Network Inference from Population-Level Observation of Epidemics

F.Di Lauro, J.C. Croix, M.Dashti, L.Berthouze, I.Z. Kiss

University of Sussex

fd212@sussex.ac.uk

April 8, 2019

Outline

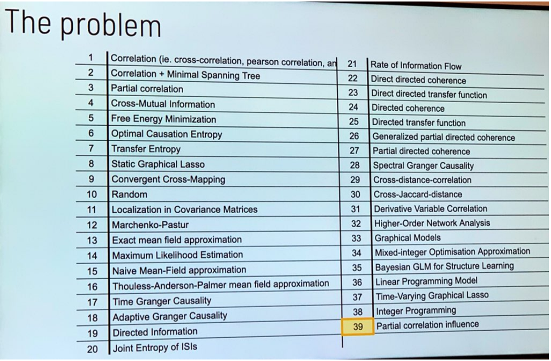
Introduction: Network inference

What do we want to **infer** about Networks?

- Link-based inference: adjacency matrix?
- Topological parameters? (i.e. $\langle k \rangle$)
- Some mesoscopic quantity related to the network structure (comunities, clustering)?

Possibilities

How many ways do we have to infer networks?



The problem

1	Correlation (ie. cross-correlation, pearson correlation, an	21	Rate of Information Flow
2	Correlation + Minimal Spanning Tree	22	Direct directed coherence
3	Partial correlation	23	Direct directed transfer function
4	Cross-Mutual Information	24	Directed coherence
5	Free Energy Minimization	25	Directed transfer function
6	Optimal Causation Entropy	26	Generalized partial directed coherence
7	Transfer Entropy	27	Partial directed coherence
8	Static Graphical Lasso	28	Spectral Granger Causality
9	Convergent Cross-Mapping	29	Cross-distance-correlation
10	Random	30	Cross-Jaccard-distance
11	Localization in Covariance Matrices	31	Derivative Variable Correlation
12	Marchenko-Pastur	32	Higher-Order Network Analysis
13	Exact mean field approximation	33	Graphical Models
14	Maximum Likelihood Estimation	34	Mixed-integer Optimisation Approximation
15	Naive Mean-Field approximation	35	Bayesian GLM for Structure Learning
16	Thouless-Anderson-Palmer mean field approximation	36	Linear Programming Model
17	Time Granger Causality	37	Time-Varying Graphical Lasso
18	Adaptive Granger Causality	38	Integer Programming
19	Directed Information	39	Partial correlation influence
20	Joint Entropy of ISIs		

Figure: A lot, apparently [1]! (picture posted on Twitter by S.V. Scarpino, from a talk by B.Klein)

What is new?

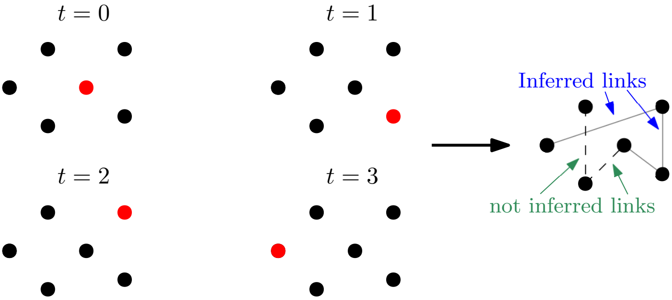
Why are we proposing a new approach then?

- Networks of interest: the ones with **dynamics** on them
- Observing the dynamics can help to infer the network...
- ...But only when detailed (node-level) information is available

Question: what can we infer when little information is available?

The idea behind: observing cascades

Example: simple SI model



Many cascades → full inference (Adjacency Matrix) [2].

Remarks

Often full recovery of the graph is not needed

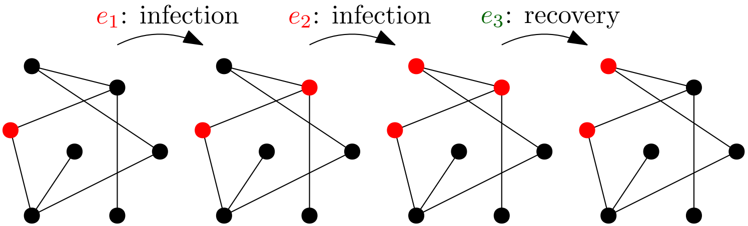
Degree distribution gives a lot of information!

With fewer observations we can infer it, but still too many details (specifically: **continuous** observations)[3]

Can we do ‘better’?

