beta}}$$

(a) Use neighborhood flow coupling to get D_i
(b) Re-weight network with $P_{ij}^{\alpha\beta}$
-> Minimize the Map Equation

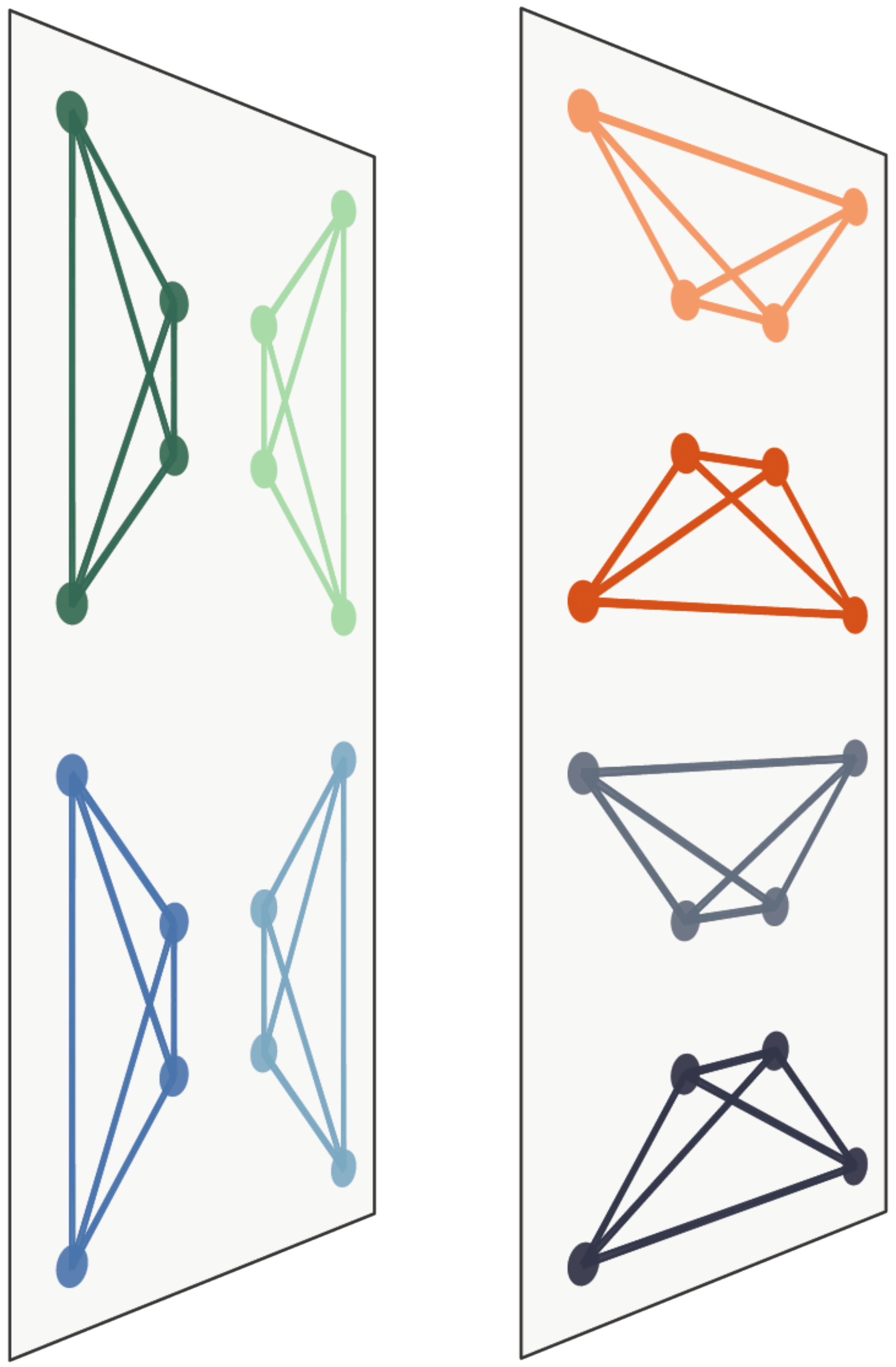


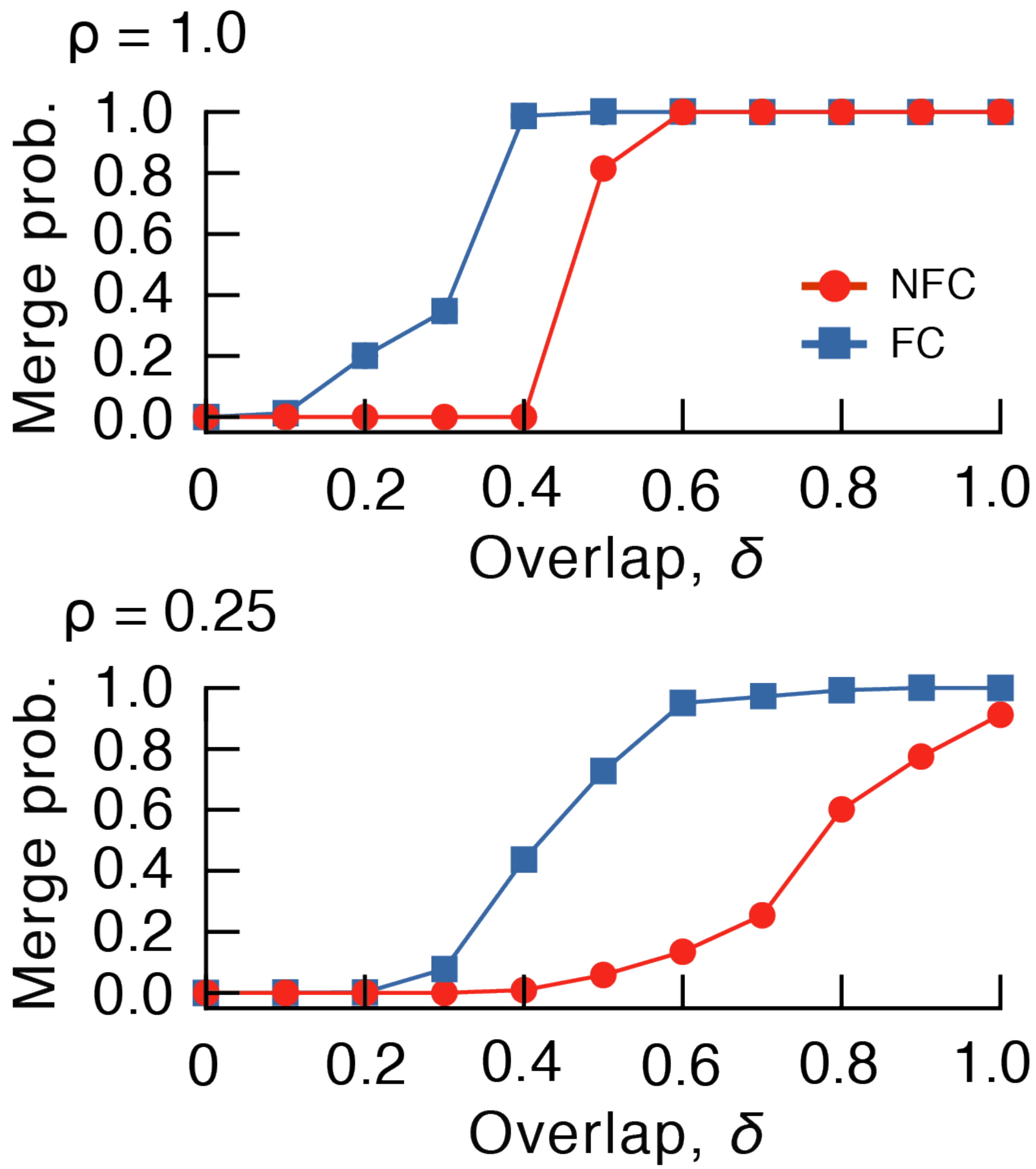
How much better is this for community detection in temporal networks?

