E Additional simulation results

The performance of the estimation method is tested using artificial data. All the computations with artificial data are performed in a standard desktop Dell Precision T7620 with 2 Intel Xeon CPUs E5-2697 v2 with 12 Dual core processors at 2.7GHZ each and 64GB of RAM. For replication purposes, there is a package in Github at https://github.com/meleangelo/netnew.[54]

Ideally, we want to compare the results of the approximate exchange algorithm with the exact algorithm. This is feasible for a special case, where preferences depend only on direct and mutual links (i.e. excluding friends of friends and popularity effects).

$$Q(g, \alpha, \beta) = \alpha \sum_{i=1}^{n} \sum_{j=1}^{n} g_{ij} + \beta \sum_{i=1}^{n} \sum_{j>i1}^{n} g_{ij} g_{ji} \tag{67}$$

For this model, described by equation (67), we can show that the constant is

$$c(\theta) = \left(1 + 2e^{\alpha} + e^{2\alpha+\beta}\right)^{\frac{n(n-1)}{2}}$$

thus we can compute the exact likelihood and we can perform inference using the *exact* Metropolis-Hastings sampler. We then compare the results of the exact algorithm with the approximate exchange algorithm.

The results of the simulations are shown in Table 1. The data were generated by parameters $(\alpha, \beta) = (-2.0, 0.5)$. The number of network simulations per each proposed parameter are $R = \{1000,\ 5000,\ 10000,\ 50000,\ 100000,\ 1000000,\ 10000000\}$. We run each algorithm for $S = 10000$ parameters iterations, and we use the output to measure the Kolmogorov-Smirnov distance and the Kullback-Leibler divergence between the posterior estimated with the exact metropolis sampler $p(\theta|g, X)$ and the posterior estimated with the approximated algorithm with R network simulations $p_R(\theta|g, X)$

$$\begin{aligned} KS &= \sup_{\theta_i \in \Theta_i} \left| \int_{-\infty}^{\theta_i} p_R(\theta_i|g, X) - \int_{-\infty}^{\theta_i} p(\theta_i|g, X) \right| \\ KL &= \int_{\Theta_i} \log \left[\frac{p_R(\theta_i|g, X)}{p(\theta_i|g, X)} \right] p_R(\theta_i|g, X) d\theta_i \end{aligned}$$

The table reports posterior mean, median, standard deviation, Monte Carlo standard errors for the posterior mean (mcse), 95% credibility intervals, Kolmogorov-Smirnov statistics, Kullback-Leibler divergence and time for computation.

The exact Metropolis-Hastings is reported in the first column of the table. The approximate exchange algorithm works very well for small to moderate networks. For a small

[54] In all estimation exercises we use independent normal priors $\mathcal{N}(0, 10)$. The proposal of the exchange algorithm is a random walk $\mathcal{N}(0, \Sigma)$. We repeat the estimation twice: the first time we use a diagonal Σ; in the second round, we use the covariance from the first round as baseline. In all simulations the probability of large steps is 0.001 and a large step updates $0.1n$ links.

Table 1: Convergence of estimated posteriors, model (67)

$n = 100$	Exact Metropolis		R=1000		R=5000		R=10000		R=50000		R=100000					
	α	β	α	β	α	β	α	β	α	β	α	β				
mean	-1.923	0.286	-1.915	0.286	-1.925	0.285	-1.922	0.275	-1.921	0.284	-1.919	0.286				
median	-1.923	0.288	-1.919	0.296	-1.923	0.292	-1.921	0.273	-1.921	0.287	-1.919	0.286				
std. dev.	0.034	0.114	0.105	0.263	0.054	0.141	0.042	0.123	0.034	0.115	0.034	0.111				
mcse	0.000	0.002	0.007	0.033	0.001	0.005	0.001	0.004	0.000	0.003	0.000	0.003				
pctile 2.5%	-1.992	0.058	-2.115	-0.257	-2.034	-0.007	-2.006	0.034	-1.987	0.039	-1.985	0.069				
pctile 97.5%	-1.857	0.506	-1.705	0.767	-1.820	0.553	-1.842	0.514	-1.853	0.505	-1.851	0.512				
KS	NA	NA	0.275	0.205	0.114	0.057	0.066	0.057	0.032	0.015	0.060	0.022				
KL	NA	NA	0.041	0.027	0.013	0.186	0.039	0.075	0.040	0.062	0.006	0.088				

