Comparing the Real-World Performance of Exponential-family Random Graph Models and Latent Order Logistic Models

Duncan A. Clark
University of California - Los Angeles, Los Angeles, USA
E-mail: duncanclark@ucla.edu
Mark S. Handcock
University of California - Los Angeles, Los Angeles, USA

Summary. Exponential-family Random Graph models (ERGM) are widely used in social network analysis when modelling data on the relations between actors. ERGMs are typically interpreted as a snapshot of a network. The recently proposed Latent Order Logistic model (LOLOG) directly allows for a latent network formation process. We assess the real-world performance of these models when applied to typical networks modeled by researchers. Specifically, we model data from the ensemble of articles in the journal *Social Networks* with published ERGM fits, and compare the ERGM fit to a comparable LOLOG fit. We demonstrate that the LOLOG models are, in general, in qualitative agreement with the ERGM models, and provide at least as good a model fit. In addition they are typically faster and easier to fit to data, without the tendency for degeneracy that plagues ERGMs. Our results support the general use of LOLOG models in circumstances where ERGM are considered.

Keywords: LOLOG, ERGM, Social Network Analysis, Degeneracy, Goodness-of-fit

1. Introduction

Social network analysis has become increasingly important in recent decades, with particular need in the social sciences to elucidate relational structure (Goldenberg et al., 2010). However developing generative models for social networks has proven challenging (Chatterjee and Diaconis, 2013). Here we consider a social network a collection of fixed nodes, each with fixed covariates and with edges stochastically present or absent between every pair of nodes. The chief problems for modelling such data being the vast space of possible networks and probable highly complex dependence structures of the network edges.

The Exponential-family Random Graph Models (ERGM) framework is widely used to represent the stochastic process underlying social networks [Frank and Strauss (1986); Hunter and Handcock (2006)]. ERGMs allow researchers to quantitatively evaluate the impact of local social processes and nodal attributes on the probability of edges between nodes forming. However these models are prone to near-degeneracy (Handcock, 2003) and can not naively be applied to large networks (Schweinberger, 2011; Chatterjee and Diaconis, 2013). Model degeneracy is the application specific tendency of the model to concentrate probability mass on a small subset of graphs, especially those which are not similar to realistic networks for that application.

Much progress has been made on managing model degeneracy by introducing local neigbourhood structures (Schweinberger and Handcock, 2015) or tapering (Fellows and Handcock, 2017). The presence of degeneracy in many fitted ERGMs motivates the search for alternative model classes with similar or complementary modelling capacity that are less susceptible to these challenges.

While ERGMs are descriptive, they are often embedded as the equilibrium distribution of a social process. The Latent Order Logistic model (LOLOG) (Fellows, 2018b) is a related model that uses an edge formation process to develop a general probability model over the space of graphs. It is motivated by using the so-called change statistics, the change in the specified graph statistics resulting from toggling an edge on or off, as predictors in a sequential logistic regression for each possible edge. Noting that an ERGM specified with dyad independent sufficient statistics, reduces to a sequential logistic regression on change statistics, ERGM and LOLOG are equivalent in the dyad independent case Fellows (2018b). LOLOG models also allow non-dyad-independent change statistics, and statistics that depend on the order of edge formation, which result in different models from ERGM.

LOLOG models have the advantage that they are straightforward to sample from, and can be used with simpler model terms, that would for an ERGM almost certainly result in near-degeneracy. This allows for a fast and user friendly fitting procedure, with easily interpretable model terms. This comes at the price of an intractable likelihood due to the necessity of summing over all possible edge orderings.

How can we assess and compare differing model classes? Both ERGM and LOLOG are fully general and able to represent arbitrary distributions over the set of graphs (Fellows, 2018b, Theorem 1). As ERGMs are the equilibrium distribution of a relatively general MCMC process, there are many mechanisms that can lead to them, as there are for LOLOG. Hence both model classes have strong theoretical and modelling motivations, although the ERGM class to this point has been much more extensively explored (Schweinberger and Stewart, 2020; Schweinberger et al., 2020). In this paper, we provide a separate and novel contribution to the assessment on the model classes. Our objective is to compare the models by a pairwise assessment on the population of networks that the research community would choose to fit them on. The idea here is to move the perspective from that of the model viewpoint (i.e., given we have a model, what can we fit with it?) to a data-centric view point (i.e., given that this is the data we have, what are the best modelling approaches?) The latter is the question facing the real-world users of these models, while the “inverse problem" addressed by the former is commonly taken as it does not require the population of networks to be specified.

However, to take the data-centric viewpoint, we need to specify the population. We operationalised this in this paper by taking a population of networks that ERGM models have been applied to in the premier journal for publishing social network analyses, *Social*

Networks (Everett and Valente, 2020). *Social Networks* is an interdisciplinary journal for those with "interest in the study of the empirical structure of social relations and associations that may be expressed in network form". While the sub-population of networks in *Social Networks* for which ERGMs have been fit is a sample of the population of interest, we believe that it is an salient and (non-statistical) representative sub-population of the broader population.

Our selection of ERGM papers was at first a census of papers in the journal *Social Networks* using the ERGM framework, published from the journal's founding in 1979 up to and including the January 2016 issue. Note that we have chosen a population of networks that are biased toward ERGM. These networks have successfully completed the peer-review process of *Social Networks*. In particular, the ERGM fit and analyses have passed peer-review and are deemed of sufficient scientific interest to appear in this premier journal. Clearly this is not sufficient to ensure the fit and models are appropriate for the data, although they represent a strongly selectiveness relative to the population of networks that researchers would consider for analysis (without regard to a model class choice). Hence a comparable or competitive fit for LOLOG models to this sub-population presents stronger evidence for the value of LOLOG models than a comparison to the broader population. In particular it seems likely that in papers published that fit an ERGM model, ERGM performs well on this data set, thus we expect a publication bias towards networks that suit ERGM well, which may not necessarily suit LOLOG well. We therefore suggest that good performance on data published with ERGM fits, is a conservative indicator that LOLOG is a useful model for analysing social networks.

Identifying, assembling and fitting ERGMs and LOLOGs to an ensemble of networks, and then analysing the goodness of fit (GOF) and discussing an interpretation, is a significant undertaking. For brevity we give the fit of networks from as case-study paper Sailer and McCulloch (2012) in detail and provide summaries for the remaining networks, providing individual comments regarding each network for clarification.

The structure of this paper is as follows. In Section 2 we briefly introduce the LOLOG model, reproducing work in Fellows (2018b). In Section 3 we discuss the theoretical similarities and differences between ERGM and LOLOG models. Section 4 gives a description of the ensemble of networks and discusses the motivation for selecting such an ensemble. Section 5 shows both the LOLOG and ERGM fit of office layout networks with the data from Sailer and McCulloch (2012). Section 6 presents a summary of all the LOLOG and ERGM fits to each of the networks in the ensemble. Section 7 discusses the results of the fitting, as well as its implications regarding the utility of the LOLOG model.

2. LOLOG Model Formulation

We introduce the LOLOG model and define the standard notation that will be used henceforth. We define y to be a graph, in particular a graph realised from the random variable Y. Since we regard any nodal covariates as fixed, for a network of size n, Y takes values in the space $\mathscr{Y} = \{a \in \mathbb{R}^{n\times n} \quad | \quad \forall i,j \quad a_{i,i} = 0 \quad a_{i,j} \in \{0,1\}\}$. For undirected networks the additional restriction that $a_{i,j} = a_{j,i} \quad \forall i,j$ can be added. Note that the sample space, even in the restricted undirected case is finite and size $2^{\frac{n(n-1)}{2}}$, this become astronomically large even for small networks.

To develop models which account for partially formed networks Fellows (2018b) introduced new notation. The random variable Y_t takes values in $\mathscr{Y}$, realisations y_t denote the partial graph formed at time t. That is, the first t edges have been considered in the LOLOG generating process. Let $\mathscr{S}_{n_d}$ be the set of possible edge orderings on a network with n_d dyads. For a time t, $s_{\leq t}$ is used to denote the first t elements of s.

The first step in specifying the LOLOG model is to state the probability of observing a graph given a specified order of edge formation s i.e.:

$$p(y|s, \theta) = \prod_{t=1}^{n_d} \frac{1}{Z_t} \exp\left(\theta \cdot C_{s,t}\right) \tag{1}$$

Where, for g a function giving specified graph statistics we define the change statistics

$$C_{s,t} = g(y_t, s_{\leq t}) - g(y_{t-1}, s_{\leq t-1}) \tag{2}$$

The change statistics are the difference in the graph statistics, between the y_t network and the y_{t-1} network. The Z_t are the usual normalising constants so the sum of the probabilities of the edge forming and not forming are 1. That is letting y_t^+ be the graph y_{t-1} with the edge s_t added and y_t^- the graph without the edge added i.e. unchanged. Then we have Z_t as follows:

$$Z_t = \exp\left(C_{s,t}^+\right) + \exp\left(C_{s,t}^-\right) = \exp\left(g(y_t^+, s_{\leq t}) - g(y_{t-1}, s_{\leq t-1})\right) + 1 \tag{3}$$

That is each edge is considered in turn as a formal logistic regression with the so-called change statistics used as predictors. To obtain the full unconditional distribution we sum over the space of possible edge permutations as follows.

$$\begin{aligned} p(y|\theta) &= \sum_s p(y|s, \theta) p(s) \\ &= \sum_s \left(p(s) \prod_{t=1}^{n_d} \frac{1}{Z_t} \exp\left(\theta \cdot C_{s,t}\right) \right) \end{aligned} \tag{4}$$

A key advantage of the LOLOG model is the ease of simulation from the model, to simulate a network we simply draw s from $p(s)$ and then do a sequential logistic regression simulation on the change statistics. The MOM method proposed uses this ease of simulation to easily carry out an approximate Newton-Raphson style optimisation to obtain the MOM estimate.