Neighborhood flow coupling enables detection of **overlapping** communities

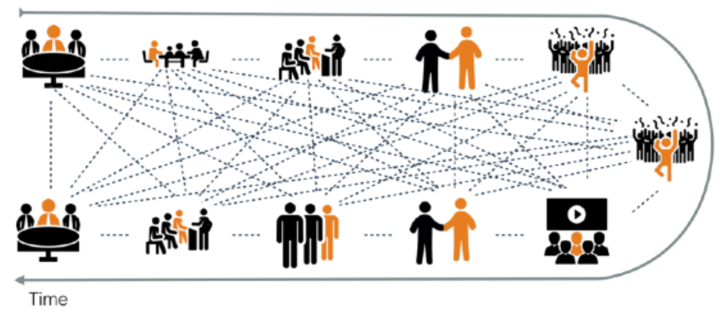


Neighborhood flow coupling enables detection of **overlapping** communities

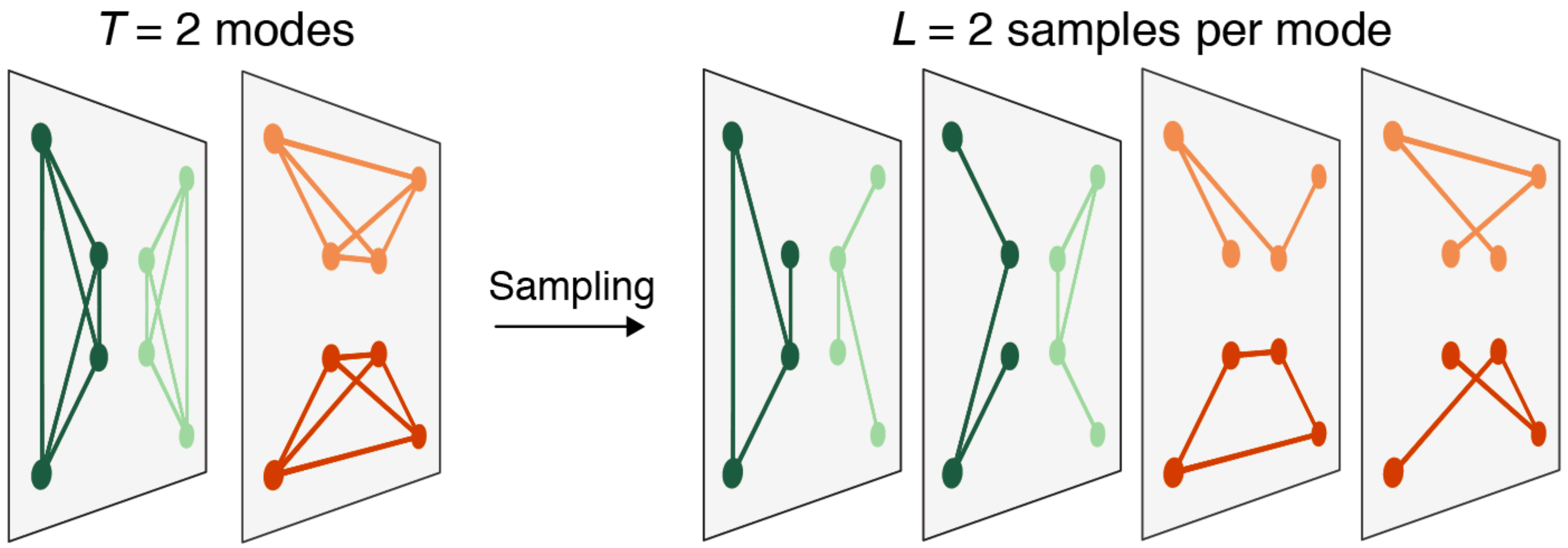




Full coupling (FC):

