Community structure and multi-wave urban epidemics

Thomas Hladish

Hoen, Hladish *et al*, J Med Internet Res 2015;17(7):e169

The Montreal network

Île Sans Fil public Wi-Fi system

- 352 Cafés, bars, restaurants, shops, etc.

200,000+ anonymous, unique users

Over 2 million connections, 2004-2010

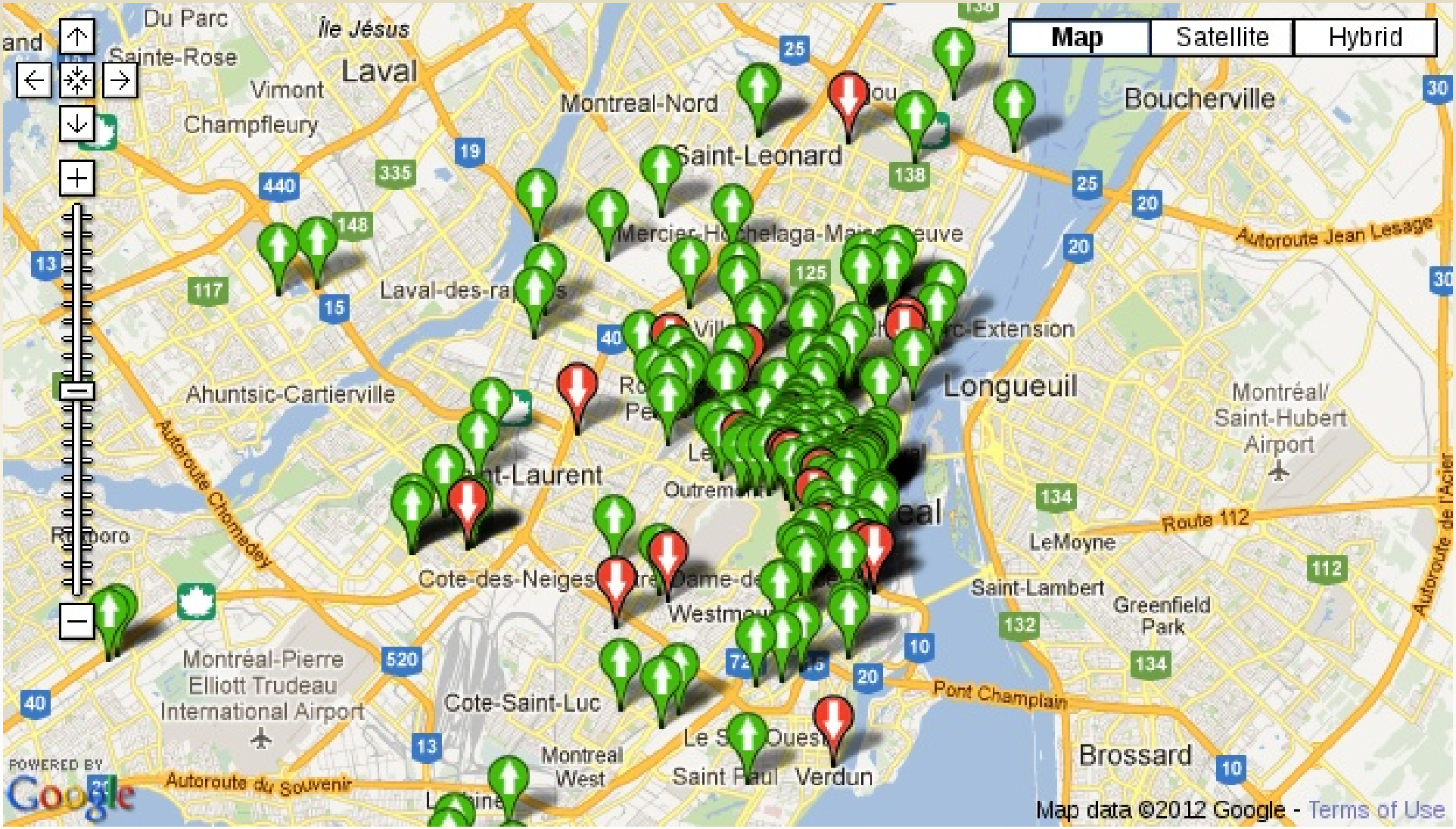
User, location, on time, off time

Nodes = users, edges = concurrent hotspot use

