

## Reading Group on Stochastic Modelling

# A Structural Model of Dense Network Formation

Mele (2017)

*Econometrica*





## Literature






### Related works

- ▶ similar paper [6]
- ▶ extensions [11], [3], [9]

### Model

-  network formation models [1]
-  ERGM theory [7], [15]

### Asymptotics

-  many networks asymptotics [14]
-  large network asymptotics [8], [12], [10]
  -  graph limits [2]
  -  large deviations [5]
  -  variational methods for the exponential family [16]

### Estimation

-  ERGM estimation [15], [4]

## Introduction

**Strategic models of network formation** provide a framework to interpret the observed network as the equilibrium of a game.

## Introduction

**Strategic models of network formation** provide a framework to interpret the observed network as the equilibrium of a game.

Estimation and identification of strategic models is challenging

- 1) **multiple equilibria**  $\Rightarrow$  links generate externalities not fully accounted for by agents
- 2) **curse dimensionality**  $\Rightarrow$  # network configs grows exponentially with # agents
- 3) **data on single graph**  $\Rightarrow$  only one network snapshot is observable

## Introduction

**Strategic models of network formation** provide a framework to interpret the observed network as the equilibrium of a game.

Estimation and identification of strategic models is challenging

- 1) **multiple equilibria**  $\Rightarrow$  links generate externalities not fully accounted for by agents
- 2) **curse dimensionality**  $\Rightarrow$  # network configs grows exponentially with # agents
- 3) **data on single graph**  $\Rightarrow$  only one network snapshot is observable

### Proposed model of network formation

- ❖ combines features from the strategic and random network formation literature
- ❖ players' utilities depend on payoffs from direct links and link externalities (e.g., reciprocity, indirect friends, popularity, . . . )
- ❖ network formation is dynamic: each period, a player meets another one and decides whether to form a new link, keep an existing link, or do nothing
- ❖ process generates a sequence of directed dense graphs

## Model of Network Formation

### Setup

- $n$  agents, with characteristics  $X_i \in \mathbb{R}^A \quad \forall i \in \mathcal{I} := \{1, \dots, n\}$
- discrete time  $t \in \mathbb{N}$
- directed, binary network  $G \in \mathcal{G}$ , realisations each time  $g^t$
- utility of network = sum utility from links, each having four components

$$U_i(g, X|\theta) = \underbrace{\sum_{j=1}^n g_{ij} u_{ij}^{\theta_u}}_{\text{direct links}} + \underbrace{\sum_{j=1}^n g_{ij} g_{ji} m_{ij}^{\theta_m}}_{\text{mutual links}} + \underbrace{\sum_{j=1}^n g_{ij} \sum_{\substack{k=1 \\ k \neq i,j}}^n g_{jk} v_{ik}^{\theta_v}}_{\text{indirect links}} + \underbrace{\sum_{j=1}^n g_{ij} \sum_{\substack{k=1 \\ k \neq i,j}}^n g_{ki} w_{kj}^{\theta_w}}_{\text{popularity}}$$

where  $u_{ij}^{\theta_u} := u(X_i, X_j|\theta_u)$ ,  $m_{ij}^{\theta_m} := m(X_i, X_j|\theta_m)$ ,  $v_{ij}^{\theta_v} := v(X_i, X_j|\theta_v)$ ,  
 $w_{ij}^{\theta_w} := w(X_i, X_j|\theta_w)$

**Assumption 1. (Preferences)**

$$m_{ij}^{\theta_m} = m(X_i, X_j | \theta_m) = m(X_j, X_i | \theta_m) = m_{ji}^{\theta_m} \quad \forall i, j \in \mathcal{I} \times \mathcal{I}$$

$$w_{ij}^{\theta_v} = w(X_k, X_j | \theta_v) = v(X_k, X_j | \theta_v) = v_{ji}^{\theta_v} \quad \forall k, j \in \mathcal{I} \times \mathcal{I}$$

- ▶ first is necessary for identification of the utility from indirect links and popularity;
- ▶ second makes another agent  $i$  internalise the externality she creates.

