A Structural Model of Dense Network Formation - Estimation and Simulation

Mele (2017) - Econometrica

Giulia Carallo
Reading Group on Stochastic Modelling

November 8th, 2018

Università Ca'Foscari Venezia

To recap...

Asymptotic results for homogeneous players, if:

1. positive externalities: model is asymptotically indistinguishable from a directed Erdös-Rényi graph + externalities are not identified
2. negative externalities with sufficiently large magnitude: asymptotically different from a directed Erdös-Rényi + externalities identified

Network Simulation

ALGORITHM 1 - Metropolis-Hastings for Network Simulation. Fix a parameter vector θ. At iteration r, with current network g_r:

1. Propose a network g' from a proposal distribution $g' \sim q_g(g'|g_r)$.
2. Accept network g' with probability $\alpha_{mh}(g_r, g')$:

$$\alpha_{mh}(g_r, g') = min\left\{1, \frac{exp[Q(g', X, \theta)]}{exp[Q(g_r, X, \theta)]}\frac{q_g(g_r|g')}{q_g(g'|g_r)}\right\}. \tag{1}$$

Problem: Which is the proposal for $q_g(\cdot|g_r)$?

Advantages and disadvantages of Algorithm 1

- Does not contain the normalizing constant $c(G, X, \theta)$
- convergence can be slow
- local sampler: at each iteration the link g_{ij} is updated according to $\alpha_{mh}(g_r, g')$
- degeneracy: large probability mass on few networks

$\Rightarrow$ **ALGORITHM 2**

Simulations with Algorithm 1

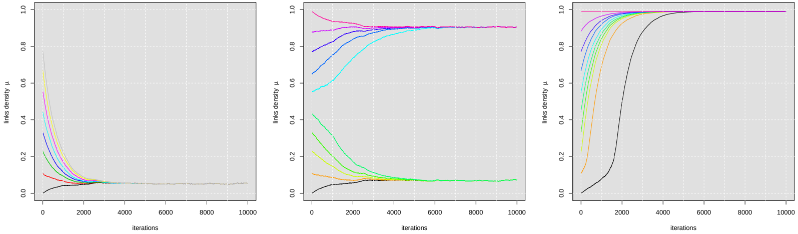

Figure: Figure 1 in Mele (2017). Simulations for a network with $n = 100$ players, $\alpha = -3$ and $\beta = \{1/n, 3/n, 7/n\}$. Simulation starts at 10 different starting network.

Further simulations with Algorithm 1.

In order to investigate better the convergence problem of Algorithm 1, further simulations are provided. Namely, I changed:

- the number of network to skip between sampled network
- the tempering function
- the size of the network (number of players, n)

Time series of the direct and indirect links, for different starting point and skips=15

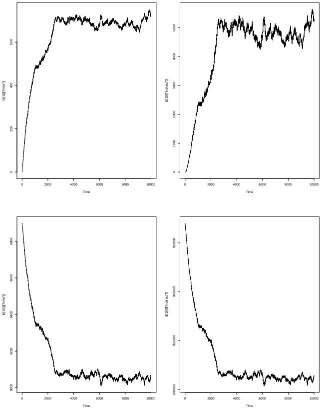

Figure: Time series of the direct and indirect links of two chain with different starting point. Skips=15

Acf of the direct and indirect links, for different starting point and skips=15

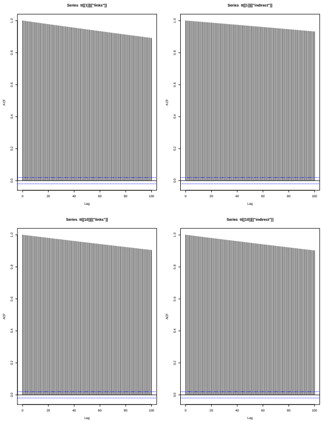

Figure: Acf of the direct and indirect links of two chain with different starting point. Skips=15

Simulations with Algorithm 1. Skip=25

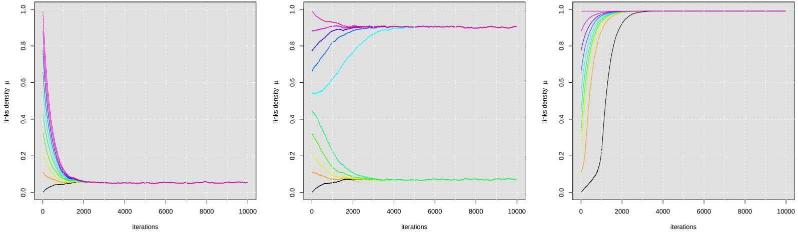

Figure: Simulations for a network with $n = 100$ players, $\alpha = -3$ and $\beta = \{1/n, 3/n, 7/n\}$. Simulation starts at 10 different starting network. Skips = 25.