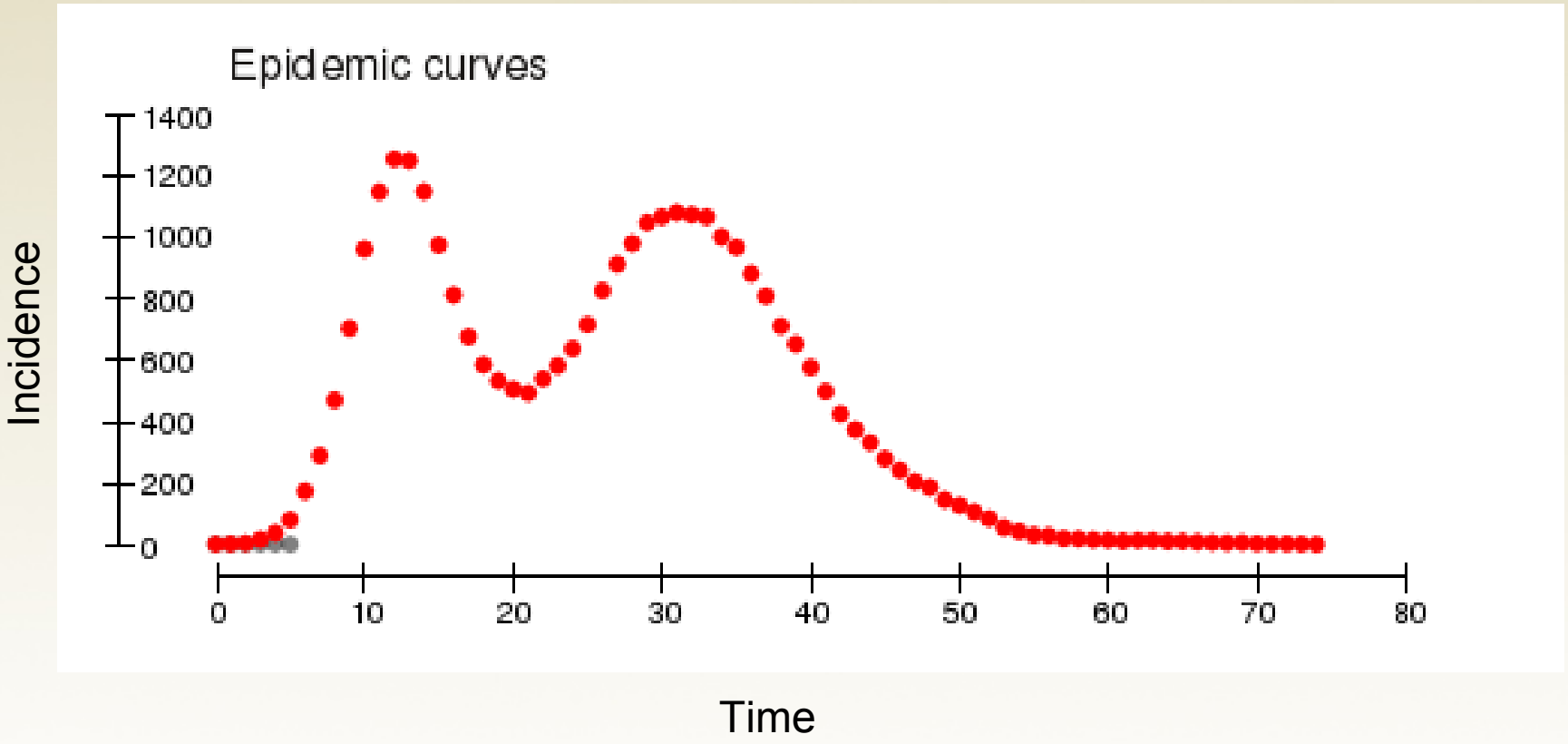
Network: 103,000 nodes, 650,000 edges

“Extra-social” contacts

ÎSF Hotspots



An unusual epidemic

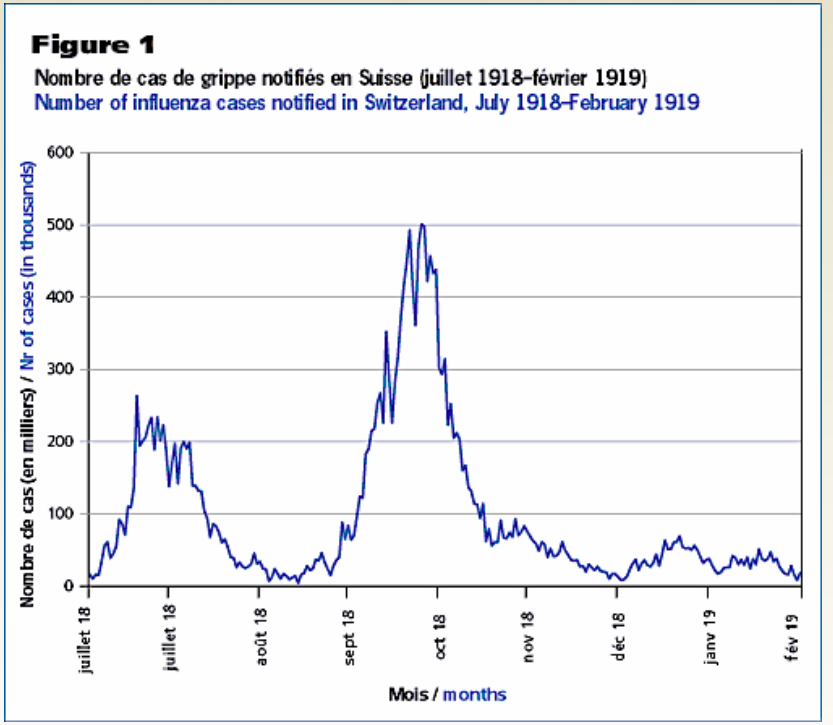


Multi-wave epidemics

Challenge for modelers and public health officials

Conventional explanations:

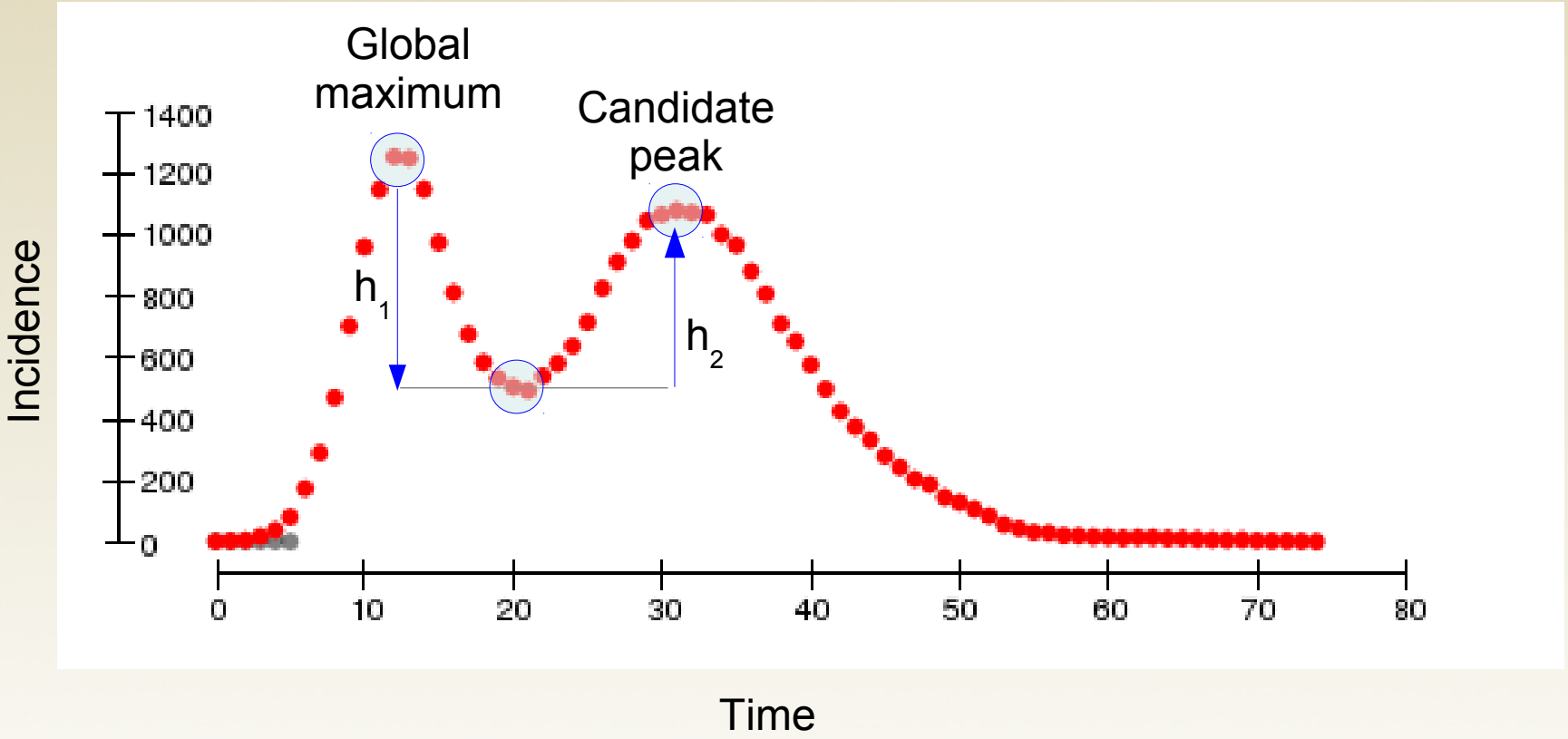
- Environmental conditions
- Behavioral changes
- Evolution
- Multiple variants



