Exponential random graph models for multilevel networks

Peng Wang [a],[*], Garry Robins [a], Philippa Pattison [a], Emmanuel Lazega [b]

[a] *Melbourne School of Psychological Sciences, The University of Melbourne, Australia*
[b] *IRISSO-ORIO, University of Paris-Dauphine, France*

ARTICLE INFO

ABSTRACT

Modern multilevel analysis, whereby outcomes of individuals within groups take into account group membership, has been accompanied by impressive theoretical development (e.g. Kozlowski and Klein, 2000) and sophisticated methodology (e.g. Snijders and Bosker, 2012). But typically the approach assumes that links between groups are non-existent, and interdependence among the individuals derives solely from common group membership. It is not plausible that such groups have no internal structure nor they have no links between each other. Networks provide a more complex representation of interdependence. Drawing on a small but crucial body of existing work, we present a general formulation of a multilevel network structure. We extend exponential random graph models (ERGMs) to multilevel networks, and investigate the properties of the proposed models using simulations which show that even very simple meso effects can create structure at one or both levels. We use an empirical example of a collaboration network about French cancer research elites and their affiliations (Lazega et al., 2006, 2008) to demonstrate that a full understanding of the network structure requires the cross-level parameters. We see these as the first steps in a full elaboration for general multilevel network analysis using ERGMs.

© 2013 Elsevier B.V. All rights reserved.

1. Introduction

Statistical analysis using multilevel models, or hierarchical linear models (HLMs) (Snijders and Bosker, 2012), is applicable where the data has a nested hierarchical structure. The units of study are categorized in two or more levels. Snijders et al. (2012) listed several multilevel examples such as schools and teachers, classes and pupils, and doctors and patients. The lower levels (i.e. teachers, pupils and patients) are commonly referred to as the *micro-* or *individual-levels*, and the upper levels (i.e. schools, classes and doctors) are known as *macro-* or *group-levels*. The interactions across levels, known as the *meso-level* structure, usually describe the affiliations between the micro- and macro-level units, typically in a nested structure where the micro-level units are affiliated with one and only one macro-level unit. The analytical focus is usually on the outcomes of micro-level units, taking into account this nested structure and partitioning the variance of the outcome variable across levels.

To date, multilevel analysis has not been commonly used in social network analysis, although attention has been drawn to the theoretical importance of multilevel concepts. In the context of organizational network theory, for example, major reviews (Kozlowski and Klein, 2000; Borgatti and Foster, 2003; Brass et al.,

2004) have highlighted the relevance of multilevel aspects of organizations. Brass et al. (2004) reviewed organizational network research at different organizational levels, but the amalgamation of multilevel and network perspectives has barely begun.

The classic HLM analysis described above assumes that links between groups at the macro-level are non-existent, and that interdependence among the individuals derives solely from shared group membership instead of other forms of endogenous network processes. Network analysis in general presents a more complex representation of interdependence. It is not plausible that groupings of individuals have no internal structure beyond the simple membership structure, nor is it always the case that such groups have no links among themselves. Yet, it is also not always plausible that network structure exists at only one level. For instance, organizations are explicitly multilevel in design and operation, yet this is not taken into account in organizational network analysis.

Network methods focus attention on relationships among individuals. For a one-mode network (nodes of one type – say, individuals), there is no nesting structure. Network methods may be used to predict individual outcomes given the network structure: for instance, when individual responses are assumed to relate to responses by network partners termed as *social influence*, *network diffusion* or *network contagion* (e.g. Friedkin, 2006; Mason et al., 2007). However, the structure itself may be the research issue. Network structure may be self-organizing into various endogenous patterns. Individual qualities may also influence the formation or dissolution of network ties, known as *social selection* (e.g. Robins et al., 2001a). Social influence and selection processes may be

* Corresponding author at: Melbourne School of Psychological Sciences, The University of Melbourne, Australia. Tel.: +61 3 83445725.
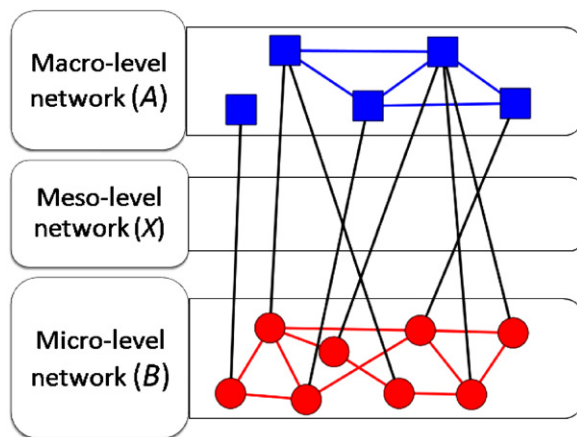*E-mail address:* pwang@unimelb.edu.au (P. Wang).

Fig. 1. A multilevel network.

studied simultaneously using the stochastic actor oriented models (SAOMs) proposed by Snijders et al. (2010) with longitudinal data. Recent methods to analyze cross-sectional network data include exponential random graph models (ERGMs) (Snijders et al., 2006; Robins et al., 2007a,b; Lubbers and Snijders, 2007).

In this paper, we present a general formulation of a multilevel network structure, and extend ERGMs to multilevel networks. We start with descriptions of the general multilevel data structure in a network context that generalizes all previous data types hitherto seen in multilevel network studies, and then compare ERGMs with the classical HLMs. Before introducing the multilevel ERGMs, we review and investigate some of the current ERGM specifications using simulations. As part of the proposed multilevel ERGMs, these simulations are crucial for the later model interpretations we give to within-level network structures. We then propose ERGM specifications for multilevel networks. Through simulation, we illustrate how simple cross-level associations can create structure at both levels. The proposed models are then applied to an empirical data set collected among French cancer research elites and their affiliations (Lazega et al., 2006, 2008). By estimating and comparing models with and without multilevel effects, we show the importance of the multilevel effects in both goodness of fit and model selection. The model interpretations reveal the dependencies among the micro-, macro- and meso-level network, and provide richer and more detailed descriptions of the empirical data. The proposed models are implemented in the program MPNet as a multilevel network extension for PNet (Wang et al., 2006).

2. Multilevel networks

Multilevel network data categorizes nodes into different levels, and the network ties represent relationships between nodes within and across different levels. A k-level network has nodes of k different types, and each type represents a different level. A one-mode network can be defined within each level, and a bipartite (two-mode) network can be defined between nodes from two adjacent levels.

For a two-level network with u nodes at the macro-level, and v nodes at the micro-level, we label the macro-level network as network A, the micro-level network as network B, and the meso-level bipartite network as network X. We refer the overall network as a (u, v) two-level network, labelled as M. This flexible data structure, as shown in Fig. 1, generalizes the multilevel networks in the literature to date:

1. In a cross-level nested structure, all B nodes have degree one in X. With both macro- and micro-level networks (A and B)

empty, we have a multilevel modelling structure. With networks among the at the micro-level B, post hoc multilevel inferences can be made after applying ERGMs to each group (Anderson et al., 1999; Snijders and Baerveldt, 2003; Lubbers, 2003; Lubbers and Snijders, 2007).

2. A fixed nested structure X, with A and B networks non-empty, is presented by Lazega et al. (2008). See also Hedström et al. (2000) and Bellotti (2011). One level is nested in another, but with networks at one or both levels. We can investigate how structures in A and B are related.

3. An empty macro-level A network, with non-empty B and X, is the data structure of Koehly et al. (2003), Torlo et al. (2010), Snijders (2002, 2009), and Lomi et al. (2011).

4. The most general form presents a range of possibilities depending on whether or not some of the three networks are assumed exogenous. For instance, with an exogenous meso-level X, we can ask: given the meso-links, how do micro and macro structures relate? Or when all three networks are endogenous, and we can examine the interdependencies among the micro-, macro- and meso-level networks. The most general form was first discussed by Iacobucci and Wasserman (1990) and Wasserman and Iacobucci (1991), but not in the context of multiple levels.

Lazega et al. (2008) presented an example with two-level network data collected among French elite cancer researchers and their affiliated research laboratories. There are individual and collective forms of agency at each level. At the researcher level, the data captures advice-seeking ties whereas at the laboratory level, the network represents collaborations among the laboratories, and the cross-level ties represent researchers' affiliations with laboratories. Based on various centrality measures, Lazega et al. (2008) categorized researchers into "little or big fish", and laboratories into "small or big pond", and inferred theory on how network strategies may affect an individual's performance; the cross-level structure is treated as exogenous and understood, whereas the within-level structure is endogenous and assumed affected by the cross-level structure. We use the data from Lazega et al. (2008) to demonstrate how the ERGMs proposed in this paper may be applied and interpreted in an empirical context in the modelling example.

3. Multilevel models

Multilevel modelling using *hierachical linear models* (HLMs) (Snijders and Bosker, 2012) has forerunners including "Within-and-between" organizational analysis (Dansereau et al., 1984). HLMs can be used for nested multilevel data analysis, when there is an underlying hierarchy that determines the categorization of the nodes such that nodes from lower level are nested or belong to nodes in the higher levels. HLMs explain the individual outcome as a result of explanatory variables at indivdal level, taking into account the nested multilevel structure. Variance is apportioned across both levels because persons are nested within groups. Let j denote the index for groups and i denote the index for individuals, the simplest form of HLMs (known as the *empty model* without any explanatory variables) can be expressed as

$$Y_{ij} = \beta_0 + u_{0j} + e_{ij}$$

where Y_{ij} is the dependent variable representing the outcome of node i from group j; β_0 is a general mean; u_{0j} is a random effect at group level; and e_{ij} is a random effect at the individual level. Both u_{0j} and e_{ij} are assumed to be independent of one another and normally distributed with means 0s and variances σ_u^2 and σ_e^2 respectively. The variance of Y_{ij} is therefore decomposed into the variances at group level and individual level. Taking into account the explanatory variable (x_{ij}) at the individual level and assuming that

the *slope* associated with the explanatory variable is group depend-ent, the HLMs have the following general form:

$$Y_{ij} = \beta_0 + u_{0j} + \beta_1 x_{ij} + u_{1j} x_{ij} + e_{ij}$$

where $\beta_{0j} = \beta_0 + u_{0j}$ defines the *intercept* and $\beta_{1j} = \beta_1 + u_{1j}$ defines the *slope*. Such a model construction assumes both the intercept and slope are group dependent with group-level residuals u_{0j} and u_{1j}. One may include several individual level explanatory vari-ables in HLMs; however, direct modelling of dependencies among micro-level network ties as endogenous effects is not possible. This approach is not aiming or designed to model network ties directly, and there are limitations when using HLMs for modelling mul-tilevel networks in general. Firstly, the HLM approach treats the nested meso-level structure as exogenous and understood, so any underlying meso-level network processes are not captured. Sec-ondly, the micro- and macro-within level network structure cannot be captured directly by HLMs, thus the interdependencies between network ties cannot be tested.