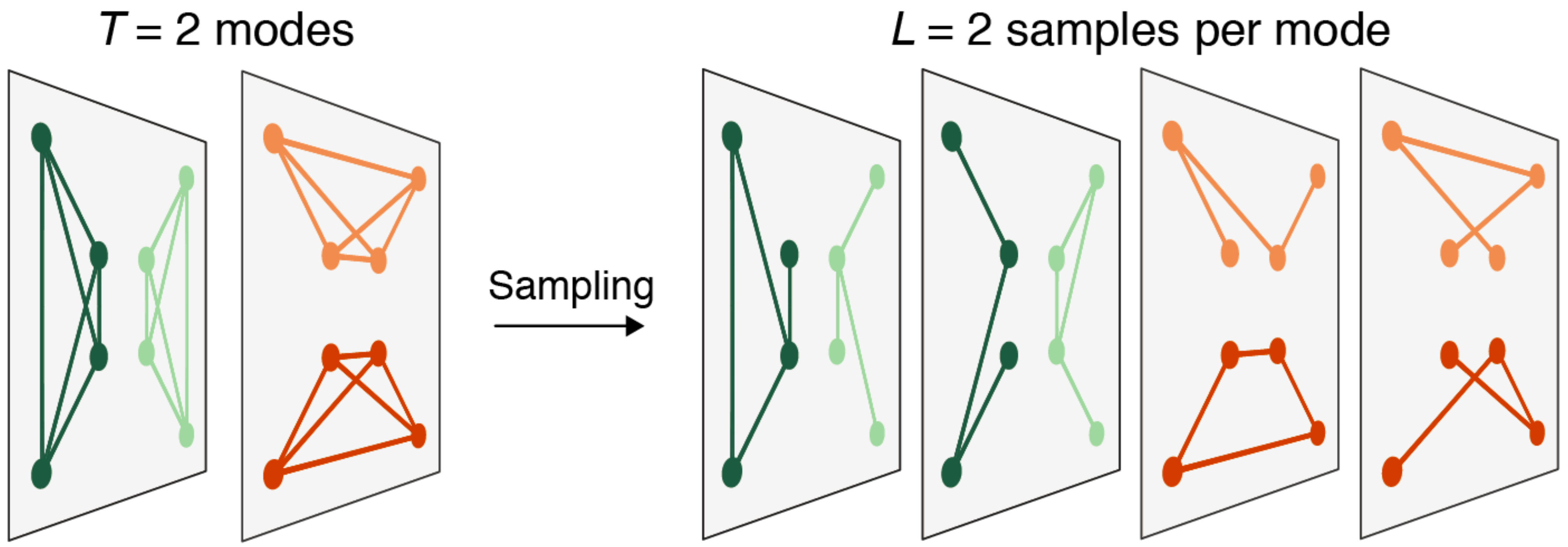


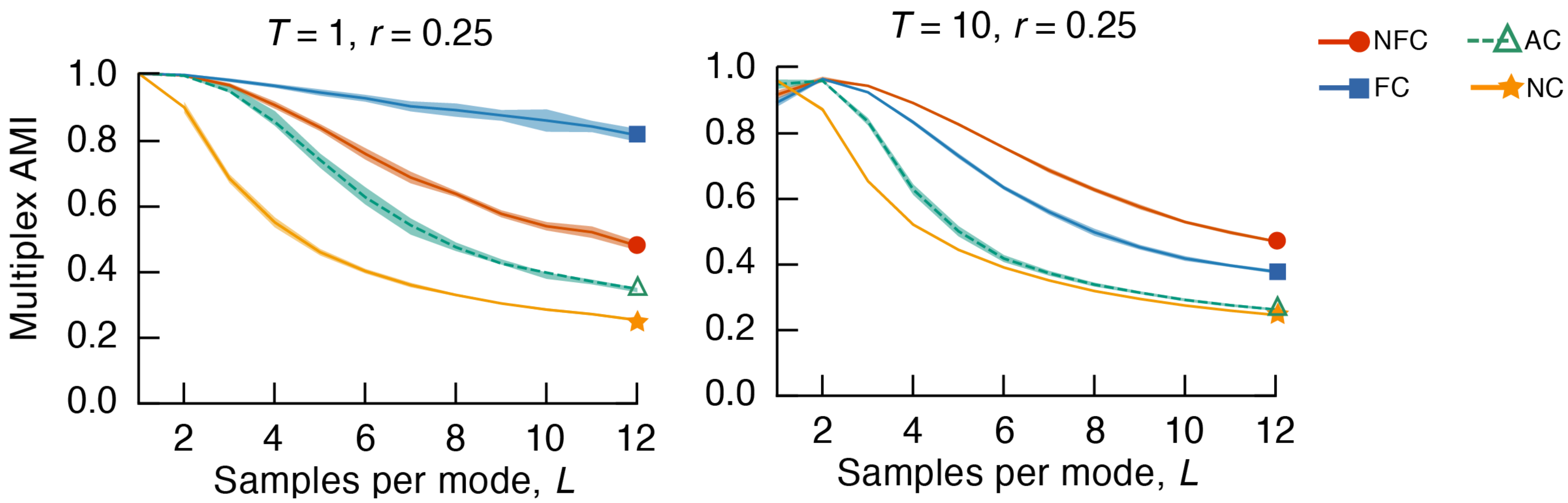
Neighborhood flow coupling enables detection of **intermittent** communities



Lancichinetti, A., Fortunato, S., & Radicchi, F. (2008). Benchmark graphs for testing community detection algorithms. Physical review E, 78(4), 046110.