**Proposition 1 (Existence Potential Function).**

*Under Assumption 1, the deterministic components of the incentives of any player in any state of the network are summarized by a **potential function**  $\mathcal{Q} : \mathcal{G} \times \mathcal{X} \rightarrow \mathbb{R}$  and the network game is a Potential Game*

$$\mathcal{Q}(g, X | \theta) = \sum_{i=1}^n \sum_{j=1}^n g_{ij} u_{ij}^{\theta_u} + \sum_{i=1}^n \sum_{j>1} g_{ij} g_{ji} m_{ij}^{\theta_m} + \sum_{i=1}^n \sum_{\substack{j=1 \\ j \neq i}}^n \sum_{\substack{k=1 \\ k \neq i, j}}^n g_{ij} g_{jk} v_{ik}^{\theta_v}$$

$\mathcal{Q}$  is an aggregate function summarising: (i) state of network; (ii) deterministic incentives of players in each state.

## Network formation process

Process follows a **stochastic best-response dynamics**, that generates a Markov chain of networks

- for each  $t$ , randomly chosen player  $i$  meets  $j$  according to meeting technology
- *meeting process* is a stochastic sequence  $\mathbf{m} = \{m^t\}_t$  supported on  $\mathcal{I} \times \mathcal{I}$ , with realisations  $m^t = ij = \{i, j\}$

$$\mathbb{P}(m^t = ij | g^{t-1}, X) = \rho(g^{t-1}, X_i, X_j)$$

### Assumption 2. (Meeting process)

The meeting probability between  $i, j$  does not depend on the existence of a link between them, and each meeting has a positive probability of occurring, that is

$$\rho(g^{t-1}, X_i, X_j) = \rho(g_{-ij}^{t-1}, X_i, X_j) > 0 \quad \forall ij$$

- ▶ guarantees any equilibrium network can be reached with positive probability
- ▶ identification: allow  $\rho$  to depend on current link  $g_{ij}$  prevents closed form likelihood



## Utility

### Players' rules

- conditional on meeting  $m^t = ij$ , player  $i$  updates link  $g_{ij}$  to maximise her utility
- existing network  $g_{-ij}^{t-1}$  is taken as given
- **complete information**: everybody known each others' attributes and whole network
- **myopia**: agents not account for effects of their linking strategy on future evolution of network
- **idiosyncratic shock** on individual preferences:  $\varepsilon \sim EV_1(\varepsilon)$  Type I extreme value distribution, *iid* among links and across time
- link established if and only if

$$U_i(g_{ij}^t = 1, g_{-ij}^{t-1}, X|\theta) + \varepsilon_{1t} > U_i(g_{ij}^t = 0, g_{-ij}^{t-1}, X|\theta) + \varepsilon_{0t}$$

Process generates a Markov chain of networks:

- ✓ transition proba determined by: (i) meeting process, (ii) agents' linking choices
- ✓ irreducible, aperiodic

## Equilibrium

### Remark 1.

Any change in utility for any agent is equivalent to change in potential  $\mathcal{Q}$ . So, any deviation from Nash (equilibrium) network must decrease the potential.

Thus, the **Nash network** is a **local maximizer** of the **potential function** over the set of networks that differ from the current network for at most one link.

### Theorem 1 (Uniqueness and Characterisation of Stationary Equilibrium).

*The network formation game, under Assumptions 1–3, converges to a **unique stationary distribution***

$$\pi(g, X|\theta) = \frac{\exp\{\mathcal{Q}(g, X|\theta)\}}{\sum_{\omega \in \mathcal{G}} \exp\{\mathcal{Q}(\omega, X|\theta)\}} \quad (1)$$

## Comments

- existence and uniqueness come from irreducibility and aperiodicity of Markov chain
- uniqueness crucial for estimation (avoid multiple equilibria)
- closed form **stationary**  $\pi(g, X|\theta)$  corresponds to the **likelihood** of observing a specific network configuration in the long run  
 $\implies$  can **estimate**  $\theta$  with **only one network**, assuming it is drawn from the stationary equilibrium
- $\pi(g, X|\theta)$  coincides with **likelihood** of **ERGM** (Exponential Random Graph Model), where probability observing a network is proportional to exponential of linear combination of network statistics