The question of the choice of $p(s)$ the PMF on the space of possible edge permutations remains. If there is no strong reason or a lack of the required data, for particular edges to considered before or after, a uniform PMF can used. However if there is a substantive reason to constrain the edge orderings, e.g. some actors in the network were introduced at a later stage to others e.g. a new year group arrives in a high school each year, edges between the upper years could reasonably be constrained to have been formed before edges between nodes in the upper years and the lower years.

The intractability of ERGMs, is due to the normalizing constant being a sum over the space of all possible graphs, the intractability in LOLOG is summing over all possible edge permutations. As the likelihood or a likelihood ratio (as is used for ERGMs), cannot be evaluated, the maximum likelihood estimate (MLE) for LOLOG is intractable, See Fellows (2018b) for details of a method of moments (MOM) approach to estimate model parameters. The graph statistics $g(y)$ are sufficient statistics for the LOLOG model, thus we seek θ_{MOM} such that $g(y) - \mathbb{E}_{\theta_{MOM}}[g(y)] = 0$. A Newton Raphson approach is available as we are able to differentiate the $\mathbb{E}_{\theta_{MOM}}[g(y)]$ with respect to θ and approximate its value using by sampling from the LOLOG model. Fellows (2018b) also provided a so called variational approximate fit, which gives a convenient starting point for the algorithm, speeding up the fitting procedure in practice.

Along with developing LOLOG models in Fellows (2018b), the `lolog` R package Fellows (2018a) has been developed. It provides a sophisticated, fast and user friendly method to fit LOLOG models to data. The package utilises the `Rcpp` package to allow meta programming; the heavy lifting in the background is carried out in C++.

However the MOM estimate approximates the MLE only asymptotically as the size of the network become large, and its finite sample distribution is unknown. As the size of a network is considered a fixed property of the network, the notion of the MOM estimate's asymptotic approximation to the MLE is not appealing. In addition since every LOLOG model (and ERGM) places non zero mass on the empty and full graph, for which both the MLE and MOM do not exist, the expectations and standard errors of the MLE and MOM parameter distributions do not exist either.

3. LOLOG Model Comparison and Relationship with ERGM

For the standard ERGM we have the probability mass function (PMF) $p(y|\theta) = \frac{\exp(\theta^\top \cdot g(y))}{Z_\theta}$, which yields the log odds interpretation of the θ parameter $\log\left(\frac{p(y^+_{i,j}|y^c_{i,j},\theta)}{p(y^-_{i,j}|y^c_{i,j},\theta)}\right) = \theta^\top \cdot C(y_{i,j}|y^c_{i,j})$ where $C(y_{i,j}|y^c_{i,j}) = g(y^+_{i,j}) - g(y^-_{i,j})$. Thus, conditional on the rest of the graph, each dyad can be thought of as a logistic regression on change statistics. This gives a helpful interpretation for the parameters, but does not help interpret the probability distribution of each edge unconditional of the rest of the graph.

Similarly for the LOLOG model, conditioning on an edge permutation s, at each step t, we have $\log\left(\frac{p(y^+_t|s_{\leq t},y_{t-1},\theta)}{p(y^-_t|s_{\leq t},y_{t-1},\theta)}\right) = \theta^\top \cdot C_{s,t}$. Thus at each time t, conditional on the network already formed by time t, each dyad is a logistic regression on change statistics.

So for ERGMs the interpretation is conditional on the entire rest of the network, whilst for LOLOG models for a specified edge ordering, the interpretation is conditional only on the network formed up until that point, though we must impress that the network formed up until that point will depend on the particular edge permutation.

We note that in the dyad independent case the edge ordering s does not matter and LOLOG reduces to logistic regression on change statistics as does ERGM. Thus ERGM and LOLOG models are equivalent in the dyad independent statistics case.

We may also consider dyad dependent ERGM and LOLOG models through the methods used to simulate from each we may gain insight as to why LOLOG often performs better and in particular is not subject to the extreme degeneracy issues ERGMs often display.

A network from an ERGM is simulated through an MCMC procedure where each dyad is considered in turn and then switched based on the probability of the dyad conditional on the rest of the network. This procedure is continued until the the MCMC chain has reached its stationary distribution, often many thousands of steps. As above the log odds of the dyad switching is equal to the inner product of the parameter and the change statistics of that dyad. The LOLOG model is formed by first sampling an dyad ordering in which to consider the dyads, and then, starting with an empty network adding an edge based on the log odds being the inner product of the parameter and the change statistic. Each dyad is considered for edge formation once and then the process is terminated leaving the simulated graph.

There is a similarity here, LOLOG considers each dyad exactly once, whereas the ERGM process can consider the dyads multiple times for both edge formation and dissolution. The disadvantage of the LOLOG model in this sense is that it makes the likelihood intractable due to the sum over all possible edge orderings. However LOLOG is easy to sample from. The ERGM likelihood is an exponential family so the likelihood is more tractable, though issues with the normalizing constant do arise. We suggest the reason that LOLOG models do not suffer from the same degeneracy, i.e. a small change in the parameters producing vastly different graphs, often only full or empty graphs, is that in the simulation each dyad is considered exactly once. This limits the scope for the explosive edge formation or deletion that often occurs when simulating from ERGM models.

More broadly we argue that the LOLOG motivated as a model with an easy simulation method, with parameters that remain interpretable, is more desirable than ERGMs where the likelihood is straightforward to write down, but requires MCMC procedures to simulate from.

4. Description of the Ensemble

We considered papers in the Social network journal, where ERGMs were fit to data. We included papers up to and including the January 2016 issue. There were 45 such papers, of which we selected 18 papers as follows. First we excluded bipartite ERGMs (5), we then included all networks with publicly available data (7) and selected a further 11 papers out of the remaining 33 based on their, subjectively assessed, novelty as well as the likely availability and ability to share data. We contacted the authors of the 11 papers and received the data for 7 of the papers. This gave an ensemble of 137 networks in 14 peer reviewed published papers, as many papers contained multiple networks. We note that 102 of these networks were from a single paper (Lubbers and Snijders, 2007), which were omitted from our analyses, leaving 35 networks. Table 1 shows a brief summary for each of the networks.

Table 1: Summary table for ensemble of networks

Description	Network	Nodes	Edges	Directed	Nodal Covariates
Add Health		1681	1236	Undirected	4
School Friends		Various	Varies	Directed	3
Kapferer's Tailors		39	267	Undirected	0
Florentine Families		16	15	Undirected	2
German Schoolboys		53	53	Directed	4
Employee Voice	1	27	104	Directed	3
Employee Voice	2	24	53	Directed	3
Employee Voice	3	30	126	Directed	3
Employee Voice	4	31	139	Directed	3
Employee Voice	5	37	149	Directed	3
Employee Voice	6	39	155	Directed	3
Office Layout	University	67	211	Directed	0
Office Layout	University	69	203	Directed	0
Office Layout	Research	109	458	Directed	0
Office Layout	Publisher	119	872	Directed	0
Disaster Response		20	148	Directed	7
Company Boards	2007	808	1997	Undirected	0
Company Boards	2008	808	1740	Undirected	0
Company Boards	2009	808	1682	Undirected	0
Company Boards	2010	808	1622	Undirected	0
Swiss Decisions	Nuclear	24	282	Directed	8
Swiss Decisions	Pensions	23	294	Directed	8
Swiss Decisions	Foreigners	20	169	Directed	8
Swiss Decisions	Budget	25	224	Directed	8
Swiss Decisions	Equality	24	248	Directed	8
Swiss Decisions	Education	20	227	Directed	8
Swiss Decisions	Telecoms	22	256	Directed	8
Swiss Decisions	Savings	19	138	Directed	8
Swiss Decisions	Persons	26	280	Directed	8
Swiss Decisions	Schengen	26	316	Directed	8
University Emails		1133	10903	Undirected	0
School Friends	grade 3	22	177	Directed	1
School Friends	grade 4	24	161	Directed	1
School Friends	grade 5	22	103	Directed	1
Online Links	Hyperlinks	158	1444	Directed	3
Online Links	Framing	150	1382	Undirected	3

Our selection of ERGM papers was at first a census of papers in the journal *Social Networks* using the ERGM framework. The conclusions drawn from this study should be considered stronger than if the networks selected were sampled at random or through convenience. We do note that we did take a convenience sample as described above as a first wave of networks to request data for, though this was also chosen based on our thoughts on which networks the authors would be able and willing to share.

The ability to recreate peer reviewed research in which statistical network models were used, irrespective those models being ERGM, gives strong empirical support for LOLOG models. This is hard if not impossible to prove theoretically. Such statements such as "the LOLOG model is useful for the types of network data sets that researchers often have substantive interest in" are difficult to quantify. Thus fitting LOLOG to networks data sets "in the wild" provides confidence that the model is useful in addition to theoretical guarantees that the model can be used to represent arbitrary probability mass functions over the space of possible networks.

We considered papers that used ERGMs for their statistical analyses as the ERGM class of models is arguable the most widely used generative statistical model for network analyses (Amati et al., 2018). LOLOG models and ERGMs are typically used to seek to model global network structure using local network structure, thus comparing the two models is appealing. Both LOLOG and ERGM are also fully general network models Fellows (2018b). That is any given PMF over the space of networks can be represented as either an ERGM or a LOLOG model with suitable sufficient statistics. However specifying interpretable models that fit the data is often the practical challenge. There is no obvious reason to suspect similar performance in terms of fit and interpretability, when fit with similar network statistics, on the same network. In particular it seems likely that in papers published that fit an ERGM model, ERGM performs well on this data set, thus we expect a publication bias towards networks that suit ERGM well, which may not necessarily suit LOLOG well. We therefore suggest that good performance on data published with ERGM fits, is a conservative indicator that LOLOG is a useful model for analysing social networks.