Ammon 2002

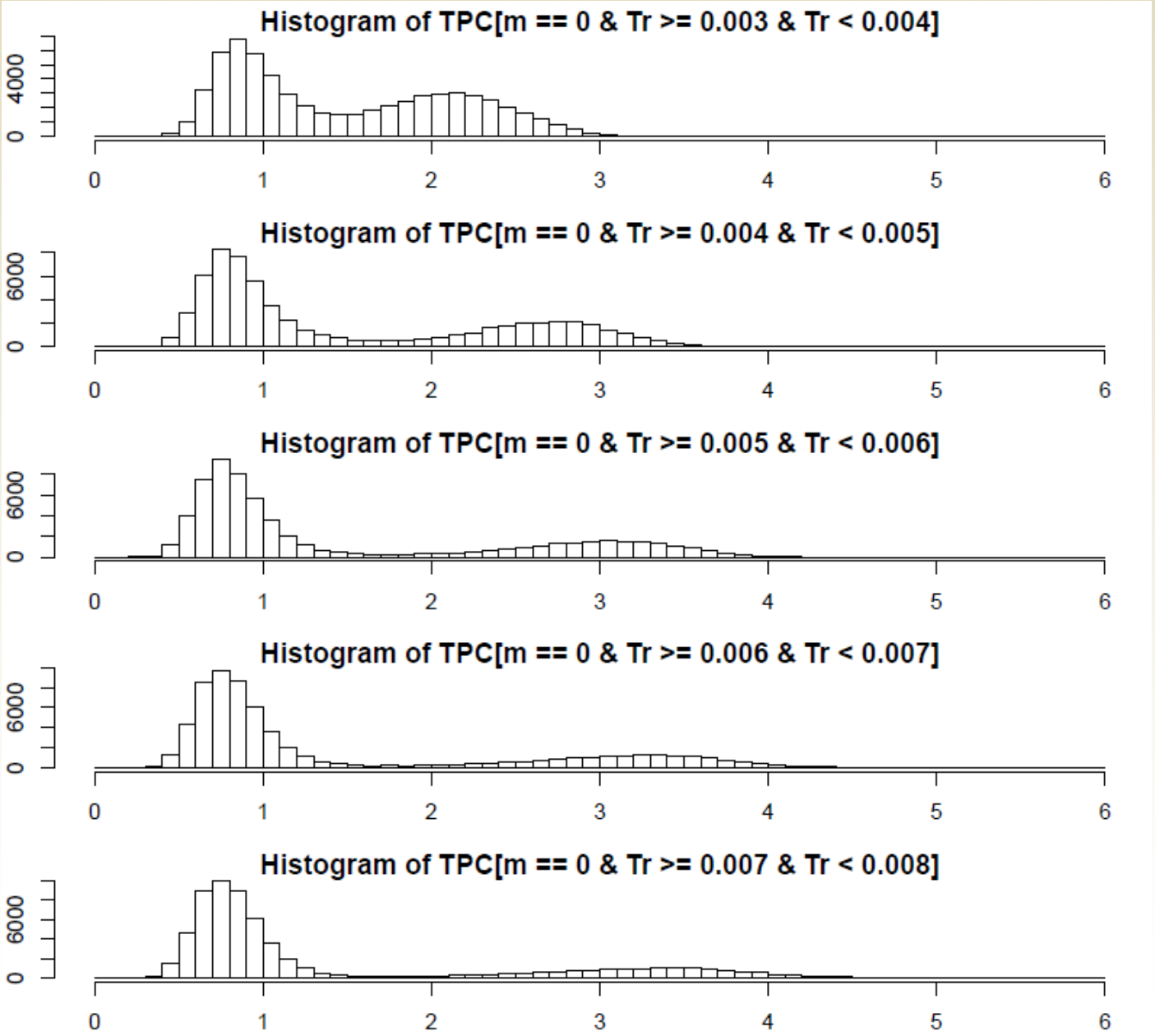
Cities are considered well-mixed

Assessing two-peakedness

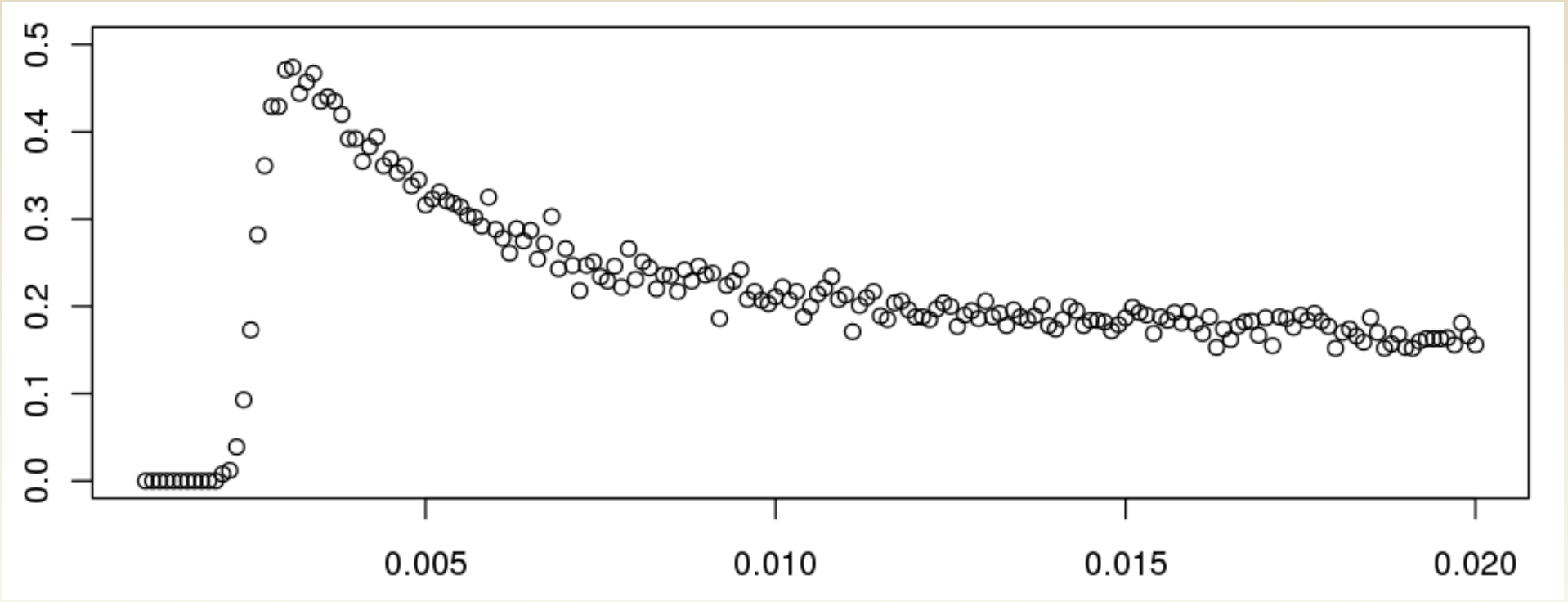


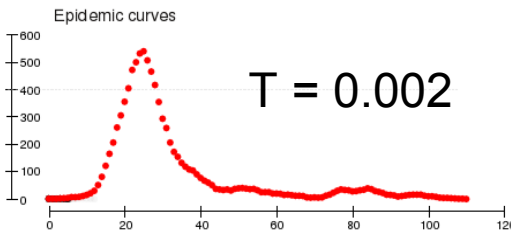
$$TP\ score = max\left(\sqrt{h_1 * h_2}\right) / FinalSize$$