### Corollary 1.1.

Let Assumptions 1–3 hold. If the utility functions are linear in parameters, the stationary distribution  $\pi(g, X|\theta)$  describes an ERGM, with  $\mathbf{t}(g, X)$  a vector of canonical statistics

$$\pi(g, X|\theta) = \frac{\exp\{\theta' \mathbf{t}(g, X)\}}{\sum_{\omega \in \mathcal{G}} \exp\{\theta' \mathbf{t}(\omega, X)\}} \quad (2)$$

## Extensions

### Utility functions

👉 possible to include **additional utility components**, as long as possible to find **restrictions on payoffs** that guarantee the existence of a potential function

### Undirected networks

👉 possible to extend existence results, characterisation of equilibrium, relation with ERGM and asymptotic results to **undirected networks**

### Sparsity

👉 model with **negative linking externalities** is compatible with a certain degree of sparsity

## Estimation and Identification

### Likelihood function

$$L(g, X|\theta) = \pi(g, X|\theta) = \frac{Q(g, X|\theta)}{\sum_{\omega \in \mathcal{G}} Q(\omega, X|\theta)} = \frac{Q(g, X|\theta)}{c(\mathcal{G}, X, \theta)}$$

whose normalizing constant  $c(\mathcal{G}, X, \theta)$  is **intractable** since it sums  $2^{n(n-1)}$  terms.

✗ standard ML infeasible

✗ MCMC with standard MH step infeasible (ratio of normalizing constants)

## Estimation Algorithm

ERGM literature  $\Rightarrow$  approximate  $c(\mathcal{G}, X, \theta)$  via MCMC (for fixed  $\theta_0$ )

---

### Algorithm 1 Metropolis-Hastings for Network Simulations

---

- 1: **procedure** MH\_NETSIM( $\theta_0, g_0, R$ )
- 2:   **for**  $r = 1, \dots, R$  **do**
- 3:     1) propose network  $g' \sim q_g(g' | g^{(r)})$
- 4:     2) accept network  $g'$  with probability

$$\alpha(g^{(r)}, g') = \min \left\{ 1, \frac{\exp\{Q(g', X | \theta_0)\} q_g(g^{(r)} | g')}{\exp\{Q(g^{(r)}, X | \theta_0)\} q_g(g' | g^{(r)})} \right\}$$

- 5:   **end for**
  - 6:   **return** sequence of  $R$  networks  $\{g^{(r)}\}_r$
  - 7: **end procedure**
- 

✓ not requires  $c(\mathcal{G}, X, \theta)$

✗ *slow convergence*

✗ *local sampler* at each iteration, update link  $g_{ij}$  according to  $\alpha(\cdot, \cdot)$

✗ *degeneracy* problem: large probability mass on few networks

## Estimation Algorithm

ERGM literature  $\Rightarrow$  approximate  $c(\mathcal{G}, X, \theta)$  via MCMC (for fixed  $\theta_0$ )

---

### Algorithm 1 Metropolis-Hastings for Network Simulations

---

```
1: procedure MH_NETSIM( $\theta_0, g_0, R$ )  
2:   for  $r = 1, \dots, R$  do  
3:     1) propose network  $g' \sim q_g(g' | g^{(r)})$   
4:     2) accept network  $g'$  with probability
```

$$\alpha(g^{(r)}, g') = \min \left\{ 1, \frac{\exp\{Q(g', X | \theta_0)\} q_g(g^{(r)} | g')}{\exp\{Q(g^{(r)}, X | \theta_0)\} q_g(g' | g^{(r)})} \right\}$$

```
5:   end for  
6:   return sequence of  $R$  networks  $\{g^{(r)}\}_r$   
7: end procedure
```

---

► how to choose  $q_g(\cdot | g^{(r)})$ ?