- Aim: infering at least the **family** a network belongs to...
- ... when only discrete observations of the process are available
- Two ingredients: SIS epidemics and Birth-Death process approximation.

SIS model as a Birth-Death process



Infection with rate τ per S-I link; recovery with rate γ.

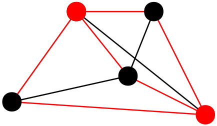
High dimensionality, many methods try to reduce it [4]

Focus on the number of infected nodes:

Space state $S = \{0, \dots, N\}$ and events are ± 1 jumps.

Birth-Death process? what rates?

Simple case: Complete network



$c_k = \gamma k$

$a_k = \tau\, k \cdot (N-k)$

$\Longrightarrow$

probability of having k infected nodes at time t:

$$\begin{aligned}\dot{p}_k(t) &= a_{k-1} p_{k-1}(t) - (a_k + c_k)\, p_k(t) \\ &\quad + c_{k+1} p_{k+1}(t)\end{aligned}$$

- a_k is known and not random ⇒ exact master equation
- What is a_k in the general case?

Maximum Likelihood Estimation

We are going to learn the rates. Assumptions:

1. Infection and recovery are independent Poisson processes of rates a_k and c_k respectively

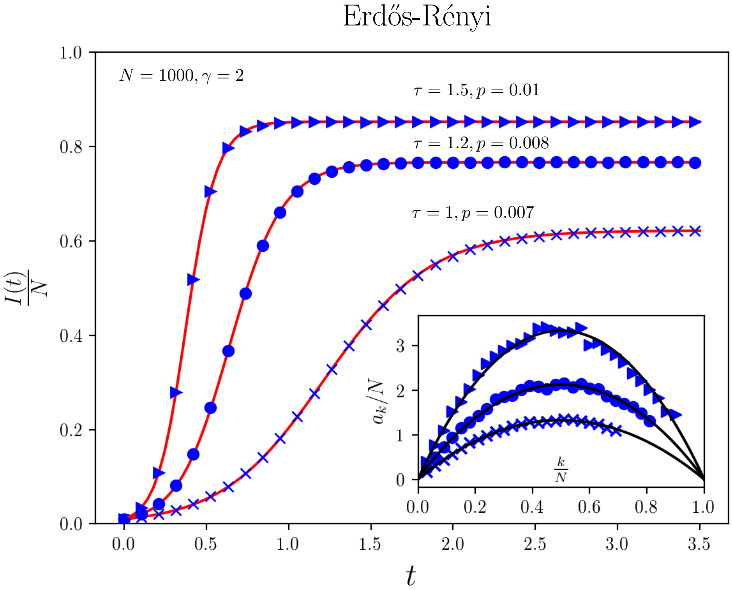
2. We have a sufficient statistics available (i.e. continuous observations)

Maximising the Likelihood [5] leads to:

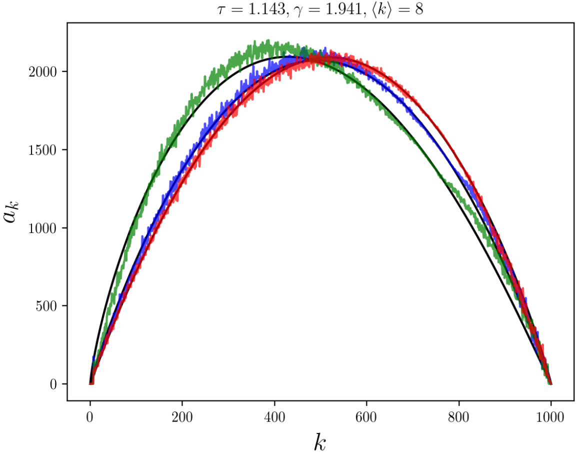
$$\max_{a_k, d_k} \mathcal{L}\left(a_k, c_k | \{Obs\}\right) = \begin{cases} \hat{a}_k = \frac{u_k}{t_k} \\ \hat{c}_k = \frac{d_k}{t_k} \end{cases}$$

u_k (d_k) is the number of up (down) jumps from k,
t_k is the time spent in k.

Is the approximation good?



how does that help?



Different network families produce different (k, a_k) curves!