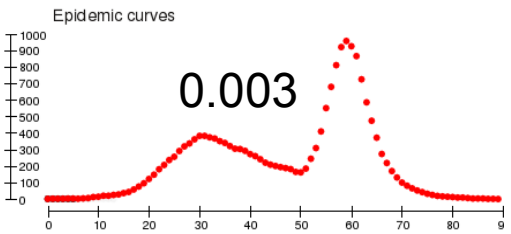
TP score distributions

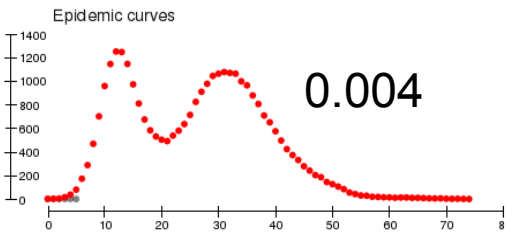


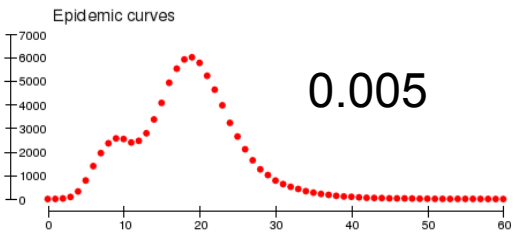
Probability of multi-wave epidemics

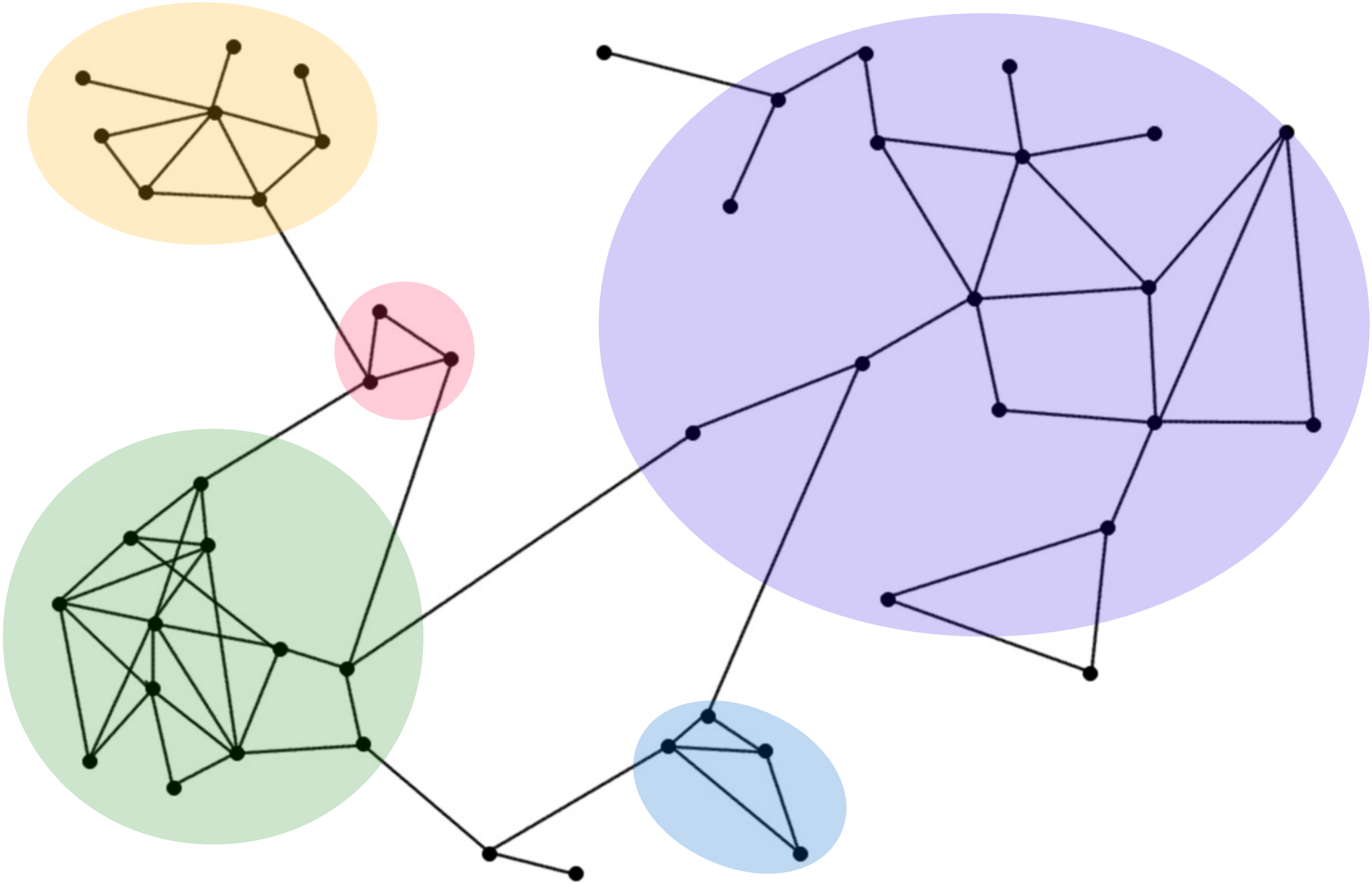




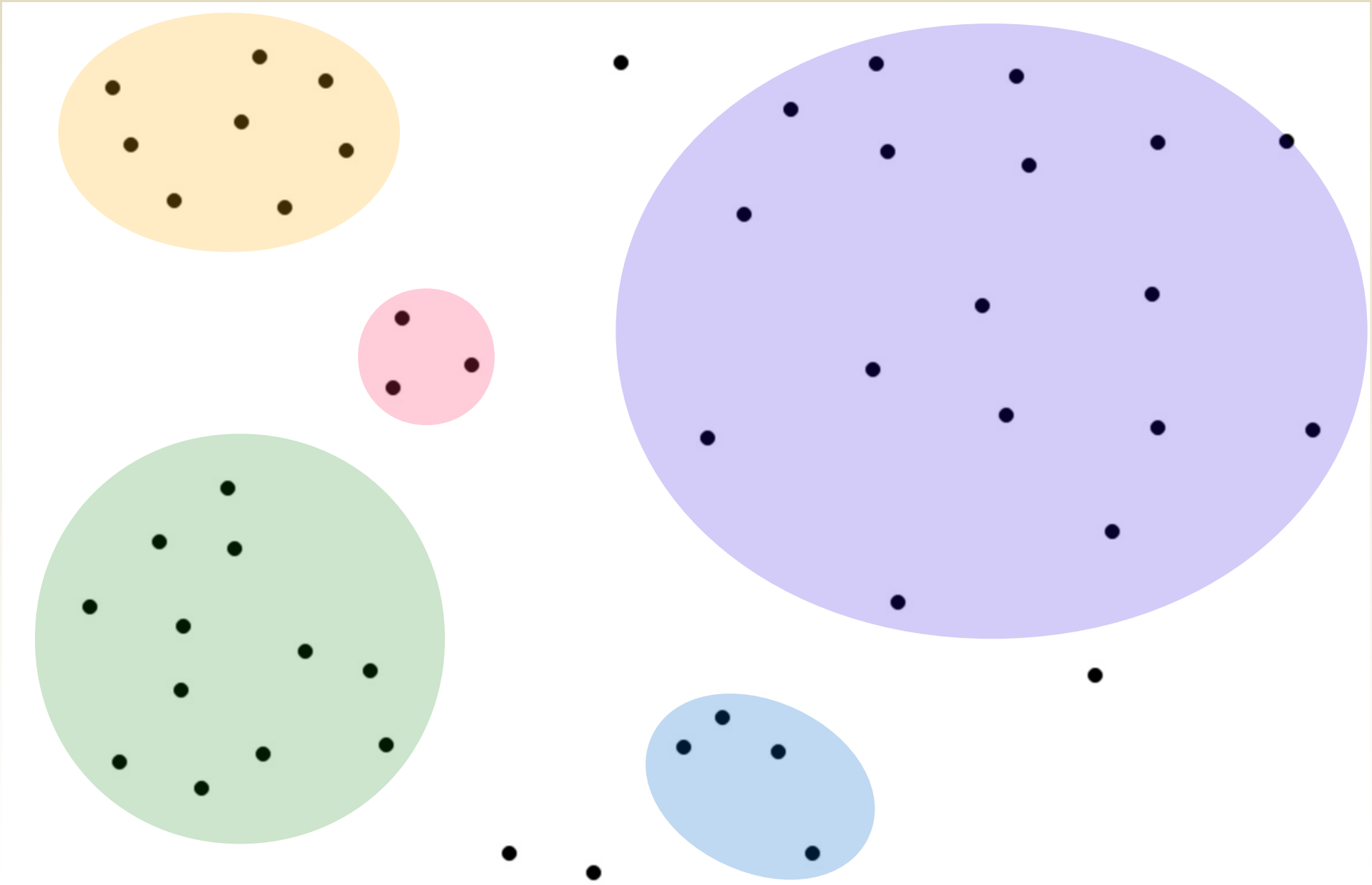




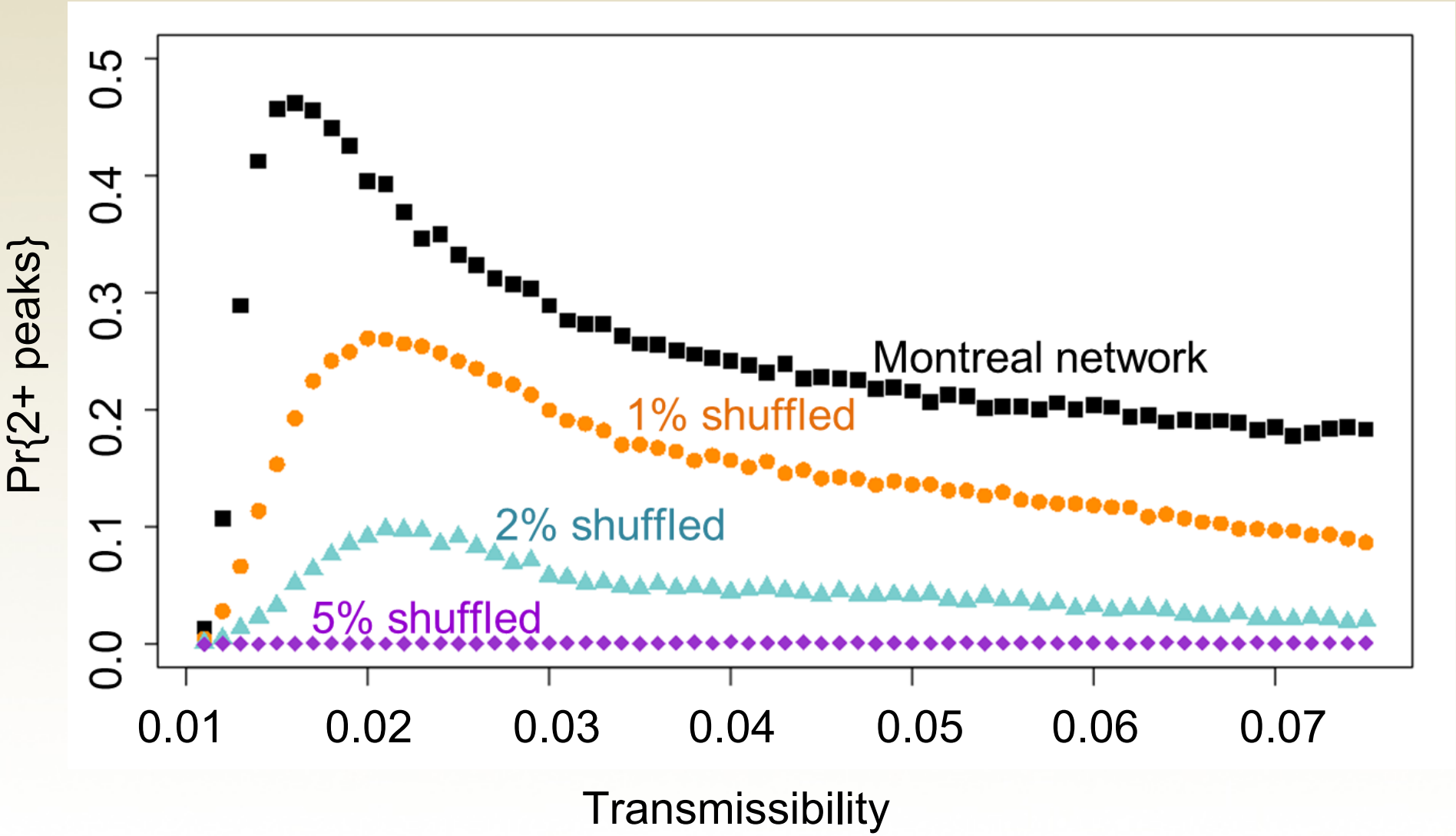




Decomposition of a Network



Shuffling stops multi-wave dynamics



Modularity Analysis

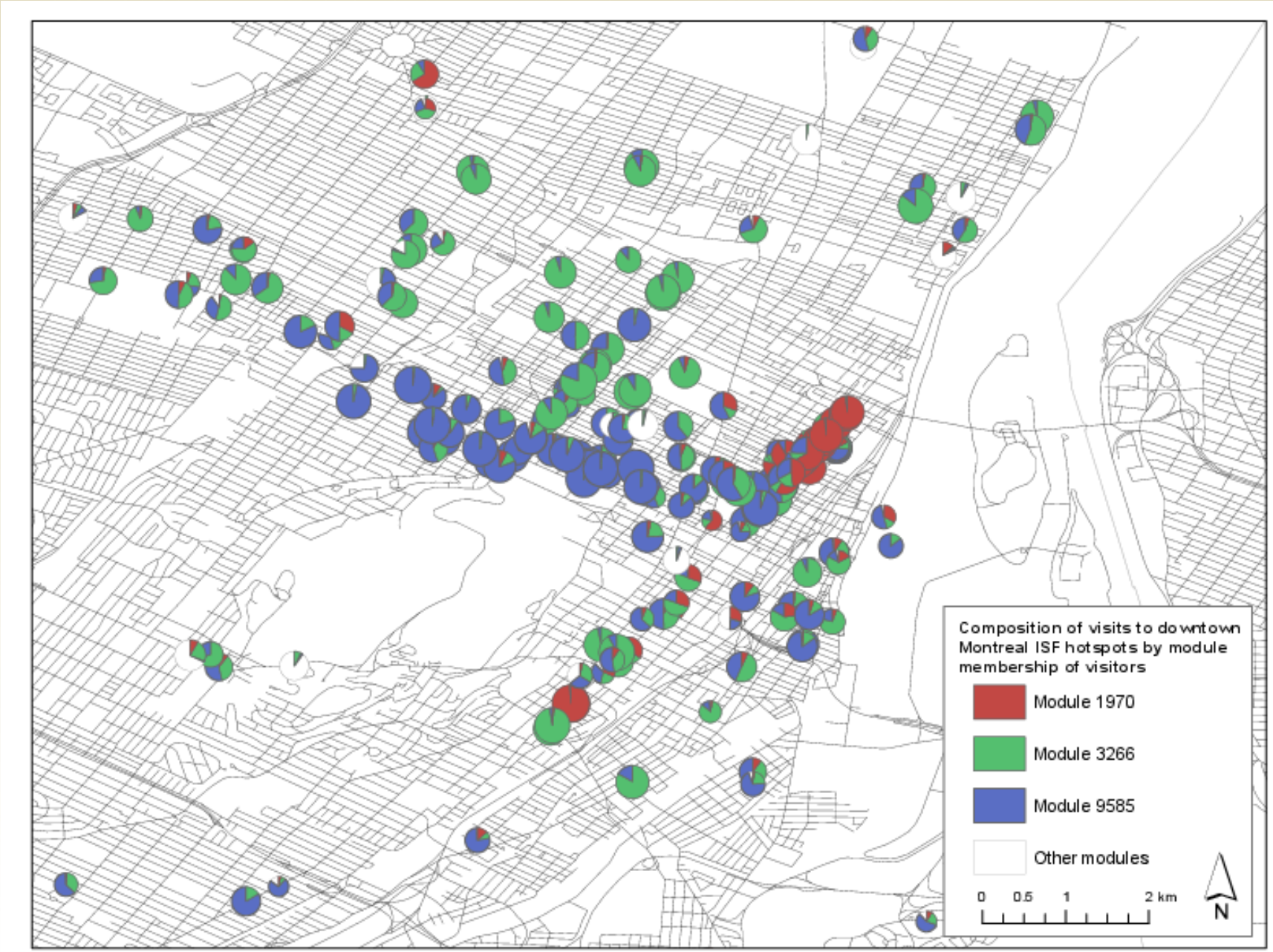