4. Exponential random graph models

Introduced by Frank and Strauss (1986) and Wasserman and Pattison (1996), exponential random graph models (ERGMs) treat the network structure as endogenous and a topic of research inter-est. Treating each network tie as a random variable, ERGMs model network ties explicitly, and present the overall network structure as a collective result of various local network processes. The local network processes are represented by graph configurations such as edges, stars, and triangles (see, e.g. Frank and Strauss, 1986; Snijders et al., 2006), where all presented ties in a particular con-figuration are assumed to be conditionally dependent reflecting hypotheses that empirical network ties do not form at random, but that they *self organize* into various patterns arising from underly-ing social processes. ERGMs have been applied to different data structures such as bipartite networks (Skvoretz and Faust, 1999; Agneessens and Roose, 2008; Wang et al., 2009, 2013) and multi-ple networks (Pattison and Wasserman, 1999; Robins and Pattison, 2006; Wang, 2013). In the generic cases, where one uni-partite (or one-mode) network (Y) with n nodes is involved, the ERGMs have the following form:

$$Pr(Y = y) = \frac{1}{\kappa(\theta)} \exp \sum_Q \theta_Q z_Q(y)$$

where

- y is a network instance.
- Q defines the network configurations which are based on the dependence assumptions of tie variables. Note that a network variable Y can be seen as a collection of tie variables (Y_{ij}) defined on each dyad (i, j) of the network. A network configuration of type Q includes tie variables that are conditionally dependent given the rest of the network.
- $z_Q(y) = \sum_y \prod_{Y_{ij} \in Q} y_{ij}$ is the network statistic for the correspond-ing network configuration of type Q.
- θ_Q is the parameter associated with $z_Q(y)$.
- $\kappa(\theta)$ is a normalizing constant defined based on the graph space of networks of size n and the actual model specification.

Since the normalizing constants of ERGMs become intractable for networks with even a small number of nodes, maximum like-lihood estimates (MLEs) cannot be derived analytically except for simple Bernoulli models. Snijders (2002) proposed an estimation algorithm which relies on Markov Chain Monte Carlo (MCMC) sim-ulations of ERGMs. The MCMC simulation also serves as a tool for model goodness of fit (GOF) test, where simulated graph distribu-tions are compared with the observed networks (Snijders, 2002; Hunter et al., 2008).

As in Frank and Strauss (1986), the model assumes network homogeneity such that isomorphic network configurations have equal parameters, and the corresponding effect is the same across the network. For networks where such assumptions do not apply, a homogeneous model may have difficulty in convergence. We may treat part of such networks as exogenous, or fit conditional models depending on exogenous covariates.

ERGM specifications for both one-mode and bipartite networks can be derived from dependence assumptions between network tie variables. There has been a rich literature on how the network ties may be conditionally dependent and their inferred ERGM specifi-cations including the dyadic independence assumption (Erdös and Rényi, 1960; Holland and Leinhardt, 1981) where all the tie vari-ables are considered independent from one another; the Markov assumption (Frank and Strauss, 1986) where a pair of tie vari-ables are considered conditionally independent unless they have a node in common; and the social circuit assumptions (Pattison and Robins, 2002, 2004; Snijders et al., 2006) where tie variables within a social circuit (four-cycle) are considered conditionally dependent. Pattison et al. (2009) and Pattison and Snijders (2013) proposed a hierarchy of dependence assumptions based on a graph theoretical framework which provides guidance for systematic development of ERGM specifications. The social circuit specifications proposed by Snijders et al. (2006) introduced alternating statistics where geometric weights with parameters (λ) are assigned to the degree distribution and the two-path (or shared partner) distribution, such that large changes of graph statistics in simulations are avoided to alleviate model degeneracy (Handcock, 2003; Rinaldo et al., 2009). The graph statistics introduced by Snijders et al. (2006) include alternating-k-stars (AS), alternating-k-triangles (AT) and alternating-k-two-paths (A2P),[1] representing the dispersion of the degree distribution, the tendency for closure, and the tendency for sharing multiple partners respectively. For bipartite networks involving nodes of two distinct sets, Wang et al. (2009) applied the same geometric weighting technique on the degree distribu-tions with alternating-stars and the two-path distributions with alternating-4-cycles (ACs) of two different types.

The current available ERGMs may model multilevel networks by parts. Without considering the cross-level network structure, ERGMs for one-mode networks can be used to analyze individual within-level networks one at a time, assuming ties from different levels are independent from each other. Such analysis are likely to be inadequate for the obvious reason that the cross level struc-tures are ignored. Explicit multilevel network structure is invoked through bipartite networks, with nodes of two different types (e.g. people and groups) and relations (associations) from persons to groups. For bipartite analysis, ERGMs, SAOMs, and other meth-ods are available (e.g. Latapy et al., 2008; Wang et al., 2009, 2013; Koskinen and Elding, 2010). In fact, all bipartite networks can be seen as special cases of two-level networks where within level ties are absent. This approach treats the cross-level structure as endogenous but there are no within-level networks. ERGMs for multivariate network analysis (Pattison and Wasserman, 1999) provided an approach that models the interdependence of several networks of different types. However, the different types of ties are defined on a common type of nodes with no level distinction. The dependencies between ties of different types from different lev-els require modelling of the within-level one-mode networks and

---

[1] Also known by Hunter and Handcock (2006) as geometrically weighted degrees, geometrically weighted edge-wise shared partners, and geometrically weighted dyadic-wise shared partners.

the cross-level bipartite networks together. We propose ERGMs for multilevel networks in the simplest form which involves nodes from two levels. It may be further extended to the general k-level ERGMs following similar data and model constructions.

5. ERGMs for two-level networks

Based on the description and labels of the two-level network structure also shown in Fig. 1, ERGMs can then be expressed in the following form:

$$Pr(A = a, \quad X = x, \quad B = b) = \frac{1}{\kappa(\theta)} \exp \sum_Q \{\theta_Q z_Q(a) + \theta_Q z_Q(x) + \theta_Q z_Q(b) + \theta_Q z_Q(a, x) + \theta_Q z_Q(b, x) + \theta_Q z_Q(a, x, b)\}$$

There are several components based on the different networks involved:

- $z_Q(a)$ and $z_Q(b)$ are network statistics for the corresponding within level network configurations. We can apply the current ERGM specifications for one-mode networks (Snijders et al., 2006; Robins et al., 2007b, 2009) to look at within level network structures, given other effects in the model.
- $z_Q(x)$ are the network statistics for structural effects within the bipartite affiliation network. The corresponding ERGM specifications were proposed by Wang et al. (2009, 2013).
- $z_Q(a, x)$ and $z_Q(b, x)$ are network statistics for configurations involving ties from one of the unipartite network (A or B) and the bipartite network (X), representing the interactions between the two networks. Using $z_Q(a, x)$ as an example, it can be expressed as $z_Q(a, x) = \sum_{a,x} \prod_{A_{ij} \in Q, X_{kl} \in Q} a_{ij} x_{kl}$.
- $z_Q(a, x, b) = \sum_{a,x,b} \prod_{A_{ij} \in Q, X_{kl} \in Q, B_{uv} \in Q} a_{ij} x_{kl} b_{uv}$ are statistics for configurations involving ties from all three networks, and express tendencies for structural effects to be associated across both levels simultaneously.

For effects that associate one type of within-level tie (A or B) with a cross level tie (X), i.e. $z_Q(a, x)$ or $z_Q(b, x)$, and the cross level effects $z_Q(a, x, b)$, we propose model specifications based on the various dependence assumptions in the literature (e.g. Erdös and Rényi, 1960; Frank and Strauss, 1986; Pattison and Robins, 2002, 2004). Although the dependence assumptions applied here are the same as in ERGMs for one- or two-mode networks, the interpretations are very different as we assume dependencies between tie-variables of different nature or types. We use the terminology "node of type A" or "A-node" to refer to the nodes that can be involved in A-ties, and similarly for "node of type B". For simplicity, we focus on the descriptions and possible interpretations of models for non-directed multilevel networks in the following simulation studies. These interpretations may be equally applied to the extensions to directed multilevel models. See Figs. 2 and 3 for a list of proposed model configurations with possible interpretations for both non-directed and directed networks.

Following the Markov assumption, several parameters representing interactions between one of the within level network and the bipartite network may be derived including stars and triangles involving different types of ties. Note that because the Markov assumption implies cliques of sizes up to three (or triangles), statistics involving all three types of tie-variables, for example (A_{ij}, B_{kl} and X_{ik}), cannot be defined. Example Markov configurations including the interaction stars (such as Star2AX and Star2BX) and the interaction triangles (TXAX, TXBX) are depicted in Fig. 2. Based on the social circuit assumption, interaction alternating-triangles (e.g. ATXAX) and cross-level four-cycles (C4AXB) may be included in the model. A further extension in dependence, the three-path assumption, allows interaction three-paths (e.g. L3XAX) and cross-level three-paths (L3AXB) to be included in the model. To understand properties of the various proposed model specifications, simulation studies for the proposed configurations were carried out in the next section, and we illustrate possible interpretations of the proposed effects in a two-level researcher–laboratory network context similar to the data we used in our empirical example.

6. Multilevel ERGM simulation studies

The simulation studies were performed on (30, 30) two-level networks with 65 ties in each of the within-level (A, B) and meso-level (X) networks. For proposed effects only involving interactions between one of the within-level network and the meso-level network, network B was kept empty for simplicity, as the interaction effects between A and X are equally applicable to interactions between B and X. The densities of the networks are fixed, and we used strongly positive (+2) and negative (−2) parameter values for each of proposed configurations and left other effects at 0s to test the non-zero main effects on the network structure. We present visualizations of sampled networks, and graph statistics such as the means and standard deviations of the global clustering coefficients (GCCs), the standard deviations (SDs) and the skewness (SKs) of the degree distributions in each of the networks A, B and X. The GCCs for the within-level networks (A, B) are calculated as the ratio between closed and open triads; whereas the GCC for the meso-level network X is calculated as the ratio between the numbers of four-cycles and three-paths (Robins and Alexander, 2004).