$n = 200$	Exact Metropolis		R=1000		R=5000		R=10000		R=50000		R=100000		R=1mil		R=10mil	
	α	β	α	β	α	β	α	β	α	β	α	β	α	β	α	β
mean	-1.988	0.463	-1.975	0.463	-1.964	0.463	-1.979	0.465	-1.989	0.455	-1.988	0.463	-1.988	0.466	-1.987	0.459
median	-1.989	0.467	-1.974	0.509	-1.968	0.468	-1.978	0.465	-1.989	0.454	-1.989	0.464	-1.988	0.467	-1.987	0.457
std. dev.	0.017	0.061	0.048	0.275	0.042	0.113	0.033	0.073	0.019	0.053	0.017	0.059	0.017	0.057	0.016	0.055
mcse	0	0.003	0.002	0.075	0.002	0.012	0.001	0.005	0	0.002	0	0.003	0	0.003	0	0.002
pctile 2.5%	-2.021	0.335	-2.071	-0.21	-2.044	0.186	-2.042	0.32	-2.024	0.353	-2.024	0.339	-2.024	0.358	-2.017	0.354
pctile 97.5%	-1.954	0.572	-1.89	0.889	-1.872	0.687	-1.921	0.614	-1.949	0.56	-1.955	0.571	-1.955	0.582	-1.955	0.577
KS	NA	NA	0.343	0.34	0.381	0.135	0.25	0.057	0.067	0.105	0.015	0.039	0.039	0.045	0.062	0.086
KL	NA	NA	0.1	0.178	0.105	0.079	0.129	0.099	0.05	0.05	0.041	0.058	0.092	0.044	0.173	0.234
time	0.124s		14.539s		21.808s		30.451s		100.761s		193.722s		1762.202s		17370.945s	

$n = 500$	Exact Metropolis		R=1000		R=5000		R=10000		R=50000		R=100000		R=1mil		R=10mil	
	α	β	α	β	α	β	α	β	α	β	α	β	α	β	α	β
mean	-2.018	0.551	-1.941	0.337	-2.014	0.562	-2.017	0.561	-2.017	0.552	-2.018	0.552	-2.018	0.55	-2.018	0.55
median	-2.018	0.552	-1.922	0.369	-2.012	0.562	-2.019	0.562	-2.016	0.553	-2.018	0.552	-2.018	0.551	-2.018	0.55
std. dev.	0.007	0.024	0.071	0.218	0.045	0.107	0.036	0.074	0.016	0.036	0.012	0.028	0.007	0.022	0.007	0.022
mcse	0	0	0.005	0.047	0.002	0.011	0.001	0.005	0	0.001	0	0	0	0	0	0
pctile 2.5%	-2.032	0.501	-2.074	-0.106	-2.105	0.335	-2.085	0.424	-2.05	0.479	-2.041	0.497	-2.032	0.508	-2.031	0.507
pctile 97.5%	-2.004	0.596	-1.838	0.666	-1.931	0.755	-1.942	0.707	-1.988	0.621	-1.994	0.606	-2.004	0.596	-2.005	0.592
KS	NA	NA	0.743	0.703	0.408	0.363	0.341	0.31	0.229	0.117	0.107	0.066	0.027	0.034	0.033	0.032
KL	NA	NA	0.466	0.196	0.121	0.081	0.081	0.019	0.026	0.009	0.02	0.008	0.061	0.036	0.049	0.041
time	0.187s		87.344s		95.831s		105.955s		181.413s		275.357s		2010.322s		19319.663s	