Clauset et al (2004). Finding community structure in very large networks. *Phys. Rev. E*

1) Start with each node in its own module
2) Join 2 modules to maximize the ratio of within-module edges: between module edges
3) Repeat until all modules are joined together
4) Find point at which ratio deviated most from expected value for randomly wired graph

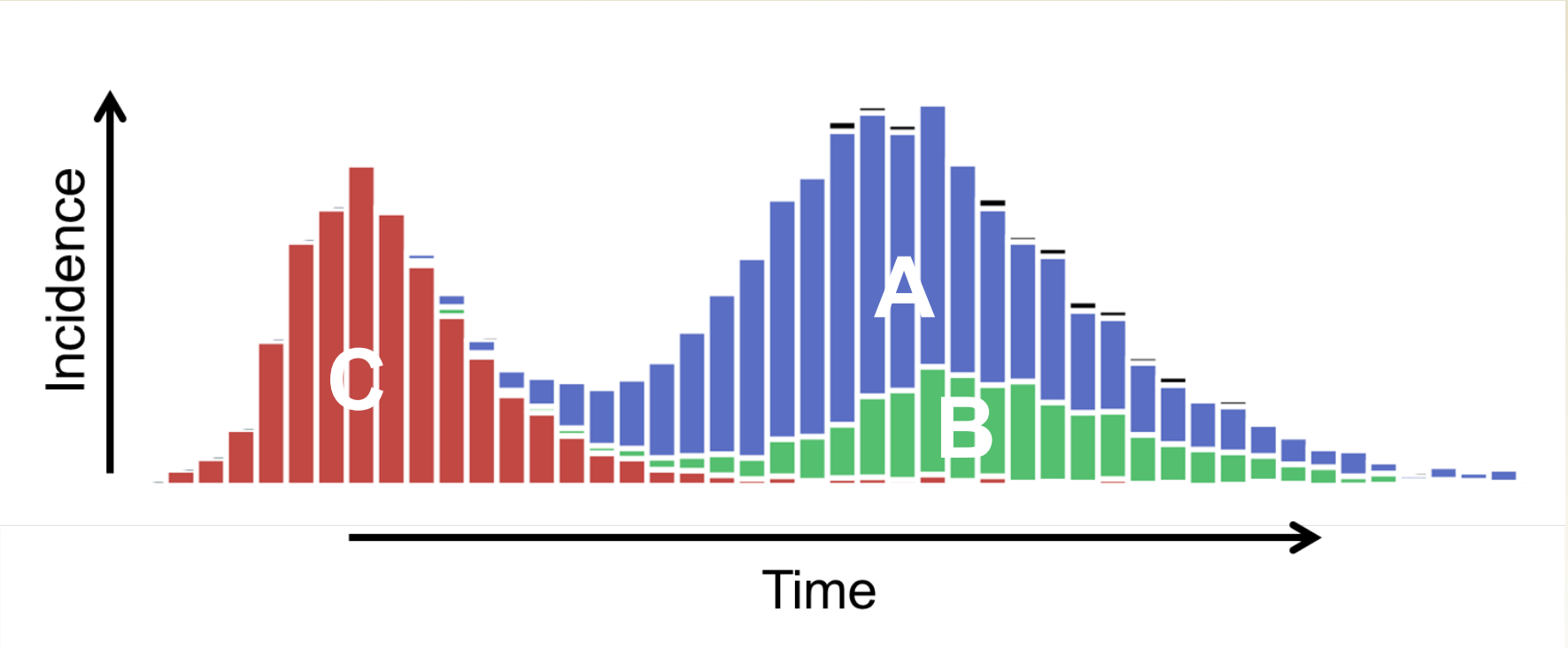
Module	Module Size
A	**38569**
B	**28101**
C	**15558**
	5256
	1928
	1897
	900
	700
	629
	608
[1410 others]	