Simulated network samples are shown from Figs. 4–8, where each row has the tested main effect listed in the first column followed by the decomposition of a simulated multilevel network M in the order of networks (A, X, B) when they apply. For the simulated graph statistics as listed from Tables 1–4, the main effects

Table 1
Simulated graph statistics for models with Star2AX and AAAXS. (*Note*: SD and SK stand for standard deviation and skewness of degree distributions. GCC stands for global clustering coefficient.).

| Statistics	Random	Star2AX+	Star2AX−	AAAXS+	AAAXS−
A					
SD	1.873(0.240)	6.863(0.008)	3.502(0.184)	2.822(0.258)	1.703(0.224)
SK	0.284(0.425)	3.422(0.010)	−0.084(0.168)	−0.108(0.253)	0.304(0.414)
GCC	0.144(0.035)	0.147(0.000)	0.334(0.043)	0.229(0.044)	0.136(0.036)
X					
SD(*A*)	1.406(0.186)	7.610(0.000)	3.345(0.194)	1.769(0.190)	1.329(0.169)
SK(*A*)	0.522(0.429)	3.607(0.000)	1.208(0.193)	0.512(0.364)	0.462(0.410)
SD(*B*)	1.394(0.178)	0.346(0.000)	1.283(0.160)	1.394(0.178)	1.407(0.180)
SK(*B*)	0.526(0.430)	2.273(0.000)	0.371(0.403)	0.524(0.409)	0.524(0.415)
GCC	0.066(0.031)	0.896(0.000)	0.207(0.037)	0.083(0.032)	0.064(0.030)
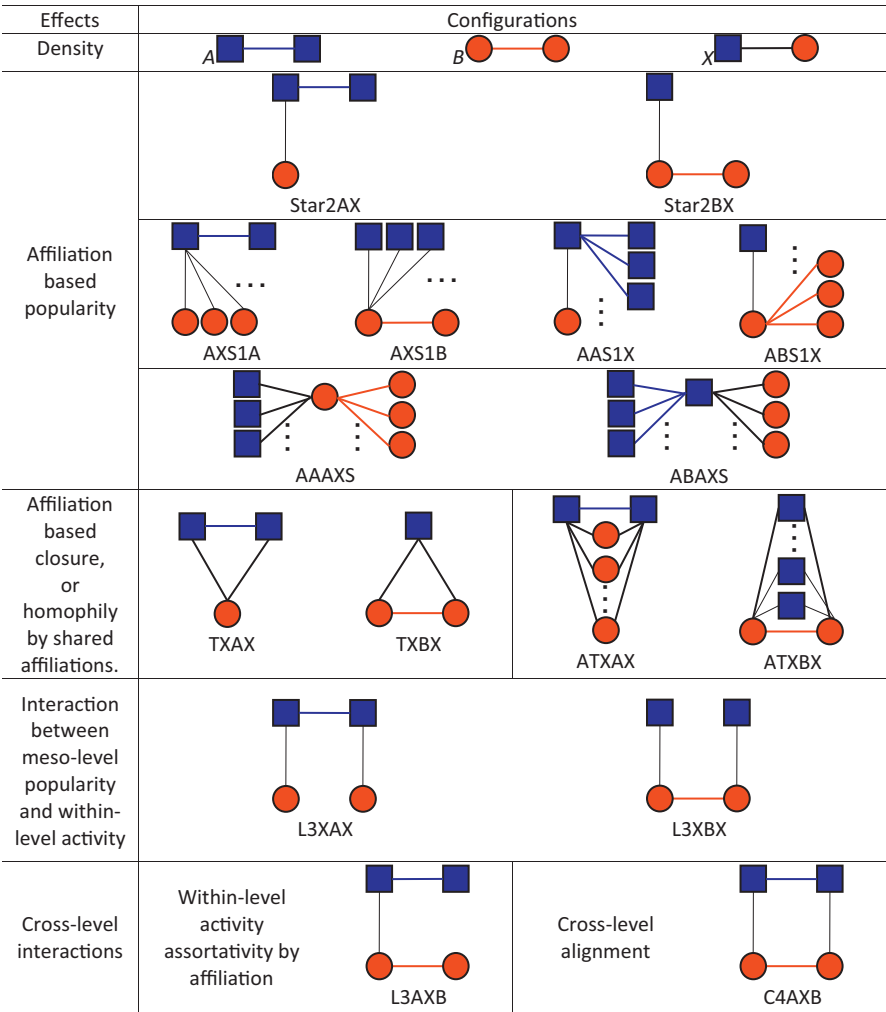
Fig. 2. Multilevel ERGM configurations for non-directed networks.

used in the simulation are listed in the header, and we compared the statistics from the simulated distributions with a random network distribution (listed under the header "random") to understand the impact on the global network structure. For each of the simulations, we picked every 10,000th graph from a 10,000,000 iteration simulation after a 1,000,000 iteration burn-in from random starting networks of the same density.

6.1. Interaction stars

The Star2AX and the Star2BX configurations as shown in Fig. 2 represent the dependence between two tie variables of different

Table 3
Simulated graph statistics for models with three-paths (L3XAX). (*Note*: SD and SK stand for standard deviation and skewness of degree distributions. GCC stands for global clustering coefficient.).

Statistics	Random	L3XAX+	L3XAX−
A			
SD	1.873(0.240)	2.554(0.310)	1.897(0.230)
SK	0.284(0.425)	0.144(0.328)	0.289(0.401)
GCC	0.144(0.035)	0.207(0.047)	0.145(0.036)
X			
SD(*A*)	1.406(0.186)	1.240(0.164)	1.423(0.188)
SK(*A*)	0.522(0.429)	0.450(0.444)	0.546(0.416)
SD(*B*)	1.394(0.178)	4.196(0.322)	1.248(0.155)
SK(*B*)	0.526(0.430)	4.313(0.514)	0.369(0.406)
GCC	0.066(0.031)	0.149(0.071)	0.069(0.033)

Table 2
Simulated graph statistics for models with interaction triangles. (*Note*: SD and SK stand for standard deviation and skewness of degree distributions. GCC stands for global clustering coefficient.).

| Statistics	Random	TXAX+	TXAX−	ATXAX+	ATXAX−
A					
SD	1.873(0.240)	5.379(0.049)	1.892(0.242)	3.628(0.257)	2.466(0.267)
SK	0.284(0.425)	0.437(0.003)	0.297(0.415)	0.730(0.110)	0.557(0.347)
GCC	0.144(0.035)	0.977(0.014)	0.144(0.034)	0.623(0.052)	0.154(0.036)
X					
SD(*A*)	1.406(0.186)	2.535(0.080)	1.420(0.190)	3.377(0.102)	2.337(0.172)
SK(*A*)	0.522(0.429)	0.428(0.036)	0.539(0.408)	0.987(0.070)	0.660(0.256)
SD(*B*)	1.394(0.178)	4.541(0.052)	1.258(0.158)	1.282(0.150)	1.359(0.181)
SK(*B*)	0.526(0.430)	1.856(0.019)	0.364(0.383)	0.359(0.380)	0.481(0.427)
GCC	0.066(0.031)	0.959(0.019)	0.068(0.033)	0.214(0.032)	0.119(0.032)
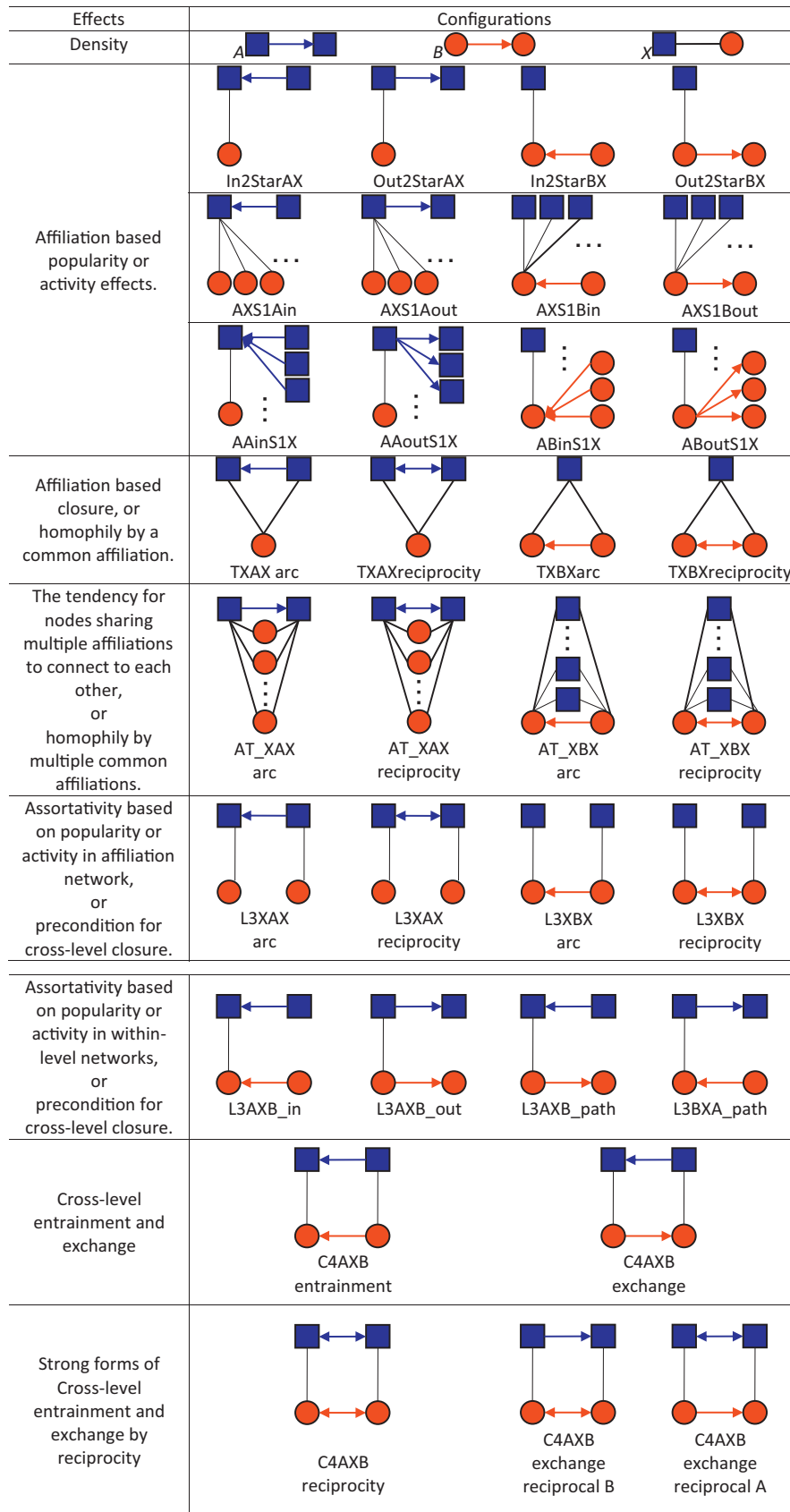
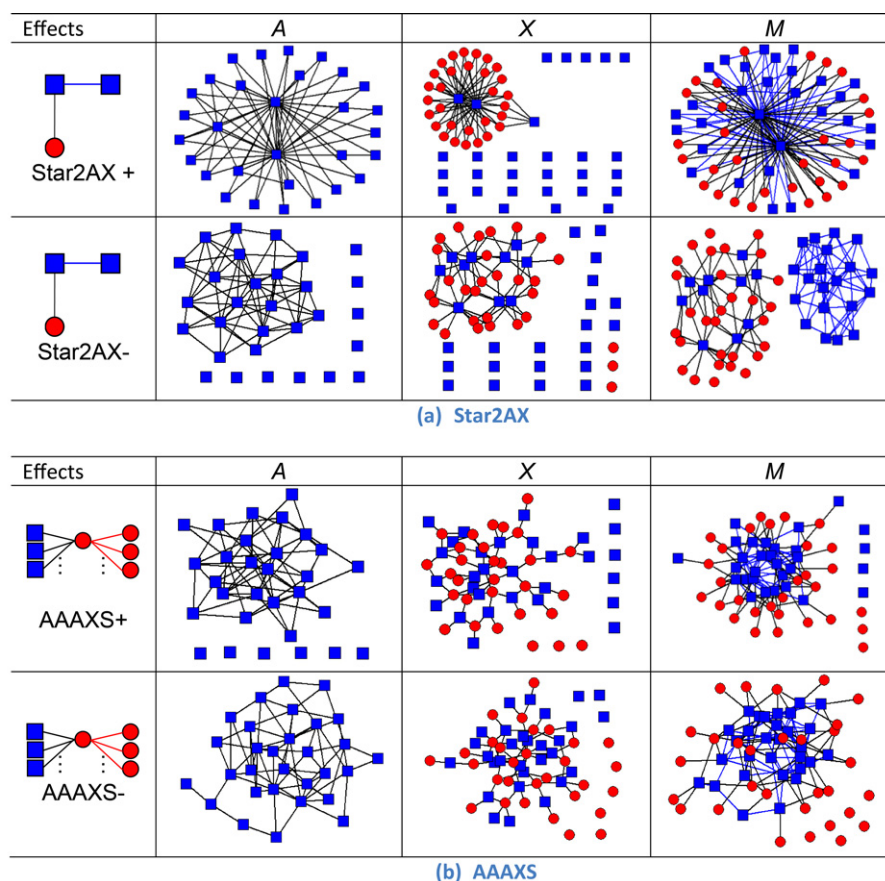Fig. 3. Multilevel ERGM configurations for directed networks.