## Asymptotic behaviour

Classes of asymptotics for networks

- 1) *many networks*  $\Rightarrow$  same players, growing number of **networks**
- 2) *large networks*  $\Rightarrow$  growing number of **players**, same network

- Hp: **homogeneous** players (i.e.  $X_i = X_j, \forall i, j$ )
- potential function re-scaled by  $n^{\nu(H)}$ , with  $\nu(H)$  # players in each utility term
- example **re-scaled likelihood**, with  $\mathcal{T}(g) = \alpha t(H_1, g) + \beta t(H_2, g)$  re-scaled potential and  $\psi_n = n^{-2} \log(c(\alpha, \beta, \mathcal{G}_n))$  log-normalising constant

$$\begin{aligned} \pi_n(g|\alpha, \beta) &= \frac{\exp \left\{ n^2 \left[ \alpha \frac{\sum_{i=1}^n \sum_{j=1}^n g_{ij}}{n^2} + \beta \frac{\sum_{i=1}^n \sum_{j=1}^n \sum_{k \neq i}^n g_{ij} g_{jk}}{n^3} \right] \right\}}{c(\alpha, \beta, \mathcal{G}_n)} \\ &= \exp \left\{ n^2 [\mathcal{T}(g) - \psi_n] \right\} \end{aligned} \quad (3)$$



## Asymptotic behaviour

### Theorem 2 (Nonnegative Link Externalities).

Model (3) with *nonnegative link externalities*  $\beta \geq 0$  exhibits the following behaviour

1) *asymptotic normalizing constant*  $\psi$  solves

$$\psi = \lim_{n \rightarrow \infty} \psi_n = \max_{\mu \in [0,1]} \left\{ \alpha\mu + \beta\mu^2 - \mu \log(\mu) - (1 - \mu) \log(1 - \mu) \right\} \quad (4)$$

2) networks generated by the model are *indistinguishable from directed Erdős–Rényi graph* with linking probability  $\mu^*$ , defined as follows:

(a) if the maximization (4) has a unique solution, then  $\mu^*$  satisfies  $2\beta\mu(1 - \mu) < 1$  for almost all  $\alpha \in \mathbb{R}$  and  $\beta \geq 0$ , and solves

$$\mu = \frac{\exp\{\alpha + 2\beta\mu\}}{1 + \exp\{\alpha + 2\beta\mu\}} \quad (5)$$

(b) if the maximization (4) has two solutions, then  $\mu^*$  picked randomly from same probability distribution over  $\mu_1^*$  and  $\mu_2^*$ , such that  $\mu_1^* < 0.5 < \mu_2^*$ , and both solve (5) and satisfy  $2\beta\mu(1 - \mu) < 1$ .

## Comments on Theorem 2

- ▶ consistent estimator of log-normalising constant – analogue of variational representation of the discrete exponential family
- ▶  $\beta \geq 0 \Rightarrow$  realisations using  $(\alpha, \beta)$  indistinguishable from those using  $(\alpha', 0) = (\log(\mu^*/(1 - \mu^*)), 0)$ , that is from Erdős-Rényi model

### Corollary 2.1.

When  $\beta \geq 0$ , the **externality cannot be identified**.

### Corollary 2.2.

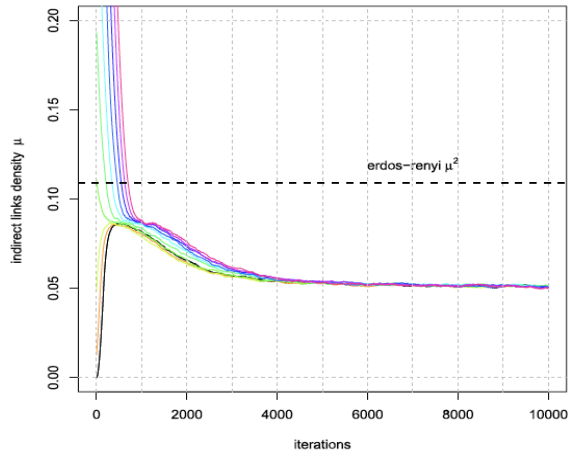
When  $\beta \geq 0$ , Algorithm 1 is not necessary since Erdős-Rényi graphs can be simulated using Bernoulli draws.

## Asymptotic behaviour

### Theorem 3 (Negative Link Externalities).

If  $\beta < 0$  and sufficiently large in magnitude, model (3) is asymptotically different from a directed Erdős-Rényi model.

