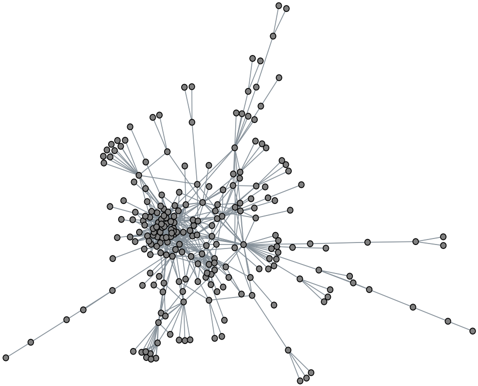
Random Walk Models of Sparse Graphs

Peter Orbanz · Columbia University

Collaborators

Benjamin Reddy · Columbia
Daniel M. Roy · Toronto
Balázs Szegedy · Alfréd Rényi Institute

GRAPH-VALUED DATA

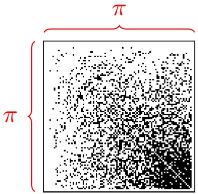


EXCHANGEABLE GRAPHS

random graph exchangeable

$\Updownarrow$

adjacency matrix jointly exchangeable



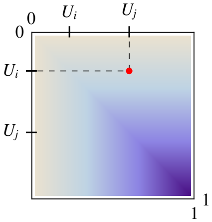
Representation theorem (Aldous, Hoover)

Any exchangeable graph can be sampled as:

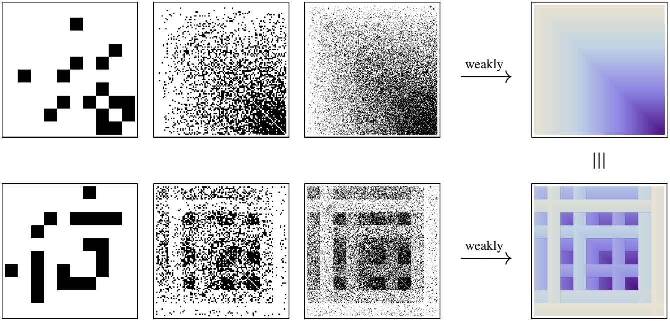
1. Sample measurable and symmetric function

$$\Theta : [0,1]^2 \longrightarrow [0,1]$$

2. Sample $U_1, U_2, \ldots \sim_{\text{iid}} \text{Uniform}[0,1]$
3. Sample edge i,j $\sim \text{Bernoulli}(\Theta(U_i, U_j))$



Law of large numbers (Kallenberg, ’99; Lovász & Szegedy, ’06)



Lifting Theorem (O. & Szegedy, 2011)

There exists a measurable mapping

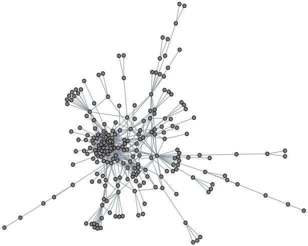
$$\xi : \widehat{\mathbf{W}} \longrightarrow \mathbf{W} \qquad \text{such that} \qquad \xi(\hat{\theta}) \in [\hat{\theta}]_{\equiv} \qquad \text{for all } \hat{\theta} \in \widehat{\mathbf{W}} \, .$$

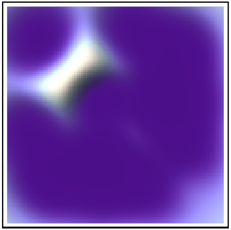
set of equivalence classes

APPLICATIONS IN STATISTICS

$$\text{Model} \subset \{P_\theta \mid \theta : [0,1]^2 \to [0,1] \text{ measurable } \}$$

Semiparametric	Estimate functional of θ	Bickel, Chen & Levina (2011)
Nonparametric (Bayesian)	Prior on θ	Lloyd et al (2012)
Nonparametric ML	ML estimator of θ	Wolfe & Olhede (2014)
		Gao, Lu & Zhou (2014)
Parametric	e.g. ERGMs	(too many)

