Edge-exchangeable graphs and sparsity

Motivation
- The traditional notion of exchangeability for random graphs is vertex exchangeability.
- Vertex-exchangeable graphs are dense or empty almost surely.
- But most real-world graphs are sparse.
- We introduce the alternative notion of edge exchangeability, and show that a wide class of edge-exchangeable models can produce sparse graphs.

Projective graph sequences: We consider graph sequences $G_1, G_2, G_3, \ldots$, where at each step of the sequence, we only add vertices and edges to the graph (instead of deleting edges).

Dense graph sequence: $[\#\text{edges}(G_n)] = \Omega((\#\text{vertices}(G_n))^2)$ (quadratic growth)

Sparse graph sequence: $[\#\text{edges}(G_n)] = o((\#\text{vertices}(G_n))^2)$ (sub-quadratic growth)

Vertex exchangeability
In vertex exchangeability, a new vertex joins the graph sequence at each step, is labeled with that step number, and instantiates any new vertices in that edge along the way.

Example realization:

1. Draw edge weights $w_{(i,j)}$ from some distribution
2. Draw edge $(i, j)$ with probability proportional to $w_{(i,j)}$
3. Discrete random measure: $W = \sum_{(i,j)} w_{(i,j)} \delta_{(i,j)}$
4. Edges drawn conditionally iid given $W$
5. Imples edge exchangeability

Graph frequency models
1. Single edge per step:
   - Draw edge weights $w_{(i,j)}$ from some distribution
   - Draw edge $(i, j)$ with probability proportional to $w_{(i,j)}$
   - Discrete random measure: $W = \sum_{(i,j)} w_{(i,j)} \delta_{(i,j)}$
   - Edges drawn conditionally iid given $W$

2. Multiple edges per step:
   - Draw edge weights $w_{(i,j)}$ from some distribution
   - Draw edges $(i, j)$ with probability proportional to $w_{(i,j)}$
   - Discrete random measure: $W = \sum_{(i,j)} w_{(i,j)} \delta_{(i,j)}$
   - Edges drawn conditionally iid given $W$

Sparsity results
Consider a graph frequency model with edge frequencies $w_{(i,j)} = w_{ij}$ and $w_{ij} \sim \text{Poisson Process}(\lambda)$, and the rate measure $\lambda$ has power law tails:

$$\int_{x} \lambda^{\alpha} e^{-\lambda x} d\lambda = \frac{\Gamma(1+\alpha)}{\Gamma(1+\alpha) \alpha} \quad \forall \alpha > 0, \lambda \in [0, \infty)$$

$\alpha$ is regularly varying with exponent $\alpha$.

Generate a multigraph $G_{(i,j)}$ and also construct binary graph:

- Multigraph
- Binary graph

Theorem: Suppose the rate measure $\lambda$ is regularly varying with exponent $\alpha$. Let $\ell : \mathbb{R}_+ \rightarrow \mathbb{R}_+$ be a slowly varying function.

$\# \text{ of instantiated vertices}$: $\Theta(n^a \ell(n))$ a.s.

Multigraph $\# \text{ of edges}$: $\Theta(n)$ a.s.

Binary graph $\# \text{ of edges}$: $O(\min\{n^{1+\alpha}/2, n^{\alpha/2} \ell(n) \ell(n/2)\})$ a.s.

Thus, the multigraph is sparse when $1/2 < \alpha < 1$.

Simulations
We generate the weights $w$ from a Poisson point process with the 3-parameter beta process rate measure:

$$\nu(dw) = \frac{\Gamma(1+b)}{\Gamma(1+a) \Gamma(a+b)} w^{a-1} (1-w)^{b-1} dw$$

Empirically, we get a range of dense and sparse behavior that agrees with the above theorem.

References