

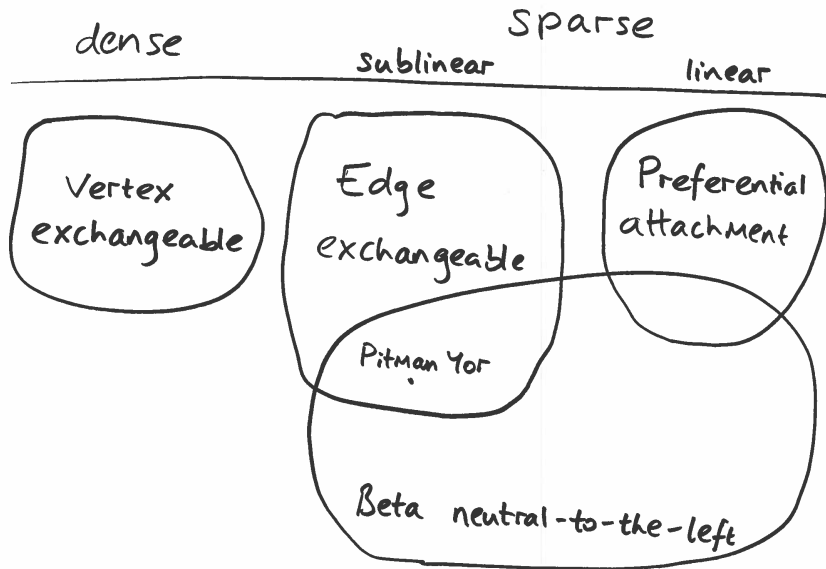
Sampling and Inference for Beta Neutral-to-the-Left Models of Sparse Networks

