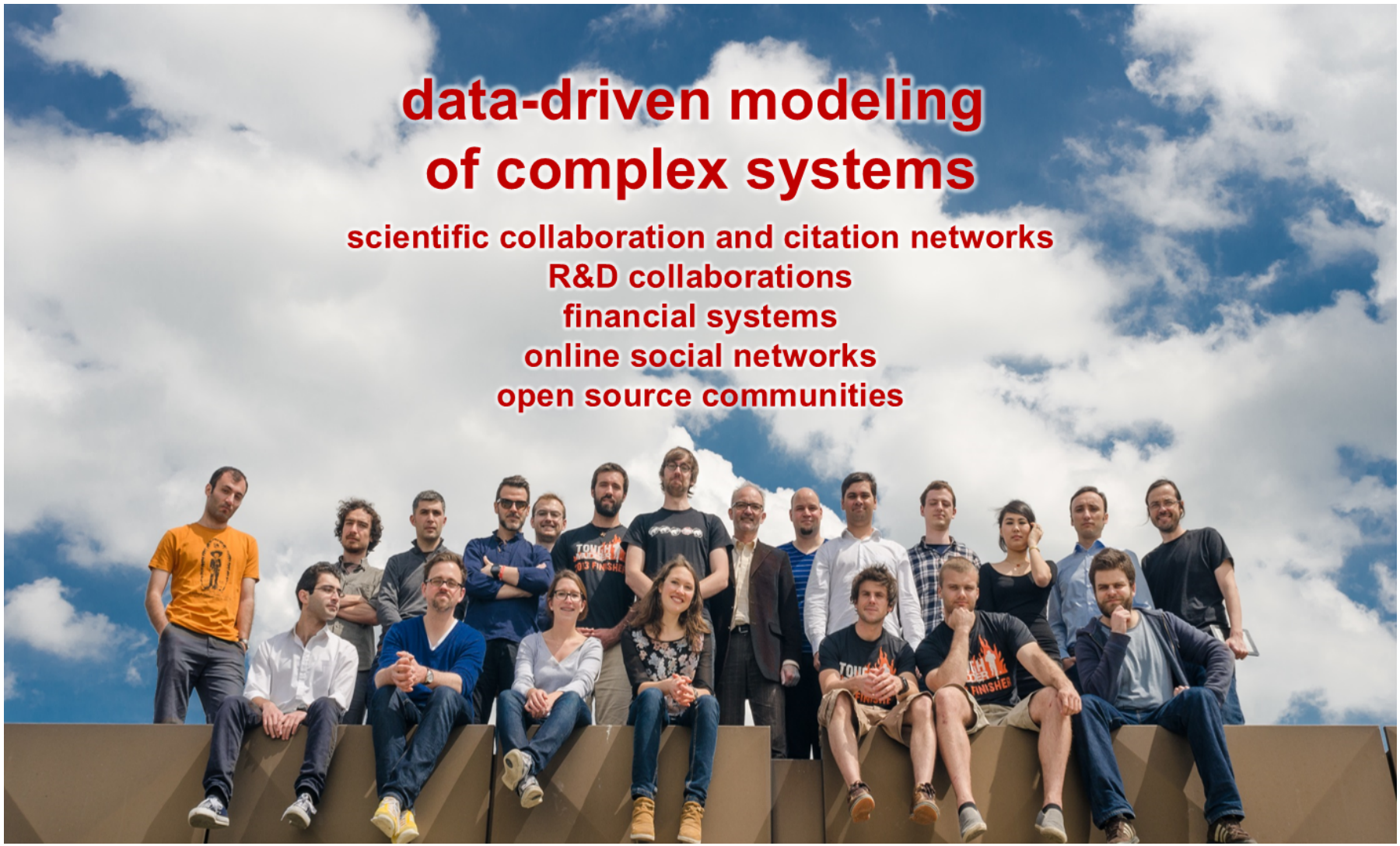
ETH zürich

Analyzing time-stamped network data: When ordering matters

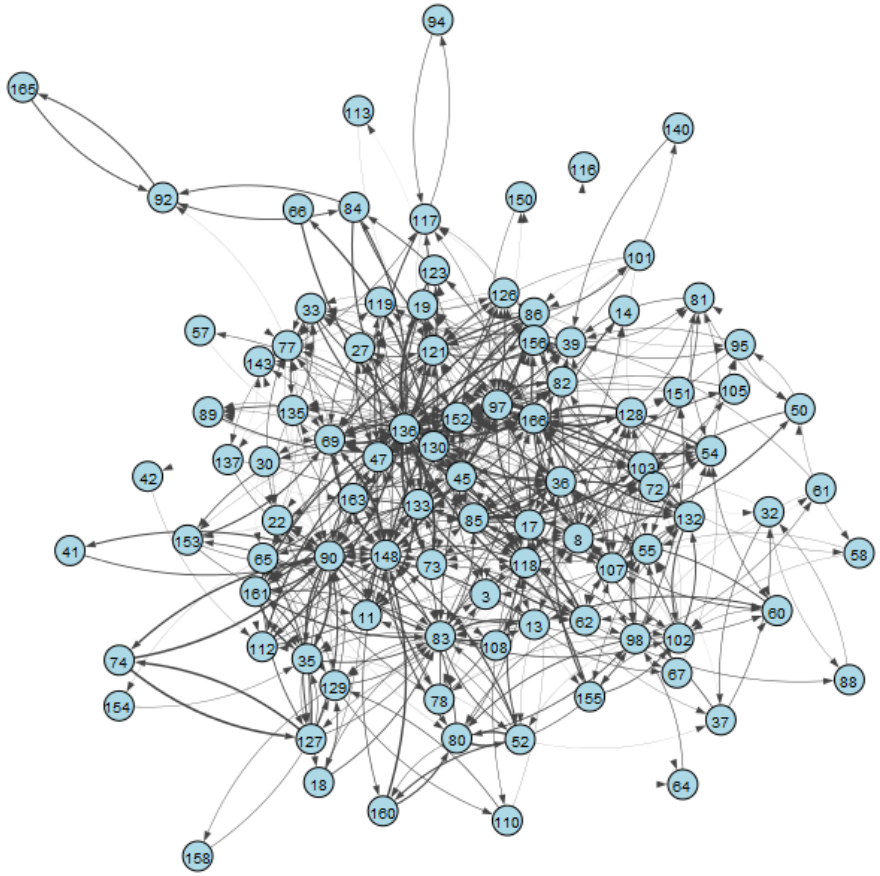
Ingo Scholtes
Chair of Systems Design, ETH Zürich
with N. Wider, R. Pfitzner, A. Garas, F. Schweitzer

Chair of Systems Design at ETH Zürich



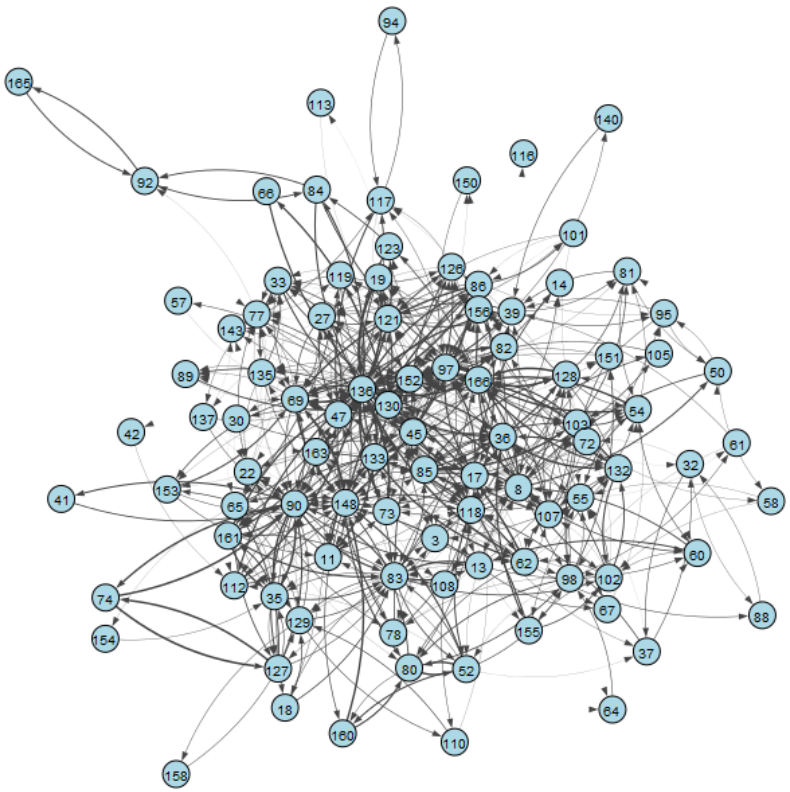
Network perspective on time-stamped data