Time series of the direct and indirect links, for different starting point and skips=25

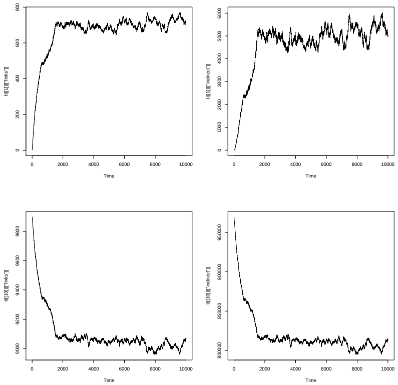

Figure: Time series of the direct and indirect links of two chain with different starting point. Skips=25

Acf of the direct and indirect links, for different starting point and skips=25

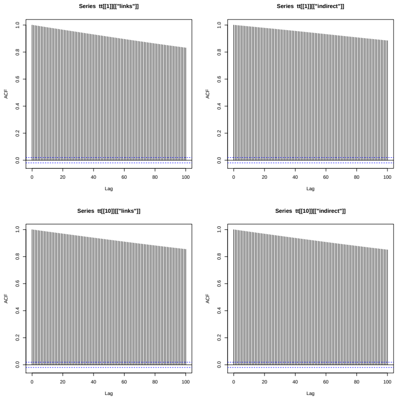

Figure: Acf of the direct and indirect links of two chain with different starting point. Skips=25

Simulations with Algorithm 1. Skips=30

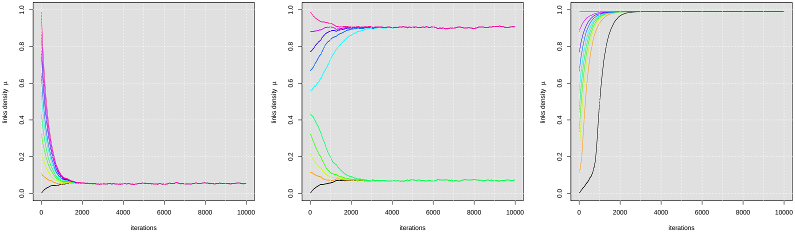

Figure: Simulations for a network with $n = 100$ players, $\alpha = -3$ and $\beta = \{1/n, 3/n, 7/n\}$. Simulation starts at 10 different starting network. Skips=30.

Time series of the direct and indirect links, for different starting point and skips=30

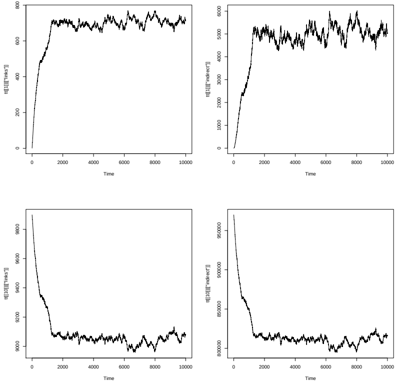

Figure: Time series of the direct and indirect links of two chain with different starting point. Skips=30

Acf of the direct and indirect links, for different starting point and skips=30

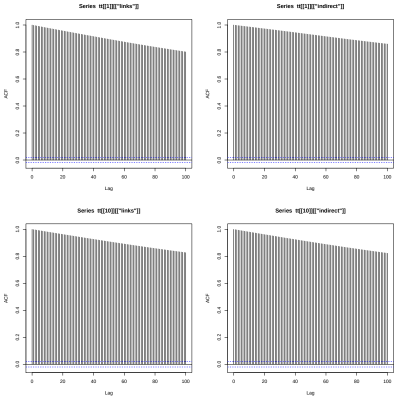

Figure: Acf of the direct and indirect links of two chain with different starting point. Skips=30