$n = 1000$	Exact Metropolis		R=1000		R=5000		R=10000		R=50000		R=100000		R=1mil		R=10mil	
	α	β	α	β	α	β	α	β	α	β	α	β	α	β	α	β
mean	-2.001	0.481	-1.986	0.456	-1.974	0.459	-1.995	0.486	-1.999	0.479	-2.001	0.479	-2.001	0.481	-2.002	0.481
median	-2.001	0.48	-1.991	0.501	-1.979	0.461	-1.993	0.479	-2.001	0.48	-2.001	0.479	-2.001	0.48	-2.002	0.482
std. dev.	0.003	0.011	0.081	0.247	0.05	0.091	0.031	0.07	0.017	0.037	0.011	0.026	0.004	0.012	0.003	0.012
mcse	0	0	0.007	0.06	0.002	0.007	0.001	0.004	0	0.001	0	0	0	0	0	0
pctile 2.5%	-2.008	0.459	-2.148	-0.007	-2.047	0.284	-2.057	0.361	-2.029	0.404	-2.024	0.425	-2.01	0.457	-2.008	0.459
pctile 97.5%	-1.995	0.503	-1.835	0.813	-1.825	0.642	-1.938	0.64	-1.966	0.55	-1.979	0.529	-1.993	0.505	-1.995	0.504
KS	NA	NA	0.506	0.484	0.63	0.47	0.502	0.355	0.351	0.268	0.271	0.216	0.078	0.018	0.027	0.036
KL	NA	NA	0.3	0.261	0.528	0.041	0.297	0.083	0.45	0.21	0.137	0.047	0.021	0.031	0.034	0.014
time	0.234s		364.761s		371.563s		381.172s		459.578s		556.330s		2304.228s		19730.772	

network with $n = 100$ players, a reasonable degree of accuracy can be reached with as low as $R = 5000$ network simulations per parameter. Simulations from over-dispersed starting values converge to the same posterior distribution. Convergence is quite fast to the high density region of the posterior.[55]

Let's now consider a model with homogeneous players where there is no utility from

Table 2: Estimated structural parameters for model (68), $n = 100$

true parameters $(\alpha, \beta) = (-3, 0.01)$								
$n = 100$	$R = 1000$		$R = 10000$		$R = 100000$		$R = 1000000$	
true $= (-3, .01)$	α	β	α	β	α	β	α	β
mean	-2.7450	-0.0233	-2.9095	-0.0035	-2.9407	-0.0009	-2.9206	-0.0025
median	-2.7575	-0.0185	-2.9171	-0.0021	-2.9498	0.0003	-2.9288	-0.0014
std. dev.	0.4782	0.0460	0.2032	0.0201	0.1860	0.0183	0.1916	0.0189
mcse	0.0975	0.0010	0.0111	0.0001	0.0094	0.0001	0.0104	0.0001
pctile 2.5%	-3.6862	-0.1208	-3.2660	-0.0462	-3.2614	-0.0400	-3.2468	-0.0412
pctile 97.5%	-1.7789	0.0545	-2.4916	0.0305	-2.5452	0.0303	-2.5158	0.0297
time (secs)	25.3800		236.2100		2485.5200		24658.1500	

true parameters $(\alpha, \beta) = (-3, 0.03)$								
$n = 100$	$R = 1000$		$R = 10000$		$R = 100000$		$R = 1000000$	
true $= (-3, 0.03)$	α	β	α	β	α	β	α	β
mean	-2.6075	0.0002	-2.7578	0.0124	-2.7618	0.0126	-2.7720	0.0134
median	-2.6425	0.0036	-2.7804	0.0140	-2.7812	0.0140	-2.7917	0.0148
std.dev.	0.4396	0.0306	0.1757	0.0122	0.1663	0.0116	0.1671	0.0116
mcse	0.0819	0.0004	0.0075	0.0000	0.0073	0.0000	0.0080	0.0000
pctile 2.5%	-3.4144	-0.0682	-3.0320	-0.0150	-3.0185	-0.0132	-3.0165	-0.0129
pctile 97.5%	-1.6856	0.0526	-2.3671	0.0299	-2.4054	0.0299	-2.3897	0.0297
time (secs)	27.3900		256.2500		2647.7600		26277.7000	