Fig. 4. Simulated network sample with Star2AX and AAAXS.

types established by a node in common. Fig. 4(a) illustrates typical networks from the simulated network distributions where the Star2AX parameter was set at 2 and −2 respectively, and network B was kept empty, so as to observe only the effects on network A. The positive Star2AX parameter is an A-hub creator, as we can see from the plot of the networks (A, X and M), two A-hubs were generated both connected to all other A-nodes in network A and all the B-nodes in the affiliation network X. The model created an almost frozen graph distribution in both network A and the affiliation network X indicated by the standard deviations close to 0 in the various graph statistics shown in Table 1. The standard deviation (SD) and skewness (SK) of A-node degree distributions in both networks A and X are much higher than the random graph distribution; the B-node degree distribution in network X is flat,

as all B-nodes are connected to the three A-hubs, hence having degrees 2 or 3.

When Star2AX is negative, the model tries to avoid any association between the within- and the meso-level ties, and created two components in the overall multilevel network. This effect created lots of isolated A-nodes in both networks A and X, as non-isolated nodes in network A became isolates in network X and vice versa to form the separated overall structure. Since more isolated nodes are involved, the available network ties had to constrain themselves to the connected component, so we have higher SDs and SKs for A-nodes in both networks A and X, and higher clustering coefficients than expected from random (see Table 1). We expect similar behaviour for the Star2BX effect, and in general, we may conclude that the within- and meso-level interaction two-star

Table 4
Simulated graph statistics for models with cross-level interaction effects (L3AXB and C4AXB). (*Note*: SD and SK stand for standard deviation and skewness of degree distributions. GCC stands for global clustering coefficient.).

| Statistics	Random	L3AXB+	L3AXB−	C4AXB+	C4AXB−
A					
SD	1.873(0.240)	5.241(0.005)	2.862(0.264)	3.706(0.195)	2.098(0.252)
SK	0.284(0.425)	0.497(0.004)	−0.217(0.243)	0.727(0.102)	0.369(0.426)
GCC	0.144(0.035)	0.857(0.003)	0.234(0.045)	0.643(0.025)	0.156(0.037)
B					
SD	1.867(0.234)	6.014(0.027)	2.854(0.272)	4.238(0.159)	2.097(0.256)
SK	0.288(0.422)	2.674(0.050)	−0.212(0.251)	0.648(0.055)	0.367(0.408)
GCC	0.145(0.035)	0.221(0.007)	0.234(0.046)	0.771(0.027)	0.156(0.036)
X					
SD(A)	1.406(0.186)	2.649(0.004)	2.267(0.341)	3.331(0.008)	1.615(0.216)
SK(A)	0.522(0.429)	0.610(0.007)	1.486(0.407)	0.954(0.016)	0.696(0.408)
SD(B)	1.394(0.178)	4.470(0.006)	2.252(0.345)	3.089(0.009)	1.615(0.217)
SK(B)	0.526(0.430)	1.797(0.003)	1.494(0.384)	0.789(0.019)	0.703(0.416)
GCC	0.066(0.031)	0.867(0.002)	0.133(0.036)	0.650(0.006)	0.082(0.032)
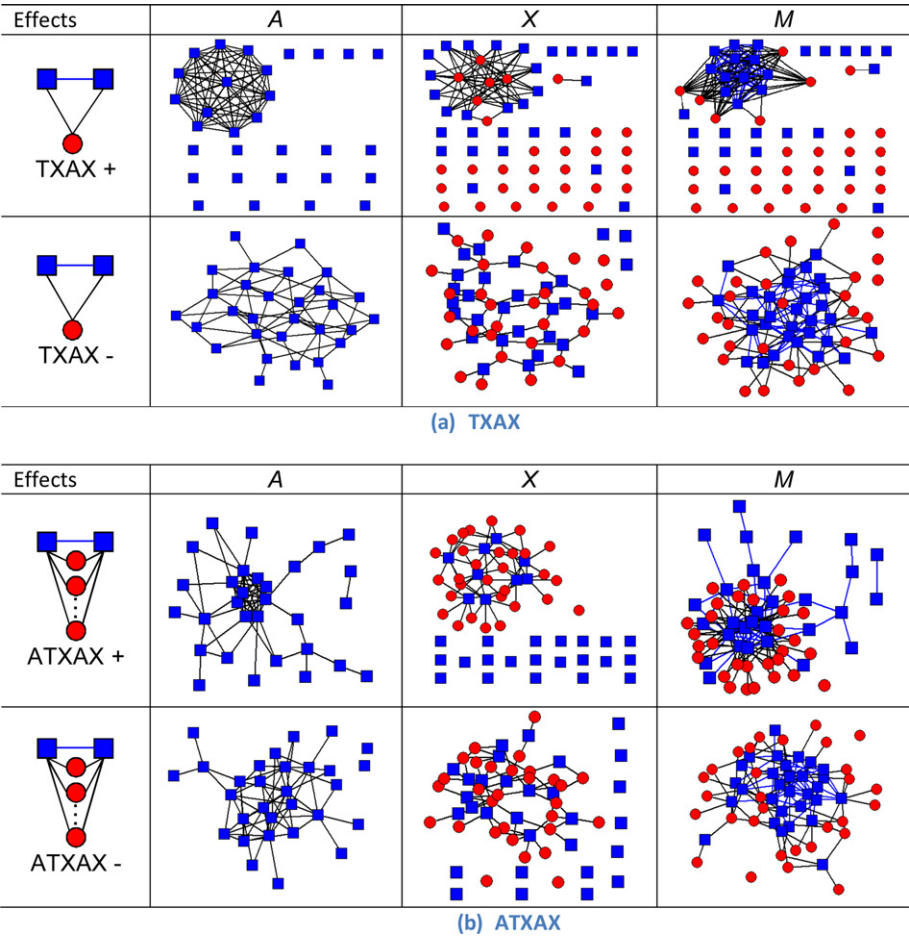
Fig. 5. Simulated network samples for models with interaction triangles.

parameters represent popularity effects in both within level and affiliation network. In a researcher and laboratory network context, where the B-level network contains advice seeking ties among the researchers, the A-level network is defined based on the collaboration ties among the laboratories, and the researcher–laboratory affiliation network X, a positive Star2BX parameter indicates that researchers popular in the advice seeking network are also popular in the affiliation network (i.e. have many affiliations).

To alleviate the issues of model degeneracy or the frozen graph distributions, we can apply the same geometric weighting techniques on either or both the affiliation star parameters and the within-level star parameters to form higher-order configurations, such as alternating-X-stars with one A-tie attached (AXS1A), or alternating-A-alternating-X-stars (AAAXS). Fig. 2 illustrates some possible alternating-star effects. The alternating-star effects generally attenuate the overwhelmingly frozen structures created by the two-star effects seen in the previous simulations, but provide similar interpretations. We followed the same simulation strategy with the same parameter values (+2 and −2) for models with AAAXS. As can been seen from Table 1, when AAAXS is positive, the graph distribution has higher SD for the A-node degree distribution than a simple random graph, but much smaller than for the previous model with positive Star2AX. The SD(A) in network X, and SKs for A-nodes for both networks A and X are not very different from
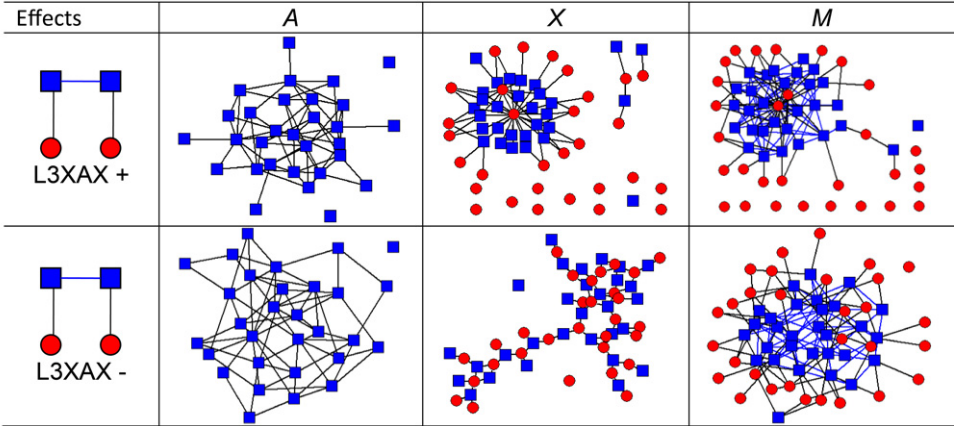

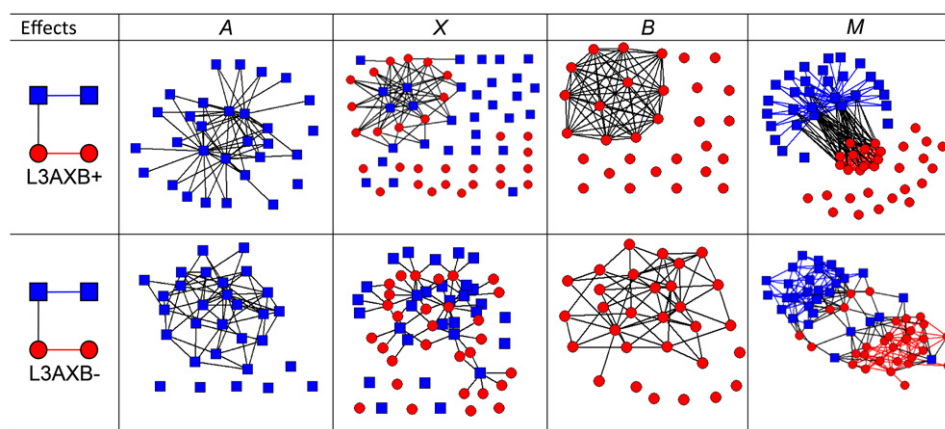**Fig. 6.** Simulated network samples for models with three-paths (L3XAX).