Benjamin Bloem-Reddy, Adam Foster, Emile Mathieu, Yee Whye Teh

Department of Statistics, University of Oxford



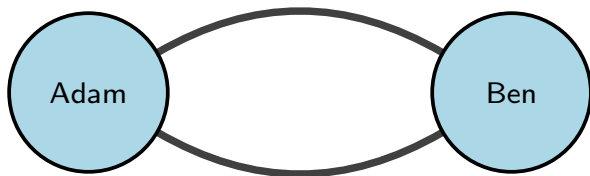
Models for sparse networks



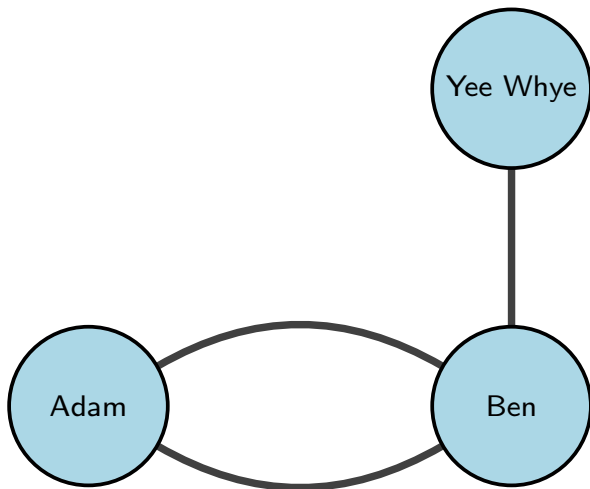
Temporal networks



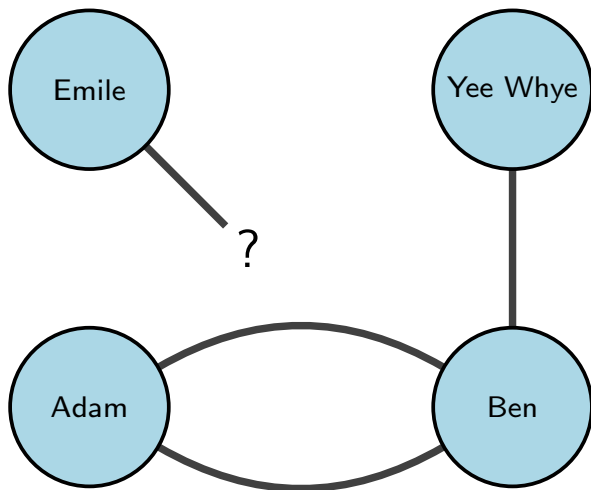
Temporal networks



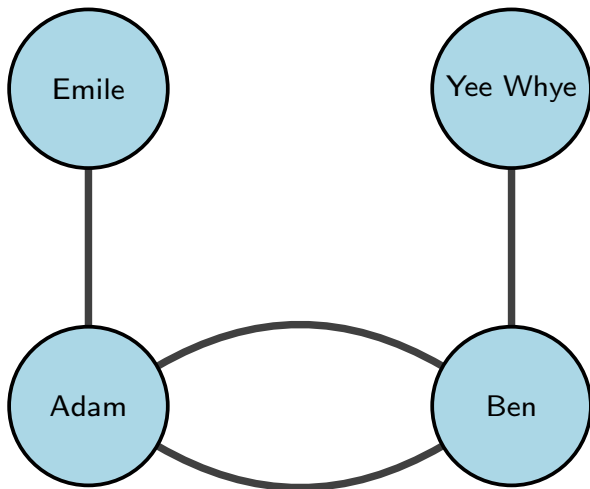
Temporal networks



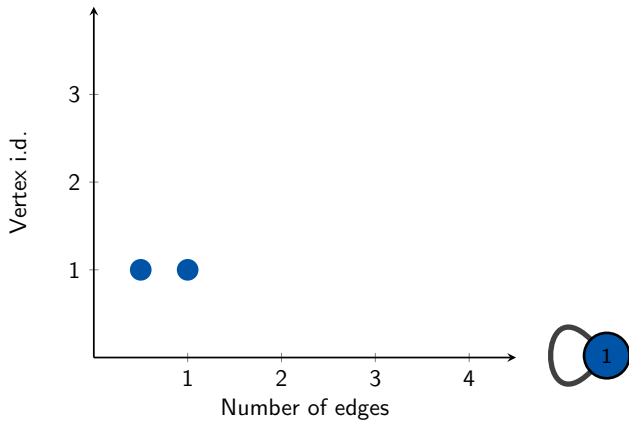
Temporal networks



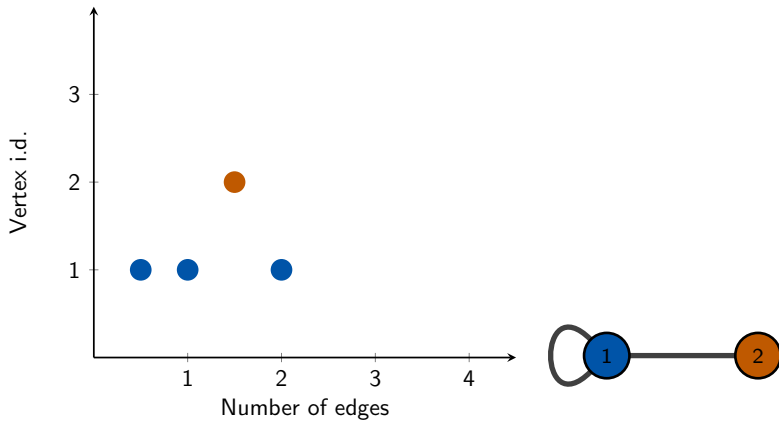
Temporal networks



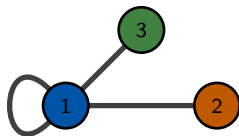
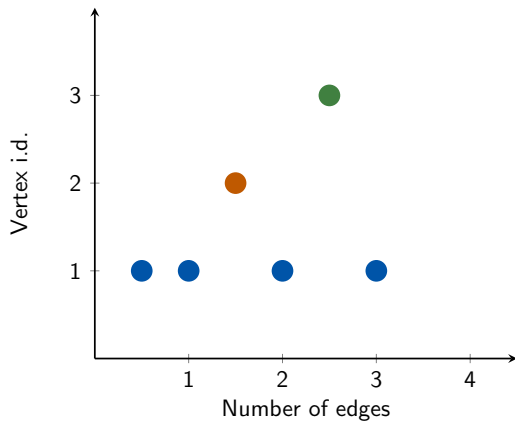
Edges and vertices