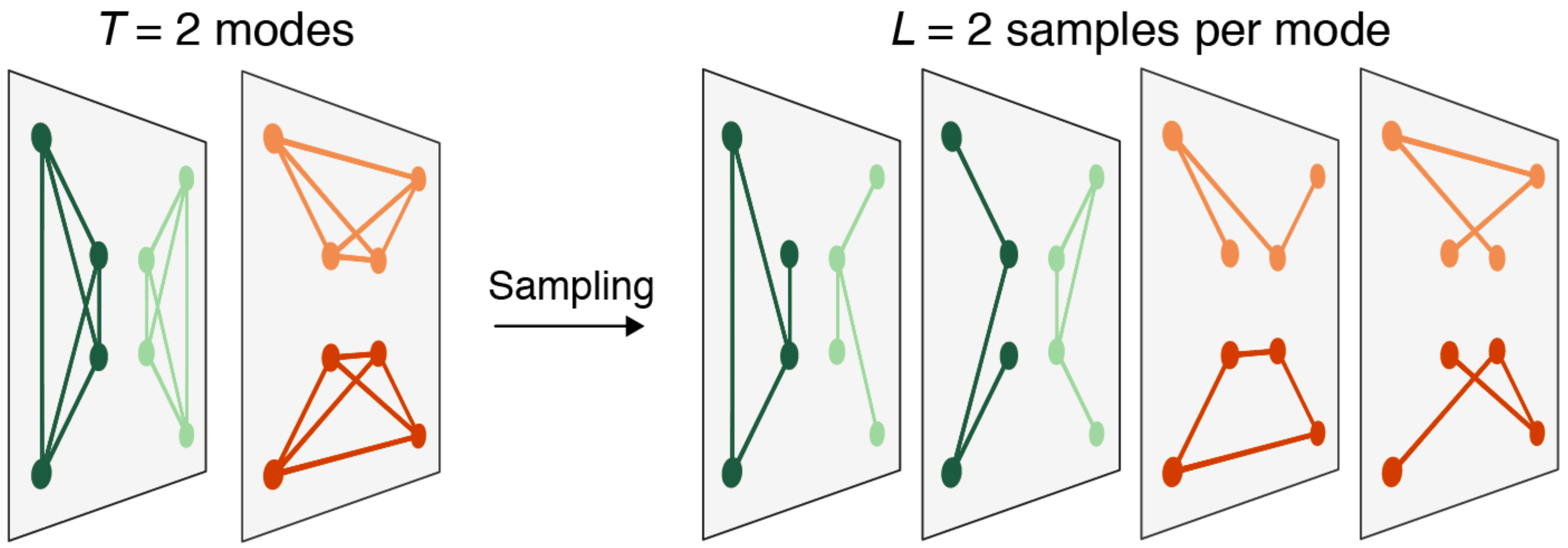
Neighborhood flow coupling enables detection of **intermittent** communities

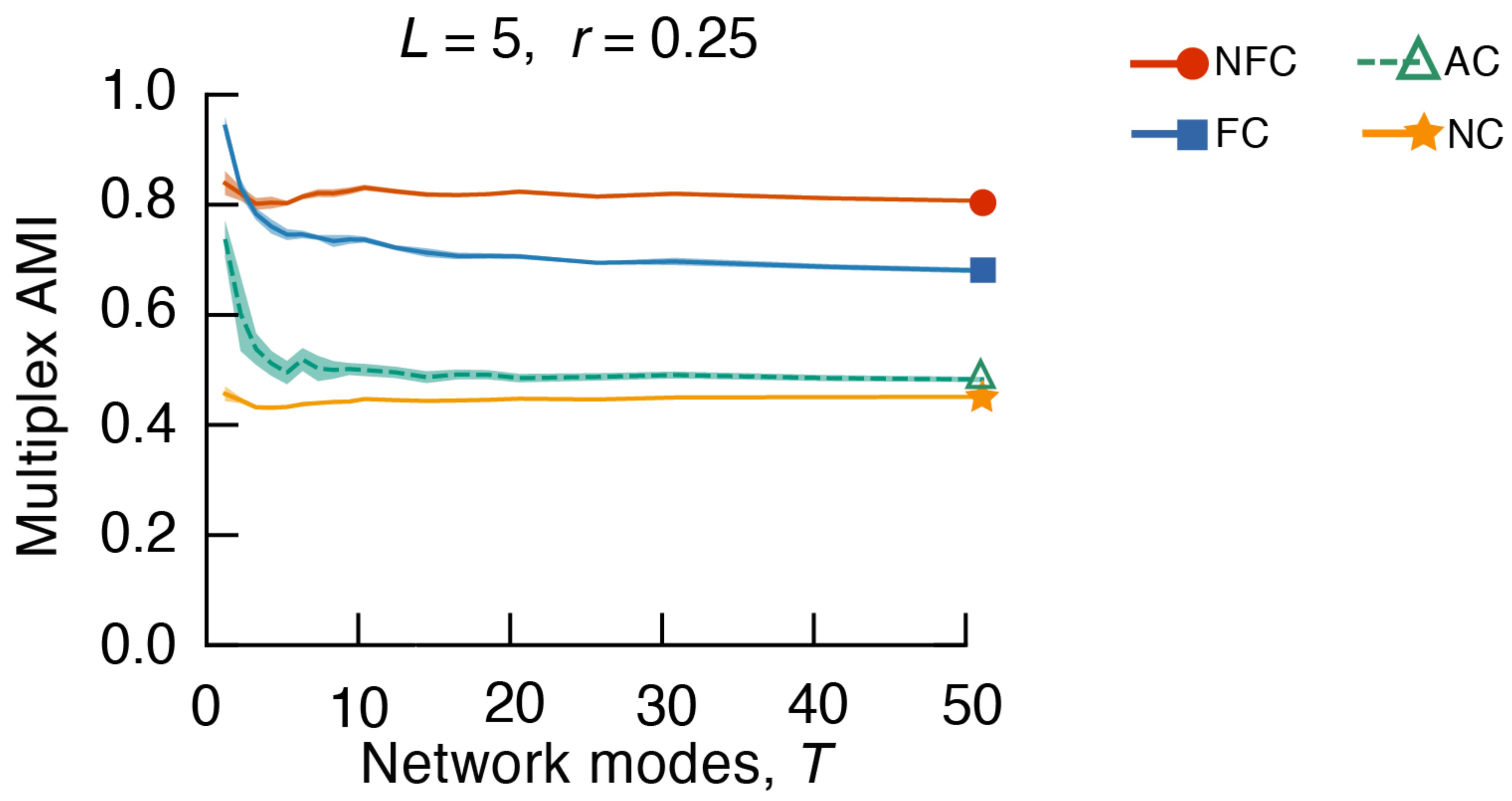




Lancichinetti, A., Fortunato, S., & Radicchi, F. (2008). Benchmark graphs for testing community detection algorithms. Physical review E, 78(4), 046110.