We also note that the LOLOG model allows for the consideration of information on the order of the edge formation within a network the researcher may have. This is currently implemented as allowing edge orderings to be constrained to those orderings compatible with the sequential adding of nodes to the network, followed by the consideration of all possible new edges. This is not possible in ERGM and few of the available networks had plausible ordering mechanisms. However this may not be entirely due the lack thereof, indeed without the ability to model such an ordering process with ERGM, it seems likely that even if there is a compelling sequential node adding process the data would not be considered or collected.

Though we are chiefly concerned with the performance of the LOLOG model, for each network we first recreated the ERGM fit, to ensure we were using the data correctly, so our comparison is valid. The fits were carried out in R using the `ergm` package (Handcock et al., 2018), and the `lolog` package (Fellows, 2018a). Where possible decay parameters were used as stated, where this was not available, the first option was to let $\alpha = 0.5$.

Table 2. Office Layout ERGM fits matching published fit. *** p-value < 0.001 , ** p-value < 0.01, * p-value < 0.5

	University 2005	University 2008	Research Institute	Publisher
Edges	-3.4 (0.37)***	-4.41 (0.2)***	-4.1 (0.12)***	-5.07 (0.15)***
Reciprocity	0.38 (0.45)	0.62 (0.31)***	2.39 (0.2)***	-1.26 (0.19)***
GWESP(0.5)	1.36 (0.14)***	1.24 (0.11)***	0.92 (0.07)***	2.09 (0.09)***
Usefulness	0.7 (0.15)***	0.54 (0.11)***	0.81 (0.04)***	1.31 (0.05)***
Team Match	0.78 (0.18)***	0.56 (0.1)***	NA	NA
Floor Match	0.15 (0.26)	0.58 (0.14)***	NA	NA
Metric Distance	-0.04 (0.01)***	-0.01 (0)***	-0.01 (0)***	NA
Topo Distance	NA	NA	NA	-0.06 (0)***

5. Case Study of LOLOG and ERGM fits: Complex networks where ERGM is insufficient

In this section we consider a case study form a single published paper where the networks in question are sufficiently complex to demonstrate that ERGM can be insufficient and LOLOG can help in modelling social network data.

We consider four networks of daily social interactions between workers within four different office spaces, an ERGM based analysis was originally carried out in Sailer and McCulloch (2012). Ties are present between node i and node j if node i reported daily social interaction with node j. Two of the networks are of a UK university faculty before and after an office refurbishment, the remaining two are a German research institute and a corporate publishing company. The networks are directed and have 69, 63, 109 and 120 nodes respectively.

The research question of interest in Sailer and McCulloch (2012) was the effect of spatial distance in the formation of social interactions within an office environment. Initially for each network the authors fit a model using the edges, geometrically weighted edgewise shared partners (GWESP) and mutual terms, together with a matching nodal covariate term on floor of the building and team. A usefulness term was added as an edge covariate term, with the value for dyad (i,j) being the node i self reported perception of the usefulness of node j. The best fitting model was then selected using the Akaike Information Criterion (AIC) and then a variety of different distance metrics were added individually as an edge covariate. The best model in terms of AIC was once again selected and analysed. Notably no analysis of the goodness of fit for the models was provided.

5.1. *Model Fits*

We were able to recreate the selected ERGM fit for all four networks, shown in Table 2. We were able to obtain LOLOG fits with the same covariates, as the ERGM fits for all networks, we summarize the fits in Table 3.

Table 3. Office Layout LOLOG fit with same terms as published ERGM. *** p-value < 0.001 , ** p-value < 0.01, * p-value < 0.5

	University 2005	University 2008	Research Institute	Publisher
Edges	-1.69 (0.38)***	-3.67 (0.36)***	-3.18 (0.13)***	-1.63 (0.09)***
Reciprocity	1.99 (0.34)***	1.96 (0.31)***	3.9 (0.25)***	0.64 (0.2)***
GWESP(0.5)	0.55 (0.12)***	0.87 (0.13)***	0.73 (0.09)***	-0.22 (0.06)***
Usefulness	1.02 (0.15)***	0.81 (0.14)***	1.21 (0.05)***	1.89 (0.06)***
Team Match	1.29 (0.19)***	0.72 (0.19)***	NA	NA
Floor Match	-0.28 (0.3)	1.08 (0.29)***	NA	NA
Metric Distance	-0.07 (0.01)***	-0.02 (0.01)***	-0.02 (0)***	NA
Topo Distance	NA	NA	NA	-0.1 (0)***

Table 4. Office Layout LOLOG fit with GWESP and 2 and 3 in and out stars. *** p-value < 0.001 , ** p-value < 0.01, * p-value < 0.5

	University_2005	University_2008	Research_Institute	Publisher
Edges	-3.2 (0.67)***	-5.04 (0.59)***	-4.04 (0.22)***	-4.87 (1.19)***
Reciprocity	2.03 (0.77)***	1.11 (0.45)***	4.7 (0.52)***	3.16 (1.27)***
GWESP(0.5)	0.33 (0.2)	0.49 (0.16)***	0.77 (0.11)***	0.01 (0.26)
Out-2-Star	1.39 (0.26)***	0.65 (0.16)***	0.41 (0.07)***	0.69 (0.15)***
Out-3-Star	-0.28 (0.07)***	-0.07 (0.03)***	-0.04 (0.01)***	-0.02 (0)***
In-2-Star	0.26 (0.22)	0.25 (0.15)	0.21 (0.12)	0.73 (0.54)
In-3-Star	-0.04 (0.05)	-0.03 (0.02)	-0.09 (0.03)***	-0.18 (0.1)
Usefulness	1.07 (0.2)***	0.75 (0.16)***	1.28 (0.07)***	2.98 (0.61)***
Team Match	1.93 (0.31)***	1.14 (0.25)***	NA	NA
Floor Match	-0.24 (0.47)	1.35 (0.43)***	NA	NA
Metric Distance	-0.09 (0.01)***	-0.02 (0.01)***	-0.02 (0)***	NA
Topo Distance	NA	NA	NA	-0.24 (0.06)***

In addition we show the LOLOG fit using GWESP, 2 and 3 in and out stars, together with all covariate matches and metric distance in Table 4.

Table 5 shows the fitted model where the nodes are added in the order of their average usefulness, as reported by the other nodes. As we suspect more useful nodes may have been in the office longer or should be the first point of contact for new employees we suggest this as a plausible ordering mechanism. We note that the fits with this ordering are comparable apart from the Publisher network where the estimated standard errors became very large.

We were also able to fit LOLOG models to each of the networks when substituting the GWESP term for a triangle term which is not possible with ERGM, we summarise this in Table 6. Though we note that the estimated standard errors for the Publisher network are very high, suggesting there is great uncertainty in the fitted data generating process, suggesting a poor model fit. The estimated standard errors for the mutual and triangle terms for the University in 2005 and 2008 also have the same problem though not as severe,

Table 5. Office Layout LOLOG fit with GWESP in and out stars, and ordering by usefulness. *** p-value < 0.001 , ** p-value < 0.01, * p-value < 0.5

	University_2005	University_2008	Research_Institute	Publisher
Edges	-3.62 (0.61)***	-5.31 (0.68)***	-4.06 (0.21)***	-8.91 (6.26)
Reciprocity	1.68 (0.59)***	1.11 (0.49)***	4.89 (0.54)***	4.32 (3.95)
GWESP(0.5)	0.23 (0.2)	0.51 (0.18)***	0.81 (0.12)***	-0.31 (0.39)
Out-2-Star	1.34 (0.26)***	0.66 (0.17)***	0.44 (0.07)***	1.05 (0.78)
Out-3-Star	-0.26 (0.07)***	-0.08 (0.03)***	-0.04 (0.01)***	-0.03 (0.02)
In-2-Star	0.3 (0.22)	0.28 (0.14)***	0.22 (0.11)	2.19 (2.12)
In-3-Star	-0.03 (0.05)	-0.03 (0.03)	-0.1 (0.03)***	-0.45 (0.43)
Usefulness	1.04 (0.2)***	0.76 (0.15)***	1.33 (0.08)***	4.65 (3.28)
Team Match	1.97 (0.33)***	1.27 (0.27)***	NA	NA
Floor Match	0.18 (0.38)	1.45 (0.44)***	NA	NA
Metric Distance	-0.1 (0.01)***	-0.02 (0.01)***	-0.02 (0)***	NA
Topo Distance	NA	NA	NA	-0.37 (0.3)

Table 6. Office Layout LOLOG fit with triangles instead of gwesp term. *** p-value < 0.001 , ** p-value < 0.01, * p-value < 0.5

	University 2005	University 2008	Research Institute	Publisher
Edges	-2.05 (0.82)***	-3.9 (0.67)***	-3.36 (0.15)***	-5.4 (49.3)
Reciprocity	-2.96 (5.83)	-0.08 (1.6)	3.34 (0.36)***	-24.8 (367.3)
Triangles	2.69 (2.95)	1.2 (0.73)	0.61 (0.13)***	3.71 (50.63)
Usefulness	1.35 (0.53)***	0.83 (0.16)***	1.21 (0.06)***	6.57 (88.39)
Team Match	1.8 (0.85)***	0.9 (0.35)***	NA	NA
Floor Match	-0.29 (0.88)	1.07 (0.46)***	NA	NA
Metric Distance	-0.1 (0.05)***	-0.02 (0.01)***	-0.02 (0)***	NA
Topo Distance	NA	NA	NA	-0.39 (5.82)

and fall out of significance for these model fits.