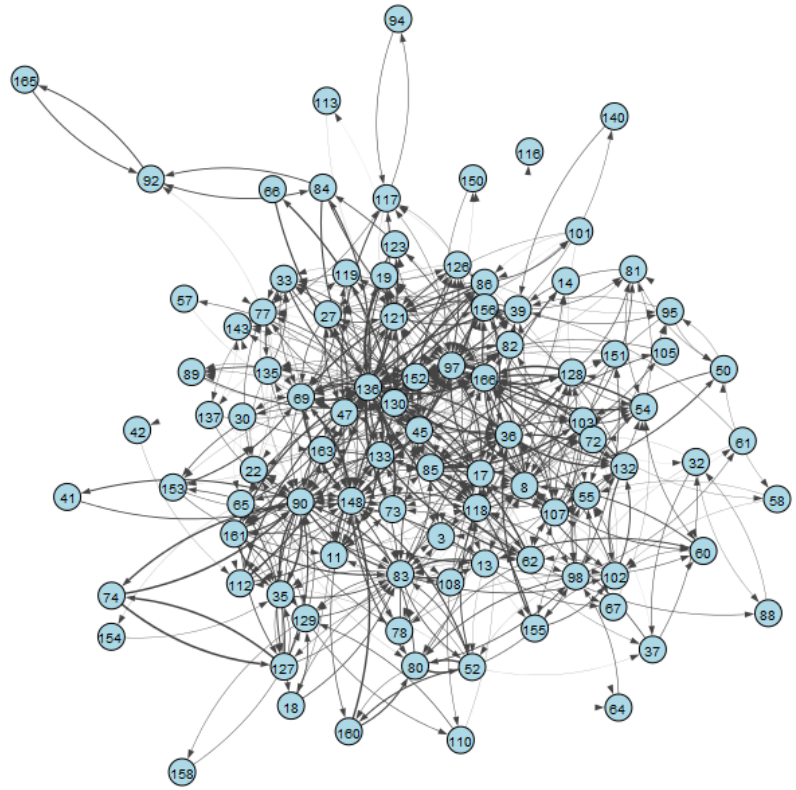
- **network perspective** on complex systems
  - provides wealth of methods
- problem: **over-simplification**
  - multi-layer structures
  - nodes with different roles
  - ...
- what about **dynamic complex systems**?
- **what do we lose** if we aggregate time-stamped links?



example: E-Mail communication network in manufacturing company

Temporal networks: two perspectives





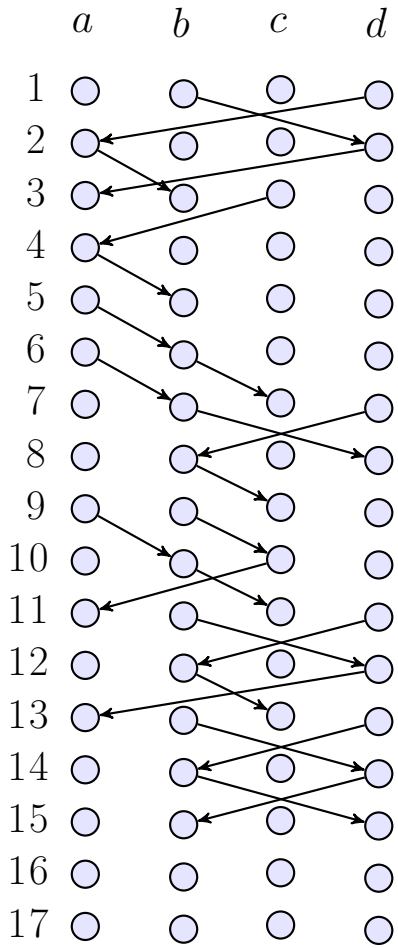
macro perspective: focus on **distribution of activities** in time

micro perspective: focus on **ordering of links** in time

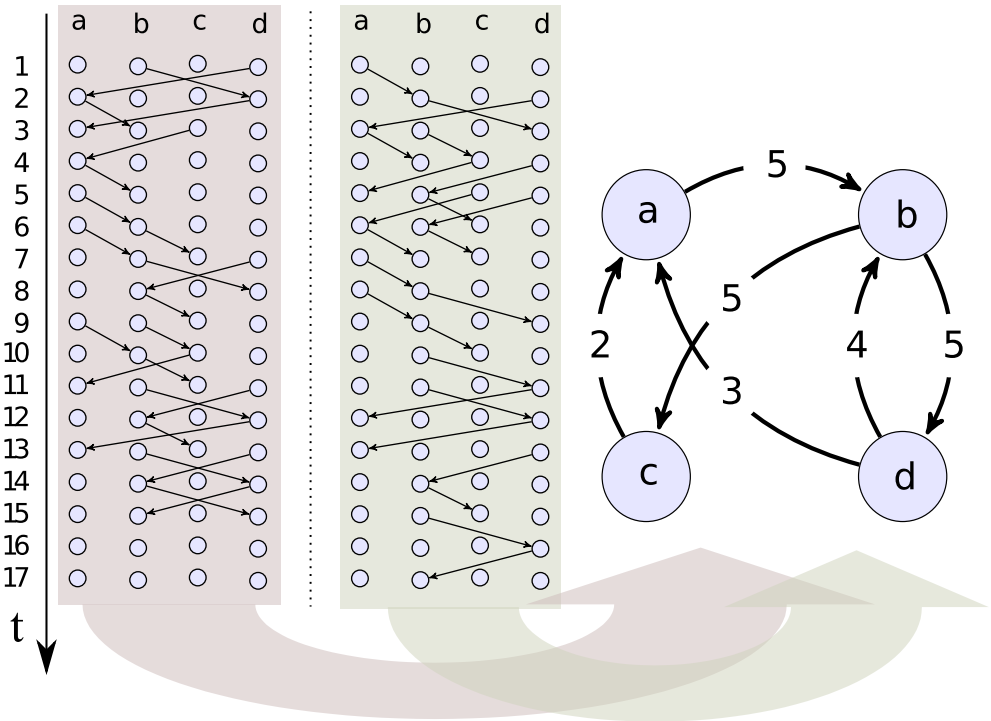
How does **ordering of links** affect network analysis?

The micro perspective: ordering of links

- we focus on the **microscopic ordering of links** and its effect on **causality** in time-varying complex systems
- we consider temporal networks with **time-stamped links** $(i, j; t)$ for discrete times t
- temporal network can be represented as **time-unfolded network**
- example: time-unfolded representation of temporal network with 4 nodes and 24 time-stamped links



Time-aggregated network



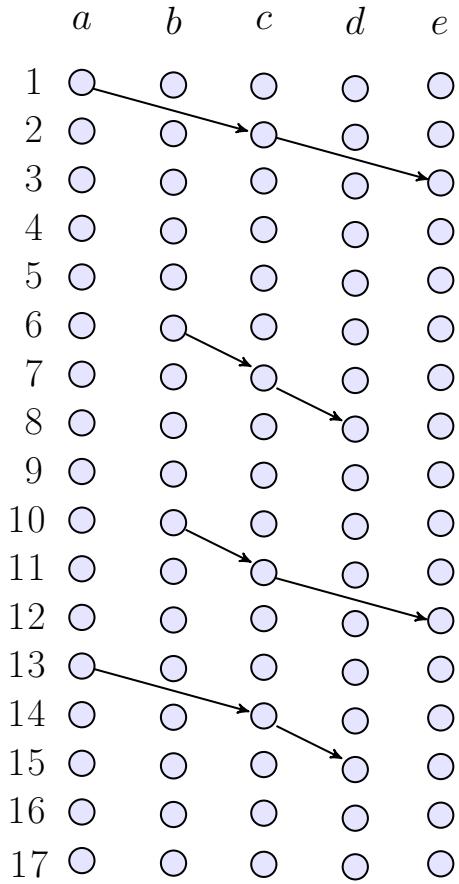
two temporal networks (left) and their corresponding time-aggregated network

- **time-aggregated network** is static representation of temporal network
- we assume that **link weights** count the number of times links have been active
- clear: different **temporal networks** are consistent with same **time-aggregated network**

  $\Rightarrow$ we **lose information on the ordering of links**

Causality: time-respecting paths

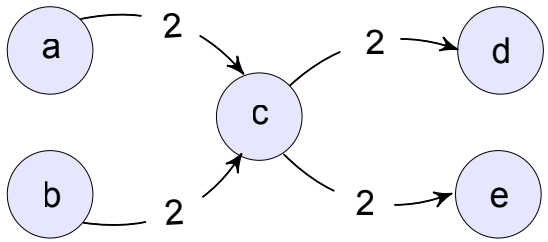
- $v_0 \to v_1 \to \ldots \to v_l$ is a **time-respecting path** if $(v_i, v_{i+1}; t_i)$ are time-stamped links which satisfy **causality**, i.e. $t_i < t_{i+1}$
- we can impose **maximum waiting time** δ, i.e. we require $t_{i+1} - t_i \leq \delta$
- example: $\delta = 1 \Rightarrow$ **four** time-respecting paths
  1. $a \to c \to e$
  2. $b \to c \to d$
  3. $b \to c \to e$
  4. $a \to c \to d$



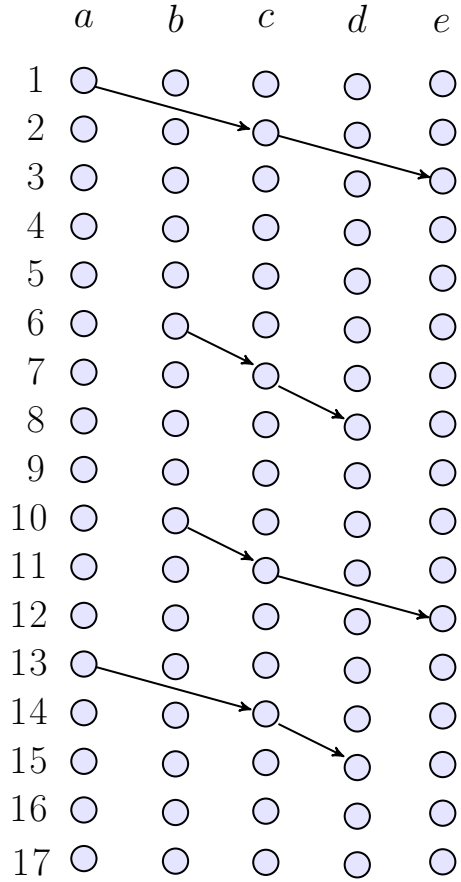
example for temporal network with four time-respecting paths

Expected paths in Markovian temporal networks

- what do we **expect** based on the **time-aggregated network**?