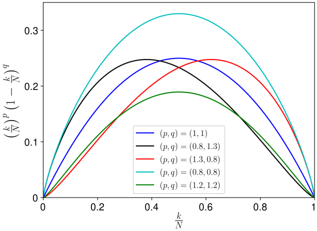
Fit proposal

$$a_k \sim a_k^\theta = Ck^p(N-k)^q$$

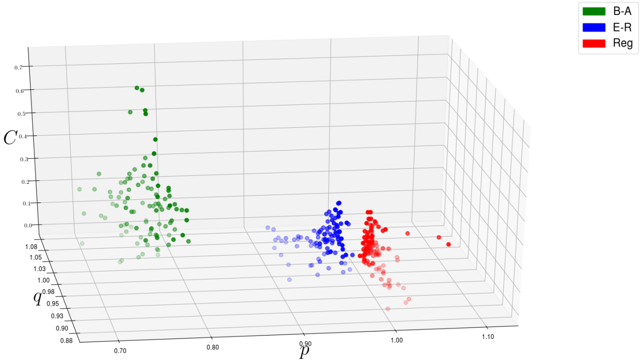
MLE:

$$\begin{cases} \dfrac{1}{C}\displaystyle\sum_{k=1}^{N-1}\left(u_k - \tau_k a_k^\theta\right) = 0 \\ \displaystyle\sum_{k=1}^{N-1} \log(k)(u_k - \tau_k a_k^\theta) = 0 \\ \displaystyle\sum_{k=1}^{N-1} \log(N-k)(u_k - \tau_k a_k^\theta) = 0 \end{cases}$$

Examples:

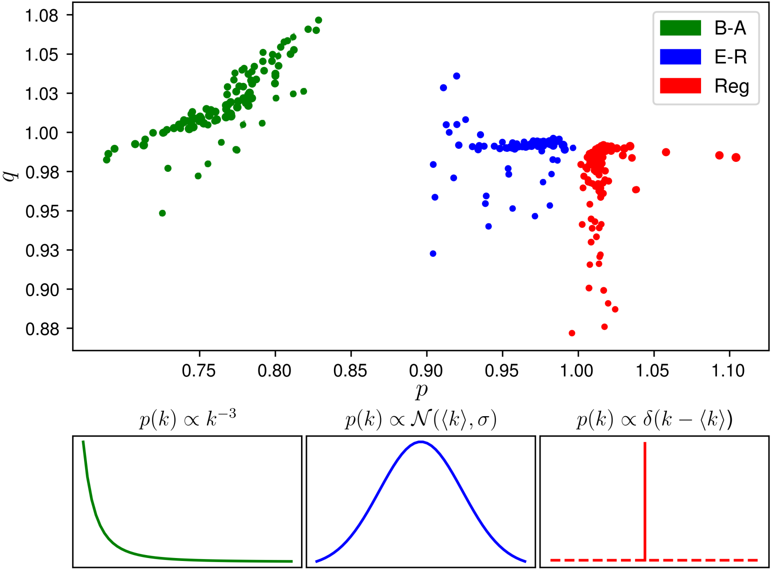


3-d plot



We concatenated 10^4 epidemics on 10^2 realisations for each point!

2-d plot



B-A
E-R
Reg
1.08
1.05
1.03
1.00
0.98
0.95
0.93
0.90
0.88
q
0.75
0.80
0.85
0.90
0.95
1.00
1.05
1.10
p
$p(k) \propto k^{-3}$
$p(k) \propto \mathcal{N}(\langle k \rangle, \sigma)$
$p(k) \propto \delta(k - \langle k \rangle)$

Towards inference based on discrete observations of the epidemic

- We used the Birth-Death process approximation of epidemics to characterise different families in the C, p, q space

- Now we will see how this helps when only discrete observations (in time) are available

Bayesian inference

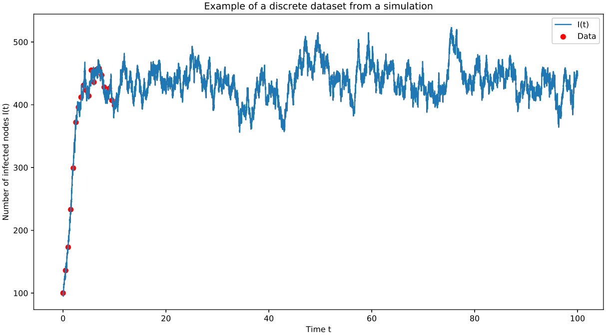
- Dataset: $D = \{I(t_1), I(t_2), \ldots, I(t_n)\} = \{k_1, k_2, \ldots, k_n\}$ discrete observations
- prior distributions using output from characterisation:

  Reg $\rightarrow (C, p, q) \sim \pi_{Reg}$
  E-R $\rightarrow (C, p, q) \sim \pi_{E-R}$
  B-A $\rightarrow (C, p, q) \sim \pi_{B-A}$

$$\mathcal{L}(C, p, q; D) = \prod_{i=2}^{n} \mathbb{P}(I(t_i - t_{i-1}) = k_i | I(0) = k_{i-1}, C, p, q)$$

- Maximum Posterior Probability $\pi(type|D)$ requires to numerically evaluate this integral(costly!): $\pi(D|type) = \int \mathcal{L}(C, p, q; D)\pi_{type}(C, p, q)dCdpdq.$

Results



Output probabilities:
$p_{Reg} = 0.01\%$, $p_{E-R} = 0.02\%$, $p_{B-A} = 99.7\%$

$$\Downarrow$$

it is a Barabási-Albert network (very likely)

To summarise

1. We introduced the Birth-Death Process approximation for SIS on networks
2. We tested that it is a good approximation
3. We proposed a model for the rates that led to characterisation

$\Downarrow$

Inference based on discrete observations of the epidemic

Conclusions

- It is possible to infer the network family from discrete observations of an epidemic

- Many directions are worth investigating:
  - More complex network models (clustering, communities,...)
  - Can we infer something more? ($\langle k \rangle$, τ)
  - Can we extend it to different dynamics? (SIR,SEIRS,$\ldots$)
  - Is there a better parametric model to fit the a_k curves?
  - Scaling with the size?
  - Real data?

- We hope that this method will become a new useful tool in network inference.

Bibliography

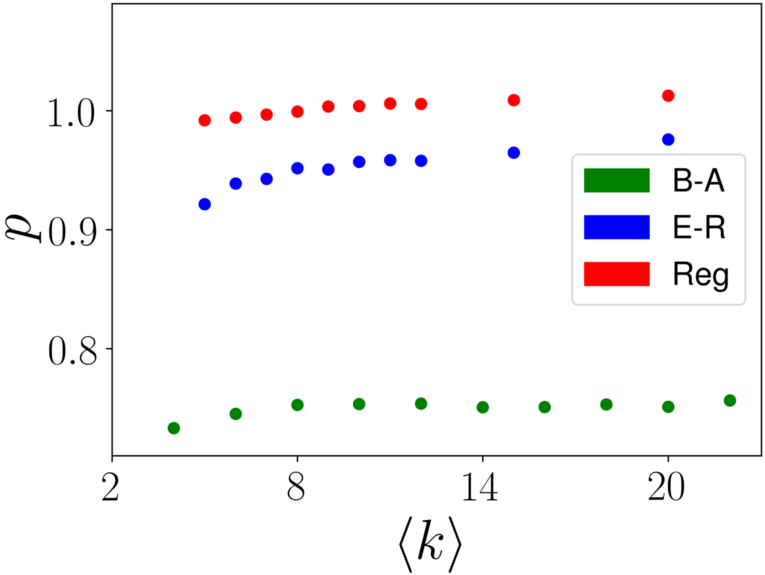
[1] Brugere, I. and Gallagher, B. and Berger-Wolf, T. Y. *Network Structure Inference, A Survey: Motivations, Methods, and Applications*, 2018, ACM Comput. Surv., 51, pp 1-39.

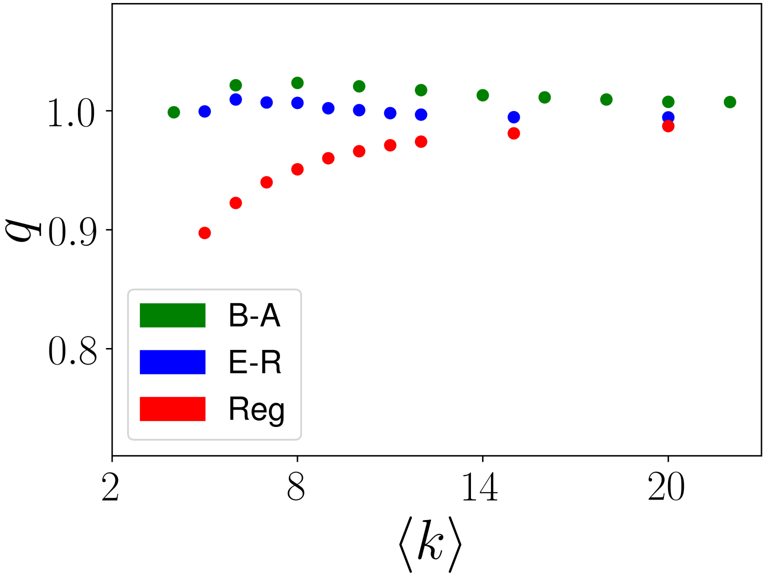
[2] Gomez Rodriguez, M. and Leskovec, J. and Krause, A. *Inferring Networks of Diffusion and Influence*, Proceedings of the 16th ACM SIGKDD, 2010, pp 1019-1028.