As the triangle term increases the estimated standard errors and does not improve the goodness of fit (see next section), we suggest using the GWESP term.

5.2. Goodness of Fit

For the interpretation of model parameters to be valid, we must show that the model is a plausible generating process for the observed network. We follow the goodness of fit procedure as in Hunter et al. (2008). Firstly we consider the goodness of fit for the published ERGM model, and the LOLOG model with the same terms. Figures 1, 2, and 3 show the goodness of fit on the in-, out-degree and ESP distributions for each network.

Table 5.2 shows comments on the goodness of fit for each network, using the recreated published ERGM and the LOLOG model with published ERGM terms. Where no comment

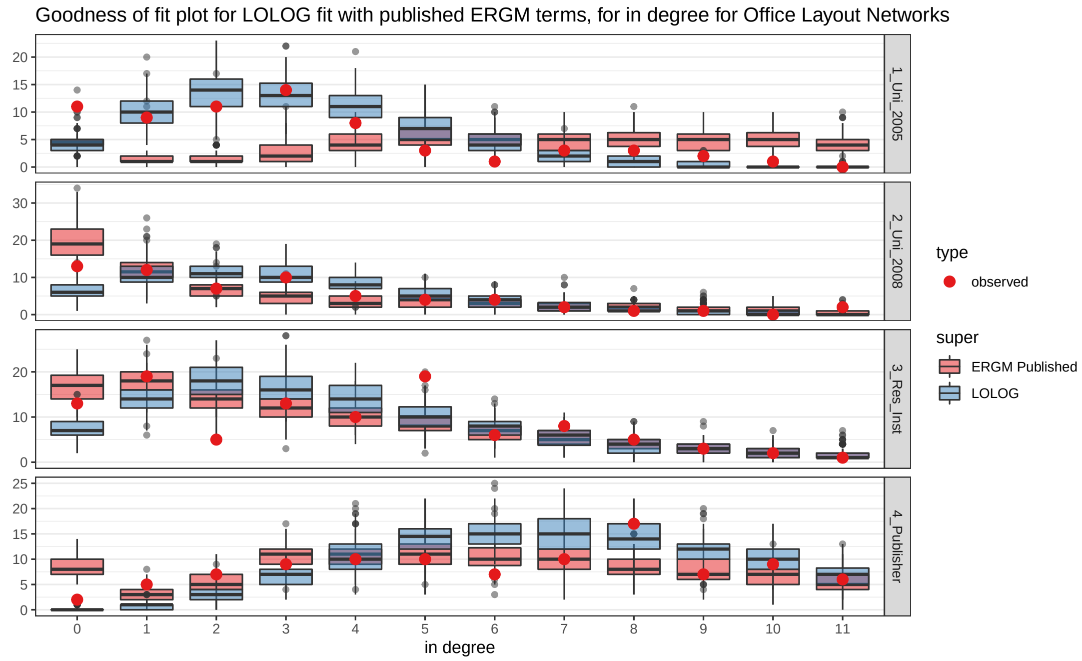

Fig. 1. Sailer's Offices ERGM and LOLOG model with published terms, in-degree Goodness of Fit

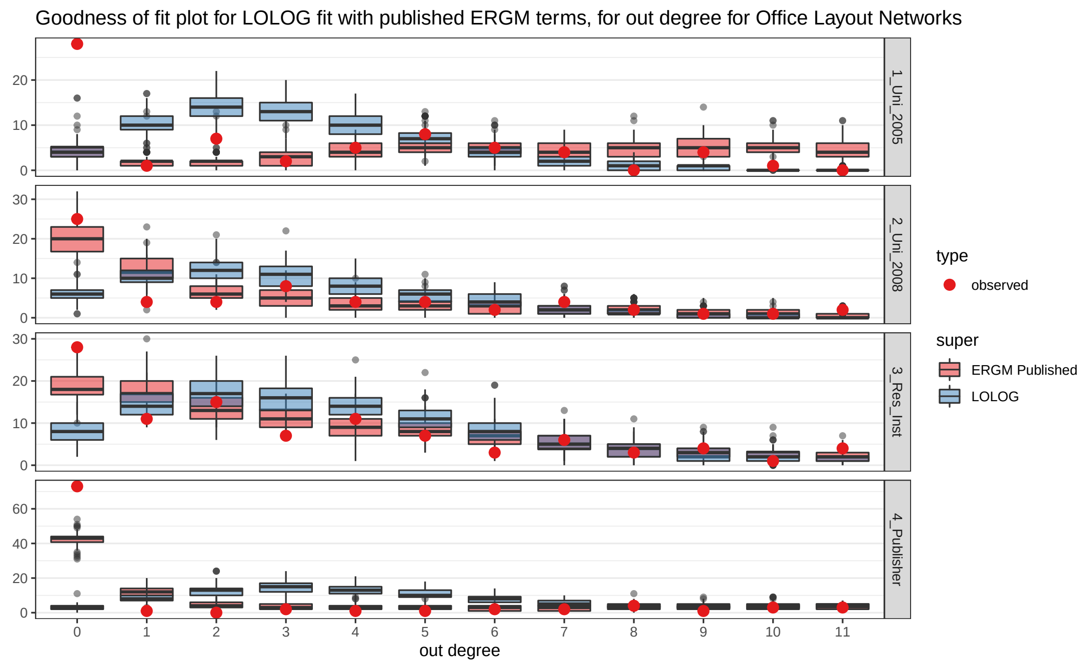

Fig. 2. Sailer's Offices ERGM and LOLOG model with published terms, out-degree Goodness of Fit

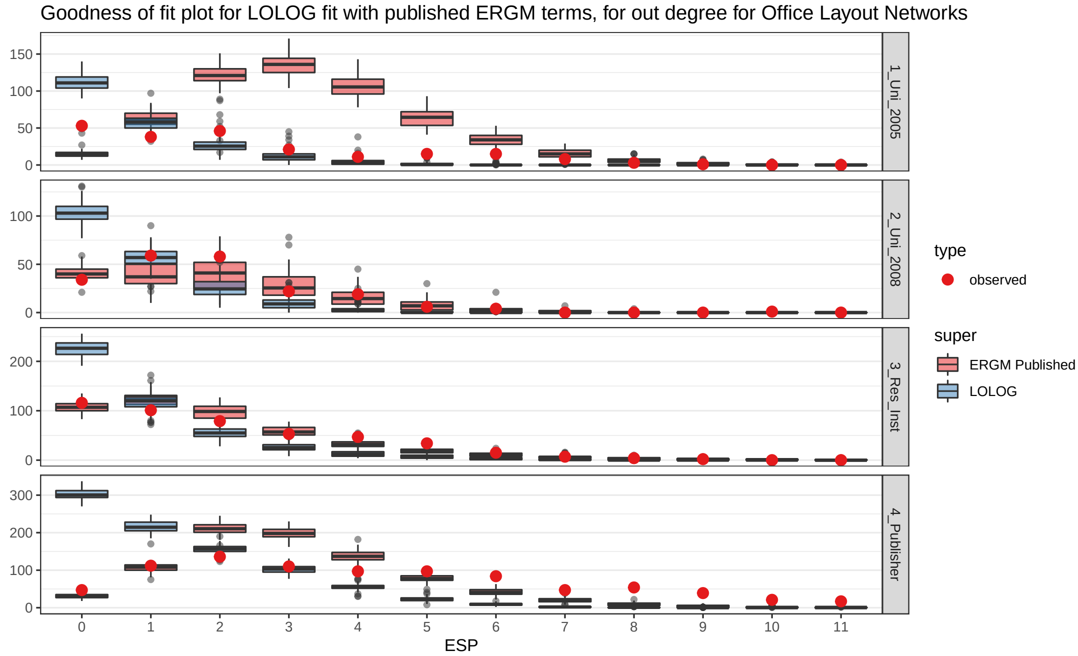

Fig. 3. Sailer's Offices ERGM and LOLOG model with published terms, ESP Goodness of Fit

is made for any of the goodness of fit terms or any model, the model fits well on that statistic.

Clearly all models for all networks have a least one of the in-degree, out-degree or ESP distribution of the observed networks not being realised as a typical value for the fitted models. As a result the models do poor job of recreating networks similar to the observed, and thus inference based on the parameter estimates and standard errors should be treated with caution. In particular we note that the LOLOG model did not seem to help improve the fit for any of the networks in question here, when using the same terms as the ERGM required to be non degenerate.

We also show the goodness of fit for the LOLOG model with GWESP and 2 and 3 in and out stars for each model in Figures 4, 5 and 6. We note here that all models fit the in degree distribution well, all models except the publisher fit the out degree distribution well and the university 2005 and 2008 models fit well on the ESP distribution. This is an improvement in all cases for versus the ERGM models published in Sailer and McCulloch (2012).

In addition we considered the goodness of fit for the LOLOG model with GWESP and 2 and 3 in and out stars but ordering by usefulness. The goodness of fit is similar to the same model without the ordering, with some effect on the 0 and 1 in and out degree distributions. However the difference is not large, and as the qualitative justification for including that particular ordering is relatively tenuous we deemphasize the fit with the ordering.