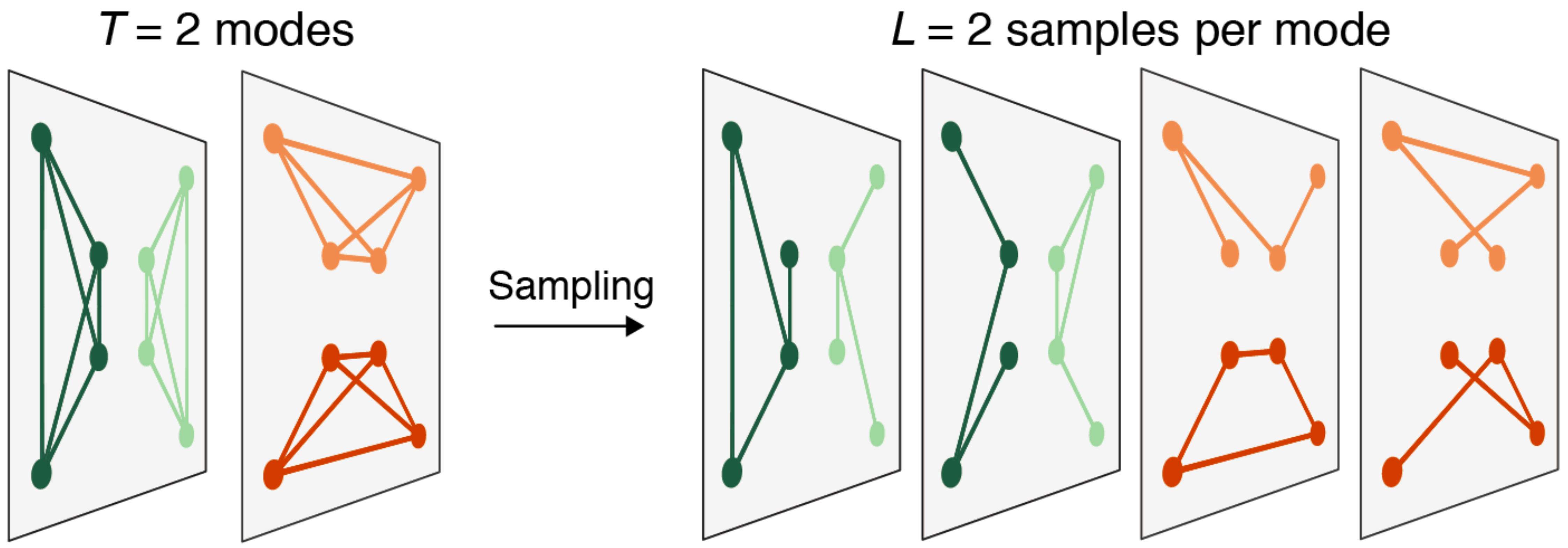
Neighborhood flow coupling enables detection of **intermittent** communities

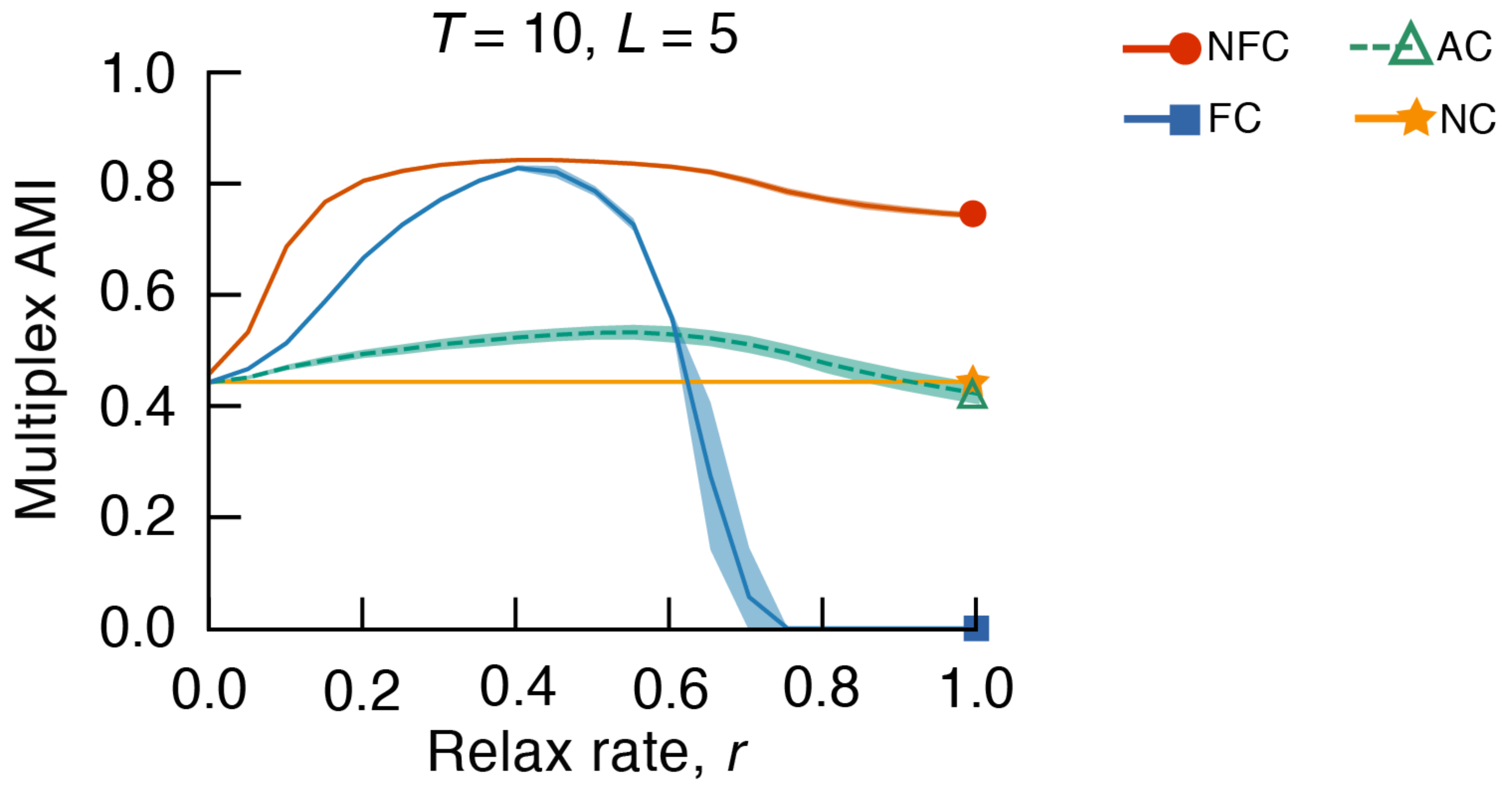




Lancichinetti, A., Fortunato, S., & Radicchi, F. (2008). Benchmark graphs for testing community detection algorithms. Physical review E, 78(4), 046110.