Size of the network n=20

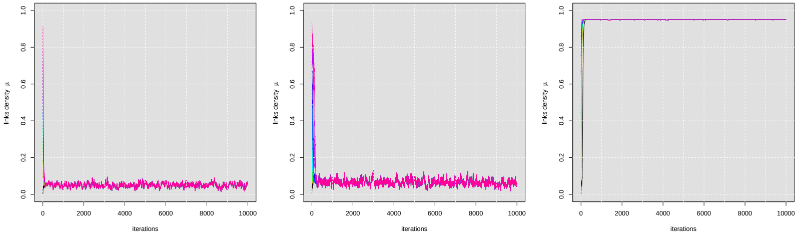

Figure: Simulation of the link density of the network by changing the size of the network. Parameters are the same of the one in Figure 1, Mele (2017).

Acf of the direct and indirect links, n=20

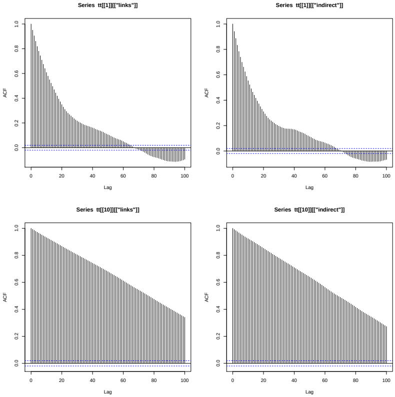

Figure: Acf of the direct and indirect links of two chain with different starting point. n=20

Algorithm 2

ALGORITHM 2 - Approximate Exchange Algorithm. Fix the number of network simulations R. At each iteration t, with current parameter $\theta_t = \theta$ and network data g:

1. Propose a new parameter θ' from a distribution $q_\theta(\cdot|\theta)$.
2. Start *Algorithm 1* at the observed network g, iterating for R steps using parameter θ', and collect the last simulated network g'.
3. Accept parameter θ' with probability $\alpha_{ex}(\theta, \theta', g', g)$:

$$\alpha_{ex}(\theta, \theta', g', g) = min\left\{1, \frac{exp[Q(g', X, \theta)]}{exp[Q(g, X, \theta)]} \frac{p(\theta')}{p(\theta)} \frac{q_\theta(\theta|\theta')}{q_\theta(\theta'|\theta)} \frac{exp[Q(g, X, \theta')]}{exp[Q(g', X, \theta')]}\right\}. \tag{2}$$

In all simulations:

- Prior: independent normal priors $N(0,10)$
- Proposal of the exchange algorithm: random walk $N(0, \Sigma)$

Local sampler versus modified sampler.

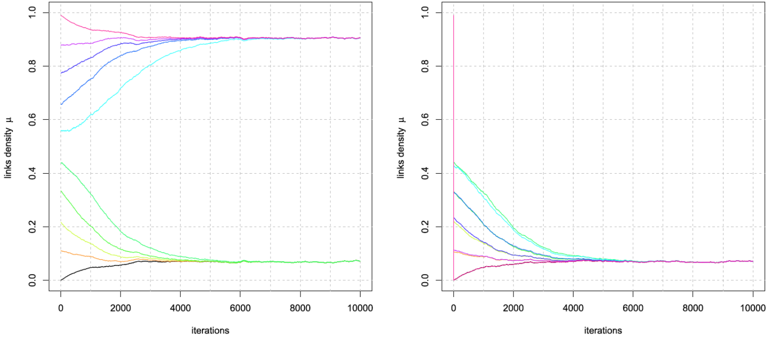

Figure: Comparison of Algorithm 1 and 2 with $(\alpha, \beta) = (-3, 3)$. For the modified algorithm $p_r = p_f = p_{inv} = 0.01$. (Figure 8, Appendix B)

Analysis for Algorithm 2

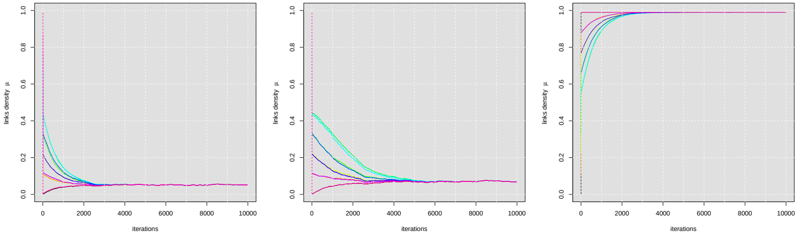

Figure: Simulation like in Figure 1, Mele (2017), re-done with Algorithm 2. All the parameters are the same as for figure 1 and $p_r = p_f = p_{inv} = 0.01$.