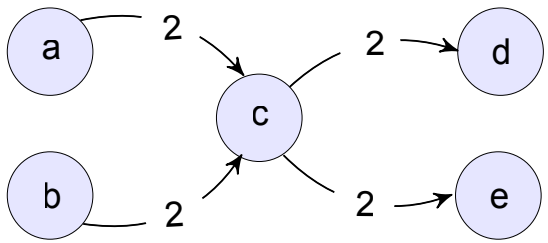
- assumption: consecutive links independent, i.e. temporal network **memoryless** or **Markovian**
- **transitivity**: $u \to v \wedge v \to w \Rightarrow u \to v \to w$
- we expect **four** time-respecting paths
  1. $a \to c \to e$
  2. $b \to c \to d$
  3. $b \to c \to e$
  4. $a \to c \to d$



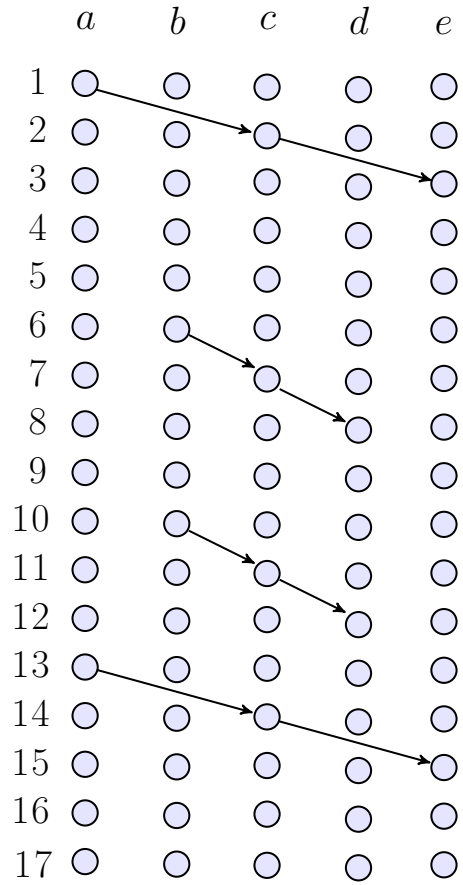
temporal network consistent with time-aggregated network

Order of links can break transitivity!

- we **change the order** of links: swap time-stamps of $(c, e; 11)$ and $(c, d; 14)$
- time-aggregated network remains the same



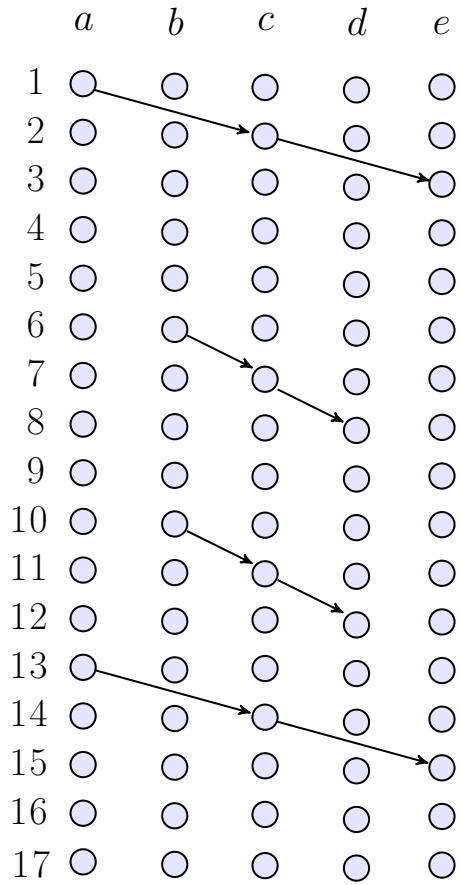
- example: **two** time-respecting paths
  1. $a \to c \to e$
  2. $b \to c \to d$
- **ordering of links can break transitivity** expected from time-aggregated network



ordering of interactions can invalidate transitivity of aggregate network

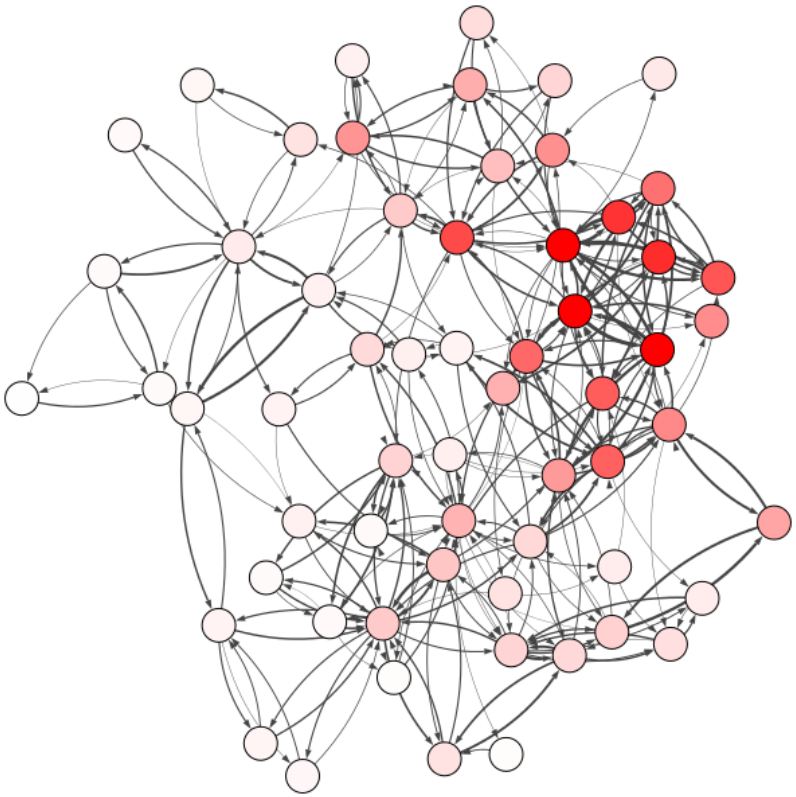
Non-Markovian Temporal Networks

- idea: calculate statistics of time-respecting paths $u \to v \to w$
  1. **Markovian** temporal network
     $P(u \to v \to w) = P(u \to v) \cdot P(v \to w)$
  2. **Non-Markovian** temporal network
     $P(u \to v \to w) \neq P(u \to v) \cdot P(v \to w)$
- different ways to **detect need for higher-order models**
  1. betweenness preference: PRL 110, 198701 (May 2013)
  2. entropy growth rate: Nature Comm. 5:5024 (Sept. 2014)
- observation: **Markovian assumption rarely holds in reality**