✓ sparser graphs than Erdős-Rényi



## Asymptotic behaviour

### Theorem 3 (Negative Link Externalities).

*If  $\beta < 0$  and sufficiently large in magnitude, model (3) is asymptotically different from a directed Erdős-Rényi model.*

► how much “sufficiently large” magnitude?

► how to know it, since we must estimate  $\beta$ ?

## Asymptotic behaviour

Consider an additional utility component (cyclic triangles):

$$T(g) = \alpha t(H_1, g) + \beta t(H_2, g) + \gamma t(H_3, g), \quad t(H_3, g) = \sum_{i=1}^n \sum_{j=1}^n \sum_{k \neq i}^n g_{ij} g_{jk} g_{ki} \quad (6)$$

### Theorem 4.

Consider model (6) as  $n \rightarrow \infty$

1) If  $\beta \geq 0$  and  $\gamma \geq 0$ , then the asymptotic normalising constant  $\psi$  solves

$$\psi = \lim_{n \rightarrow \infty} \psi_n = \max_{\mu \in [0,1]} \left\{ \alpha \mu + \beta \mu^2 + \gamma \mu^3 - \mu \log(\mu) - (1 - \mu) \log(1 - \mu) \right\} \quad (7)$$

and model is asymptotically indistinguishable from directed Erdős-Rényi graph, with  $\mu^*$  maximising (7). If the maximisation problem has multiple solutions, then  $\mu^*$  picked randomly from some distribution on maximisers.

2) If at least **one externality is negative** (i.e.  $\beta < 0$  or  $\gamma < 0$ ) and sufficiently large, then model (6) not converge asymptotically to directed Erdős-Rényi graph and **externalities can be identified**.

## Summary of Asymptotics

### Remark 2.

**Homogeneous** players ( $X_i = X_j, \forall i, j$ ):

(a) positive externalities

- asymptotically indistinguishable from Erdős-Rényi graph
- externalities not identified
- can **approximate likelihood** of model via likelihood of Erdős-Rényi graph

(b) at least one externality negative and large

- asymptotically **sparser** than Erdős-Rényi graph
- externalities **identified**

**Heterogeneous** players ( $\exists i, j$  such that  $X_i \neq X_j$ ):

- no results
- preliminary study in Mele & Zhu (2017) - working paper

## Sampler Convergence

### Theorem 5 (Convergence of Local Sampler with Nonnegative Externalities).

Model (6), with probability of meeting  $\rho_{ij} = 1/(n(n-1))$ . Fix  $\gamma \geq 0$ . Then, in the case of nonnegative externalities  $\beta \geq 0$ , there exists a V-shaped region of the parameter space delimited by functions  $S_\gamma(\phi_1(\alpha)), S_\gamma(\phi_2(\alpha))$  such that

- 1) if  $(\alpha, \beta)$  belongs to the V-shaped region, then model converges to stationarity in  $e^{Cn^2}$  steps,  $C > 0$ . This results holds for any local sampler.
- 2) otherwise, model converges in  $Cn^2 \log(n)$  steps,  $C > 0$ .

Intuition:

- (1a) in the V-shaped region problem (7) has 2 *local maxima*, the sampler spend exponential time at one of them (i.e. probability  $e^{-Cn^2}$  to escape from local max)
- (1b) increasing  $\gamma \implies$  increase area of exponentially slow convergence
- (2a) when convergence is quadratic  $\implies$  sampler feasible for  $n < 500$
- (2b) this happens when model is indistinguishable from directed Erdős-Rényi graph

## Simulation and Estimation in Finite Networks

Posterior inference via approx version of exchange algorithm of MGM06 [13]

- ▶ double Metropolis-Hastings step to avoid computing  $c(\mathcal{G}, X, \theta)$
- ▶ data augmentation via auxiliary network  $g'$
- ▶ higher  $R \implies$  better approximation of posterior, but higher rejection rate

---

### Algorithm 2 Approximate Exchange Algorithm

---

```

1: procedure AEA( $\theta, g, M, R$ )
2:   for  $m = 1, \dots, M$  do
3:     1) propose parameter  $\theta' \sim q_\theta(\cdot|\theta)$ 
4:     2) run Algorithm 1 for  $R$  iterations using  $\theta'$ . Keep last simulated network  $g'$ 
5:     3) accept parameter  $\theta'$  with probability