Table 7. Summary of GOF for ERGM and LOLOG with published terms for Office Layout networks

Network	ERGM	LOLOG
2005 University	Fits poorly on out-degree Fits poorly on ESP	Fits poorly on in-degree Fits poorly on ESP but much better than ERGM in-degree plausible with ERGM or LOLOG
2008 University	Fits poorly on out-degree Fits poorly on ESP	ERGM convex, LOLOG concave Fits poorly on out-degree Fits poorly on ESP
Research Institute	Fits poorly on out-degree Fits poorly on ESP	Fits poorly on in-degree Fits poorly on out-degree Fits poorly on ESP
Publisher	Fits poorly on in-degree Fits poorly on out-degree Fits poorly on ESP	Fits poorly on in-degree Fits poorly on out-degree Fits poorly on ESP

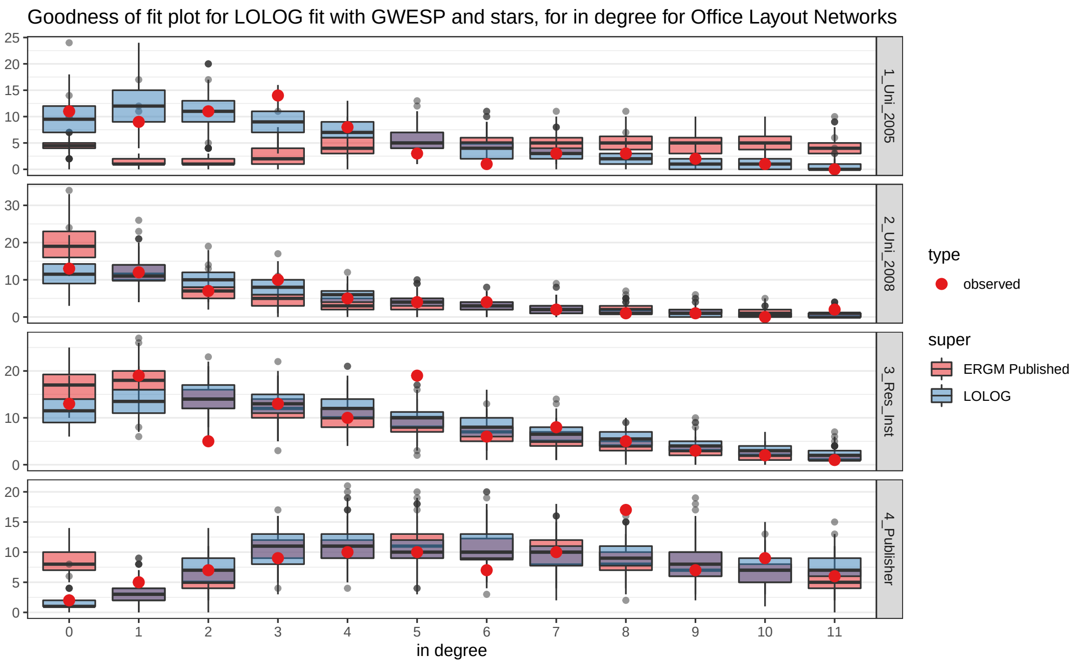

Fig. 4. Sailer's Offices LOLOG model with GWESP and stars, in degree Goodness of Fit

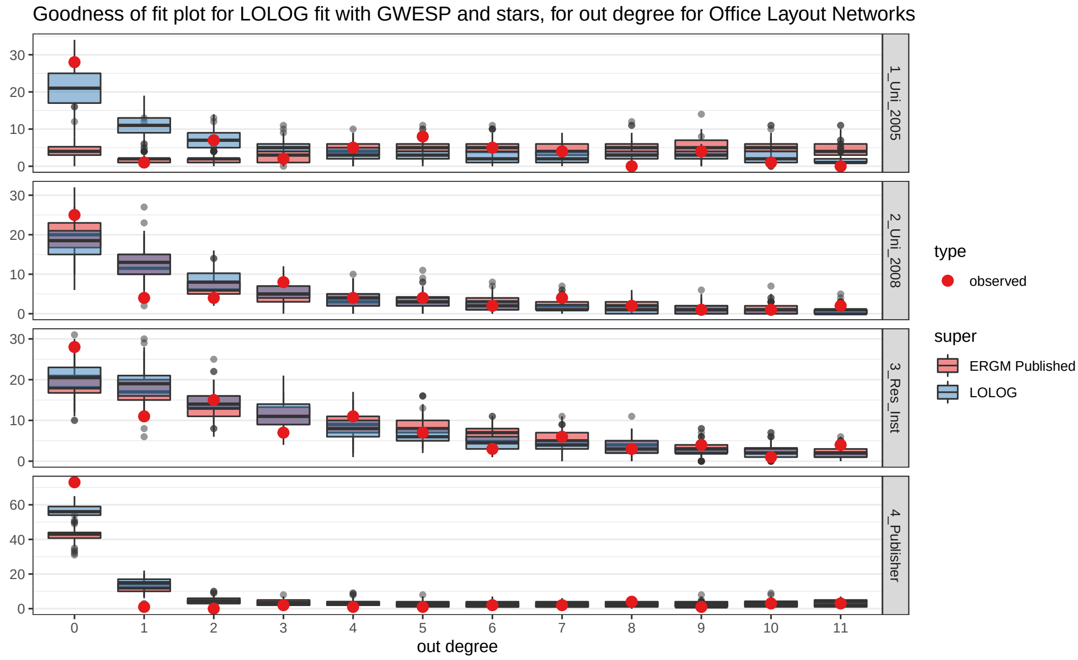

Fig. 5. Sailer's Offices LOLOG model with GWESP and stars, out degree Goodness of Fit

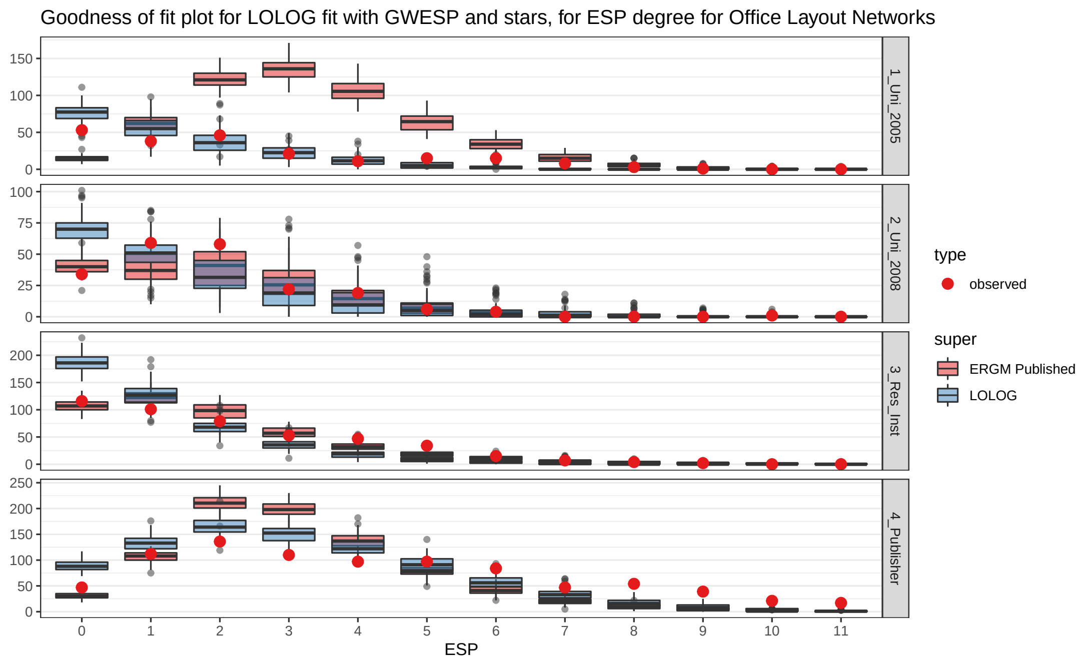

Fig. 6. Sailer's Offices LOLOG model with GWESP and stars, ESP Goodness of Fit

5.3. Interpretation

These networks are of particular interest as they represents a real world example of applied researchers seeking a statistical tool to explain their thoughts on a subject and analyse their collected data. Good performance in such a setting for the LOLOG model suggest the model could be of real use to the applied social network research community. Using these four non trivially complex networks as an example, help us to present the the utility of the LOLOG model. The ERGM model and LOLOG model with the terms as in Sailer and McCulloch (2012) produced models with the same qualitative interpretation. That is the network of daily social interactions in each office space, exhibits a tendency for mutual edges and edges that facilitate social closure. Ties are also more likely to form in dyads where there is a high level of usefulness of the receiver to the sender as well as along matching teams. Matching floors play a roll in some models, however the distance measure plays a roll in all models which is correlated with the matching floors, with dyads more likely to form when the distance between nodes spatial locations is lower.

However it is important to note that neither the LOLOG or the ERGM with the fitted terms fitted the data well in terms of all of in and out degree and ESP distribution. Therefore the models are not capturing basic aspects of the observed network data, and the above interpretation should be treated with caution. In particular the Publisher network proved especially hard to fit.

Using the triangle term in the LOLOG model in place of the GWESP term did not improve the fit. Including 2 and 3 in and out star terms yields models that fit much better on the in and out degree distribution as well and the ESP distribution. We therefore have more belief that inferences from these models are valid. They show similar conclusions to the published ERGM, though in addition we observe a positive significant out-2-star term and a negative significant out-3-star term, suggesting that there is a tendency for some nodes to nominate many more people that others. This tendency for super daily interactors was not captured in the ERGM fit. We also note that the lack of significant in-2-star parameter suggests that there is not a corresponding tendency for some nodes to attract more interactions, when their usefulness had already been accounted for. Thus we can infer that perhaps there is a surplus of unwanted daily interaction due to nodes with a tendency for high out degree. Thus the LOLOG model allowed for a better fitting model, as well as deeper interpretations to be made in this case.