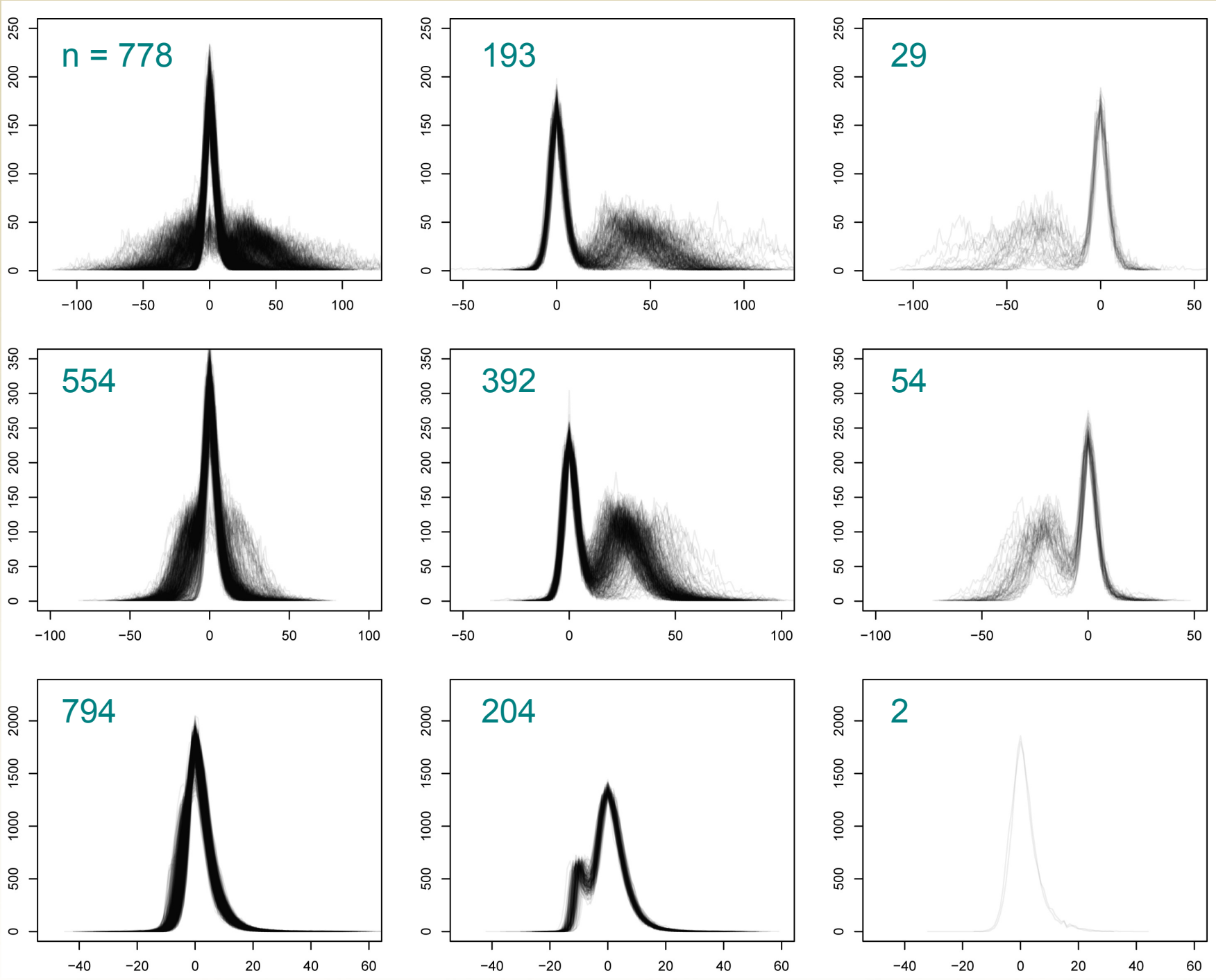
$$Modularity\ Q = \frac{1}{2\mathrm{m}} \sum_{u,v} \left(A_{u,v} - \frac{k_u k_v}{2\mathrm{m}} \right) \delta\left(M_u, M_v \right) = 0.6$$

Geographical structure

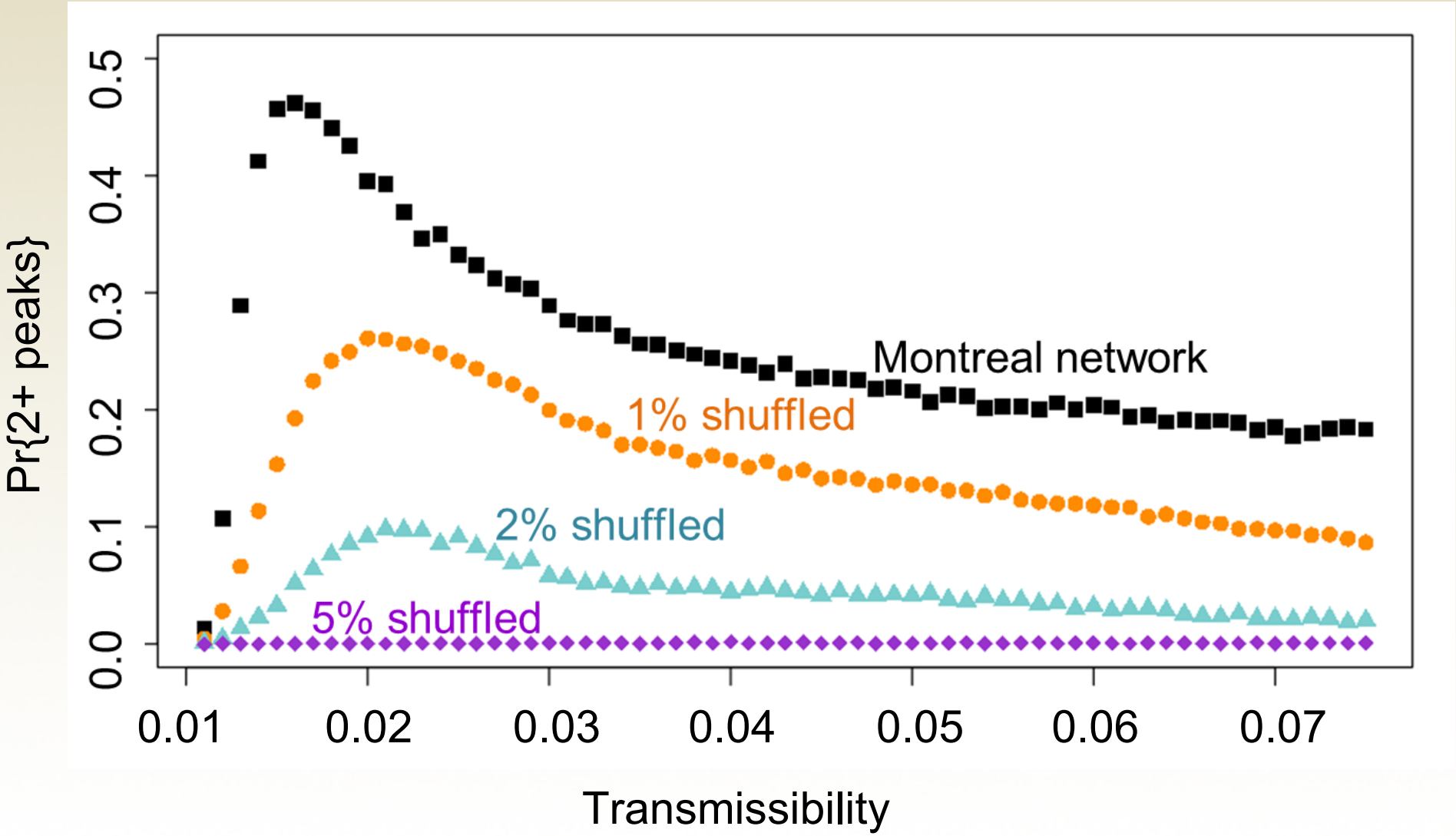


An anecdotal epidemic

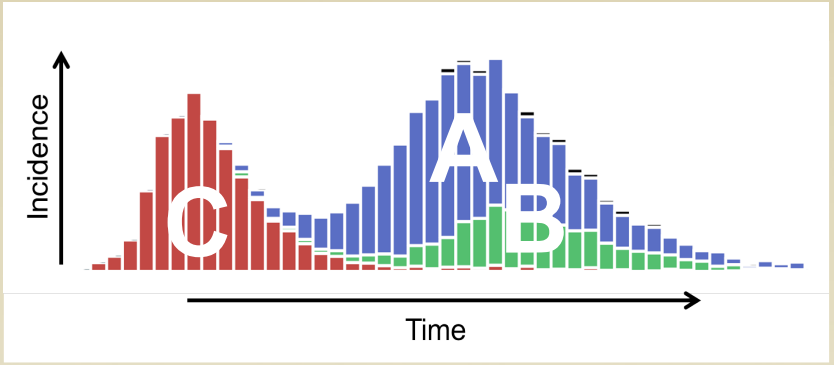



Low T
Moderate T, Maximizing Pr{2 peaks}
High T
n = 778
193
29
554
392
54
794
204
2
One peak
Red peak first
Red peak second