Neighborhood flow coupling is parameter insensitive





Lancichinetti, A., Fortunato, S., & Radicchi, F. (2008). Benchmark graphs for testing community detection algorithms. Physical review E, 78(4), 046110.

Constrained information flows in temporal networks reveal intermittent communities

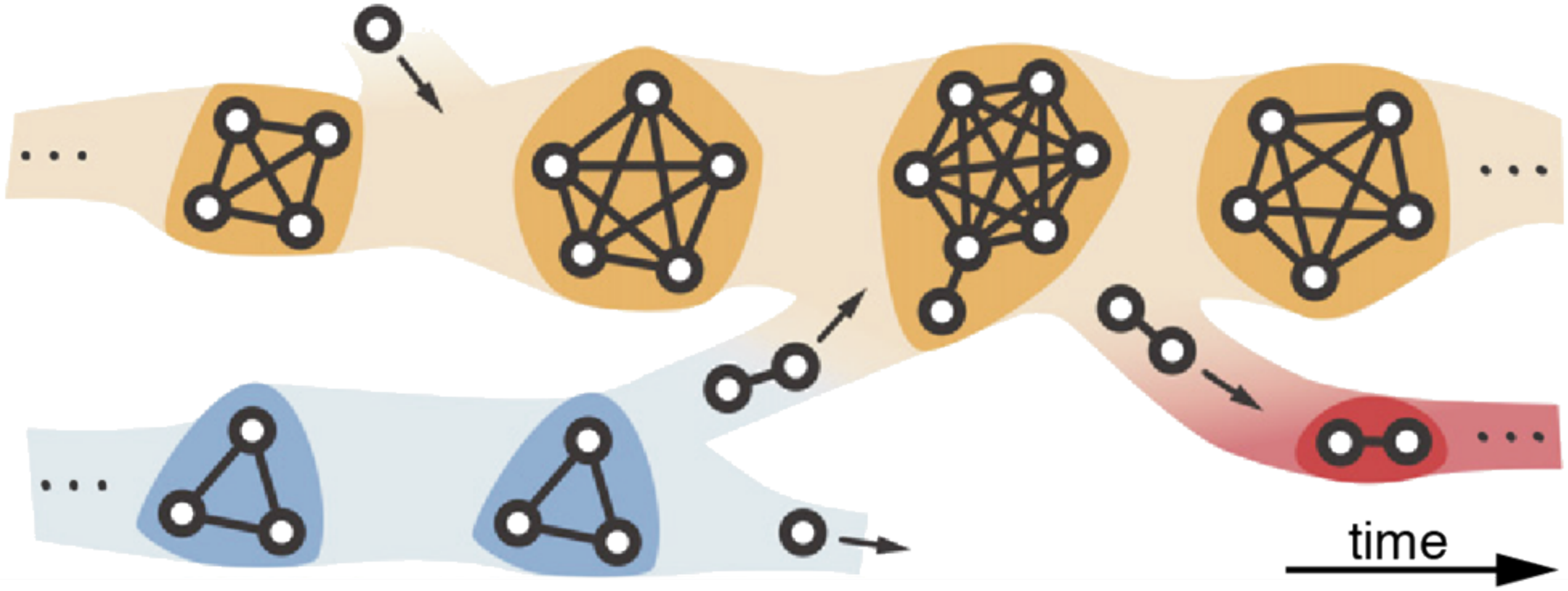
Ulf Aslak*
Centre for Social Data Science, University of Copenhagen, DK-1353 København K and DTU Compute, Technical University of Denmark, DK-2800 Kgs. Lyngby

Martin Rosvall†
Integrated Science Lab, Department of Physics, Umeå University, SE-901 87 Umeå, Sweden

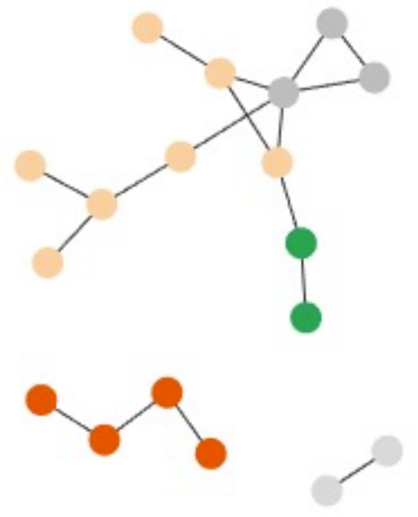
Sune Lehmann‡
*DTU Compute, Technical University of Denmark, DK-2800 Kgs. Lyngby
Niels Bohr Institute, University of Copenhagen, DK-2100 København Ø and
Department of Sociology, University of Copenhagen, DK-1353 København K*
(Dated: November 22, 2017)

Many real-world networks are representations of dynamic systems with interactions that change over time, often in uncoordinated ways and at irregular intervals. For example, university students connect in intermittent groups that repeatedly form and dissolve based on multiple factors, including their lectures, interests, and friends. Such dynamic systems can be represented as multilayer networks where each layer represents a snapshot of the temporal network. In this representation, it is crucial that the links between layers accurately capture real dependencies between those layers. Often, however, these dependencies are unknown. Therefore, current methods connect layers based on simplistic assumptions that cannot capture node-level layer dependencies. For example, connecting every node to itself in other layers with the same weight can wipe out essential depen-