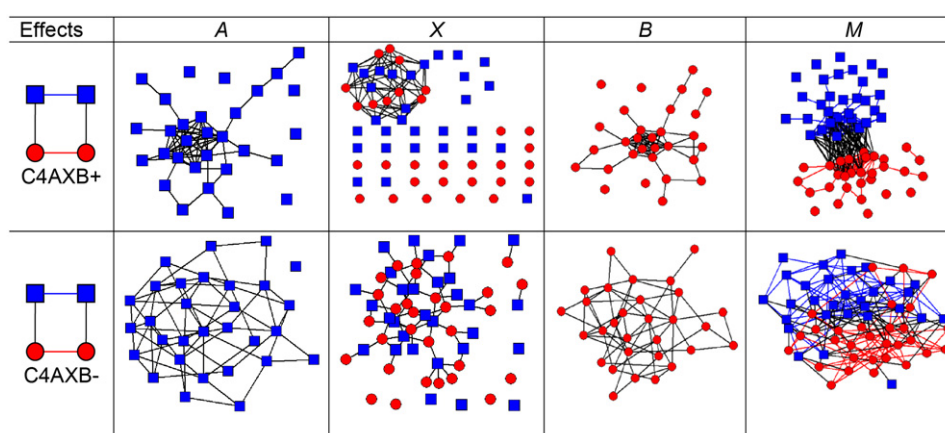**Fig. 7.** Simulated network samples for models with three-paths (L3AXB).

Fig. 8. Simulated network samples for models with cross-level four-cycles (C4AXB).

random. From the visualization in Fig. 4(b), we do not see obvious hubs in the network *A* but more nodes with degree greater than 2 or 3 compared with positive Star2AX model. There are also more isolates in network *A*, as nodes are not forced to be part of hubs anymore; and four out of the six isolates are also isolates in the overall network *M*, i.e. the unpopular nodes in within level network are also unpopular in the affiliation network. When AAAXS is negative, we actually see marginally smaller *A*-node SDs in both networks *A* and *X*, and smaller SK in network *A* than random networks representing more flat *A*-node degree distributions in both networks *A* and *X*, suggesting a tendency against centralization.

6.2. Interaction triangles

The interaction triangle effects (TXAX, TXBX, ATXAX and ATXBX) have two connected nodes within the same level, sharing one or more nodes from the other level through affiliations. They can be seen as affiliation based closures, or within level homophily effects based on common affiliations (i.e. the homophily can be construed as arising from a shared affiliation). We conducted simulations using positive and negative values (+2 and −2) for TXAX and ATXAX similarly as before, and the results are shown in Fig. 5 and Table 2. When TXAX is positive, it created clique-like structures in the within-, meso- and overall two-level networks.[2] In turn, we

see greater than random SD and SK in network *A*, and greater a SD(*A*) in network *X*, as there are centralization effects on network ties among nodes involved in the clique; and the global clustering coefficients for both networks *A* and *X* are also much greater than expected from random networks. These clique-like formations are to be expected from positive Markov triangle parameters, given that similar effects are evident in unipartite models. When TXAX is negative, the clique-like structures disappear, and most of the previously isolated nodes merge into the big components of the networks. Since the negative TXAX is trying to separate bipartite two-paths from the within level *A* ties, or in other words, the connected *A*-nodes tend not to share common affiliations, we can see long circular-like structures in network *X*. Compared with random networks, there is no obvious differences on the degree distribution and clustering measures we obtained.

The ATXAX attenuates the TXAX effect in a similar fashion as in the star configurations. When ATXAX is positive, we still have clique-like core structures in both network *A* and *X*. However, instead of being isolated, the rest of the *A*-nodes are connected to the core in network *A*; and only the core-members in network *A* are sharing multiple *B*-nodes in the bipartite network *X* and the overall multilevel network *M*. The GCCs and SDs for *A*-node degree distributions in both networks *A* and *X* are higher than random networks but smaller than the case when TXAX is set at the same positive value. Comparing the negative effects of TXAX and ATXAX, the attenuating ATXAX had higher values in most of the statistics in Table 2, but generally lower compared with the model with positive ATXAX. In the cases of researcher–laboratory networks for instance,

---

[2] Note that for bipartite networks, sub-graphs with two nodes sharing one or more nodes of the other type are all considered as cliques.

a positive TXBX or ATXAX effect may indicate that researchers from the same laboratories tend to seek advice from each other.

6.3. Interaction three-paths

Defined by two X-ties and one within level tie (either A or B), we label the three-path statistics involving the interactions between the within- and meso-level networks as L3XAX and L3XBX. They can be seen as the interaction between meso-level popularity and within-level activity, or meso-level degree assortativity through within level ties. From the simulation results as shown in Fig. 6 and Table 3 using the same settings as previously, we can see that the positive L3XAX parameter creates as many XAX three-paths as possible such that non-isolated A-nodes in network A are all involved in the bipartite network X, as an A-tie is an essential part of the L3XAX. Compared with the random GCCs for both networks A and X, the networks with positive L3XAX have higher GCCs as shown in Table 3. Because all bipartite ties are associated with non-isolated A-nodes in network A, they create more TXAXs, hence more L3XAXs as the TXAX configuration is also part of L3XAX. However, the effect is not as strong as a positive TXAX effect where the GCCs are much higher, as shown in Table 2. When L3XBX is negative, there are very long paths in the bipartite network to make bipartite clustering less likely; and the clustering in network A is very similar to random networks. Together, they create fewer XAX three-paths.

In the researcher–laboratory network context, a positive L3XAX parameter may suggest that researchers (B) popular in the affiliation network tend to be associated with collaborating laboratories (A); or there is a degree assortativity effect such that labs tend to collaborate with other labs having a similar number of researchers.

6.4. Cross-level three-path

The cross-level three-path L3AXB configuration involves ties from all three (A, B and X) networks, and is very different from L3XAX or L3XBX; it represents a tendency for nodes that are popular within levels to be affiliated through the meso-level network X, and can be seen as degree assortativity effect between networks A and B through meso-level network X. We conducted simulations for the L3AXB effect using the same positive and negative parameter values. Since L3AXB involve tie variables from all three networks (A, X and B), network B is now set to have 65 ties with fixed density; other simulation settings are set the same as before. The simulated sample plots and graph statistics are presented in Fig. 7 and Table 4. When L3AXB is positive, hubs emerge from one of the within level networks (in our case, network A); and ties from the other network (B) form a clique-like strongly connected component. All nodes that are part of the connected component in network (A) are connected to the hubs in network (B) through the bipartite network; and unsurprisingly, all of the SDs, SKs and GCCs are higher than random networks. Note that starting the simulation from random networks of the same densities, the structures for networks A and B may swap depending on the network in which the hubs emerge first, and the other network in turn forms the clique-like structure. When L3AXB is negative, to form as few as possible AXB three-paths, the parameter avoids any affiliations between nodes from the connected components of the within-level networks. As a result, the isolated nodes in within level networks become the popular nodes in network X. We can see from the visualization of the overall multilevel network M that the two within-level networks are loosely connected by few bipartite ties, and the isolates are working as agents or brokers between the two components. The statistics listed in Table 4

indicate that the global structures of the within-level networks are very similar.

In a researcher–laboratory network context, a positive L3AXB parameter suggests high degree researchers are affiliated with high degree labs.

For directed within level networks, we can define four types of L3AXB as L3AXB-in, L3AXB-out, L3AXB-path and L3BXA-path as shown in Fig. 3. The L3AXB-in/-out, where the within level incoming or outgoing ties represent within level popularity or expansiveness, may be interpreted similarly to the non-directed L3AXB as effects indicating the tendency for popular or expansive nodes from the two levels to form affiliations. L3AXB-path and L3BXA-path however, represent the interaction between popularity of nodes at one level and expansiveness of nodes at the other level through their affiliations. In the researcher–laboratory network context, a positive L3AXB-in effect may indicate that the popular laboratories (A) in the collaboration network tend to have resourceful researchers (B) from whom lots of other researchers seek advice; whereas a positive L3AXB-path effect may indicate that popular laboratories tend to have researchers seeking more advice from others.

6.5. Cross-level four-cycle

The cross-level four-cycle C4AXB involves ties from all three networks, representing a cross-level "mirroring" or alignment effect such that members of connected groups are themselves connected (Lazega et al., 2010, 2011). Through simulations using similar settings as before, we can see from Fig. 8 that when C4AXB is positive, it generated clique-like cores in both networks A and B, and the members of the cores are affiliated with each other. Statistics from Table 4 suggest there is centralization on ties from both of the within-level networks with greater than random SDs and SKs; with the clustering of high-degree nodes reflected by the high GCCs in all networks A, B and X. The negative C4AXB effect however, is trying to misalign A- and B-ties through affiliation, and can be generally understood as a decentralization effect, with fewer isolated nodes especially in the affiliation network X. In terms of degree distributions and global clustering, the SDs, SKs and GCCs for all three networks are not very different from random.

In the researcher–laboratory network context, a positive C4AXB may suggest researchers from collaborating laboratories tend to seek advice from each other.

For directed networks, depending on the directions of the within-level ties and whether reciprocal ties are involved, C4AXB may have several forms as shown in Fig. 3.

6.6. Models for caveman graphs

As an illustrative example to show how the proposed multilevel model parameters may interact with each other, we compare ERGMs with and without the cross level interaction effects to simulate caveman graphs, where small clique like components are sparsely connected to form a large component. Caveman graphs were introduced in a discussion of the small world phenomenon by Watts (1999). In this case, we continue to use the multilevel terminology, although now the 'levels' may simply be construed as two different types of nodes as Wasserman and Iacobucci (1991) originally described.