We tried to fit an ERGM model with the in and out geometrically weighted degree (GWDEG) terms but this was degenerate for the University 2005 and 2008 networks. For the Research institute and Publisher out GWDEG was negative and significant, at odds with the LOLOG model, though the fit was still poor and inferior to the LOLOG model. We do not comment further on this, though it is concerning that the ERGM may give different interpretations, the GWDEG terms were not discussed in Sailer and McCulloch (2012).

We note the performance differential in the LOLOG and ERGM fit. On our departmental server we ran each with a single core with Intel(R) Xeon(R) Platinum 8160 CPU @ 2.10GHz processor. The recreated ERGM took around 35 seconds, and the LOLOG took around 8 seconds. For larger networks, we found parallelisation in the network simulation step of

[!h] **Table 8.** Summary table glossary

Key	Feature
1	Able to recreate the published ERGM qualitatively
2	Recreation of published ERGM fits the network well
3	Able to fit LOLOG with published ERGM terms
4	LOLOG with published ERGM terms fits well
5	Able to fit LOLOG with ERGM Markov terms usually degenerate in ERGM
6	Better fit achieved with LOLOG than published ERGM
7	Published ERGM and best fitting LOLOG have consistent interpretation
8	Which model is more useful subjectively

the fit to be extremely helpful for both the LOLOG and ERGM models. The LOLOG model is easy to simulate from whereas ERGM requires MCMC for the simulation step. From our experience for larger networks the performance differential between LOLOG and ERGM can be much greater, in particular when the ERGM MCMC simulation is computationally expensive.

6. Summary of Results for the Ensemble

Table 9 provides a summary of the ERGM and LOLOG fits for the networks in our ensemble, descriptions of the column names in Table 9 are given in Table 8. We provide brief commentary of the results overall, and provide more detailed modelling comments for each network in Appendix A

Table 9: Summary table for LOLOG and ERGM Fits

Description	Network	Nodes	1	2	3	4	5	6	7	8
Add Health		1618	Yes	No	Yes	No	Yes	Yes	Yes	LOLOG
School Friends		Various								
Kapferer's Tailors		39	Yes	No	Yes	No	Yes	Yes	No	LOLOG
Florentine Families		16	Yes	Yes	Yes	Yes	Yes	Yes	No	ERGM
German Schoolboys		53	Yes	Yes	No	NA	Yes	Yes	Yes	Both
Employee Voice	1	27	No	NA	Yes	Yes	Yes	Yes	NA	LOLOG
Employee Voice	2	24	Yes	Yes	No	NA	No	No	NA	ERGM
Employee Voice	3	30	No	NA	Yes	Yes	Yes	Yes	NA	LOLOG
Employee Voice	4	31	No	NA	Yes	Yes	Yes	Yes	NA	LOLOG
Employee Voice	5	37	No	NA	Yes	Yes	Yes	Yes	NA	LOLOG
Employee Voice	6	39	No	NA	Yes	Yes	Yes	Yes	NA	LOLOG
Office Layout	University	67	Yes	No	Yes	No	Yes	Yes	Yes	LOLOG
Office Layout	University	69	Yes	Yes	Yes	No	Yes	Yes	Yes	LOLOG
Office Layout	Research	109	Yes	Yes	Yes	No	Yes	Yes	Yes	LOLOG
Office Layout	Publisher	119	Yes	No	Yes	No	Yes	Yes	Yes	LOLOG

Table 10. Citation summary table

Network	Citation
Add Health	Harris et al. (2007)
School Friends	Lubbers and Snijders (2007)
Kapferer's Tailors	Robins et al. (2007)
Florentine Families	Robins et al. (2007)
German Schoolboys	Heidler et al. (2014)
Employee Voice	Pauksztat et al. (2011)
Office Layout	Sailer and McCulloch (2012)
Disaster Response	Doreian and Conti (2012)
Company Boards	Wonga et al. (2015)
Swiss Decisions	Fischer and Sciarini (2015)
University Emails	Toivonen et al. (2009)
School Friends	Anderson et al. (1999)
Online Links	Ackland and O'Neil (2011)

Disaster Response		20	No	No	No	No	Yes	Yes	No	LOLOG
Company Boards	2007	808	No	No	No	No	Yes	Yes	NA	LOLOG
Company Boards	2008	808	No	No	No	No	Yes	Yes	NA	LOLOG
Company Boards	2009	808	No	No	No	No	Yes	Yes	NA	LOLOG
Company Boards	2010	808	No	No	No	No	Yes	Yes	NA	LOLOG
Swiss Decisions	Nuclear	24	No	Yes	No	NA	Yes	Yes	Yes	ERGM
Swiss Decisions	Pensions	23	No	Yes	Yes	No	Yes	No	No	ERGM
Swiss Decisions	Foreigners	20	No	Yes	No	NA	Yes	No	No	ERGM
Swiss Decisions	Budget	25	No	Yes	No	NA	Yes	Yes	No	ERGM
Swiss Decisions	Equality	24	No	No	No	NA	Yes	Yes	No	LOLOG
Swiss Decisions	Education	20	No	No	Yes	No	Yes	Yes	NA	LOLOG
Swiss Decisions	Telecoms	22	No	No	No	NA	Yes	Yes	NA	LOLOG
Swiss Decisions	Savings	19	Yes	Yes	No	NA	Yes	Yes	No	ERGM
Swiss Decisions	Persons	26	No	Yes	No	NA	Yes	Yes	No	ERGM
Swiss Decisions	Schengen	26	No	No	No	No	Yes	Yes	NA	LOLOG
University Emails		1133	No	No	No	No	No	No	NA	Neither
School Friends	grade 3	22	Yes	No	No	No	Yes	Yes	NA	LOLOG
School Friends	grade 4	24	Yes	No	No	No	Yes	Yes	NA	ERGM
School Friends	grade 5	22	Yes	No	No	No	Yes	Yes	NA	ERGM
Online Links	Hyperlinks	158	Yes	No	Yes	No	Yes	Yes	Yes	LOLOG
Online Links	Framing	150	Yes	No	Yes	No	Yes	No	Yes	LOLOG

6.1. Overall Comments

We make some general comments regarding the significant amount of information on the hundreds of models fitted to the data that we gathered, more detailed summaries for each individual network are contained in Appendix A. More detailed overall comments on the study are in the discussion in Section 7.

Overall we see that in many cases, we were not able to recreate the published ERGM, and often when we could, the model did not fit the data well using the GOF methodology of Hunter et al. (2008). We were sometimes able to use the same terms as the published ERGM to fit a LOLOG model, however there were also some networks where we could not fit the LOLOG model with ERGM terms.

Where a LOLOG model with ERGM terms was able to be fit it usually did not fit the data well. However in almost all cases we were able to fit the LOLOG model, with terms that usually result in degenerate ERGMs e.g. triangles and stars, and usually could achieve at least as good a fit as the published ERGM.

In general our experience in fitting the LOLOG model was that it was faster to fit that ERGM, with the MOM estimation typically requiring little to no tuning, in contrast to ERGM models which can require some tuning. In addition the triangles and star terms that can be readily fit with LOLOG models provide a simple and intuitive interpretation for users of the model.

7. Discussion

We have shown that the LOLOG model can be fit to most members of an ensemble of network data sets that have published ERGM fits in the journal *Social Networks*. We report a case-study of a complex data set and show that the LOLOG model is at least the equal of the ERGM, in terms of goodness of fit and interpretability. We carried out fits to 35 networks in total and gave a summary of each of the networks' fits. We regard this as strong evidence that the LOLOG model is a useful model for modelling real social network data, as journal articles with published ERGM fits likely have a bias towards data sets that are well suited to ERGMs.

In carrying out this study we have gained a great deal of practical experience in the types of tasks for which ERGMs are used, as well as practical problems in fitting them, in particular code run time and degeneracy issues. We have found the LOLOG model to be in general more user friendly and faster to fit, leading to easier identification of poor models, and a much faster data analysis procedure. The benefits of this should not be overlooked, in particular when social network analyses are often of interest to applied researchers whose expertise is not statistical modelling. As a result LOLOG models seem particularly better suited to larger networks, increasing the maximum size of networks that can be feasibly analysed with generative social network models.

LOLOG models can usually be fit with terms that are almost always degenerate for ERGMs on even small networks. Using this greater flexibility of specification, we were often able to achieve a better fit. In addition the need to use complex geometrically weighted statistics is reduced, aiding interpretability of the LOLOG model. In practice we also believe LOLOG models could facilitate more robust model selection procedures. The degeneracy issues of ERGM as well as the time taken to fit the model, can result in researchers omitting terms based on their degeneracy, as well as considering fewer models than they would want. The fast fit and robust to degeneracy properties of the LOLOG model should help alleviate these practical issues.

We have also seen that qualitative interpretations of analyses carried out with both ERGMs and LOLOG models are generally in agreement. We do note, however, from our experiences that the LOLOG model applied to small networks can result in parameter estimates with high variance, where the ERGM model parameters have lower variances, more amenable to interpretation.

Goodness of fit of LOLOG models also compares favourably with the ERGMs, with little drop in quality, for the same terms. In particular with the ability to use simpler terms for the LOLOG model we were often able to achieve improved fit over the published ERGMs in the ensemble of networks that we fit.

The LOLOG model has the advantage of being able to account for edge orderings, we believe that this may be helpful for analysing network data, however we have not seen clear benefits in the ensemble of network data in this study. We suggest that as statistical models for analysing edge ordering processes are not available, such processes and data to model them are often not considered by applied researchers. We also emphasise that although including an ordering mechanism in the edge formation process, may be novel and at odds with some researcher's philosophical beliefs around how network problems should be approached, the LOLOG model is a fully general model over the space of networks. That is it can represent any probability mass function over the space of networks. Therefore even if it is hard to justify such an edge formation procedure, the LOLOG model may still be a useful approach to understanding the social processes producing network data.