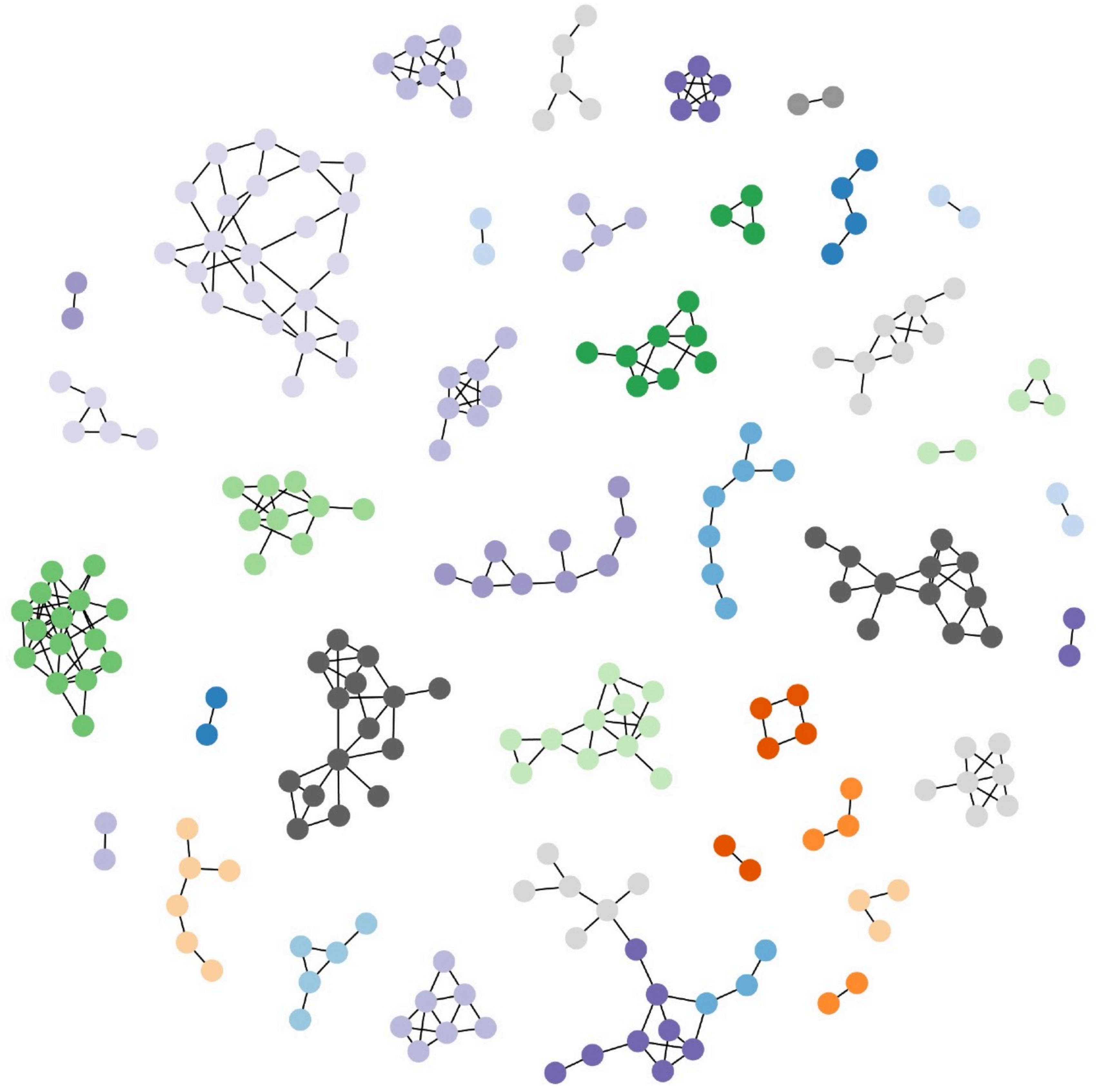
How does this method help us understand social contact networks?



We use two fundamentally different temporal contact networks



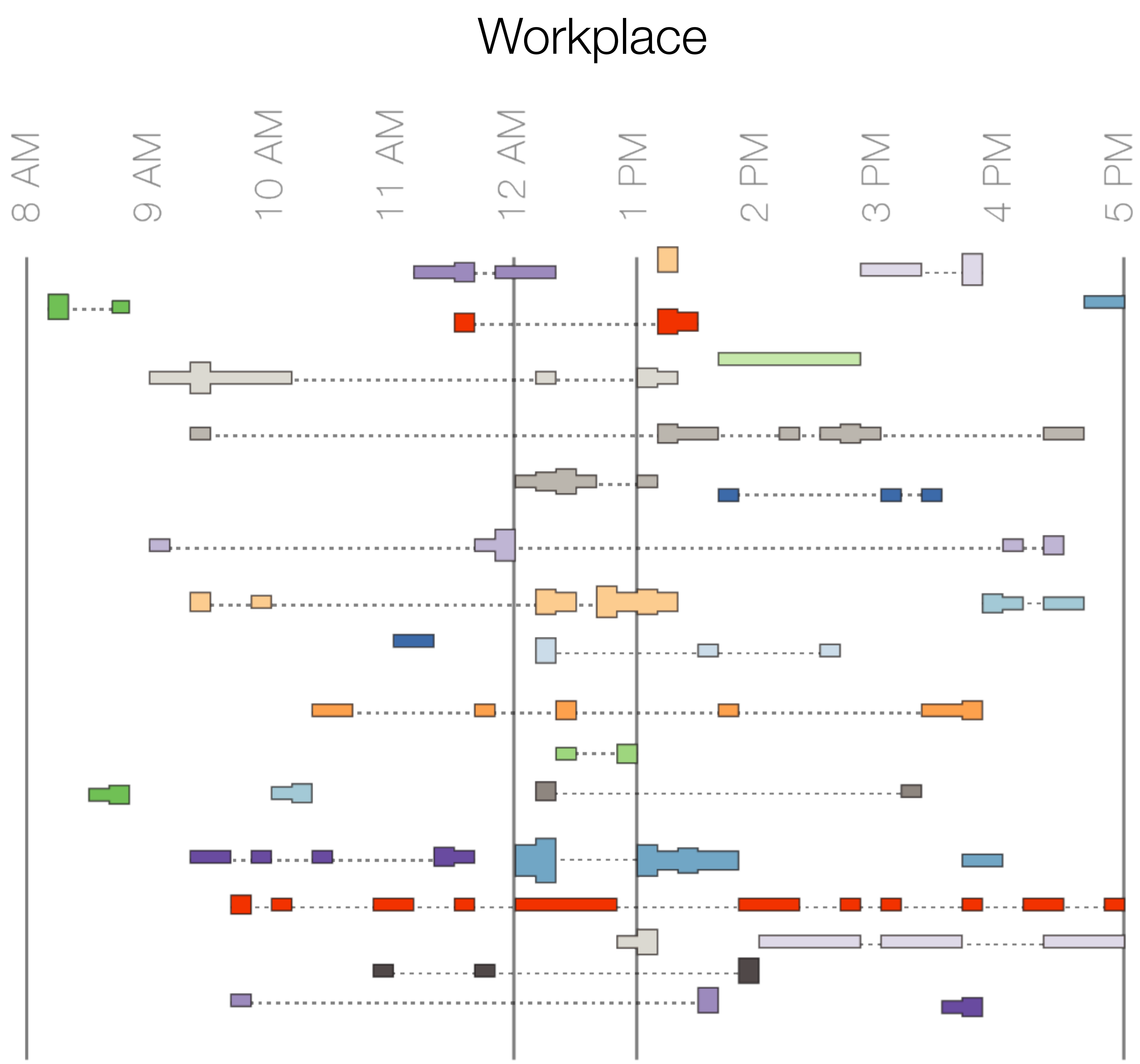
Workplace (N = 145)