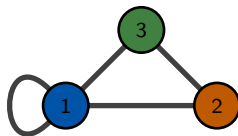
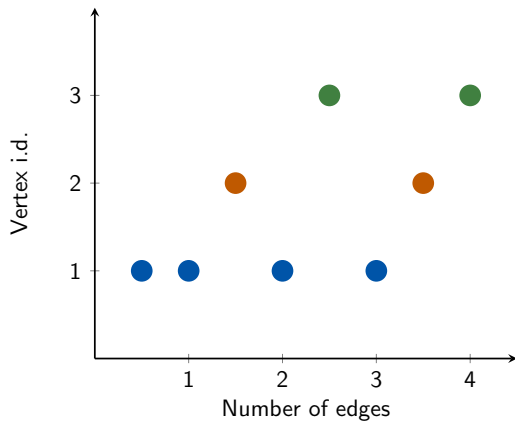
Edges and vertices



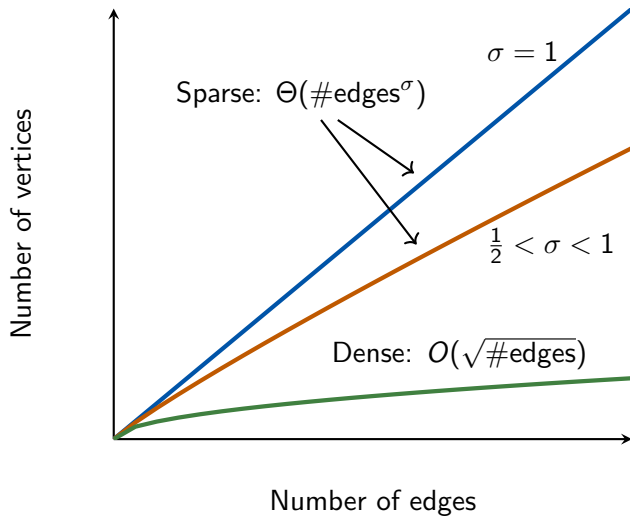
Edges and vertices



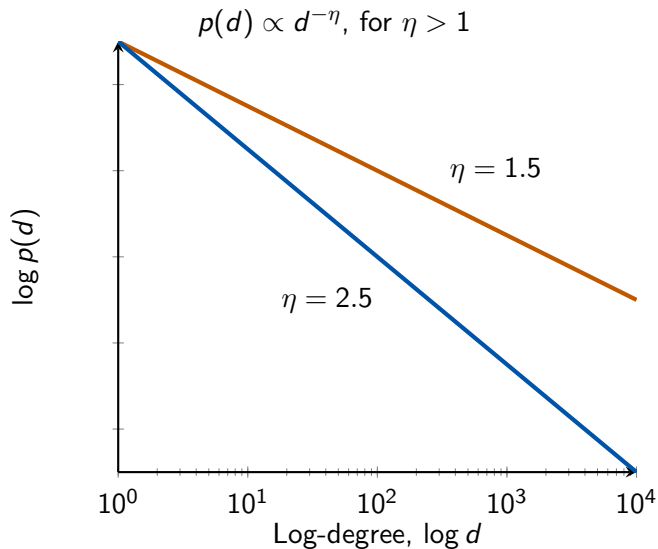
Edges and vertices



Sparsity



Power law degree distribution

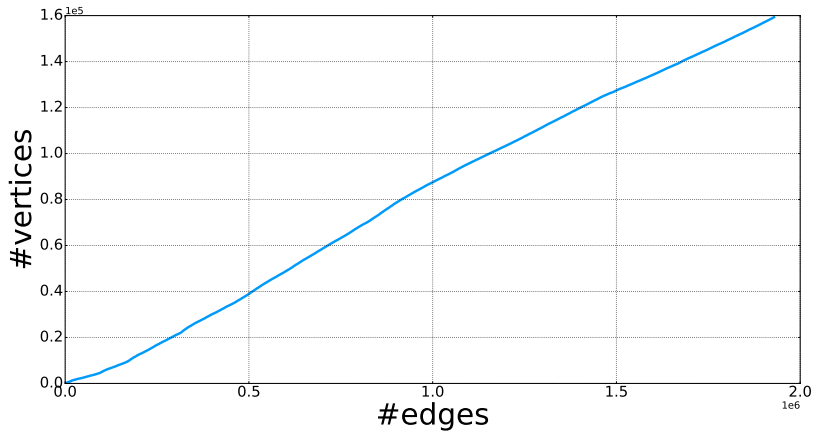


Sparsity and power law

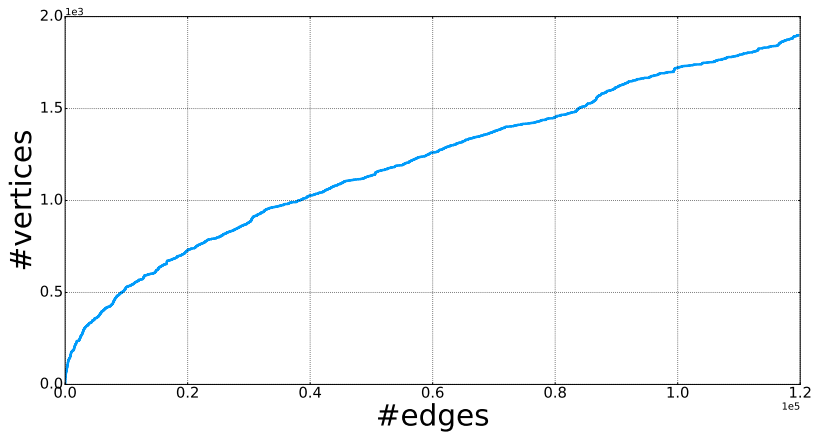
Sublinear sparsity $\iff \eta \in (1, 2)$

Linear sparsity $\iff \eta > 2$

Empirical study: Ask Ubuntu



Empirical study: UCI social network



Exchangeable models

A sequence of random variables X_1, X_2, \dots is *exchangeable* if for a finite permutation σ ,

$$X_{\sigma(1)}, X_{\sigma(2)}, \dots$$

has the same distribution as the original sequence.

Exchangeable models lead to tractable inference (de Finetti).