To date, there is not a typical set of one-mode ERGM parameters that could reproduce caveman graphs.[3] However, it is not difficult to reproduce caveman like graphs using two-level ERGMs where

---

[3] Robins et al. (2005) presented some simulation studies involving frozen graphs of different forms representing small and other worlds.
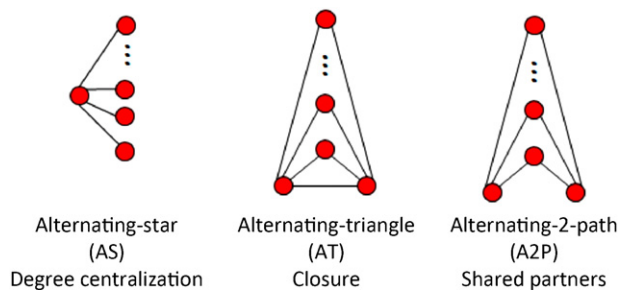
Fig. 9. ERGM configurations following Snijders et al. (2006).

Fig. 11. A caveman graph with cross-level effects.

within level networks form the clique like structures or "caves", and the cross level bipartite network can be used to connect the caves. Simulations were conducted on (20, 20) two-level networks with fixed within level densities of 0.15, and a fixed bipartite density of 0.06. For within level networks (*A* and *B*), we use the one-mode ERGM specification following Snijders et al. (2006), including alternating-stars (AS), alternating-triangles (AT) and alternating-two-paths (A2P) as shown in Fig. 9.

To have the highly clustered but separated components or "caves" in the within level networks, we use the following set of parameters for both within level networks *A* and *B* where:

AS = −1, AT = 2 and A2P = −1. The positive alternating-triangles ensure high clustering; the negative alternating-stars decentralize the degrees, i.e. discourage formation of hubs; and the negative alternating-two-paths break down the network into separate components. The bipartite network (*X*) is left at random, and all between or cross-level interaction effects are set at 0s. We use this model as a reference where no between-level interaction effects are imposed. Fig. 10(a) and (b) show a typical set of within level networks simulated from this model. Fig. 10(c) is the simulated random bipartite network (*X*) that connects the caves from both networks *A* and *B* to form the caveman like two-level graph in Fig. 10(d). We can observe some random overlapping between the caves where connected dyads from caves of different types are forming cross-level four-cycles (C4AXBs). Depending on the density of the bipartite networks, the random overlapping may have some of the caves tightly interlocked.
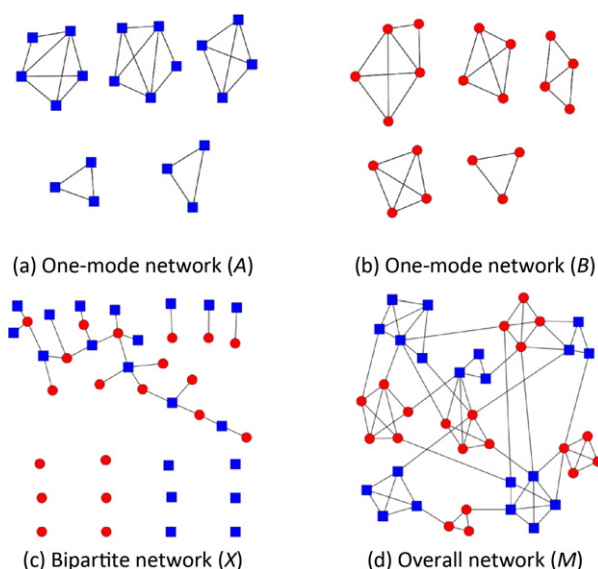
For the interaction model, we use the same set of within level parameter values to create the "caves". The within bipartite network (*X*) effects were left at 0s, but the following interaction effects were imposed where: AAAXS = 2, ABAXS = −2 and C4AXB = −2. A simulated graph sample is presented in Fig. 11. These parameters in effect imply that popular nodes in the *A* network will be popular with *B* nodes as well (i.e. centrality in the *A* network is associated with centrality in the *X* network); but that popular *B* nodes are not popular with *A* nodes; and that cross-level 4 cycles are less likely to be present.

The opposite signs of the interaction stars (AAAXS and ABAXS) affect the popularities of nodes as well as the sizes of the clique-like components of different types. To generate more AAAXS but fewer ABAXS, the bipartite ties are more centralized on nodes of type *A*. Comparing Fig. 11(a) to Fig. 10(a), we have fewer but large clique like components, as well as more isolates, indicating stronger degree centralization in network *A*. Compare Fig. 11(c) to Fig. 10(c), there are more isolates of type *A* than *B*, hence the average degree of *A* nodes are higher in the bipartite network. As nodes from larger cliques have higher average degrees than smaller cliques, and the bipartite ties are centralized on *A* nodes that are not isolates (as shown in Fig. 11(d)), the interaction effects effectively created more AAAXS and less ABAXS than the previous model. The negative C4AXB limits the alignment or entrainment of ties from different cliques such that only a single member within a clique may connect to multiple members of a clique from the other level. In other words, we have brokerage effects between different types of 'caves' where one member of one cave tends to be the gatekeeper in regra to the other cave. The combined effect of the three interaction parameters created caveman graphs where smaller caves from one level are centralized on larger caves of the other level as shown in Fig. 11(d). Such networks are not uncommon, e.g. in organizational networks, suppose that larger cliques represent a small number of central management divisions, and smaller cliques represent other business divisions. This model represent networks where within each division (large or small) nodes are highly connected, but between divisions, there is only one or two nodes worked as brokers or representatives liaising with other divisions.

As we can see from these simulation studies, even very simple interaction effects can create structure at one or both levels. Our example in the next section demonstrated the importance of these parameters in capturing the dependencies between the macro-, micro- and meso-level networks.
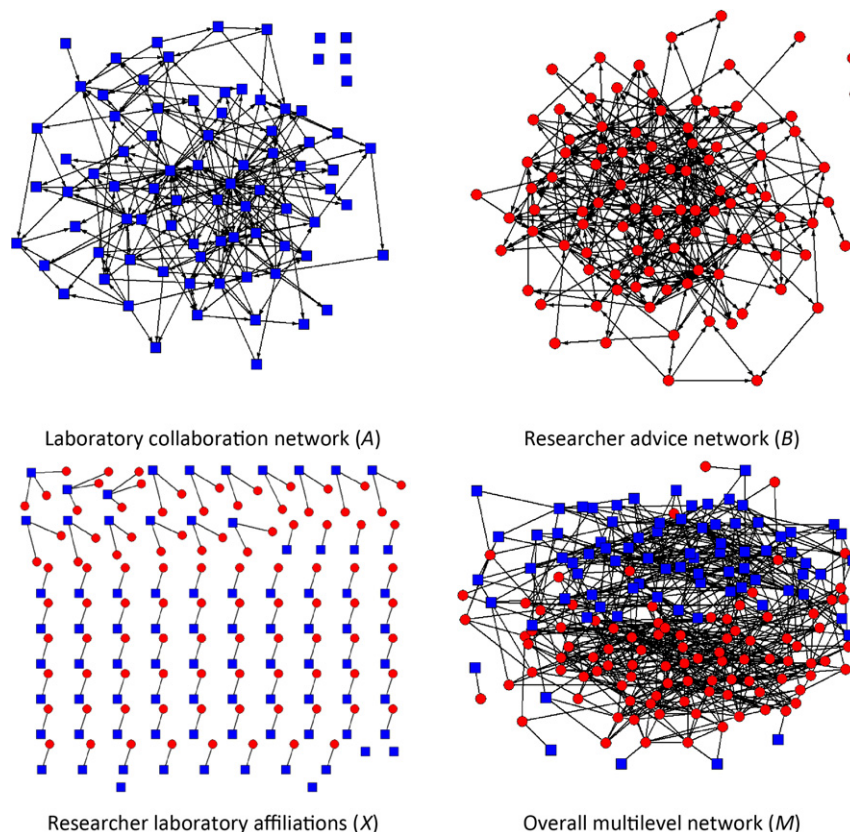

Fig. 10. A caveman graph without cross-level effects.

Fig. 12. The multilevel network among French cancer research elites (Lazega et al., 2008).

7. Models for French cancer researchers and their institutions

7.1. The network data

The network data was collected from French cancer researcher elites[4] and their institutions in 1999. Several studies have been carried out based on this set of data (Lazega et al., 2004, 2006, 2008, 2010, 2011). The data set consists of 97 researchers and 82 laboratories. The corresponding two-level network consists of a directed collaboration network among the laboratories (*A*) as nominated by the laboratory directors; a directed advice network (*B*) where researchers nominated from whom they seek advice; and the bipartite network (*X*) representing the affiliation between the researchers and laboratories. Fig. 12 displays the multilevel network as three separate networks.

This data set has several features. First of all, the affiliation network is based on public information about the researchers' affiliations, each individual researcher was only affiliated with one laboratory, and most laboratories only had one researcher who was a member of these elites. It is hard to infer more interesting bipartite structures apart from a list of isolated dyads and stars of sizes two or three, so we treated the bipartite network as exogenous in our models. This implies that the bipartite network is fixed throughout the model estimations and goodness of fit tests. Secondly, there are two laboratories with much higher

out-degrees (29 and 36) compared with the rest of the laboratories (with a maximum out-degree of 13). Both of these laboratories were providers of experimental equipment and materials for other laboratories. Given their special position, we treated ties associated with these two laboratories as exogenous in our models. Note that the isolated nodes in the laboratory or researcher networks are not isolates in the affiliation network.

7.2. Modelling results and interpretations

We firstly present models for within level networks, and then compare them with a model for the overall multilevel network in both parameter estimates and model goodness of fits. In other words, the within-level model simply models the *A* and *B* (researcher and laboratory) network structures under the assumption that the bipartite affiliations are irrelevant, whereas the interaction models takes the bipartite structure into account. Such comparisons can reveal the importance of the cross level network processes that shape the structure of both the individual within level networks and the overall multilevel network.