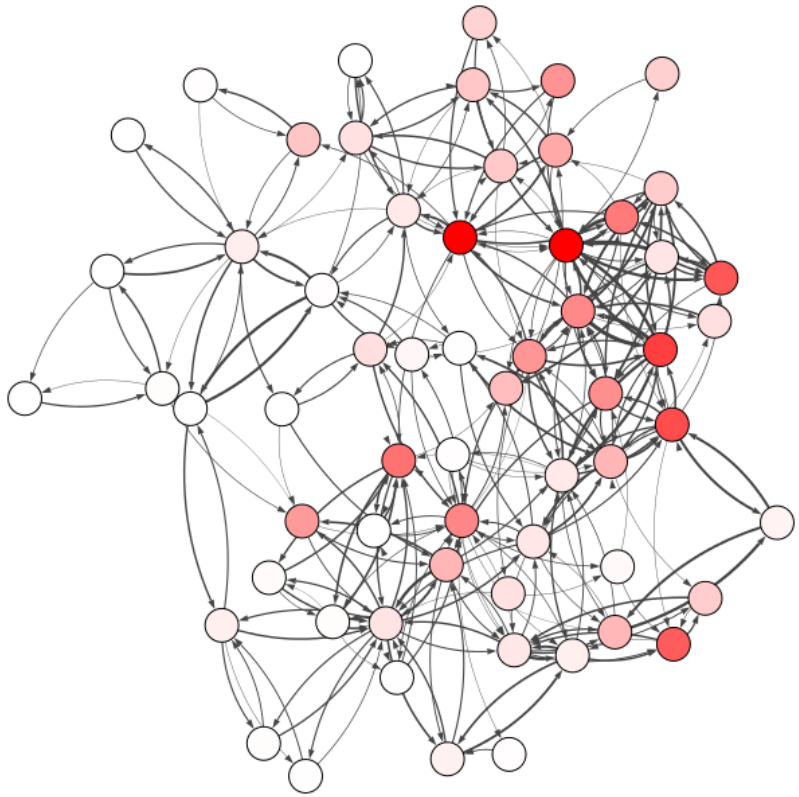


toy example for a non-Markovian temporal network

Non-Markovianity affects diffusion: ant colony



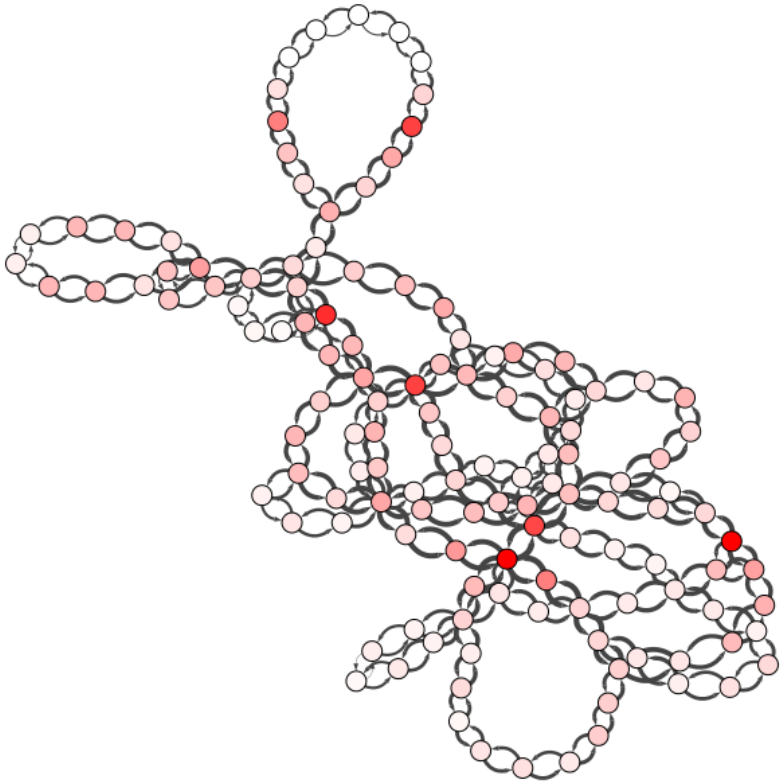
actual diffusion dynamics in
non-Markovian temporal network



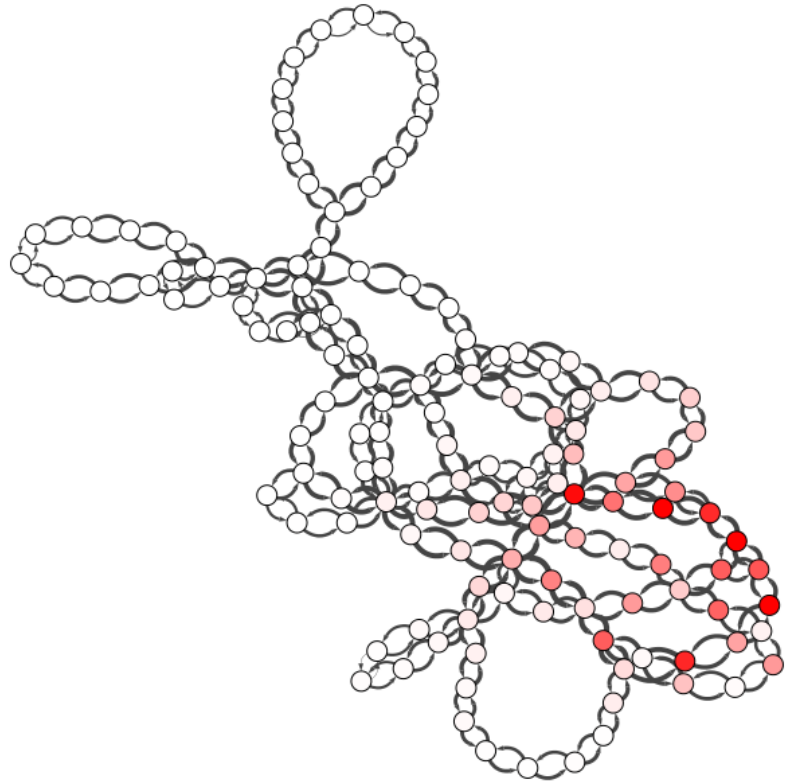
expected diffusion dynamics in
Markovian temporal network

non-Markovianity **slows down** dynamical process

Non-Markovianity affects diffusion: London Tube



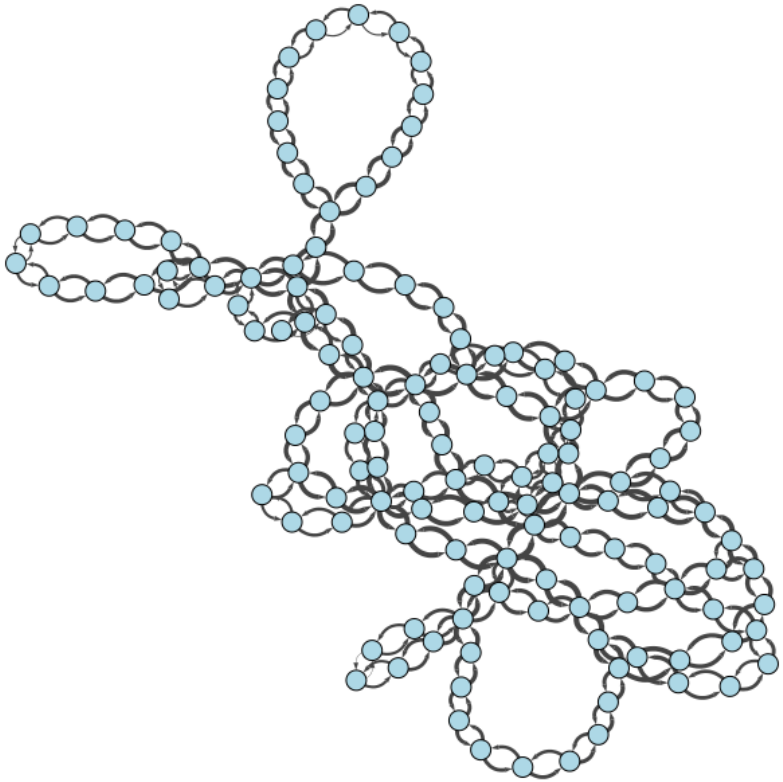
actual diffusion dynamics in
non-Markovian temporal network



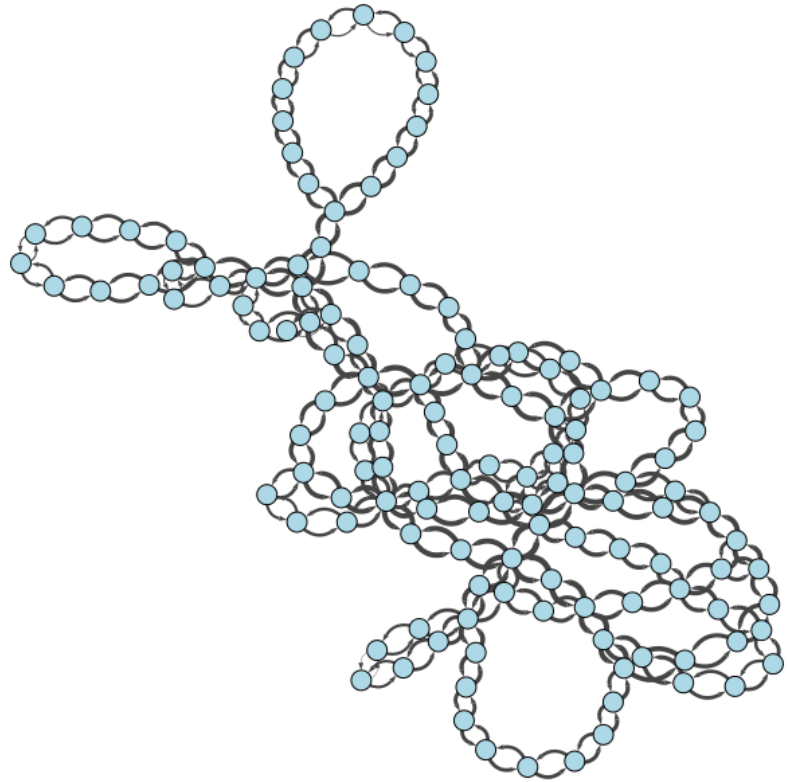
expected diffusion dynamics in
Markovian temporal network