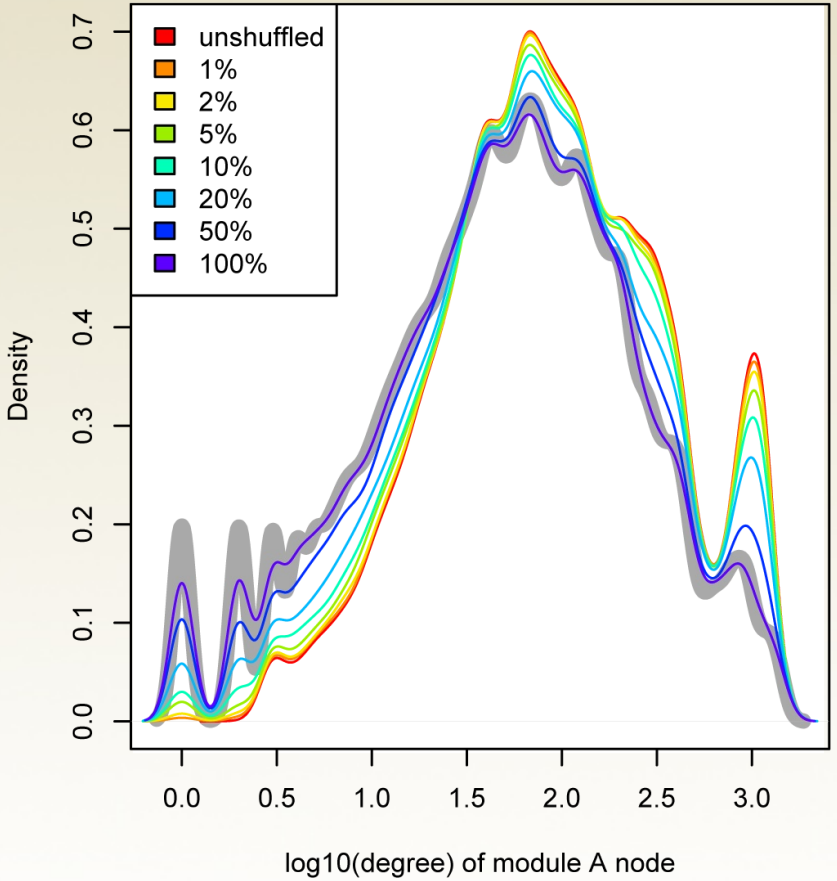
Why is shuffling so effective?



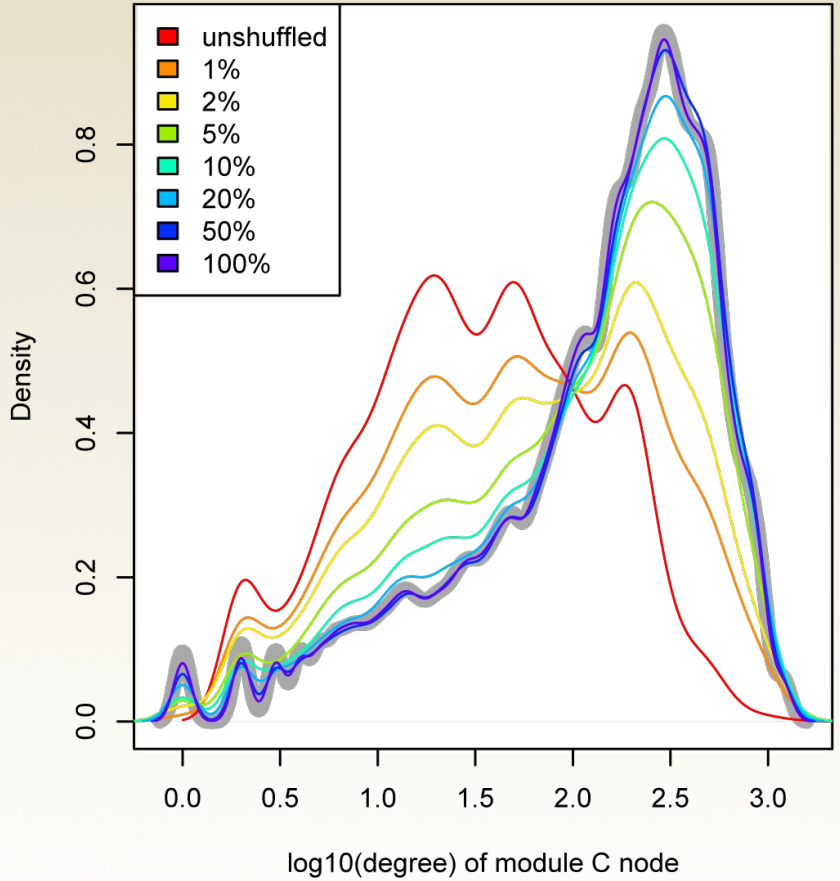
High degree nodes determine synchrony



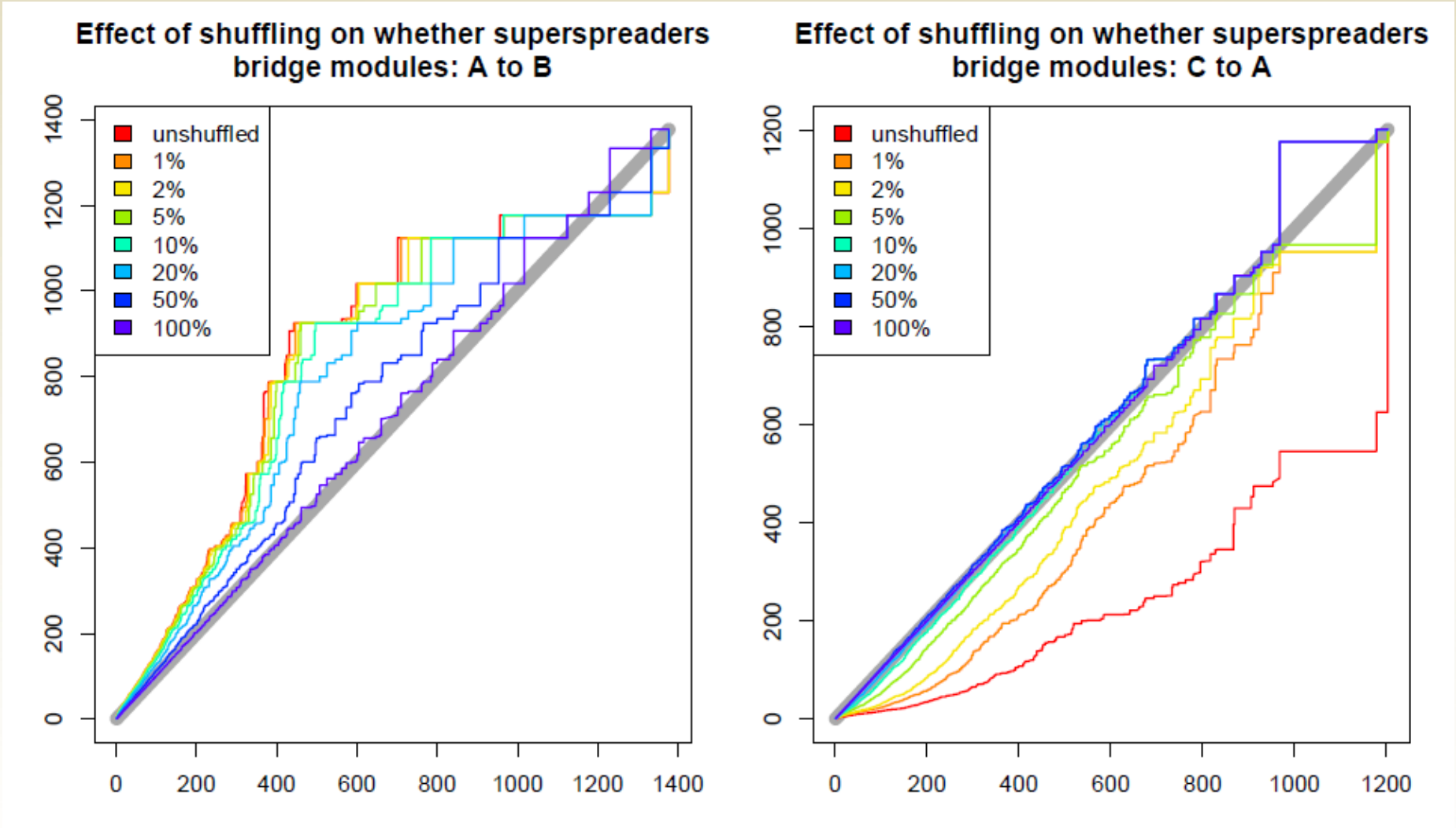
Bridges from A to B
(synchronized)