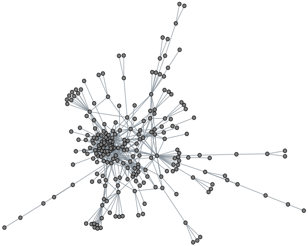


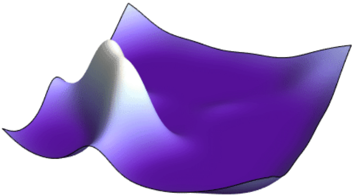


APPLICATIONS IN STATISTICS

$$\text{Model} \subset \{P_\theta \,|\, \theta : [0,1]^2 \to [0,1] \text{ measurable } \}$$

Semiparametric	Estimate functional of θ	Bickel, Chen & Levina (2011)
Nonparametric (Bayesian)	Prior on θ	Lloyd et al (2012)
Nonparametric ML	ML estimator of θ	Wolfe & Olhede (2014)
		Gao, Lu & Zhou (2014)
Parametric	e.g. ERGMs	(too many)





MISSPECIFICATION PROBLEM

Dense vs sparse

A random graph is called

- **dense** if $\#\text{edges} = \Omega((\#\text{vertices})^2)$.
- **sparse** if $\#\text{edges} = O(\#\text{vertices})$.

Most real-world graph data ("networks") is sparse.

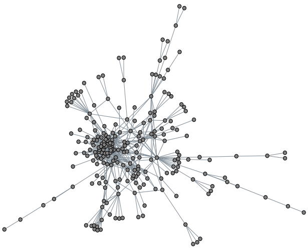
Exchangeable graphs are dense (or empty)

$$p = \int_{\triangledown} \Theta(x, y) dx dy \qquad \hat{p}_n = \frac{\#\text{ edges observed in } G_n}{\#\text{ edges in complete graph}} = p \cdot \Omega(n^2) = \Omega(n^2)$$

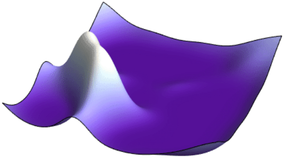
Sparsification (Bollobas & Riordan)

Generating sparse graphs: Choose decreasing rate function $\rho : \mathbb{N} \to \mathbb{R}_+$

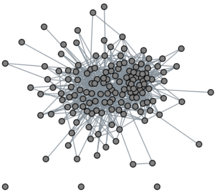
1. $U_1, U_2, \ldots \sim_{\text{iid}} \text{Uniform}[0, 1]$
2. $\text{edge}(i, j) \sim \text{Bernoulli}(\rho(\#\text{vertices})\theta(U_i, U_j))$



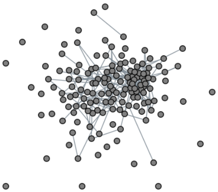
Input data



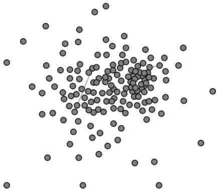
Estimate of parameter function



Exchangeable sampling
(rate $\rho(n) = 1$)



Rate $\rho(n) = \frac{1}{\log(n)}$



Rate $\rho(n) = \frac{1}{n}$

Sampling with random walks

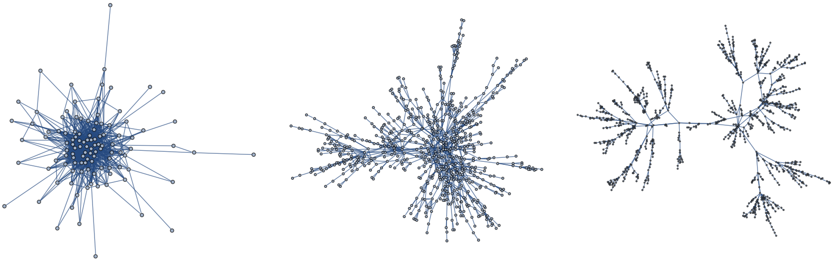
- Independent subsampling only adequate in dense case
- Sparse case: Need to follow existing link structure
- This means edges are not conditionally independent

A simple random walk model

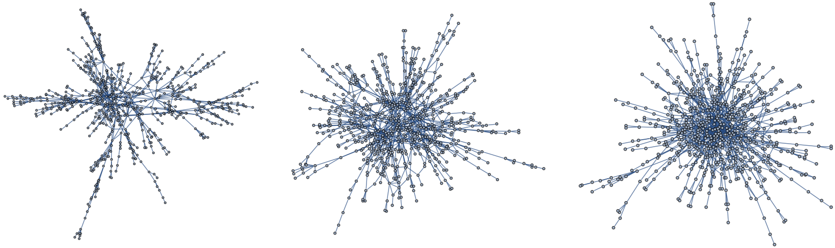
For $n = 1, \ldots, \#$edges:

1. Choose a vertex v_n uniformly at random from the current graph.
2. With probability α, create a new vertex attached to v_n.
3. With probability $1 - \alpha$: Start a simple random walk at v_n.
   Stop after $\text{Poisson}(\lambda) + 1$ steps.
   Connect v_n to terminal vertex.