non-Markovianity **speeds up** dynamical process

non-Markovian vs. Markovian temporal network

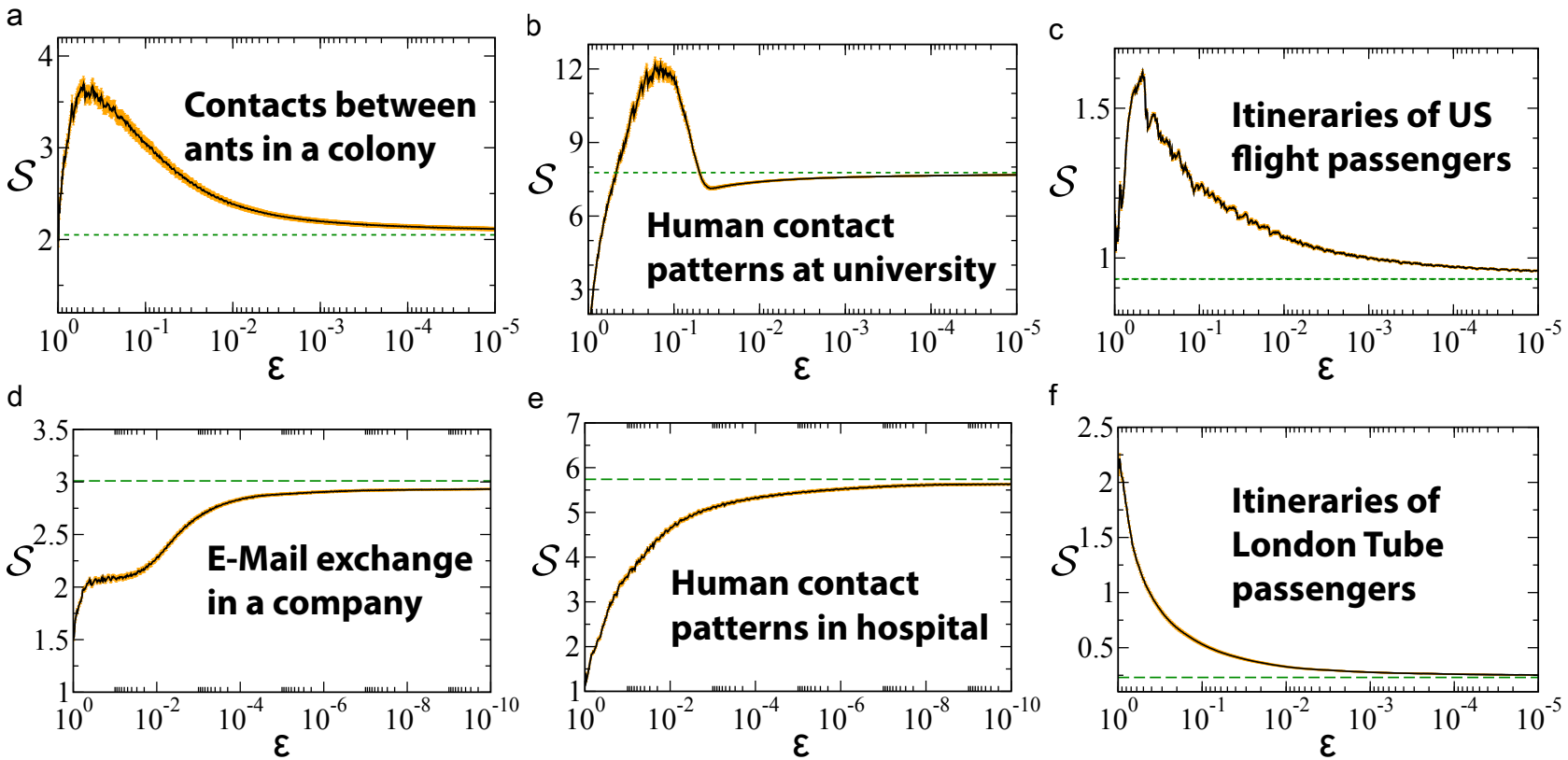


empirical sequence of passenger trips: **non-Markovian temporal network**



shuffled sequence of passenger trips: **Markovian temporal network**

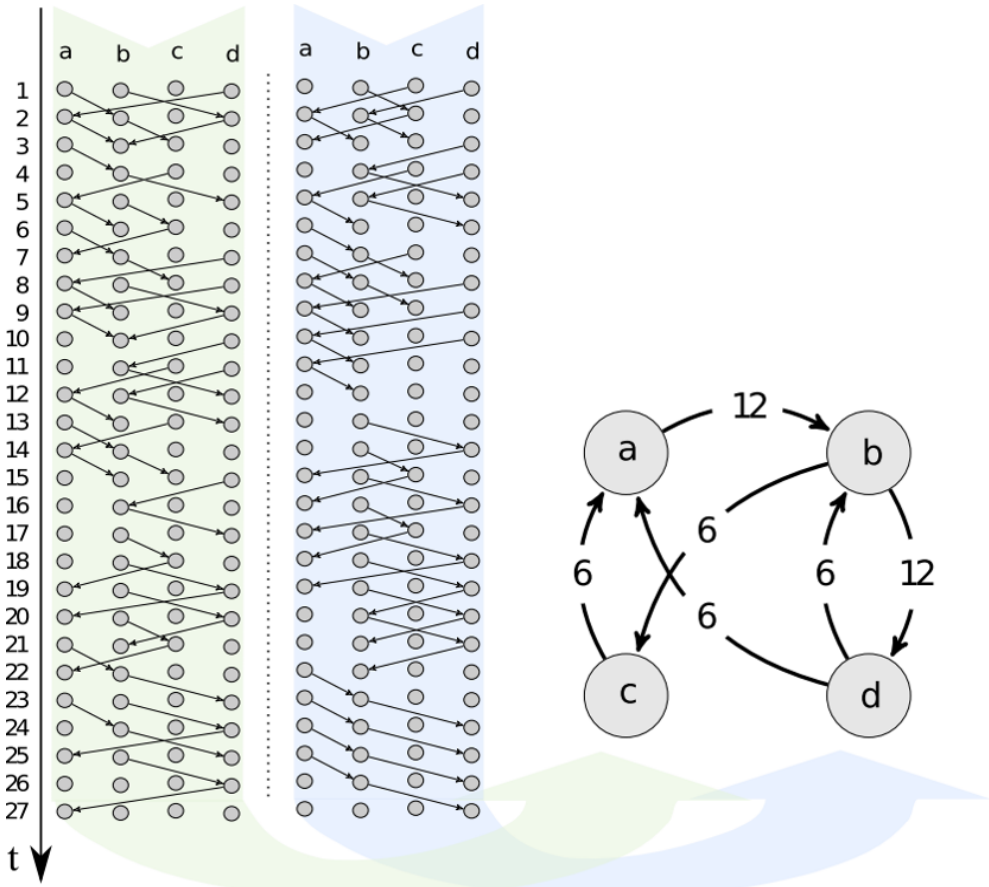
Non-Markovianity: effect on diffusion speed



London Tube: $S \approx 0.25$, i.e. non-Markovianity **speeds up** diffusion

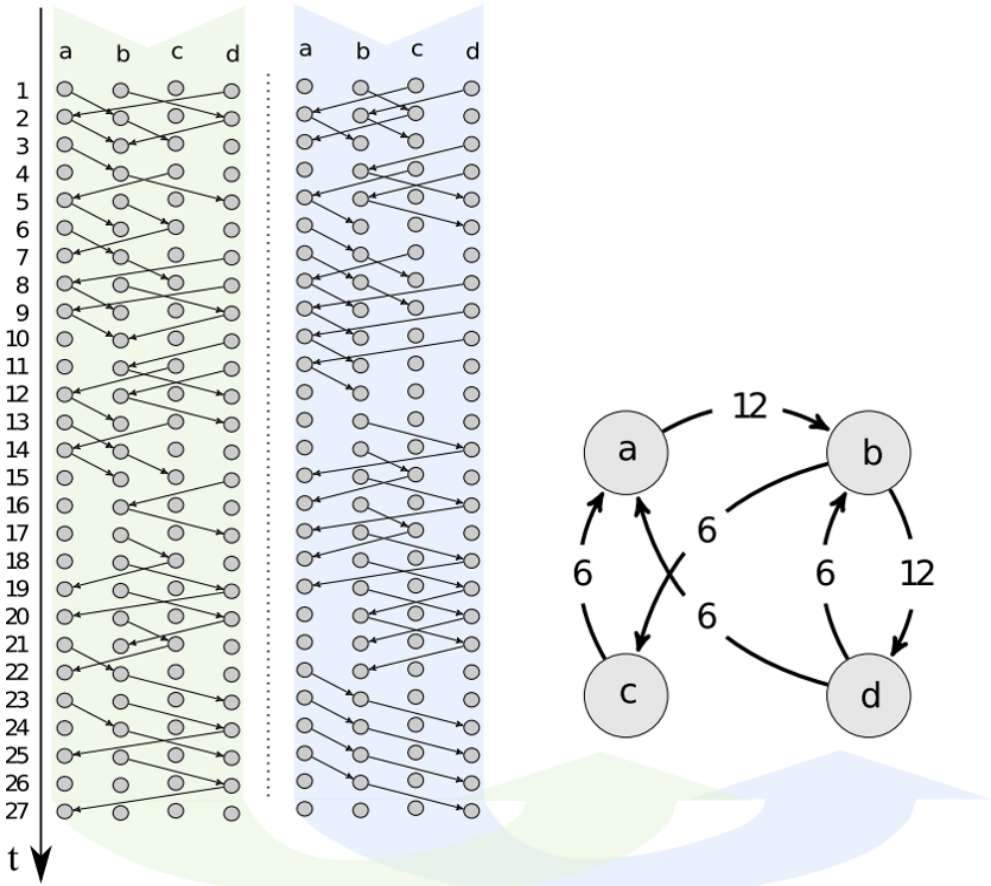
Ant colony: $S \approx 2$, i.e. non-Markovianity **slows down** diffusion