true parameters $(\alpha, \beta) = (5, -0.1)$								
$n = 100$	R=1000		R=10000		R=100000		R=1000000	
true $= (5, -0.1)$	α	β	α	β	α	β	α	β
mean	4.8397	-0.0964	4.8722	-0.0968	4.8743	-0.0968	4.8856	-0.0970
median	4.8265	-0.0963	4.8674	-0.0968	4.8682	-0.0968	4.8846	-0.0970
std. dev.	0.4031	0.0067	0.1550	0.0026	0.1188	0.0018	0.1137	0.0018
mcse	0.0427	0.0000	0.0064	0.0000	0.0041	0.0000	0.0039	0.0000
pctile 2.5%	4.0677	-0.1101	4.5688	-0.1020	4.6493	-0.1005	4.6645	-0.1006
pctile 97.5%	5.6615	-0.0836	5.1707	-0.0920	5.1202	-0.0933	5.1156	-0.0936
time (secs)	49.3300		433.5100		4254.5800		41218.3700	

reciprocated links, but only from indirect connections and popularity, i.e.

$$Q(g, \alpha, \beta) = \alpha \sum_{i=1}^{n} \sum_{j=1}^{n} g_{ij} + \beta \sum_{i=1}^{n} \sum_{j=1}^{n} \sum_{k=1}^{n} g_{ij} g_{jk} \tag{68}$$

The estimated parameters for this model are shown in Table 2 and 3. We generate the network data using different parameter vectors. The first panel correspond to parame-

[55]This result is common with the class of exchange algorithms. See Caimo and Friel (2010), Atchade and Wang (forthcoming) for examples. Computations can be faster if we embed sparse matrix algebra routines in the codes. The results in Table 1 are obtained with codes that do no use sparse matrix algebra, thus representing a worst case scenario in computational time.

Table 3: Estimated structural parameters for model (68), $n = 200$

true parameters $(\alpha, \beta) = (-3, 0.005)$								
$n = 200$	$R = 1000$		$R = 10000$		$R = 100000$		$R = 1000000$	
true = $(-3, .005)$	α	β	α	β	α	β	α	β
mean	-2.6694	-0.0112	-2.9707	0.0042	-3.0529	0.0083	-3.0603	0.0086
median	-2.7045	-0.0086	-2.9972	0.0056	-3.0675	0.0089	-3.0784	0.0095
std. dev.	0.5631	0.0254	0.1841	0.0083	0.1137	0.0053	0.1113	0.0052
mcse	0.1341	0.0003	0.0089	0.0000	0.0035	0.0000	0.0032	0.0000
pctile 2.5%	-3.7109	-0.0665	-3.2574	-0.0148	-3.2202	-0.0035	-3.2244	-0.0036
pctile 97.5%	-1.5002	0.0332	-2.5689	0.0159	-2.8044	0.0159	-2.8007	0.0158
time (secs)	173.7800		1651.4400		16248.9400		149962.1800	