```

$$\alpha(\theta, \theta', g', g) = \min \left\{ 1, \frac{\exp\{\mathcal{Q}(g', X|\theta)\}}{\exp\{\mathcal{Q}(g, X|\theta)\}} \frac{p(\theta')}{p(\theta)} \frac{q_\theta(\theta|\theta')}{q_\theta(\theta'|\theta)} \frac{\exp\{\mathcal{Q}(g, X|\theta')\}}{\exp\{\mathcal{Q}(g', X|\theta')\}} \right\}$$

```

6:   end for
7:   return sequence of  $M$  parameters  $\{\theta^{(m)}\}_m$ 
8: end procedure

```

---



## Simulation and Estimation in Finite Networks

---

**Algorithm 2** Approximate Exchange Algorithm

---

```
1: procedure AEA( $\theta, g, M, R$ )  
2:   for  $m = 1, \dots, M$  do  
3:     1) propose parameter  $\theta' \sim q_\theta(\cdot|\theta)$   
4:     2) run Algorithm 1 for  $R$  iterations using  $\theta'$ . Keep last simulated network  $g'$   
5:     3) accept parameter  $\theta'$  with probability  

$$\alpha(\theta, \theta', g', g) = \min \left\{ 1, \frac{\exp\{\mathcal{Q}(g', X|\theta)\}}{\exp\{\mathcal{Q}(g, X|\theta)\}} \frac{p(\theta')}{p(\theta)} \frac{q_\theta(\theta|\theta')}{q_\theta(\theta'|\theta)} \frac{\exp\{\mathcal{Q}(g, X|\theta')\}}{\exp\{\mathcal{Q}(g', X|\theta')\}} \right\}$$
  
6:   end for  
7:   return sequence of  $M$  parameters  $\{\theta^{(m)}\}_m$   
8: end procedure
```

---

► what prior distribution  $p(\theta)$ ?

► what proposal distribution  $q_\theta(\cdot|\theta)$ ?

## Simulation results

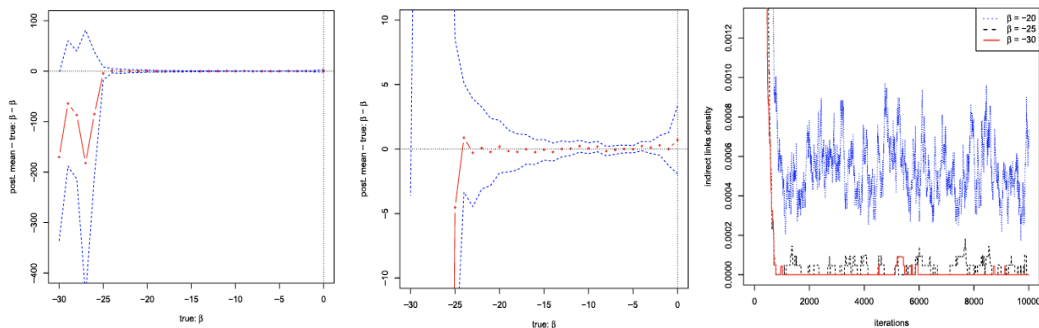


Figure: Left: Estimates of  $\beta < 0$ , with 95% credibility intervals (middle: zoom-in). Right: indirect links density.

- $\beta \geq 0 \implies$  Erdős-Rényi case, not identified
- $\beta < 0 \implies$  identified
- $\beta \ll 0 \implies$  estimation impossible: # indirect links close to 0

## Simulation results

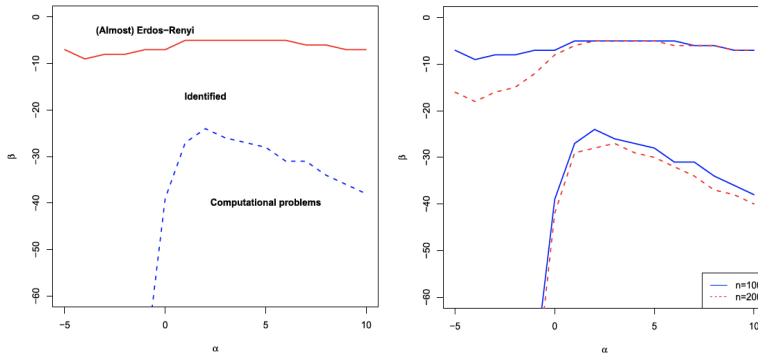


Figure: Left: approximate regions of identified parameters, for  $n = 100$ . Right: comparison of regions for  $n = 100$ ,  $n = 200$ .