Time-aggregated networks: problem



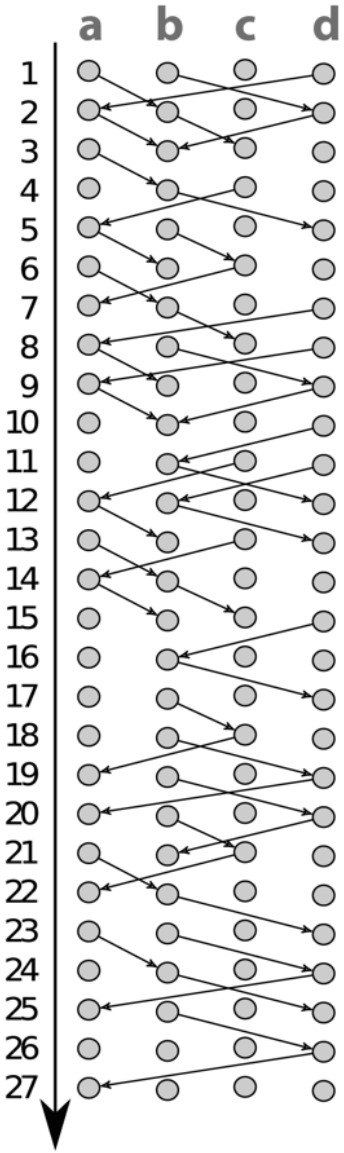
- weights in time-aggregated network capture **link statistics** in temporal network
- information on **ordering of links** is lost
- **limits our understanding** of dynamical processes, node centralities, community structures, etc.

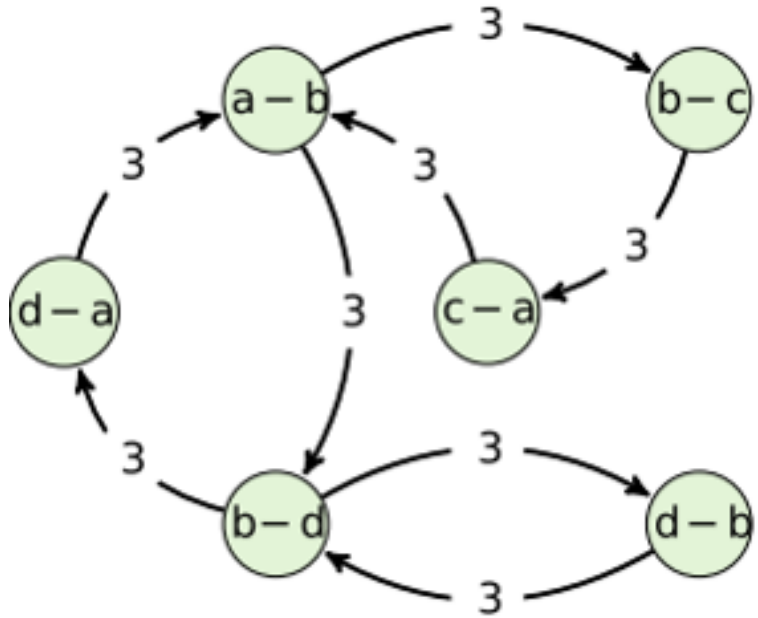
Higher-order aggregate networks?



- link = **time-respecting path of length one**
- time-aggregated networks are **first-order static representation** of temporal networks
- can we construct **higher-order static representations** which capture statistics of longer time-respecting paths?

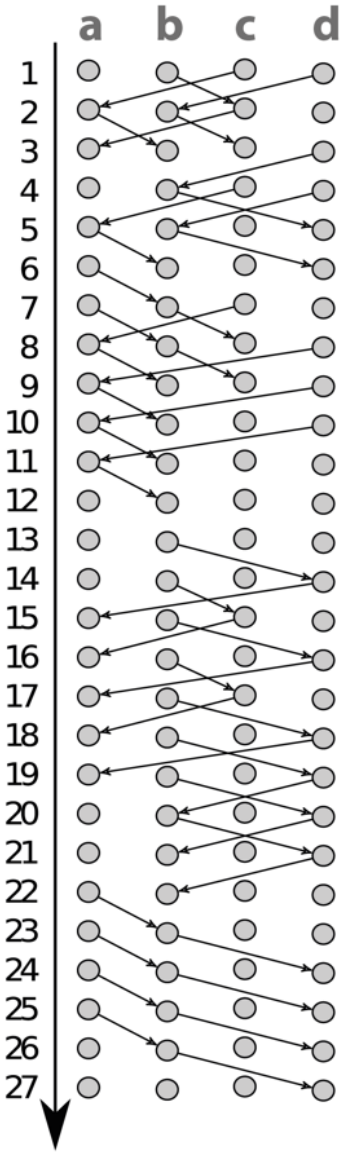
Second-order aggregate network

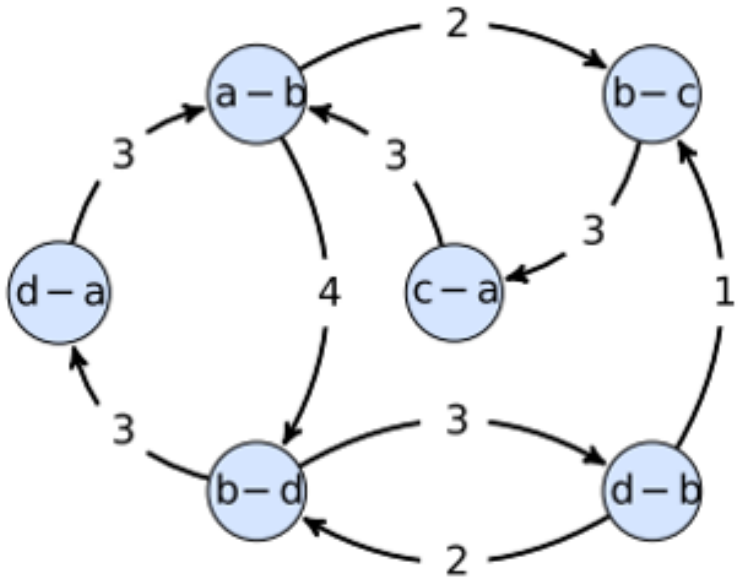




- nodes = links in temporal network
- link weight = statistics of time-respecting paths of length two
- link (e_1, e_2) implies: e_1 before e_2

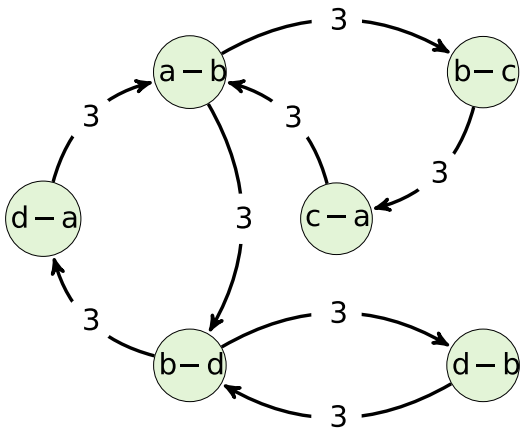
Reordering links . . .



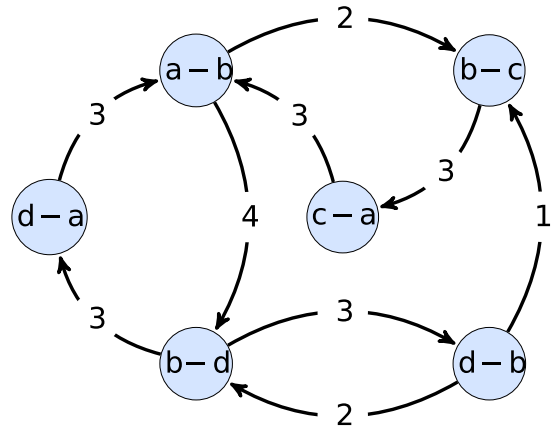


- second-order aggregate networks are **sensitive to ordering** of interactions
- we obtain the **causal topology** of temporal networks
- how does this help us?