Models



Vertex
exchangeable

Models

dense

sparse

Vertex
exchangeable

Models

dense

sparse

Vertex
exchangeable

Preferential
attachment

Models

dense

sparse

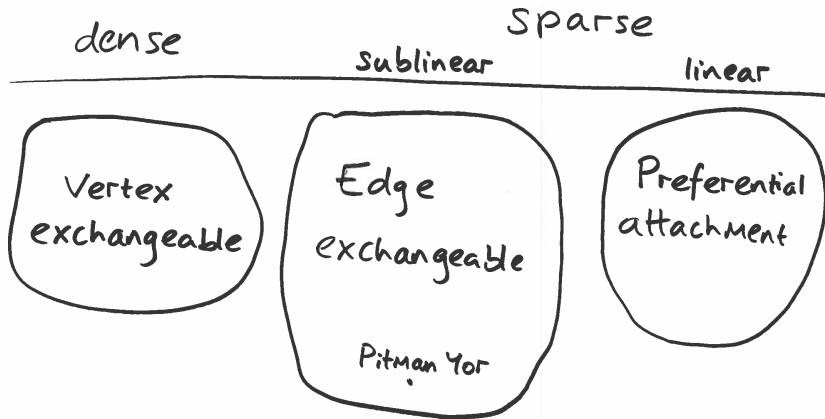
Vertex
exchangeable

Edge
exchangeable

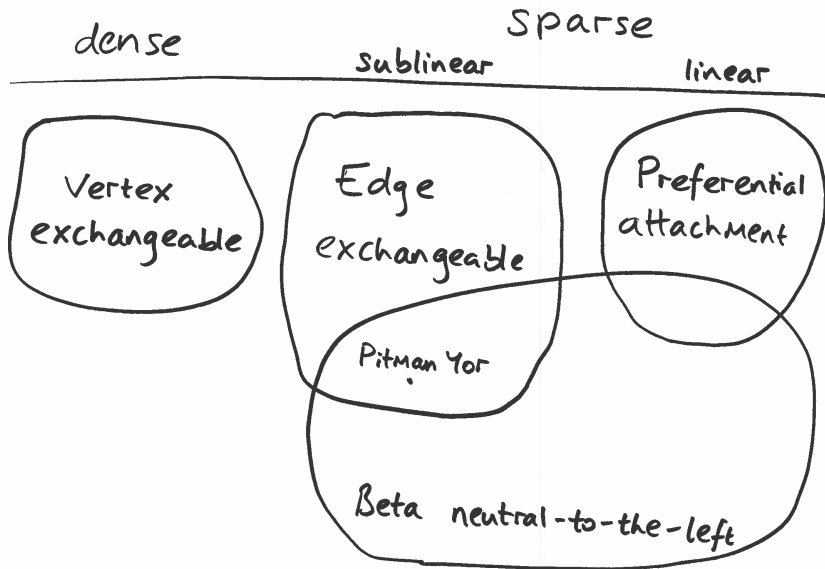
Pitman Yor

Preferential
attachment

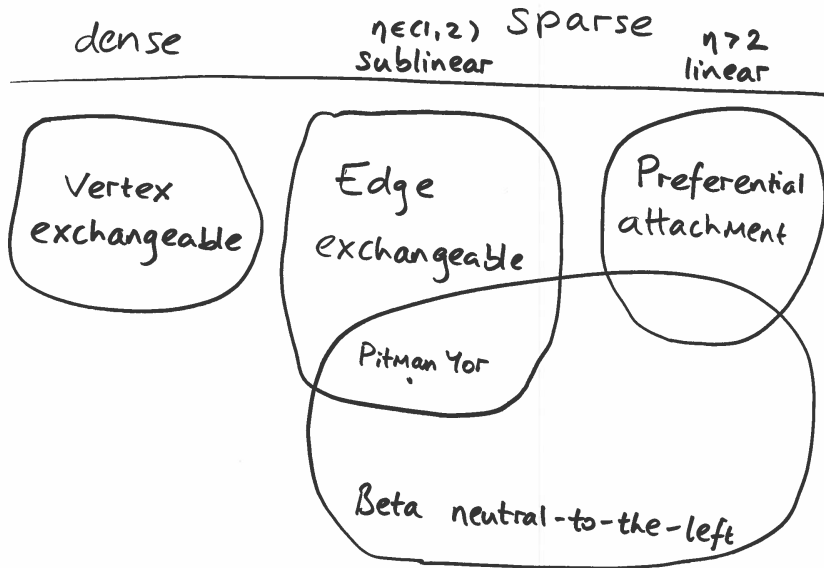
Models



Models

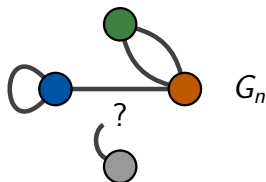


Models



Beta Neutral-to-the-left Model [4]