true parameters $(\alpha, \beta) = (-3, 0.015)$								
$n = 200$	$R = 1000$		$R = 10000$		$R = 100000$		$R = 1000000$	
true = $(-3, 0.015)$	α	β	α	β	α	β	α	β
mean	-2.4770	-0.0033	-2.7773	0.0075	-2.8601	0.0104	-2.8518	0.0101
median	-2.5002	-0.0019	-2.8042	0.0083	-2.8785	0.0111	-2.8703	0.0108
std. dev.	0.5828	0.0200	0.1627	0.0055	0.1012	0.0035	0.1028	0.0035
mcse	0.1012	0.0001	0.0078	0.0000	0.0028	0.0000	0.0028	0.0000
pctile 2.5%	-3.6184	-0.0474	-3.0206	-0.0054	-3.0026	0.0024	-2.9961	0.0020
pctile 97.5%	-1.2515	0.0346	-2.4080	0.0148	-2.6267	0.0150	-2.6149	0.0149
time (secs)	190.5800		1783.1900		17496.5900		161462.5300	

true parameters $(\alpha, \beta) = (5, -0.05)$								
$n = 200$	R=1000		R=10000		R=100000		R=1000000	
true = $(5, -0.05)$	α	β	α	β	α	β	α	β
mean	5.0734	-0.0504	5.0782	-0.0503	5.0528	-0.0501	5.0477	-0.0501
median	5.0475	-0.0502	5.0718	-0.0503	5.0539	-0.0501	5.0478	-0.0501
std. dev.	0.4791	0.0039	0.1535	0.0012	0.0713	0.0005	0.0644	0.0005
mcse	0.0765	0.0000	0.0068	0.0000	0.0017	0.0000	0.0011	0.0000
pctile 2.5%	4.1775	-0.0587	4.7926	-0.0529	4.9129	-0.0512	4.9220	-0.0511
pctile 97.5%	6.0765	-0.0431	5.3987	-0.0480	5.1904	-0.0490	5.1781	-0.0491
time (secs)	361.5600		2996.3300		29001.6900		257572.3700	

ters $(\alpha, \beta) = (-3, 1/n)$. This is a model that generates a sparse network and the likelihood has a unique mode. The second panel shows estimates for a model with parameters $(\alpha, \beta) = (-3, 3/n)$, with a variational problem with two local solutions that generates problems of convergence with a local sampler. The last panel is a model with negative externalities $(\alpha, \beta) = (5, -10/n)$ that does not converge to an Erdos-Renyi model. We also simulated a model with parameters $(\alpha, \beta) = (-3, 7/n)$. However, if we solve the variational problem with these parameters, we can show that the solution is an Erdos-Renyi model with probability of linking $\mu^* = 1$, i.e. the full network. Therefore a model with parameters $\alpha = -3$, for any $\beta > 7/n$ would also generate a full network. Any attempt to estimate β with data consisting of a full network is futile.

The estimates using the non-local sampler are precise for a moderate amount of network simulations. Clearly, estimates in Table 2 are less precise than the ones in Table 3, since the number of links is smaller. Because of our modified network sampler, there is no need to have a large number of network simulations. The reason is that the non-local sampler can jump quickly to the correct mode(s) of the likelihood: once it reaches an area close to the global maximum, convergence is in quadratic time, since it will reject jumps to local maxima of the variational problem that are not global maxima. While the number of iterations may be lower, the computational time of each iteration is higher, because the large steps are computationally expensive. In Table 2, the precision gain from additional network simulations is negligible when $R > 10000$. Notice that the computational time is higher for the model with negative externality. The reason is that the equilibrium network generated at the true parameters $(\alpha, \beta) = (5, -10/n)$ is denser than the ones in the previous panels, and therefore the large steps are more computationally expensive than for the other two models.

In the next table we consider a model where players are homogeneous and they receive utility from direct links, reciprocated links, indirect links and popularity. The potential function of such model is defined as

$$Q(g, \alpha, \beta, \gamma) = \alpha \sum_{i=1}^{n} \sum_{j=1}^{n} g_{ij} + \beta \sum_{i=1}^{n} \sum_{j>i}^{n} g_{ij} g_{ji} + \gamma \sum_{i=1}^{n} \sum_{j=1}^{n} \sum_{k=1}^{n} g_{ij} g_{jk} \tag{69}$$

The data are generated by parameters $(\alpha, \beta, \gamma) = (-2.00, 0.50, 0.01)$. The pattern of Table 4 is similar to the previous analysis: the increase in precision for $R > 10000$ is minimal with respect to the increased cost of sampling networks.