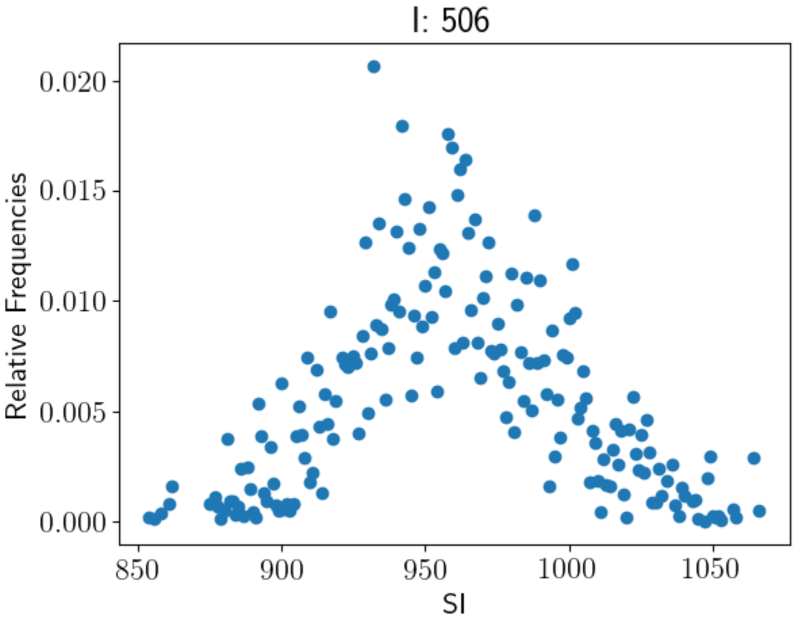
[3] Britton, T. and O'Neill, P. D. *Bayesian Inference for Stochastic Epidemics in Populations with Random Social Structure*, 2002, Scandinavian Journal of Statistics, pp 375-390.

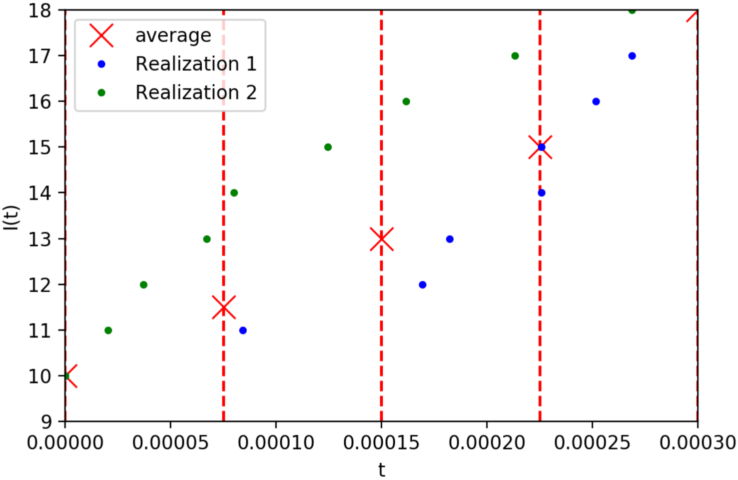
[4] Kiss, I. Z. and Miller, J. C. and Simon, P. L. *Mathematics of Epidemics on Networks: from exact to approximate models*,2017, Springer.

[5] Wolff, R. W.. *Problems of Statistical Inference for Birth and Death Queuing Models*,1965, Operations Research, pp 343-357.


B-A
E-R
Reg
p
1.0
0.9
0.8
2
8
14
20
$\langle k\rangle$


q
1.0
0.9
0.8
B-A
E-R
Reg
2
8
14
20
$\langle k \rangle$


I: 506
Relative Frequencies
0.020
0.015
0.010
0.005
0.000
850
900
950
1000
1050
SI


average
Realization 1
Realization 2
I(t)
t
18
17
16
15
14
13
12
11
10
9
0.00000
0.00005
0.00010
0.00015
0.00020
0.00025
0.00030