Time series of the direct and indirect links with Algorithm 2

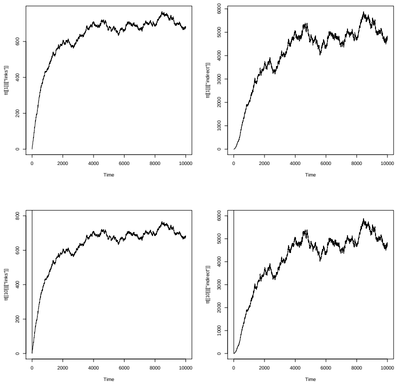

Figure: Time series of the direct and indirect links of two chain with different starting point.

Acf of the direct and indirect links with Algorithm 2

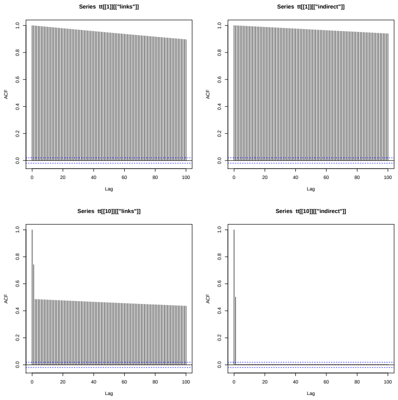

Figure: Acf of the direct and indirect links of two chain with different starting point.

Analysis for Algorithm 2: skips=25

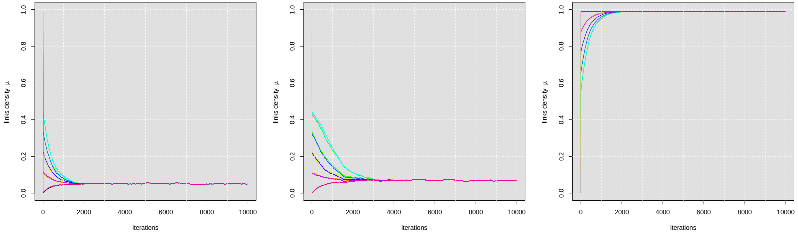

Figure: Simulation like in Figure 1, Mele (2017), re-done with Algorithm 2. All the parameters are the same as for figure 1, with skips=25 and $p_r = p_f = p_{inv} = 0.01$.

Conclusion

- Algorithm 1 problems are persistent with several changes: number of skips, tempering function;
- the acf and the time series of the degree of the network consistently show the non-convergence for different starting point by using Algorithm 1;
- the size of the network seem to have some effect on the convergence with Algorithm 1 (too small?);
- Algorithm 2 seems to perform better with respect to Algorithm 1;
- Changing the tempering function and the number of skips do not affect the performance of Algorithm 2, namely chains started at dense networks converge to the correct sparse network.

Other simulations with Algorithm 1. Skips=100

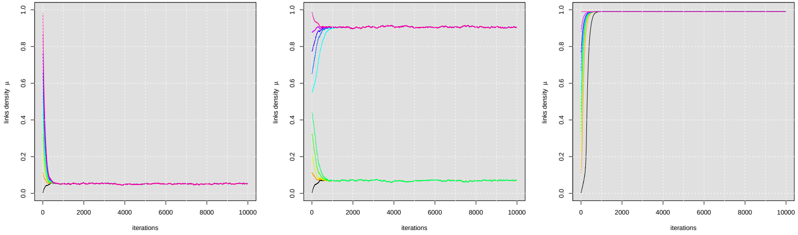

Figure: Simulations for a network with $n = 100$ players, $\alpha = -3$ and $\beta = 1/n, 3/n, 7/n$. Simulation starts at 10 different starting network. Skips=100.

Size of the network n=10

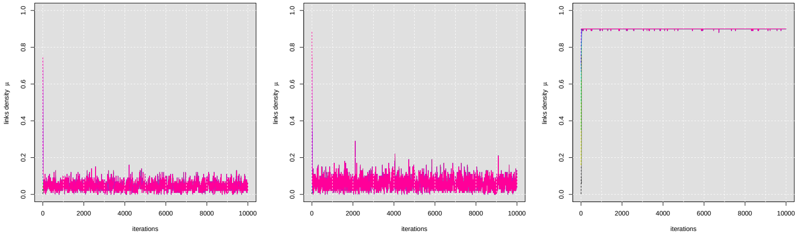

Figure: Simulation of the link density of the network by changing the size of the network. Parameters are the same of the one in Figure 1, Mele (2017).