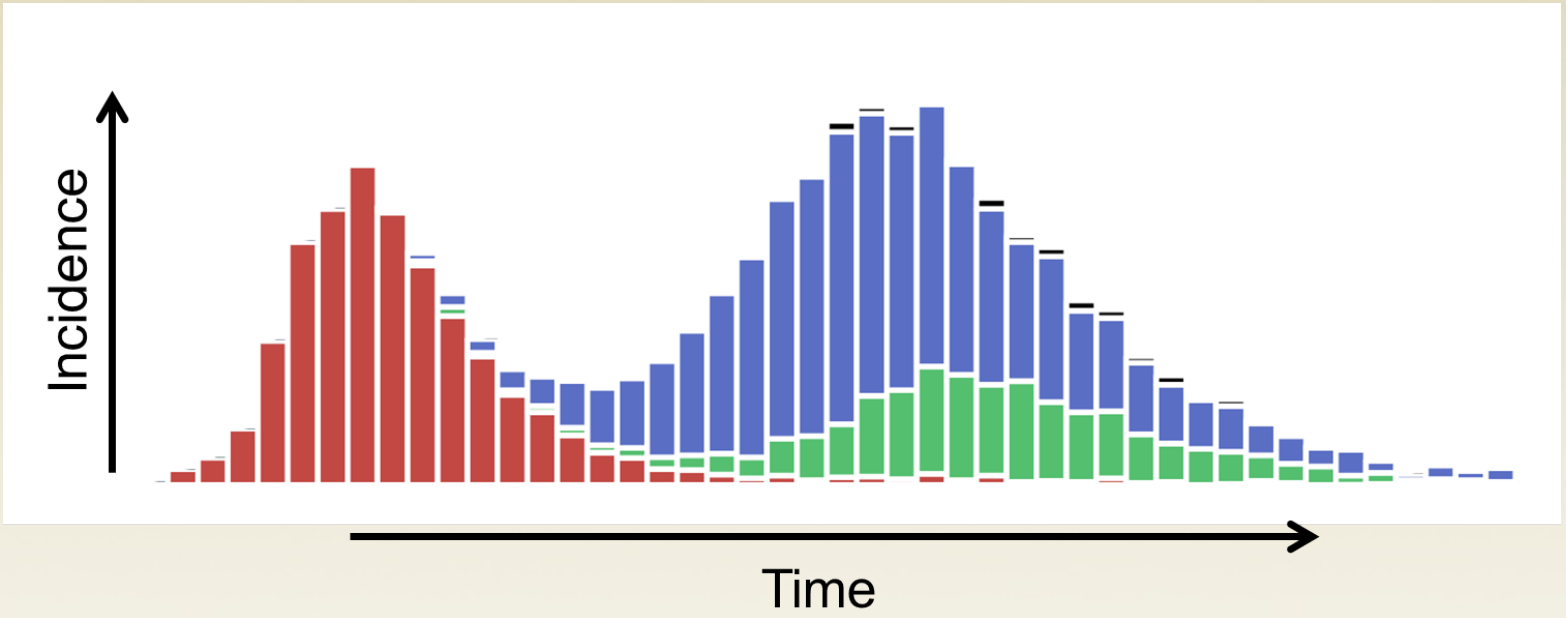
Bridges from C to A
(not synchronized)

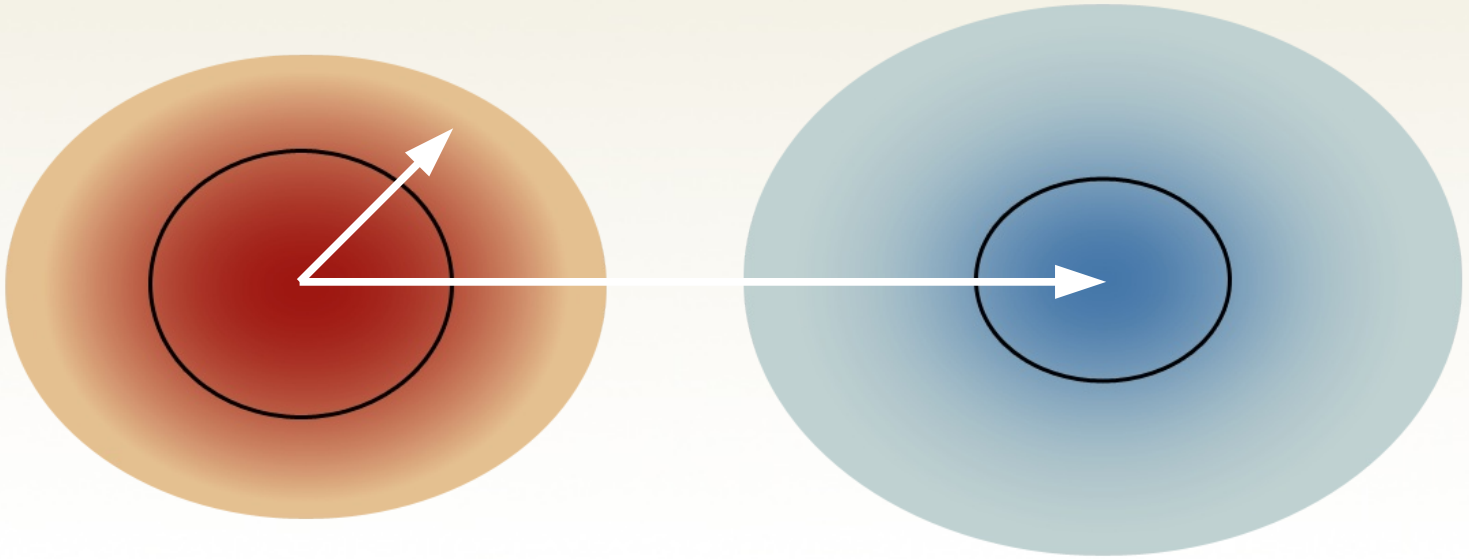


Type of bridges is important

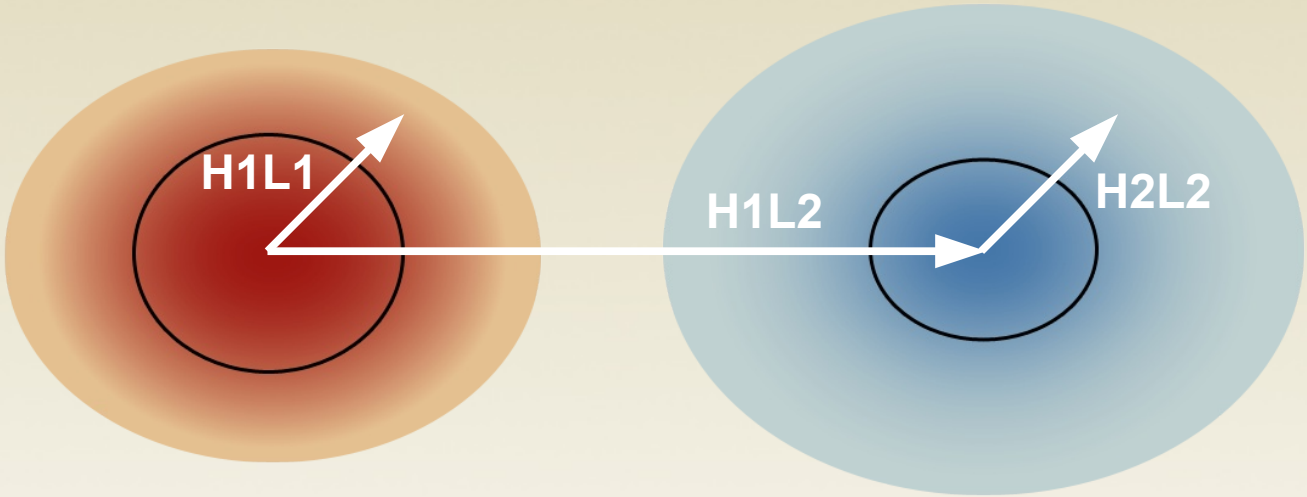


Distance ∝ time





Multiwave metric



Are high degree core nodes in different communities closer than core and non-core nodes within communities?

MWM = 2 * H1H2 / (H1L1 + H2L2)

Multiwave metric

MWM = 2 * H1H2 / (H1L1 + H2L2)

	A&B	A&C	B&C	Freq(MW epis)
0%	1.06	**1.54**	**1.47**	**46%**
1%	1.04	**1.24**	**1.20**	**19**
2%	1.08	**1.14**	**1.11**	**5**
5%	0.98	1.01	0.98	0
100%	0.85	1.01	1.00	0

What are “core nodes?”

Possibilities:

- Degree cutoff
- Percentile cutoff
- 80/20 rule
- 50/X rule
- Maximum clique
- Maximum degree clique

Conclusions

Communities are not all the same

Potential understood, but not usually modeled

Epidemic forecasting

Targeted interventions