Table 4: Estimated structural parameters for model (69), n=100

true parameters $(\alpha, \beta, \gamma) = (-2.00, 0.50, 0.01)$												
$n = 100$	R=1000			R=10000			R=100000			R=1000000		
	α	β	γ	α	β	γ	α	β	γ	α	β	γ
mean	-1.9321	0.5098	0.0074	-2.1182	0.5168	0.0133	-2.1034	0.5115	0.0129	-2.0938	0.5196	0.0126
median	-1.9756	0.5080	0.0089	-2.1382	0.5214	0.0139	-2.1251	0.5134	0.0136	-2.1066	0.5207	0.0131
std. dev.	0.4677	0.2330	0.0135	0.1899	0.0997	0.0054	0.1877	0.0894	0.0054	0.1832	0.0882	0.0053
mcse	0.0967	0.0209	0.0001	0.0121	0.0037	0.0000	0.0110	0.0028	0.0000	0.0132	0.0027	0.0000
pctile 2.5%	-2.7241	0.0459	-0.0224	-2.4259	0.3186	0.0014	-2.4002	0.3341	0.0012	-2.3871	0.3416	0.0013
pctile 97.5%	-0.9492	0.9699	0.0300	-1.7149	0.7115	0.0216	-1.6983	0.6896	0.0213	-1.7014	0.6894	0.0211
time (secs)		42			355			3545			35806	

Finally, we estimate a simple model with heterogeneous players. There is only one binary covariate X and the players receive utility from direct links, and indirect links and popularity. The covariate is generated as a Bernoulli variable with $P(X_i = 1) = 0.3$. The utility from indirect links/popularity is positive if both i and k belong to type-1; and it is negative if they belong to different types. The potential of this model is

$$\begin{aligned} Q(g, \alpha, \beta, \gamma) = \alpha \sum_{i=1}^{n} \sum_{j=1}^{n} g_{ij} \quad &+ \quad \beta \sum_{i=1}^{n} \sum_{j=1}^{n} \sum_{k=1}^{n} g_{ij} g_{jk} \mathbb{1}_{\{X_i = X_k = 1\}} \\ &+ \quad \gamma \sum_{i=1}^{n} \sum_{j=1}^{n} \sum_{k=1}^{n} g_{ij} g_{jk} \mathbb{1}_{\{X_i \neq X_k\}} \end{aligned} \tag{70}$$

and the data are generated with parameters $(2, 11/n, -5/n)$.

The estimation results in Table 5 for $n = 100$. The estimates are again very precise for a moderate amount of simulations.

Table 5: Estimated structural parameters for model (70), $n = 100$

true parameters $(\alpha, \beta, \gamma) = (2.00, 0.11, -0.05)$									
$n = 100$	R=1000			R=10000			R=100000		
	α	β	γ	α	β	γ	α	β	γ
mean	2.2144	0.1056	-0.0530	2.0718	0.1053	-0.0503	2.0636	0.1052	-0.0501
median	2.0828	0.1064	-0.0506	2.0443	0.1054	-0.0498	2.0396	0.1055	-0.0497
std. dev.	0.8575	0.0227	0.0155	0.3348	0.0084	0.0061	0.2723	0.0066	0.0049
mcse	0.4485	0.0002	0.0001	0.0711	0.0000	0.0000	0.0487	0.0000	0.0000
pctile 2.5%	0.8456	0.0598	-0.0881	1.4803	0.0875	-0.0636	1.5959	0.0914	-0.0608
pctile 97.5%	4.1495	0.1473	-0.0286	2.8115	0.1222	-0.0397	2.6445	0.1174	-0.0418
time (secs)	97			314			2913		