Algebraic connectivity of temporal networks



real causal topology



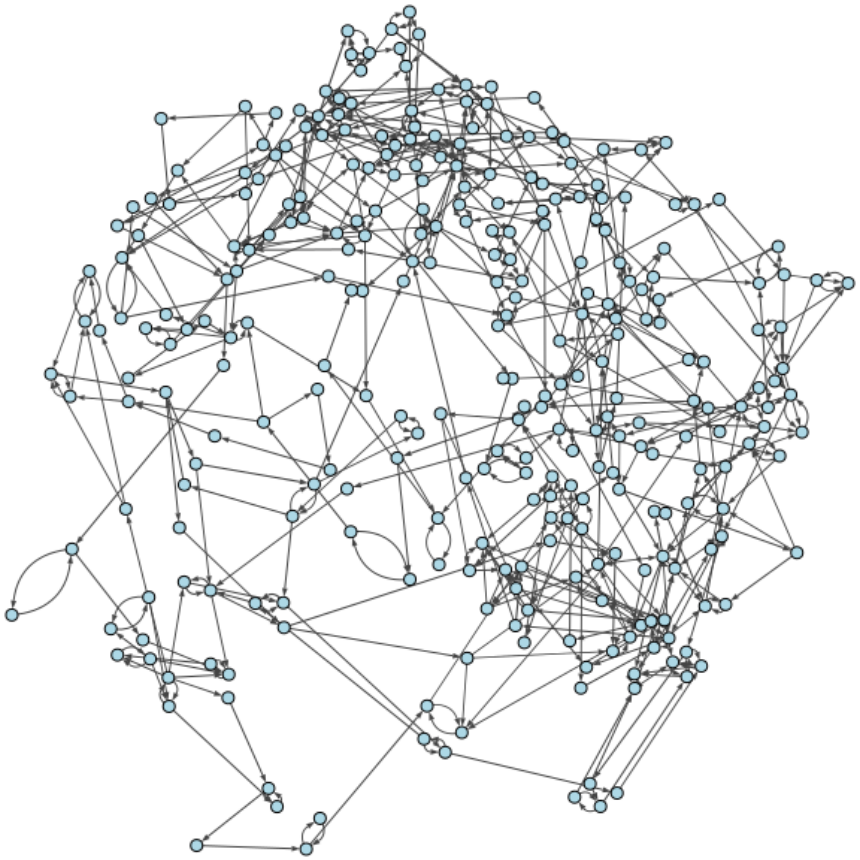
expected causal topology

$$\mathcal{L} = \begin{array}{c} \\ a;b \\ b;c \\ d;a \\ c;a \\ b;d \\ d;b \end{array} \begin{array}{c} \begin{array}{cccccc} a;b & b;c & d;a & c;a & b;d & d;b \end{array} \\ \begin{bmatrix} 1 & -\frac{1}{2} & 0 & 0 & -\frac{1}{2} & 0 \\ 0 & 1 & 0 & -1 & 0 & 0 \\ -1 & 0 & 1 & 0 & 0 & 0 \\ -1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & -\frac{1}{2} & 0 & 1 & -\frac{1}{2} \\ 0 & 0 & 0 & 0 & -1 & 1 \end{bmatrix} \end{array}$$

$$\mathcal{L} = \begin{array}{c} \\ a;b \\ b;c \\ d;a \\ c;a \\ b;d \\ d;b \end{array} \begin{array}{c} \begin{array}{cccccc} a;b & b;c & d;a & c;a & b;d & d;b \end{array} \\ \begin{bmatrix} 1 & -\frac{1}{3} & 0 & 0 & -\frac{2}{3} & 0 \\ 0 & 1 & 0 & -1 & 0 & 0 \\ -1 & 0 & 1 & 0 & 0 & 0 \\ -1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & -\frac{1}{2} & 0 & 1 & -\frac{1}{2} \\ 0 & -\frac{1}{3} & 0 & 0 & -\frac{2}{3} & 1 \end{bmatrix} \end{array}$$

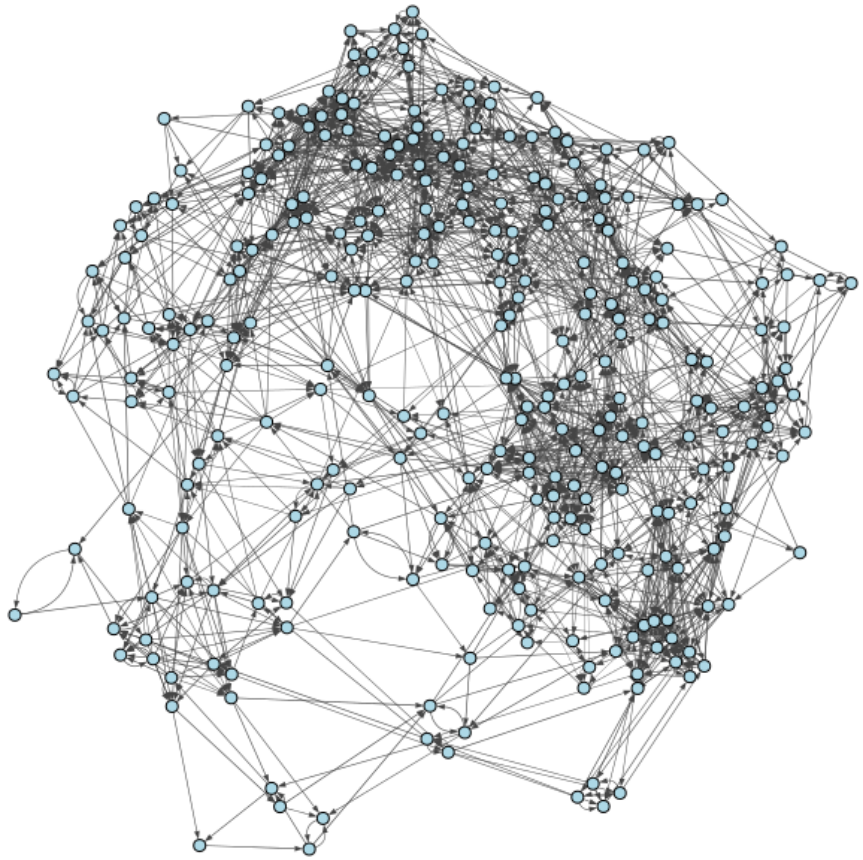
$\Rightarrow$ **spectral analysis of causal topology** induced by order of links