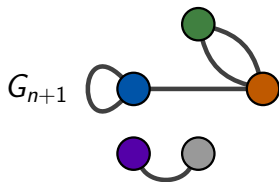
1. Generate arrival times $1 = T_1 < T_2 < T_3 < \dots$ in any way.
2. Generate ends of edges sequentially:



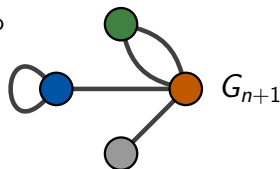
Is the next arrival time equal to $n + 1$?

yes

no



New vertex



$$\mathbb{P}[\rightarrow j] \propto \deg_{j,n} - \alpha$$

Sampling and inference

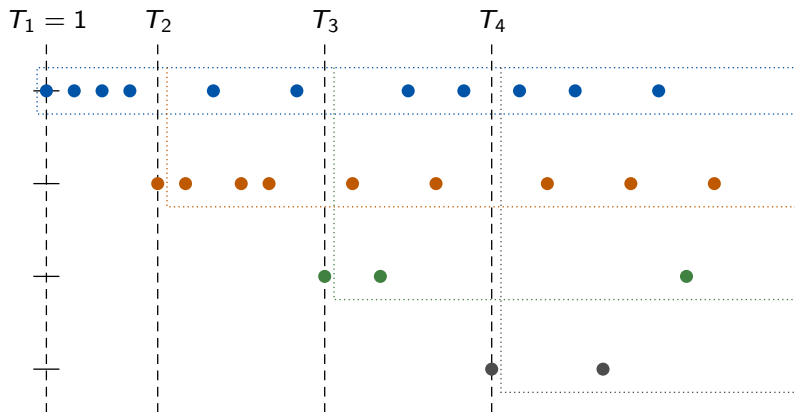
Why a paper on sampling and inference for BNTL models?