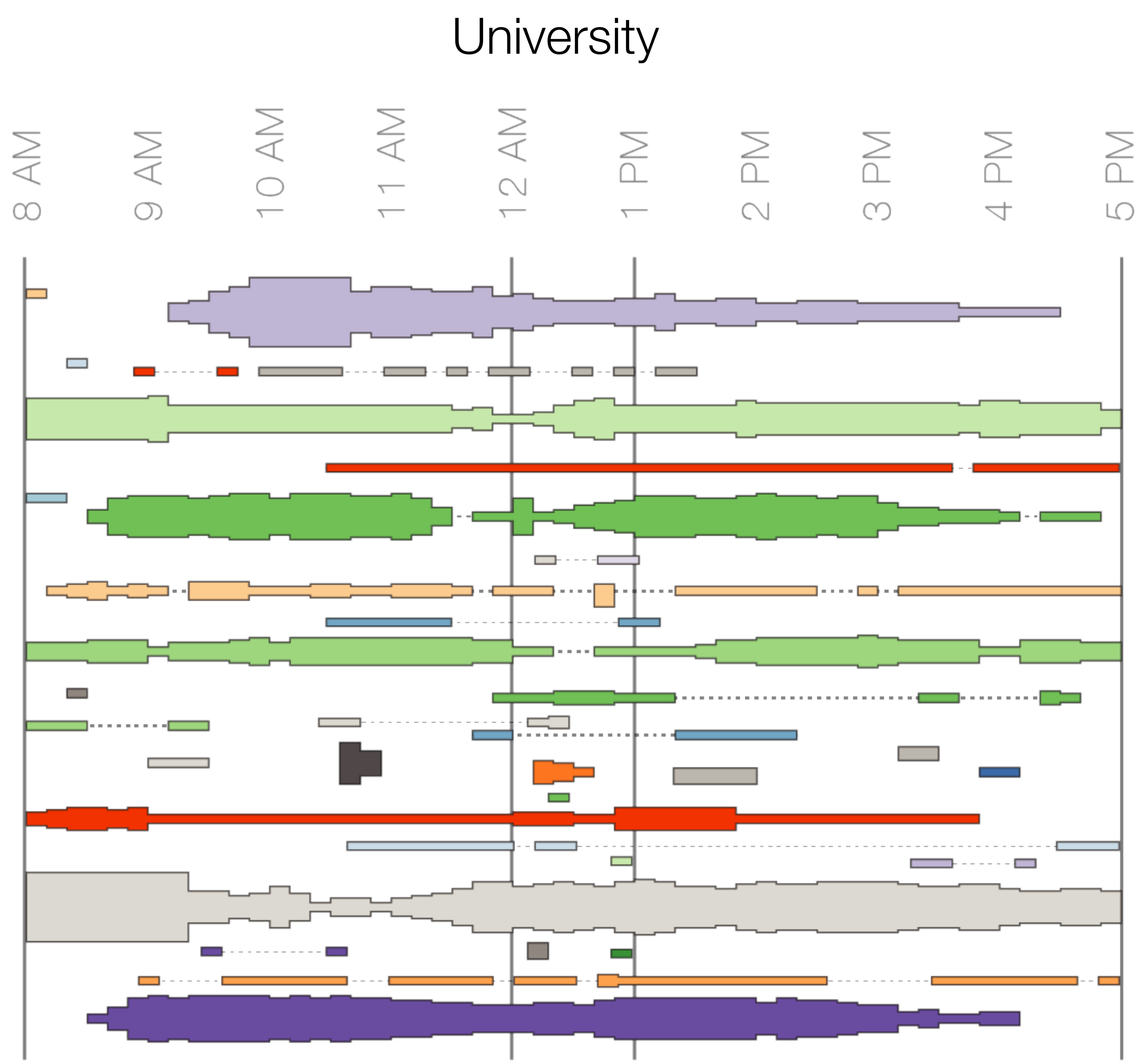


University (N = 665)

Génois, M., Vestergaard, C. L., Fournet, J., Panisson, A., Bonmarin, I., & Barrat, A. (2015). Data on face-to-face contacts in an office building… *Network Science*, *3*(3), 326-347.
Stopczynski, A., Sekara, V., Sapiezynski, P., Cuttone, A., Madsen, M. M., Larsen, J. E., & Lehmann, S. (2014). Measuring large-scale social networks with high resolution. PloS one, 9(4), e95978.

Visualization gives us an intuitive understanding of how our method behaves





- **We introduce *neighborhood flow coupling*: a principle for inferring intertemporal dependencies in networks**
- We provide an open source implementation
- NFC makes flow based community detection less parameter sensitive
- **NFC highlights interesting structure in social networks!**

fin