Ant colony: ordering slows down diffusion



real causal topology

$\lambda_2(\mathbf{L}) \approx 0.043$

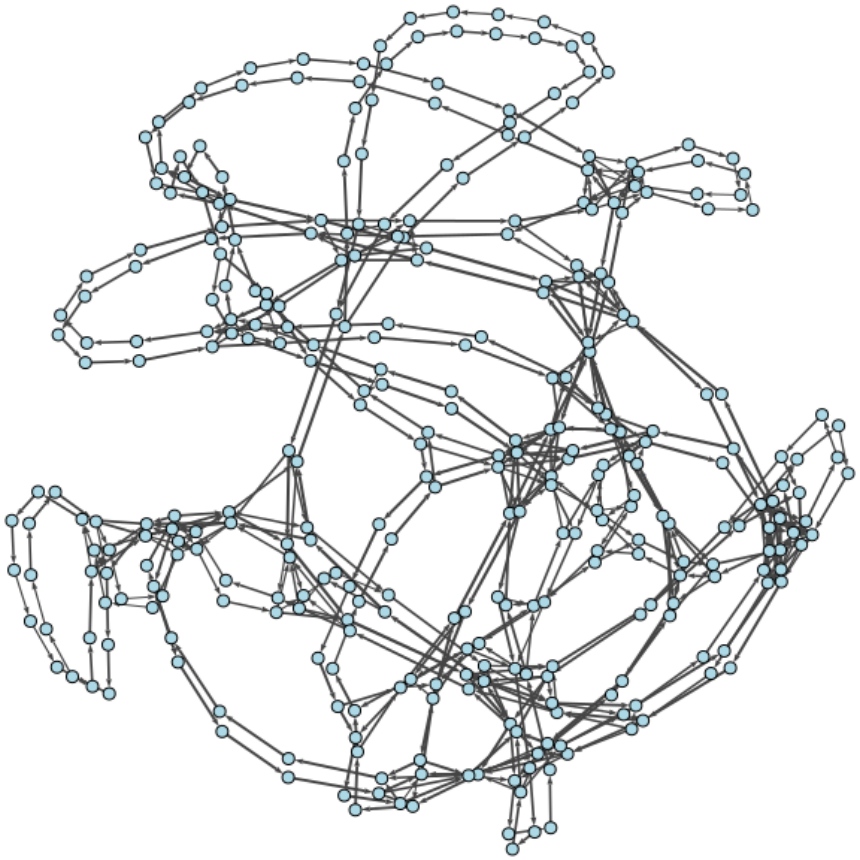


expected causal topology

$\tilde{\lambda}_2(\mathbf{L}) \approx 0.086$

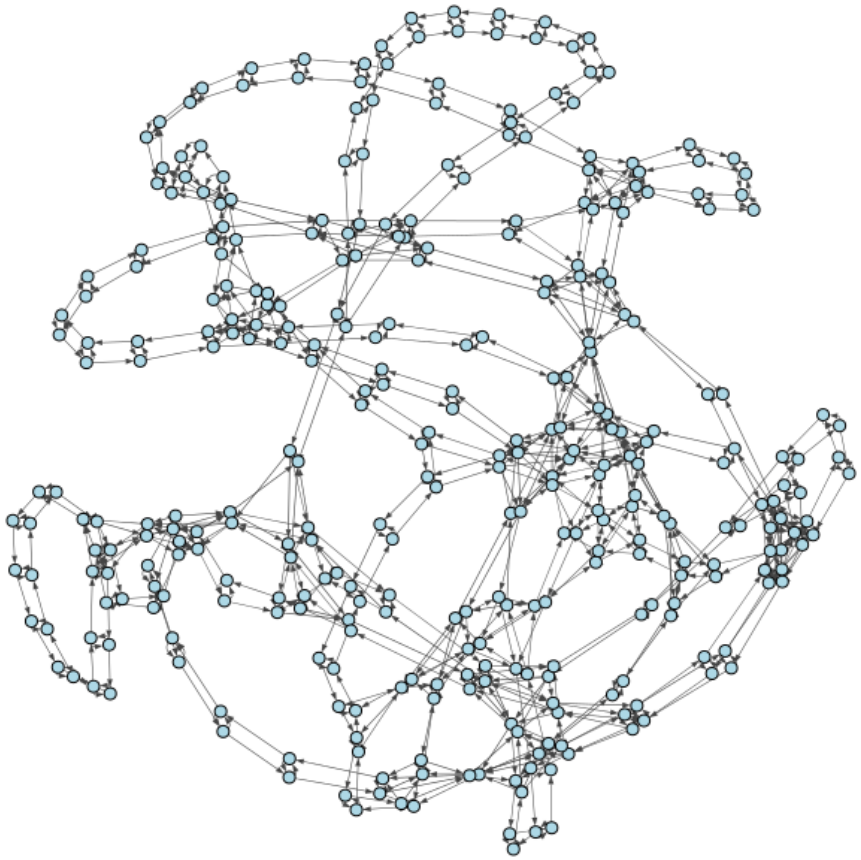
$$S = \frac{\ln|\tilde{\lambda}_2(\mathbf{T})|}{\ln|\lambda_2(\mathbf{T})|} = 2.05$$

London Tube: ordering speeds up diffusion



real causal topology

$\lambda_2(\mathbf{L}) \approx 0.125$



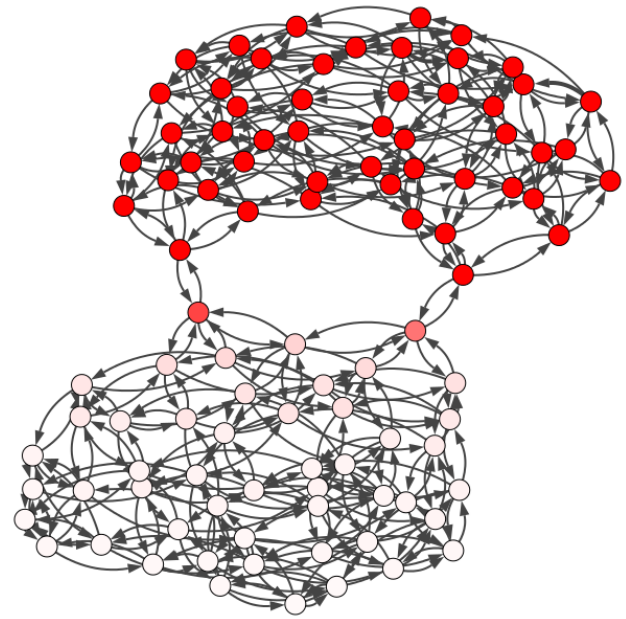
expected causal topology

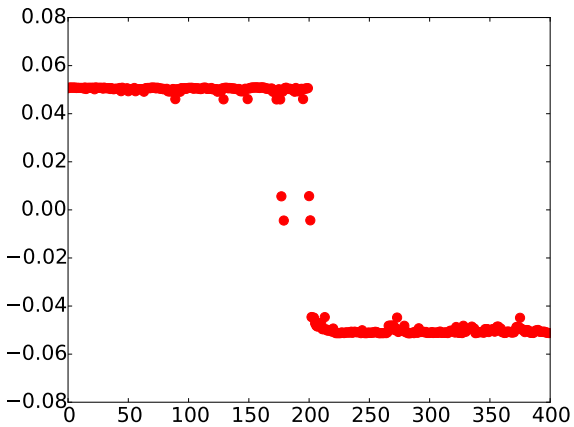
$\tilde{\lambda}_2(\mathbf{L}) \approx 0.011$

$$S = \frac{\ln|\tilde{\lambda}_2(\mathbf{T})|}{\ln|\lambda_2(\mathbf{T})|} = 0.24$$

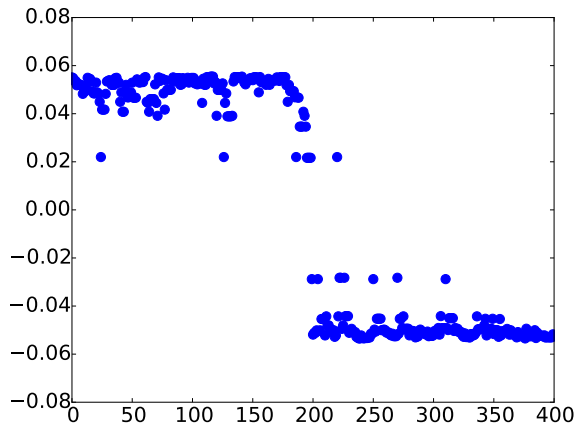
Shameless plug: spectral partitioning

- higher-order networks allow **spectral partitioning**
- example: Fiedler vector of second-order aggregate network
- order correlations can **enforce** or **mitigate** community structures
- details: temporal networks session (THU)

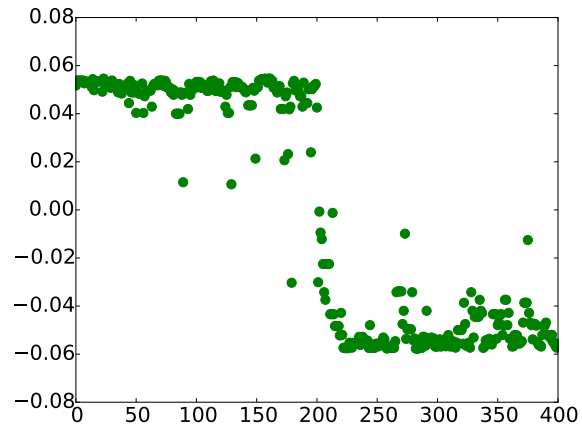




non-Markovian



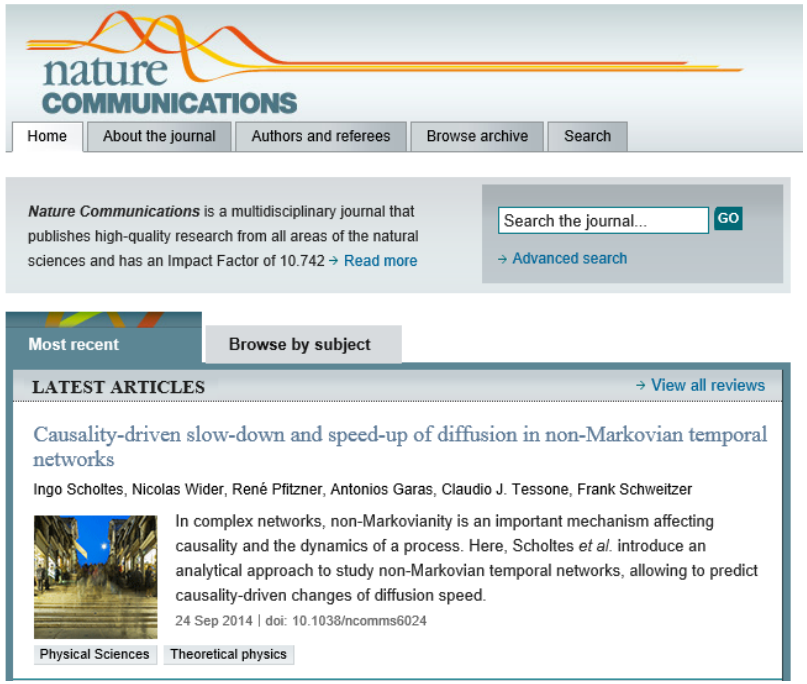
Markovian



non-Markovian

Conclusion

- **ordering matters in temporal networks!**
- **complex interplay with network topology** can lead both to slow down and speed up of dynamical processes
- interplay can be studied analytically using **higher-order aggregate networks**
- allows to explore **additional dimension of complexity** in temporal networks
- **pyTempNets**: full-fledged python module to reproduce results



Scholtes, I., Wider, N., Pfitzner, R., Garas A., Tessone, C.J., Schweitzer, F. (2014) **Causality-driven slow-down and speed-up of diffusion in non-Markovian temporal networks**. Nature Communications 5, September 24 2014

Thank You!

R Pfitzner, I Scholtes, A Garas, CJ Tessone, F Schweitzer: **Betweenness Preference: Quantifying Correlations in the Topological Dynamics of Temporal Networks**, Phys Rev Lett, Vol. 110, 198701, May 10 2013

I Scholtes, N Wider, R Pfitzner, A Garas, CJ Tessone, F Schweitzer: **Slow-down and Speed-up of Diffusion in Non-Markovian Temporal Networks**, Nature Communications, Vol. 5, September 24 2014

I Scholtes et al.: **Force-Directed Layout of Non-Markovian Temporal Networks**, working paper, available online at http://www.sg.ethz.ch, May 06 2014

http://www.sg.ethz.ch
http://www.ingoscholtes.net

IngoScholtes / **pyTempNets**

Follow **Chair of Systems Design**

@ingo_s

ischoltes@ethz.ch