Sampling and inference

Why a paper on sampling and inference for BNTL models?

- ▶ Inference is notoriously difficult for *non-exchangeable* structures
- ▶ Need to identify *exchangeable substructures* for tractable inference

Exchangeable substructure



Gibbs structure

The joint probability has **Gibbs structure**

$$P(\text{graph} | T_1, T_2, \dots) = \prod_{j \in \text{vertices}} P(\text{choose } j \text{ } d_j - 1 \text{ times out of } T - T_j \text{ trials})$$

- ▶ d_j = degree of vertex j
- ▶ T = final time
- ▶ T_j = arrival time of vertex j

Available data

Observation	Unobserved variables
Entire history	Model variables
Degrees in arrival order	Model variables, arrival times
Snapshot	Model variables, arrival times, arrival order

Gibbs sampler

Variable	Gibbs sampling scheme
Model variables	Analytic using Gibbs structure
Arrival times	Update each arrival time separately
Arrival order	Initialise in descending degree order use M-H with adjacent swap proposal

Experiments

- ▶ Gibbs: parameter recovery
- ▶ Gibbs: scalability
- ▶ Point estimation with massive graphs

Parameter recovery

- Small graph
- Need to learn model variables, arrival times and arrival order

Gen. arrival distn.	Inference model	$ \hat{\alpha} - \alpha^* $	Pred. log-lik.
$\mathcal{PYP}(1.0, 0.75)$	$(\tau, \mathcal{PYP}(\theta, \tau))$	0.046 ± 0.002	-2637.0 ± 0.1
$\mathcal{PYP}(1.0, 0.75)$	$(\alpha, \text{Geom}(\beta))$	0.049 ± 0.004	-2660.5 ± 0.7
$\text{Geom}(0.25)$	$(\tau, \mathcal{PYP}(\theta, \tau))$	0.086 ± 0.002	-2386.8 ± 0.1
$\text{Geom}(0.25)$	$(\alpha, \text{Geom}(\beta))$	0.043 ± 0.003	-2382.6 ± 0.2

Scalability

- ▶ Runtime linear in $\#edges$
- ▶ Most expensive Gibbs update is for arrival times

MLEs for real data

Ask Ubuntu

- ▶ Edge exchangeable models misspecified
- ▶ Non-exchangeable BNTL provide better fit

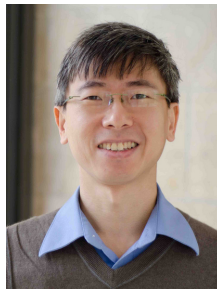
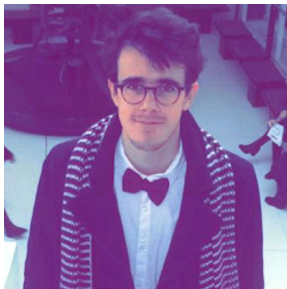
Conclusion

- ▶ BNTL models are *flexible*
- ▶ Inference was challenging due to non-exchangeability

Conclusion

- ▶ BNTL models are *flexible*
- ▶ Inference was challenging due to non-exchangeability
- ▶ BNTL models are *tractable* due to exchangeable substructure

Thank you



- [1] Jure Leskovec and Andrej Krevl. SNAP Datasets: Stanford large network dataset collection. <http://snap.stanford.edu/data>, June 2014.
- [2] Diana Cai, Trevor Campbell, and Tamara Broderick. Edge-exchangeable graphs and sparsity. In *Advances in Neural Information Processing Systems*, pages 4249–4257, 2016.
- [3] Harry Crane and Walter Dempsey. Edge exchangeable models for interaction networks. *Journal of the American Statistical Association*, 2017. In press.
- [4] Benjamin Bloem-Reddy and Peter Orbanz. Preferential attachment and vertex arrival times. *arXiv preprint arXiv:1710.02159*, 2017.