Both of the within level networks in our data are directed. Fig. 13 shows some of the directed ERGM configurations involved in the within level models. The alternating-in-stars (AinS) represent the in-degree centralization or popularity spread; the alternating-out-stars (AoutS) represent out-degree centralization or activity spread (Robins et al., 2009). Combining the in and out degree alternating statistics, we can also derive a statistic known as the alternating-in-alternating-out-star (AinAoutS) representing the correlation between in and out degrees. While this is seldom used in unipartite networks (sometimes a simple two-path parameter is used to control for this correlation), it will be important in

---

[4] The selection criterion for inclusion in this population was having published at least eight papers per year for three consecutive years – 1996–1998 – in international cancer research journals listed in the CANCERLIT database (nowadays included in the PUBMED database).
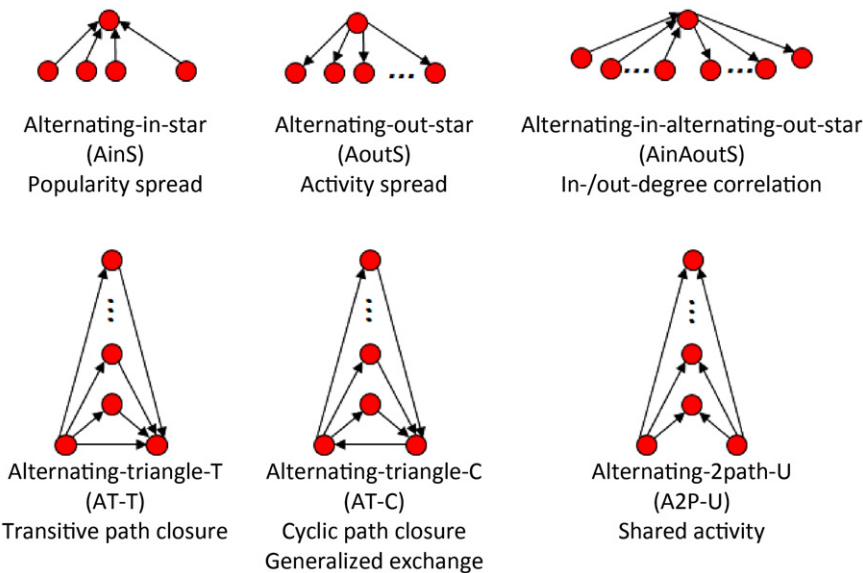
Fig. 13. Directed ERGM configurations used in the modelling example.

the modelling examples. Robins et al. (2009) applied the same dependence assumption and techniques using geometric weightings and alternating signs as in Snijders et al. (2006) for directed networks, and introduced alternating-*k*-triangles/-*k*-two-paths of different forms. In our models, we include the AT-T configuration representing the tendency for multiple transitive paths to form closure; AT-C for cyclic path closure, which may be interpreted as a form of generalized exchange; and A2P-U for shared activity.

All models presented here are successfully converged, so convergence statistics are not presented. For model goodness of fit (GOF) test, simulated graph distributions using the converged model were compared with the observed statistics, and *t*-ratios are used as heuristic GOF test statistics (Snijders, 2002; Hunter et al., 2008). We tested each within-level models against 54 within-level graph statistics, and for the overall multilevel goodness of fit, 131 graph statistics were tested. We only present the statistics that are not well fitted, defined as *t*-ratios greater than 2.0 in absolute values.

7.2.1. Model for laboratory collaboration network only

Despite treating the two high out-degree equipment/material providers as exogenous, the in- and out-degree distributions of the laboratory remain skewed as shown in Fig. 14, and there are still a small group of laboratories with high in-/out-degrees compared with the rest of the nodes, making the degree distributions long tailed, or even multimodal. Despite the robustness of the social circuit model specifications (Snijders et al., 2006), the long tailed degree distributions remain a challenge for ERGMs in model goodness of fit. The model estimates for the laboratory network shown

in Table 5 fitted all 54 graph statistics adequately, including the standard deviations and skewness of the in-/out-degree distributions. To provide a good fit to the long-tailed in-degree distribution, we needed two alternating-in-star (AinS) parameters with different λ-values (4.0 and 2.0). (Appendix presents some simulation studies on how combining different star or AS parameters with different λ-values within the same model may help capturing the long-tailed degree distribution.) Even though the AinS with λ = 2.0 is not

Table 5
Parameter estimates for laboratory collaboration network only.

Effects	Estimates	Standard errors
Arc	−3.811	0.530[*]
Reciprocity	1.784	0.345[*]
Two-path	−0.100	0.025[*]
Isolates	2.051	0.834[*]
AinS(4.00)	0.702	0.255[*]
AoutS(4.00)	0.315	0.082[*]
AinS(2.00)	−1.019	0.582
AT-T(2.00)	0.493	0.133[*]

[*] indicates significant effect where the estimated effect exceeds twice the standard error in absolute value.

Laboratory in-degree distribution
Standard deviation = 2.58
Skewness = 1.12

Laboratory out-degree distribution
(without the two high degree nodes with out-degrees 29 and 36)
Standard deviation = 4.96
Skewness = 4.20
In/Out degree correlation = 0.24
Mean degree = 3.34

Fig. 14. In- and out-degree distributions of the laboratory collaboration network.

Researcher in-degree distribution
Standard deviation = 3.76
Skewness = 0.67

Researcher out-degree distribution.
Standard deviation = 5.23
Skewness = 1.40
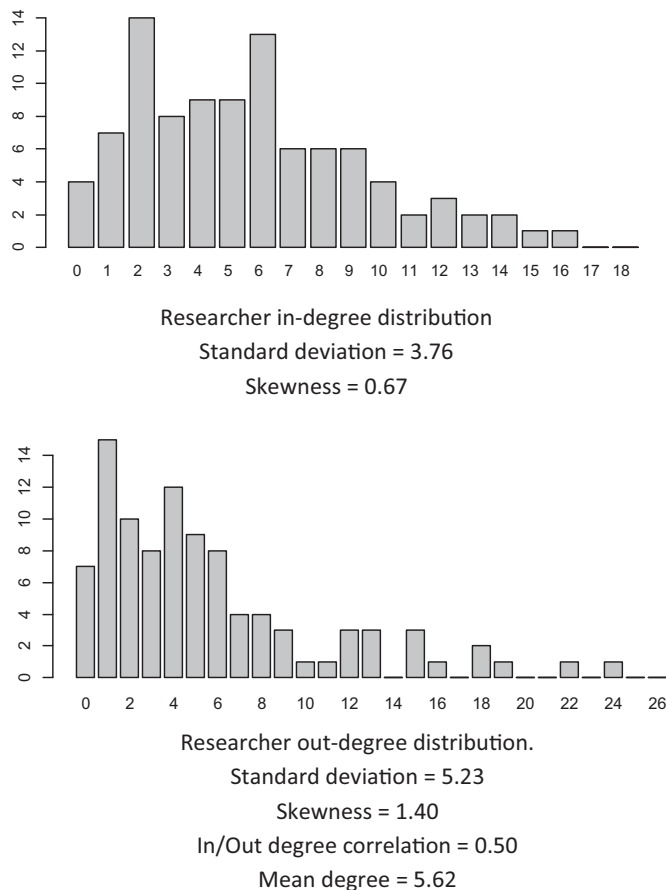In/Out degree correlation = 0.50
Mean degree = 5.62

Fig. 15. In- and out-degree distributions of the researcher advice network.

significant, a GOF comparison between the presented model and a model without AinS(2.0) showed that by including the AinS(2.0) in the model, the t-ratio for the skewness of the in-degree distribution reduced from 2.74 to 1.26. For many practical modelling applications, seeking to obtain such excellent fit may not be necessary, so these models may seem overly complex. However, we seek good fit here on many indices in order to illustrate the relevance of the full multilevel structure.

From the parameter estimates, we can see that the collaborations between laboratories tended to be reciprocated. The positive and significant alternating in-/out-stars with $\lambda = 4.0$, indicate both the in- and out-degree degree distributions are more dispersed than expected from random networks. In other words, both the popularity and expansiveness of the laboratory collaborations are centralized. The positive AT-T suggests the laboratory collaboration tended to be clustered. Combining the clustering effect and the negative two-path suggests a path-shortening effect whereby the collaborations are more direct, rather than going through multiple paths. Given the other effects, there are more isolated nodes than expected (although those isolates are not isolated nodes in the affiliation network).

7.2.2. Model for the researcher advice network only

The advice network contains advice seeking ties among 97 researchers. Both the in- and out-degree distributions are skewed, especially for the out-degree distribution with long tails with two expansive researchers nominating more than 20 ties, as shown in Fig. 15. To provide adequate fit to the degree distributions along with the other statistics, we needed several star parameters as shown in Table 6. Some of the star parameters are not significant, but they are important in fitting the degree distributions and

Table 6
Parameter estimates for researcher advice network only.

Effects	Estimates	Standard errors
Arc	−3.213	1.024[*]
Reciprocity	3.534	0.213[*]
2-Out-star	0.358	0.146[*]
3-Out-star	−0.018	0.009
Two-path	−0.135	0.010[*]
AinS(4.00)	0.596	0.159[*]
AoutS(4.00)	−0.722	0.599
AinS(2.00)	−1.164	0.450[*]
AoutS(2.00)	0.384	0.787
AinAoutS(2.00)	−0.233	0.369
AT-T(2.00)	0.932	0.067[*]

[*] indicates significant effect where the estimated effect exceeds twice the standard error in absolute value.

the clustering coefficient as well as the correlation between in- and out-degrees. The same model specification without two- and three-out stars underestimated the skewness of the out-degree distribution (t-ratio = 2.12), and failed to capture the global clustering coefficient (t-ratio = 3.48). The same model specification without AinAoutS underestimated the in- and out-degree correlation (t-ratio = −4.66).

Besides the various star parameters, the interpretation of this model is similar to the laboratory collaboration network. The strongly positive reciprocity effect indicates advice seeking tended to be reciprocated in this network. Through the positive AT-T, we see researchers seek advice in clusters; combined with a negative two-path parameter, we can infer the path-shortening effect also existed in the advice network. In other words, network closure is the predominant effect in this model, suggesting advice seeking among researchers in a clustered, team-like way.

The interpretations of the various star parameters generally infer a long-tailed or multi-mode degree distribution, as we have discussed in the simulation study of AS with different λ-values. The positive and significant 2-out-star suggests there are hub-like advice seekers who ask a lot more advice than others. The positive AinS(4.0) allow some variations for nodes with high in-degrees but not hubs; and the negative AinS(2.0) suggests there was little variation among researchers in being popular sources of advice.

7.2.3. Within level model GOF with respect to between-level statistics

The separate models for laboratories and researcher fitted within level statistics well with all t-ratios less than 2.0. However, by testing the models against multilevel interaction statistics, they fail to capture a list of statistics listed in Table 7.

The cross-level triangulation effects (TXBXarc and TXBXreciprocity) are quite basic: they represent researchers from the same laboratory seeking advice from each other. Comparing the observed statistics with the simulated means based on the within level models, the observed advice seeking within laboratories is much higher than expected from the models that do not explicitly cater for cross-level effects. The within level models also failed to capture the

Table 7
Poorly fitted statistics for models with within-level effects only.

Configurations	Observed	Mean	Standard deviations	GOF (t-ratio)
TXBXarc	30	2.391	1.775	15.55
TXBXreciprocity	14	0.612	0.770	17.38
C4AXBentrainment	202	43.918	14.893	10.62
C4AXBexchange	196	43.532	13.768	11.07
C4AXBexchangeBreciprocity	136	23.700	9.209	12.20
C4AXBexchangeAreciprocity	60	6.844	5.091	10.44
C4AXBreciprocity	11	0.942	0.993	10.13

Table 8
A multilevel ERGM for French cancer research elites.