### Remark 3.

Regions of identified parameters  $(\alpha, \beta)$  vary with  $n$ , the number of players.

## Simulation results

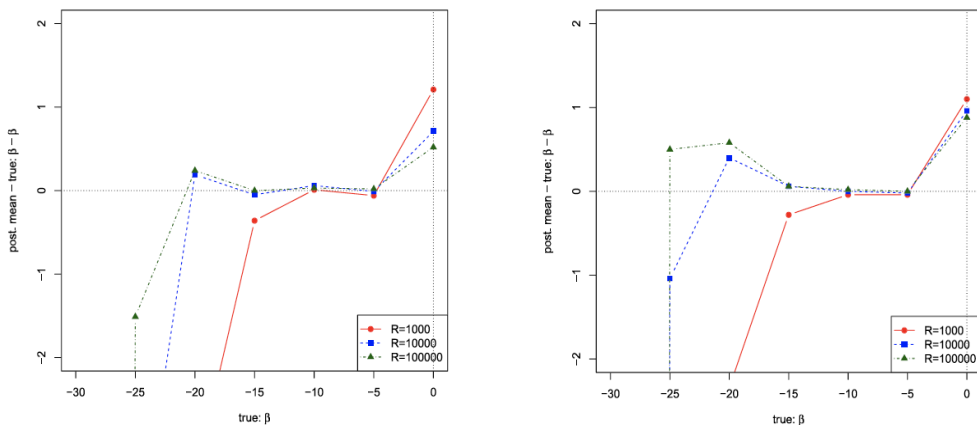


Figure: Difference between posterior estimates and true, for varying number of network simulations  $R$ :  $n = 100$  (left) and  $n = 200$  (right).

## Simulation results

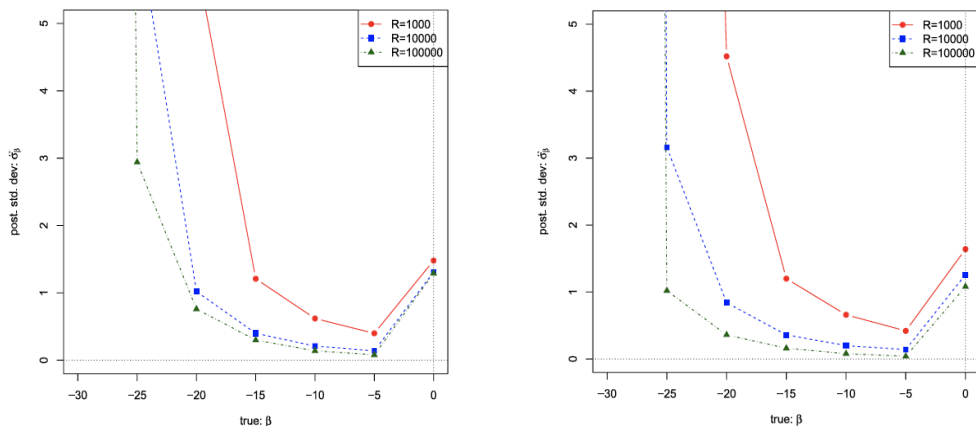


Figure: Posterior standard deviation for varying number of network simulations  $R$ :  $n = 100$  (left) and  $n = 200$  (right).

## Simulation results

- ▶  $R = 1000 \implies$  imprecise estimates
- ▶ no significant difference between  $R = 10,000$  and  $R = 100,000 \implies$  suggest rule-of-thumb  $R = 10,000$
- ▶ cost of increasing network simulations  $\implies$  almost linear  $\mathcal{O}(R)$
- ▶ results suggest convergence is almost quadratic  $\mathcal{O}(n^2)$  in this area of parameter space

## Simulation results

- ▶  $R = 1000 \implies$  imprecise estimates
- ▶ no significant difference between  $R = 10,000$  and  $R = 100,000 \implies$  suggest rule-of-thumb  $R = 10,000$
- ▶ cost of increasing network simulations  $\implies$  almost linear  $\mathcal{O}(R)$
- ▶ results suggest convergence is almost quadratic  $\mathcal{O}(n^2)$  in this area of parameter space

▶ what was the **computing time**?