8. Acknowledgements

The project described was supported by grant number 1R21HD075714-02 from NICHD, and grant numbers SES-1230081 and IIS-1546300 from the NSF.

We would like to acknowledge and thank all of the authors that provided data that made this study possible. We would like to thank the following, for taking the time to correspond with us and for providing their data: Greetje Van Der Werf, Lotte Vermeij, Miranda Lubbers, Mikko Kivelä, Riitta Toivonen, Jari Sarimäki, Jukka-Pekka Onnela, Robert Ackland, Birgit Pauksztat, Kerstin Sailer, Dean Lusher, André Gygax, Roger Guimera, and Manuel Fischer. We note that there is uncertain personal benefit as well as some risk in doing so. We greatly appreciate their time and effort in preserving their data and providing it when we requested. They have made significant contributions to reproducibility of research in its many forms.

Appendices

A. Individual Network Modelling Comments

A.1. Add Health

This was a network of high school students, obtained from the well studied National Longitudinal Study of Adolescent to Adult Health (Harris et al., 2007). Networks from the survey have been fit using ERGMs [Goodreau (2007), Hunter et al. (2008)]. There are multiple networks available but the particular network in this case obtained from ?? (Add) has 1681 nodes with covariates for grade, gender and race provided, this matches the data used in the detailed ERGM fit in Goodreau (2007).

We were able to fit ERGMs and LOLOG with the published ERGM terms but the models did not fit the data well as noted extensively in Goodreau (2007).

Goodreau (2007) provided extensive commentary on the goodness of fit of many ERGM models, the authors considered the degree and ESP distributions as well as the distribution of geodesic distances between nodes. A good fit on the degree distribution was only able to be achieved by the authors by including terms that sacrificed the fit on the ESP distribution. The LOLOG models exhibited similar problems, however we were able to fit a LOLOG model with triangles and stars to achieve an improved fit, but did not eliminate this issue.

A.2. Junior High

These data are 102 friendship networks in junior high school. Lubbers and Snijders (2007) performed reanalysis of 102 networks consider pseudo-likelihood and the then recent MCMCMLE methods. We omitted this from our study due to its size, and the fact that is atypical of the usual applied social network analyses that ERGM is used for.

A.3. Kapferer's Tailors

The paper that fit this ERGM was Robins et al. (2007). The authors in this paper were investigating applying novel specifications to a range of networks available through the UCINET software. We note that the models were fitted using `pnet` software.

In the Kapferer Tailor Shop networks, the nodes are workers in a Zambian tailor shop, with two different interactions, social and "instrumental"(work or assistance related). These were collected at two distinct time points giving 4 networks. Robins et al. (2007) stated the ERGM fit of the kapfts1 - the first social interaction network, of 39 nodes, which is what concerns us here.

Our estimated coefficients were different to those stated by the authors, though the results are not qualitatively different, the use of `pnet` instead of the `ergm` package may contribute to this. We were unable to recreate the fit when including a 2-star parameter unlike the authors who state a result for this. We were able to fit an ERGM with high decay parameters, but this neither matched the published ERGM, nor provided any extra insight. We were also unable to fit the network to an ERGM with triangle and star parameters as stated by the authors.

We were able to fit LOLOG models to the network using the geometrically weighted terms. However in contrast to ERGM we were able to fit LOLOG just with triangle, 2 and 3 star parameters.

Using higher order terms as in the published fit, the fit of the LOLOG and ERGM models were poor on the degree and edgewise shared partners (esp) distribution of the network. Fitting the LOLOG model with triangle and star terms was a slight improvement.

The authors commented on the ERGM fit that the significant and positive geometrically weighted edgewise shared partner parameter, suggests clustering into dense regions of overlapping triangles. In addition the alternating k star triangle not being significant suggests core periphery structure due to triangulation not popularity effects. This interpretation was consistent with the LOLOG triangle and star model interpretation.

A.4. Florentine Families

This network was the second network fit in Robins et al. (2007). In this network the nodes are 16 influential families in Florence in the 1500s. Marital networks and business tie networks are available with the fit published being the business network.

The published fitting focused on structural terms, including nodal covariates did not have a large effect on the coefficients.

We were able to recreate the published ERGM and fit LOLOG model with the same terms. Both models fitted the observed network well. However the LOLOG model parameters had high estimated variance, suggesting that the model fit could be sensitive to variation in the data. This limits the interpretation possible from the LOLOG model. We do however note that variance estimates for both the LOLOG and ERGM model are only asymptotically valid, so are likely not valid for such a small network.

A.5. German Schoolboys

This network is a directed network of friendships between German schoolboys in class from 1880 to 1881, collected by Johannes Delitsch, in one of the earliest studies to engage a network based approach. This was reanalysed in Heidler et al. (2014) with an ERGM approach, and compared with similar friendship networks in schools today.

Nodal covariates were available for academic class rank, whether the student was repeating class rank, whether the student gave sweets out, and whether the student was handicapped or not. Note that academic rank also has a spatial component since the schoolboys were sat in order of their academic rank in the classroom.

We were able to match the models in the paper, which used a wide array of network terms. We found models fitted using star and triangle parameter to be degenerate. We noted that the models in the paper did not include geometrically weight degree terms as is usual to account for social popularity processes.

We were not able to fit LOLOG models using the terms in the published ERGM fit. However substituting the geometrically weighted ESP term for triangle term allowed for the fitting of the LOLOG model. The published ERGM and LOLOG model with triangle term substituted both fit the observed network well.

The LOLOG model interpretation was broadly consistent with the ERGM interpretation with some small differences on various nodal covariate terms.

We also experimented with constraining the orderings by nodal covariates for this network. Introducing rank based ordering i.e. considering edges involving higher ranked nodes first (least academically able) increases the up-rank effect and produces a highly significant nodal rank effect. As we are considering high rank nodes first, in the generating process, if all else were equal they become “filled up”, i.e. highly connected before the lower rank nodes are added. However we observe in the data low rank nodes nominating high rank nodes as friends. To counteract the negative effect on tie formation between low and the “filled up” high rank nodes, the up-rank effect increases. This impresses upon us the need to interpret LOLOG fits conditional on the specified ordering process, in particular when the ordering process is based on nodal covariates.

A.6. Employee Voice

This data set contained 6 directed networks of between 24 and 39 nodes of employee voice, i.e. making a suggestion or voicing a problem Pauksztat et al. (2011). The data was collected from employees of three Dutch preschools, each with two waves of data. Since there was significant longitudinal incompleteness, authors of paper, treated each network separately, and carried out a meta analysis for each wave to test their hypotheses.

We were able to replicate published ERGM in only 1 case, however removing the out-2-star term allowed us to fit a further 4 cases, and removing the in-2-star term sufficed to allow a model for the final case. We note that the decay parameters were not specified in the paper, though we tried possible combinations without being able to match the published fit. The results were not qualitatively. It seems likely the effect due to the omission of the 2-star terms, was absorbed by other terms somewhat.

The authors also did not include an edge parameter in their tables of their fits. We included an edges parameter, as measure of the baseline propensity to form edges

We were able to fit the LOLOG model the published ERGM terms in 5 out of 6 networks, where the this were possible the fit to the observed network was good. For each of these 5 networks we were also able to fit the LOLOG model with triangle and star terms, which improved the goodness of fit also.

A.7. Office Layouts

As this was a complex example, we showed a detailed fit as our main example in Section 5.

A.8. Disaster Response

This network is a 20 node directed communication network formed between various agencies in the search and rescue operation in the aftermath of a tornado striking a boat on Pomona Lake in Kansas. Because the tornado destroyed much communication equipment, an important feature of this network was that the state's highway patrol was the only organisation having functioning communication equipment. The local sheriff took control of the operation, and the highway patrol was used for communication purposes, therefore there are two nodes that are very highly connected in the observed network. An ERGM was fit in Doreian and Conti (2012) and the data was obtained through ?? (Dis).

The authors goal in fitting the ERGM was to consider whether local or global processes lead to the formation of the network. The fit only with structural ERGM terms and then compared this to a fit using a block model parameter, it is not specified exactly how this is achieved. The authors comment that adding the block model parameter yields a superior fit. The authors did not include nodal covariates in their network.

We were not able to reproduce the ERGM fit stated in the paper. We were able to fit an ERGM only when omitting the out and mixed star parameters and including a geometrically weighted in star parameter. We were able to fit a LOLOG model using the terms in the published ERGM. With the omission of nodal covariates these models fit the observed network poorly.

On including nodal covariates we were able to find an ERGM that fit the data well, as well as a LOLOG model with the same terms that also fit the observed network well.

The authors did not provide a detailed interpretation of their fit mainly using the the ERGM with the block model covariate to argue that both global and local processes drove the formation of the network.

As the ERGM the LOLOG model with nodal covariates fits well, we argue that the network and in particular its formation can be explained using local processes. We also note that the LOLOG with structural terms fits similarly well to the LOLOG using nodal covariates. This may suggests that structural social processes are sufficient to explain the network formation. We note that the LOLOG significant parameter of the in 2 star, and lack of the significant triangles parameter, suggests the network is driven by a popularity process. This is consistent with the ERGM fit.

A.9. Company Boards

Here we consider the 808 node, undirected networks of interlocking boards in S&P 500 companies in the years 2007, 2008, 2009 and 2010. The nodes in the network are companies, with a tie being present if the company's board shares members. The network approach using ERGMs to understand the network, was presented in Wonga et al. (2015).