Acf of the direct and indirect links, n=10

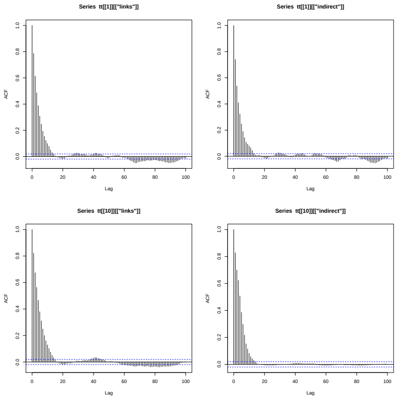

Figure: Acf of the direct and indirect links, for two chains with different starting point.

Size of the network n=200

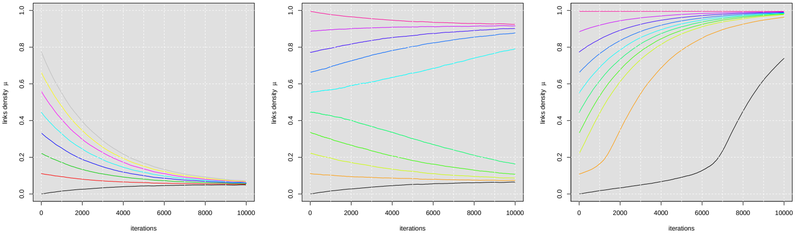

Figure: Simulation of the link density of the network by changing the size of the network. Parameters are the same of the one in Figure 1, Mele (2017).

Size of the network n=300

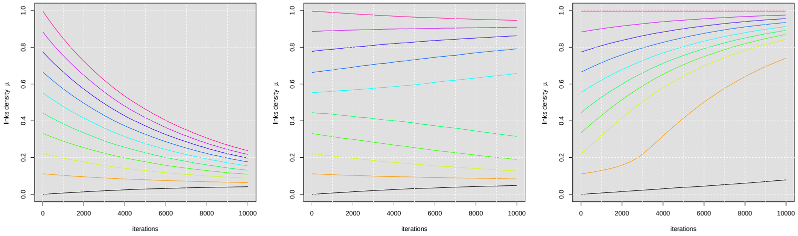

Figure: Simulation of the link density of the network by changing the size of the network. Parameters are the same of the one in Figure 1, Mele (2017).

Algorithm 3

ALGORITHM 3 - Exact Exchange Algorithm. Start at current parameter $\theta_t = \theta$ and network data g.

1. Propose a new parameter vector θ', $\theta' \sim q_\theta(\cdot|\theta)$.
2. Draw an exact sample network g' from the likelihood, $g' \sim \pi(\cdot|X, \theta')$.
3. Compute the acceptance ratio:

$$\begin{aligned} \alpha_{ex}(\theta, \theta', g', g) &= min\left\{1, \frac{exp[Q(g', X, \theta)]}{exp[Q(g, X, \theta)]} \frac{p(\theta')}{p(\theta)} \frac{q_\theta(\theta|\theta')}{q_\theta(\theta'|\theta)} \frac{exp[Q(g, X, \theta')]}{exp[Q(g', X, \theta')]} \frac{c(\theta)}{c(\theta)} \frac{c(\theta')}{c(\theta')}\right\} \\ &= min\left\{1, \frac{exp[Q(g', X, \theta)]}{exp[Q(g, X, \theta)]} \frac{p(\theta')}{p(\theta)} \frac{q_\theta(\theta|\theta')}{q_\theta(\theta'|\theta)} \frac{exp[Q(g, X, \theta')]}{exp[Q(g', X, \theta')]}\right\}. \end{aligned} \tag{3}$$

4. Update the parameter according to

$$\theta_{t+1} = \begin{cases} \theta', \text{with prob. } \alpha_{ex}(\theta, \theta', g', g) \\ \theta, \text{with prob. } 1 - \alpha_{ex}(\theta, \theta', g', g) \end{cases} \tag{4}$$