▶ what about **real data applications**?

## Conclusions

The paper in a nutshell:

- ✿ network formation model, combining **strategic** and **random networks** features
- ✿ payoffs depend on links: **direct** + indirect (**externalities**)
- ✿ **homogeneous** players meet sequentially at random, **myopically** updating links
- ✿ network formation process is a potential game and **converges to ERGM**, generating **directed dense** networks
- ✿ **identification**: only if at least **1 externality negative and sufficiently large**
- ✿ standard estimation for ERGMs exponentially slow  $\Rightarrow$  Bayesian MCMC (almost quadratic time)



## Conclusions

Unclear points and questions:

- 👉 theoretical **quantification** of “sufficiently large” (negative) magnitude of  $\beta$ ?
- 👉 choice of prior for parameters  $p(\theta)$ ?
- 👉 choice of proposal for network  $q_g(\cdot|g)$ ?
- 👉 choice of proposal for parameters  $q_\theta(\cdot|\theta)$ ?
- 👉 duration of **computing time** in simulations?
- 👉 **real data** applications?









Thanks for your attention!

Any question?

## References I

-  Jackson, M. O. (2008), *Social and Economics Networks*. Princeton
-  Lovasz, L. (2012), *Large Networks and Graph Limits*. American Mathematical Society Colloquium Publications, Vol. 60. American Mathematical Society
-  Badev, A. (2017), "Discrete Games in Endogenous Networks: Equilibria and Policy", Working Paper. Available at arXiv <https://arxiv.org/abs/1705.03137v1>
-  Caimo, A., & N. Friel (2011), "Bayesian Inference for Exponential Random Graph Models", *Social Networks*, 33(1):41–55
-  Chatterjee, S., & S. R. S. Varadhan (2011), "The Large Deviation Principle for the Erdős-Rényi Random Graph" *European Journal of Combinatorics*, 32(7):1000–1017. Homomorphisms and Limits
-  Christakis, N., J. Fowler, G. W. Imbens, & K. Kalyanaraman (2010), "An Empirical Model for Strategic Network Formation" Harvard University
-  Frank, O., & D. Strauss (1986), "Markov Graphs", *Journal of the American Statistical Association*, 81:832–842
-  Graham, B. (2017), "An Econometric Model of Network Formation With Degree Heterogeneity", *Econometrica*, 85(4):1033–1063

## References II

-  Hsieh, C.-S., & L.-F. Lee (2016), "A Social Interactions Model With Endogenous Friendship Formation and Selectivity", *Journal of Applied Econometrics*, 31(2):301–319
-  Leung, M. (2018), "A weak law for moments of pairwise-stable networks", Working Paper. Available at SSRN: <https://ssrn.com/abstract=2663685>
-  Mele, A., & L. Zhu (2017), "Approximate Variational Estimation for a Model of Network Formation", Working Paper. Available at arXiv: <https://arxiv.org/abs/1702.00308>
-  Menzel, K. (2017), "Strategic Network Formation With Many Agents", Working Papers, NYU
-  Murray, I. A., Z. Ghahramani, & D. J. C. MacKay (2006), "MCMC for Doubly-Intractable Distributions", *Uncertainty in Artificial Intelligence*
-  Miyauchi, Y. (2016), "Structural Estimation of a Pairwise Stable Network Formation With Nonnegative Externality", *Journal of Econometrics*, 195(2):224–235
-  Snijders, T. A. B. (2002), "Markov Chain Monte Carlo Estimation of Exponential Random Graph Models", *Journal of Social Structure*, 3(2)
-  Wainwright, M. J., & M. Jordan (2008), "Graphical Models, Exponential Families, and Variational Inference", *Foundations and Trends® in Machine Learning*, 1(1–2):1–305