The authors supplied the data set without nodal covariates, therefore we were unable to replicate the reported ERGM fit.

We were able to fit LOLOG models for each of the 4 networks, both with geometrically weighted degrees and ESP parameters and triangle and star parameters. We expect a structural fit to fit the data well because the effect size of the nodal covariates in the data was small. However fitting the LOLOG with structural terms alone provided a much better fit than the ERGM using structural terms alone.

A.10. Swiss Decisions

The authors in Fischer and Sciarini (2015) investigated directed reputational trust networks of between 19 and 26 nodes among actors in 10 decision making processes in Switzerland in the 2000s. A node in this networks is an actor in the decision process, with a tie from node i to node j being i nominating j as being influential in the decision making process. The authors argue that aggregating reputational power, and then proceeding with the analysis, ignores the inherent relational nature of the data. They argue that to fully model the concept of reputational power explicitly accounting for the social structure with ERGM is important.

We were able to fit the ERGM with the published parameters in 9 out of 10 cases, but the parameter estimates were often inconsistent. Despite the signs and significance of our estimated parameters not always being consistent with the published models, our fitted ERGMS in general fitted the network data well. We were unable to fit the LOLOG model with the published ERGM terms in 8 out of 10, we suspect this is due to the correlation between the GWESP and GWDSP (geometrically weighted dyadwise shared partners) terms. As these were small networks with between 19 and 26 nodes with complex models fit to them we believe the LOLOG models with triangles and stars were potentially over fitting, achieving a good fit, yet providing large parameter estimate standard deviations. We suggest that inference based on such models should be treated with caution. In general in such small networks it seems that ERGM is often a preferable model.

A.11. University Emails

This is a undirected network of 1133 nodes within a university, with a connection defined based on a specified frequency of email contact. We suspect this is not a typical social network, as a connection based on an email is a very weak social interaction. We note that the authors did not fit an ERGM using an MLE approach, they selected parameters that yielded networks that fit on some subjective quantities, the statistical properties of their analysis are therefore unknown.

We were able to fit an ERGM with the standard MCMC MLE approach however, this fit the observed network data very poorly, so we do not discuss it further. We were able to fit a LOLOG model with triangle and star terms however we were not able to obtain a good fit to the observed network and the model had limited interpretability.

In general we do not regard this network as a good example for fitting a generative social network model based on simple local structures, as the social connection is very weak, which likely means most of the complex social structure is not reflected in the data.

A.12. Elementary School Friendships

These networks were directed networks of friendships in middle schools classed on between 22 and 24 nodes. The paper that fit this model was published before MCMC methods for fitting ERGM were widespread and available. The authors used pseudo likelihood to estimate the models.

The authors' approach was non standard in the context of modern methods. They first fit a single network with "expansiveness" and "attractiveness" parameters for each individual node, essentially a unique parameter governing the number of friends a node is likely to nominate as well as the number of times they are likely to be nominated by other nodes.

Another model was next fit, regarding the 3 classes as a single model with no edges between nodes in different classes. The authors then fit ERGM with pseudo likelihood with various constraints regarding the parameters for each of the classes.

As this was a non standard modelling approach we did not recreate the published ERGM fits directly. We were able to fit the ERGM model with MCMC MLE methods, with GWESP and GWDEG terms for the grade 4 and 5 models, but needed to omit the GWESP terms to be able to fit the grade 3 network. All models showed strong homophily on grade, with the GWESP term significant and positive and the GWDEG terms not significant for grades 4 and 5. The simpler grade 3 model had significant and negative terms for GWDEG terms suggesting that the network was not driven by super friendship nominators or nomination receivers. These models fitted the observed network data well.

We were not able to fit LOLOG models to these networks using the published ERGM terms, however using triangle and star terms we were able to achieve a better fit with the LOLOG model. However the LOLOG model parameters had large standard errors in line with our experiences with very small ~20 node networks, so for the grade 4 and 5 networks the ERGM model with modern terms was preferable. As we were unable to fit an ERGM to the grade 3 model with the GWESP term and the ERGM with GWDEG terms did not fit this network well, so we suggest the LOLOG model was more suitable for modelling the grade 3 network.

A.13. Online Links

These networks are directed and undirected networks of websites with hyperlinks and similar "framing" of issues respectively. The hyperlink network had 158 nodes whereas the framing network had 150 nodes.

We were able to recreate the published fits in both cases however found that the models did not fit the observed networks well. We found the recreated ERGM for the hyperlink network in particular fit very poorly. We were able to fit the LOLOG models with the ERGM terms but both models had similarly poor fit on the observed data.

We were able to fit LOLOG models using triangle and star terms to achieve a good fit to the observed data, and therefore recommend LOLOG as a better model for explaining these networks.

References

Addolescent Health Survey.

Pajek Datasets.

Ackland, R. and M. O'Neil (2011). Online collective identity: The case of the environmental movement. *Social Networks 33*(3), 177 – 190.

Amati, V., A. Lomi, and A. Mira (2018). Social network modeling. *Annual Review of Statistics and Its Application 5*(1), 343–369.

Anderson, C. J., S. Wasserman, and B. Crouch (1999). A p* primer: logit models for social networks. *Soc. Networks 21*, 37–66.

Chatterjee, S. and P. Diaconis (2013, 10). Estimating and understanding exponential random graph models. *Ann. Statist. 41*(5), 2428–2461.

Doreian, P. and N. Conti (2012). Social context, spatial structure and social network structure. *Social Networks 34*, 32–46.

Everett, M. and T. Valente (2020). *Social Networks: An International Journal of Structural Analysis.* Elsevier.

Fellows, I. (2018a). Latent order logistic (lolog) graph models. https://github.com/statnet/lolog.

Fellows, I. and M. Handcock (2017, 20–22 Apr). Removing Phase Transitions from Gibbs Measures. Volume 54 of *Proceedings of Machine Learning Research*, Fort Lauderdale, FL, USA, pp. 289–297. PMLR.

Fellows, I. E. (2018b). A new generative statistical model for graphs: The latent order logistic (lolog) model.

Fischer, M. and P. Sciarini (2015). Unpacking reputational power: Intended and unintended determinants of theassessment of actors' power. *Social Networks 42*, 60–71.

Frank, O. and D. Strauss (1986). Markov graphs. *Journal of the American Statistical Association 81*(395), 832–842.

Goldenberg, A., A. X. Zheng, S. E. Fienberg, and E. M. Airoldi (2010). A survey of statistical network models. *Foundations and Trends® in Machine Learning 2*(2), 129–233.

Goodreau, S. (2007, 06). Advances in exponential random graph (p*) models applied to large social networks. *Social networks 29*, 231–248.

Handcock, M. S. (2003). Assessing degeneracy in statistical models of social networks. Working paper.

Handcock, M. S., D. R. Hunter, C. T. Butts, S. M. Goodreau, P. N. Krivitsky, and M. Morris (2018). *ergm: Fit, Simulate and Diagnose Exponential-Family Models for Networks.* The Statnet Project (http://www.statnet.org). R package version 3.9.4.

Harris, K., C. Halpern, A. Smolen, and B. Haberstick (2007, 01). The national longitudinal study of adolescent health (add health) twin data. *Twin research and human genetics : the official journal of the International Society for Twin Studies 9*, 988–97.

Heidler, R., M. Gampner, A. Herz, and F. Esser (2014). Relationship patterns in the 19th century: The friendship network in a german boys' school class from 1880 to 1881 revisited. *Social Networks 37*, 1–13.

Hunter, D. R., S. M. Goodreau, and M. S. Handcock (2008). Goodness of fit of social network models. *Journal of the American Statistical Association 103*(481), 248–258.

Hunter, D. R. and M. S. Handcock (2006). Inference in curved exponential family models for networks. *Journal of Computational and Graphical Statistics 15*(3), 565–583.

Lubbers, M. J. and T. A. Snijders (2007). A comparison of various approaches to the exponential random graph model: A reanalysis of 102 student networks in school classes. *Social Networks 29*(4), 489 – 507.

Pauksztat, B., C. Steglich, and R. Wittek (2011). Who speaks up to whom? a relational approach to employee voice. *Social Networks 33*(4), 303–316.

Robins, G., T. Snijders, P. Wang, M. Handcock, and P. Pattison (2007). Recent developments in exponential random graph (p) models for social networks. *Social Networks 29*(2), 192–215.

Sailer, K. and I. McCulloch (2012). Social networks and spatial configuration—how office layouts drive social interaction. *Social Networks 34*, 47–58.

Schweinberger, M. (2011). Instability, sensitivity, and degeneracy of discrete exponential families. *Journal of the American Statistical Association 106*(496), 1361–1370. PMID: 22844170.

Schweinberger, M. and M. S. Handcock (2015). Local dependence in random graph models: characterization, properties and statistical inference. *Journal of the Royal Statistical Society: Series B (Statistical Methodology) 77*(3), 647–676.

Schweinberger, M., P. N. Krivitsky, C. Butts, and J. Stewart (2020). Exponential-family models of random graphs: Inference in finite-, super-, and infinite population scenarios. *Statistical Science*.

Schweinberger, M. and J. Stewart (2020, 02). Concentration and consistency results for canonical and curved exponential-family models of random graphs. *Ann. Statist. 48*(1), 374–396.

Toivonen, R., L. Kovanen, M. Kivelä, J.-P. Onnela, J. Saramäki, and K. Kaski (2009). A comparative study of social network models: Network evolution models and nodal attribute models. *Social Networks 31*(4), 240 – 254.

Wonga, L. H. H., A. Gygax, and P. Wang (2015). Board interlocking network and the design of executive compensation packages. *Social Networks 41*, 85–100.