PARAMETER EFFECTS



Increasing probability of new vertex →



Increasing probability of long random walk →

Probability of random walk in graph G

degree matrix

graph Laplacian

$$\mathbb{P}(v \to u | G, \lambda) = \Big(\sum_{k=0}^{\infty} \underbrace{e^{-\lambda} \frac{\lambda^k}{k!}}_{\text{Poisson weight}} (D^{-1} A)^{k+1} \Big)_{vu} = \big((\mathbf{I} - \Delta) e^{-\lambda \Delta} \big)_{vu}$$

adjacency matrix

Parameter estimation

- **Case 1: History observed**

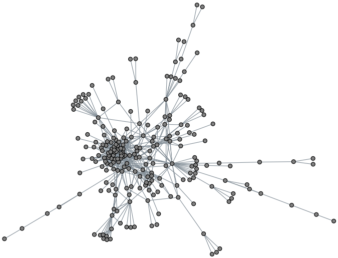
$$P(G_N | \alpha, \lambda) = \prod_{n=1}^{N-1} P(G_{n+1} | G_n, \alpha, \lambda)$$

  Maximum likelihood estimator derivable from random walk probability.

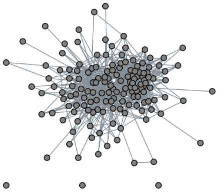
- **Case 2: Only final graph observed**
  Impute history by sampling. Importance sampling possible.

Results: Cartoon

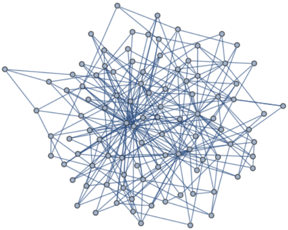
Input data



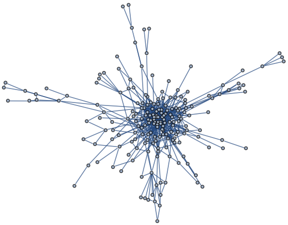
Reconstruction



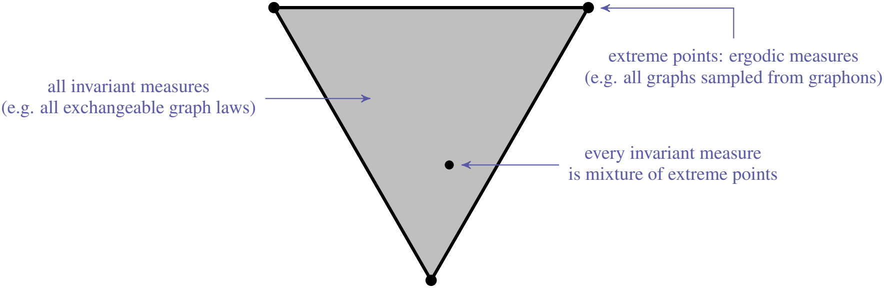
exchangeable model



preferential attachment model



random walk model



Consequences

- Representation theorem (Aldous-Hoover, de Finetti, etc).
- Invariant statistical models are subsets of extreme points.
- Law of large numbers: Convergence to extreme points.

Open problem

What is a statistically useful notion of invariance in sparse graphs?

CONCLUSIONS

Approach

Statistical theory of graphs as statistical theory of random structures.

What we have

- Exchangeable graphs: Some technical problems aside, statistics carries over.
- "Dependent" edge case:
  - Inference possible in principle.
  - No general theory.
  - Excellent prediction results with very simple model.
  - We cannot prove much about the model.

Well-studied invariance properties

1. Exchangeability: Only dense case.
2. Involution invariance: Too weak for statistical purposes.