Effects	Estimates	Standard errors
Laboratory collaboration network		
Arc	−3.831	0.556*
Reciprocity	1.679	0.381*
Two-path	−0.079	0.029*
Isolates	2.017	0.760*
AinS(4.00)	0.640	0.268*
AoutS(4.00)	0.320	0.086*
AinS(2.00)	−0.889	0.614
AT-T(2.00)	0.446	0.127*
Researcher advice network		
Arc	−4.084	0.118*
Reciprocity	3.313	0.212*
AT-T(2.00)	1.085	0.072*
AT-C(2.00)	−0.384	0.068*
A2P-U(2.00)	−0.071	0.020*
Laboratory collaboration and affiliation		
AXS1Ain(2.00)	0.240	0.131
AXS1Aout(2.00)	−0.324	0.129*
Researcher advice and affiliation		
TXBXarc	1.958	0.275*
Cross level interactions		
L3AXBin	−0.006	0.018
L3AXBout	−0.012	0.008
L3BXApath	−0.003	0.010
L3AXBpath	−0.051	0.015*
C4AXBentrainment	0.634	0.104*
C4AXBexchange	0.639	0.109*
C4AXBexchange reciprocal A	−0.293	0.065*
C4AXBexchange reciprocal B	−0.295	0.136*

* indicates significant effect where the estimated effect exceeds twice the standard error in absolute value.

cross level exchange and various forms of cross level entrainment effects. These effects represent tendencies for researchers from collaborating laboratories to ask advice from each other. Evidently, these tendencies are much stronger than expected from the within level models. These poorly fitting statistics indicate that the cross level interaction parameters are required to provide good fit to the network data.

7.2.4. Model for the overall multilevel network

The model for the overall multilevel network is presented in Table 8. It has several sections including effects for both within level networks, interaction effects between the advice network and the affiliation network, interaction effects between the collaboration network and the affiliation network, and the cross level interactions involving all three networks. In this case, all 131 GOF statistics were less than 2.0 in absolute values, indicating a good fit to the observed two-level network on a very wide range of graph statistics.

We interpret this model by sections while keeping in mind that all parameters in the model are dependent on each other, hence the interpretation of one particular section is always conditional on the rest of the model. We start with the sections involving cross level network effects and then move on to the interpretations of the within level effects. Comparing the changes in the within level effects with the models without cross level effects, we discuss the impact and importance of the cross level effects.

The Interaction effects between the laboratory collaborations and the affiliation network include alternating laboratory affiliation stars with an incoming collaboration tie (AXS1Ain) and an outgoing collaboration tie (AXS1Aout). Note that the AXS1Ain parameter is not significant (defined here as an absolute value greater than twice the size of the estimated standard error) but close to significance. Its removal from the model leads to poor fit on the in-degree distribution of the within level laboratory collaboration network. Having the alternating affiliation star as part of the statistics explicitly models the laboratory with more than one researcher, which

are among the popular laboratories or the "big ponds" in the network as described in Lazega et al. (2008). This positive estimate suggests that bigger laboratories tended to receive more nominations as collaborators from other labs. However, in contrast, they tended to nominate fewer collaborating labs, as indicated by the negative and significant AXS1Aout parameter. This interpretation makes empirical and theoretical sense, as the "big pond" labs might be the dominant research centres that had good reputations and more research resources that other labs were chasing, but they themselves did not feel the same urgency to collaborate with other laboratories in the same country.

The only significant interaction effect between the advice network and the affiliation network in the model is a triangle: researchers affiliated with the same laboratories seek advice from each other (TXBXarc). Not surprisingly, such effect is positive and strongly significant. From the model goodness of fit test, this effect, together with other parameters, provide adequate fit to the reciprocated version of TXBXarc, or the TXBXreciprocity which were not captured by the model without cross level effects.

The cross level interaction parameters involving all three networks include the in-/out-degree assortativity effects represented by various forms of three-paths, and the cross level entrainment and exchange effects represented by different four-cycles. Only one of the three-path effects is significant, but removing the non-significant three-paths made the model convergence difficult. The negative and significant L3AXBpath parameter estimate suggests there was an anti degree assortativity effect between the laboratories' in-degree and researchers' out-degree. In other words, the popularity of labs in the collaboration network and expansiveness of researcher advice seeking are compensating each other rather than forming a joint core-periphery structure. The negative L3AXBpath does not mean such a configuration does not exist in the network. By combining the negative L3AXBpath effect with the positive and significant closure parameter C4AXBexchange, we see that L3AXBpath is more likely to exist within closure, which means that researchers from labs who receive more collaboration tend to seek advice from researchers from collaborating laboratories but not from other researchers.

The different types of cross level four-cycle effects present an interesting interpretation of the multilevel network structure. (Note that another possible cross level four-cycle parameter (C4AXBreciprocity) was not included in the final model, as it was not significant, and the current model fits the statistic well.) The cross level entrainment and exchange four-cycle effects are both positive, indicating a general co-occurrence of advice at one level and collaboration at another, such that researchers within collaborating laboratories were more likely to seek advice from each other. This is consistent with the 'fusional' strategies described in Lazega et al. (2008) where researchers only work within the boundaries defined by their affiliated laboratories and the laboratories' collaborators; or a tendency for researchers not to seek advice from researchers affiliated with non-collaborating laboratories.

However, such co-occurrence was less likely to happen when reciprocation occurs in either the collaboration network or the advice network, as indicated by the negative and significant (C4AXBexchange reciprocal A) and (C4AXBexchange reciprocal B) parameters. Note that the reciprocity effects for both the laboratory and researcher within networks are strong and positive. But these reciprocations are less likely to occur when cross-level closure is involved. As the reciprocal ties can be considered as strong ties, we suggest a possible interpretation for these negative effects. It is possible that a strong form of collaboration among laboratories may result in shared knowledge and resources generally, so the researchers no longer have to seek advice from one another in those laboratories. On the other hand, reciprocal advice seeking among researchers may be more likely to occur across non-collaborating

laboratories, perhaps as they may have distinct but complementary knowledge or skills, such that advice in rather different areas may provide mutual benefits (Lazega et al., 2011). This is also consistent with the argument that the strong forms of fusional strategies are not ideal for the researchers' performance (Lazega et al., 2008).

The within level effects for the laboratory collaboration network are largely consistent with the within level only model, as we have the same parameter specifications and compatible scales on the parameter values. Therefore, we may use the same interpretations as previously.

For the within level researcher advice network, however, the multilevel model greatly simplified the previous specifications. No star parameters were required to capture the researcher degree-distributions. We may now interpret the apparently multi-modal or long tailed degree-distributions as largely an artefact of the cross level interactions. The interpretations of the positive and significant AT-T confirm that advice seeking in this system tended to happen in clusters; together with the negative and significant generalized exchange effect (AT-C), we see the advice network as locally hierarchical. The negative alternating-two-path (A2P-U) representing shared advice seeking may also help shape the out-degree distribution as it indicates that active advice seekers tended not to seek advice from common others.

In summary, the multilevel ERGM reveals some interesting structural features of the researcher multilevel network, and intuitive interpretations consistent with previous work are available. Collaboration does take place in a clustered manner for both researchers and laboratories; advice seeking is stronger within laboratories; and collaborating laboratories tend to have affiliated researchers seeking advice from one another; however, reciprocated or strong collaboration did not encourage advice exchange between researchers. More importantly, we see that some complicated features of the within-level network structure are explained solely by the cross-level interactions: in this case the degree distribution of researcher advice-seeking.

8. Conclusion

In this paper, we proposed a generalized data structure which captures both within-level and meso-level networks. The data structure is flexible, and accommodates most of the multilevel network studies to date. We proposed a new multilevel ERGM to model this structure.

Based on the hierarchy of tie dependence assumptions as summarized by Pattison et al. (2009) and Wang et al. (2013), we have formulated the general form of ERGMs for two-level network data, and proposed model specifications for both non-directed and directed networks. They are designed to reveal the interdependencies among the micro-, macro- and meso-level networks. Our simulation studies demonstrated the properties of each of the proposed configurations, and showed that even simple cross-level effects can create highly structured within-level networks. When the cross-level effects are not included, the models are no different from fitting three independent ERGMs, two for within-level one-mode networks, and one for the bipartite ties. Our empirical example shows that these cross-level effects not only provide better fits to the data, but also simplified otherwise complicated within-level models. By comparing the models with and without the interaction effects, our example showed that an apparently multi-modal long-tailed degree distribution can be seen as an artefact of the dependence among micro-, macro- and meso-level ties.

Depending on the context and the research question, the definition of the levels in a multilevel network can be determined by either the different nature of the nodes, individuals and groups, for example; or the classifications of nodes by their attributes such as gender, occupation, and education. In the first instance, network ties of different types may be defined for the nodes of different levels, such as the example we presented in the paper. For the second case where the levels are defined based on a nodal attribute, but the same type of relation is defined for nodes within and cross different categories. This is the type of data originally envisaged by Iacobucci and Wasserman (1990) and Wasserman and Iacobucci (1991). The research question is then based on the assumption that the network effects are heterogeneous among nodes with different attributes. In fact, any network data involving binary categorization of nodes can be seen as a "two-level" network. For example, in a friendship network context, a two-level network may be constructed by the genders of the nodes assuming that friendship network processes are different within and between genders. When more than two categories are defined among a common set of nodes, we can analyze the data as "multilevel" networks assuming network processes are different within and between categories. The ERGMs proposed in this paper can be equally applied to both cases.

Multilevel data opens interesting possibilities for sampling. The cluster sampling method, where one first samples units and then samples within units, may be relevant to multilevel networks. We cannot take a simple random sample of nodes in order to estimate a network model for ties within units, but Handcock and Gile (2010) and Pattison et al. (2011) have proposed conditional estimation strategies for snowball sampled network data where ERGM parameters can be obtained based on a snowball sample of the (small or large) network. More work needs to be done to consider bow best such approaches may be applied to multilevel networks.

Future elaboration of the proposed multilevel ERGMs may include social selection models (Robins et al., 2001a) and autologistic actor attribute models (ALAAMs) (Robins et al., 2001b; Daraganova and Robins, 2013). Social selection models include actor attributes as covariates, and allow the tests of how nodal attributes may affect multilevel structure, for example, how individuals' performance affect their network positions. ALAAMs for multilevel networks on the other hand assume network structures are exogenous, and test how nodal attributes are affected by individual's multilevel network position; they may share similar configurations or network statistics as in social selection models but these are treated as exogenous.

Regardless of the forms of the possible models, multilevel networks have more complicated dependencies among ties from different levels. These interactions make the models and their interpretations more complex. As there are almost unlimited number of possible graph configurations, and no formal "step-wise" model selection strategies, to find the best model for an empirical multilevel network will require theoretical guidance as well as model fitting experience. We see the proposed models and examples presented in this article as the first steps in a full elaboration of an ERGM approach to multilevel network analysis.

Acknowledgements

Appendix.

A.1. Alternating stars and different λ-values

To fit long-tailed degree distributions well, we may need to include two or more star parameters with different λ values. We illustrate how different star parameters may be combined in unipartite models to fit the degree distribution better. The
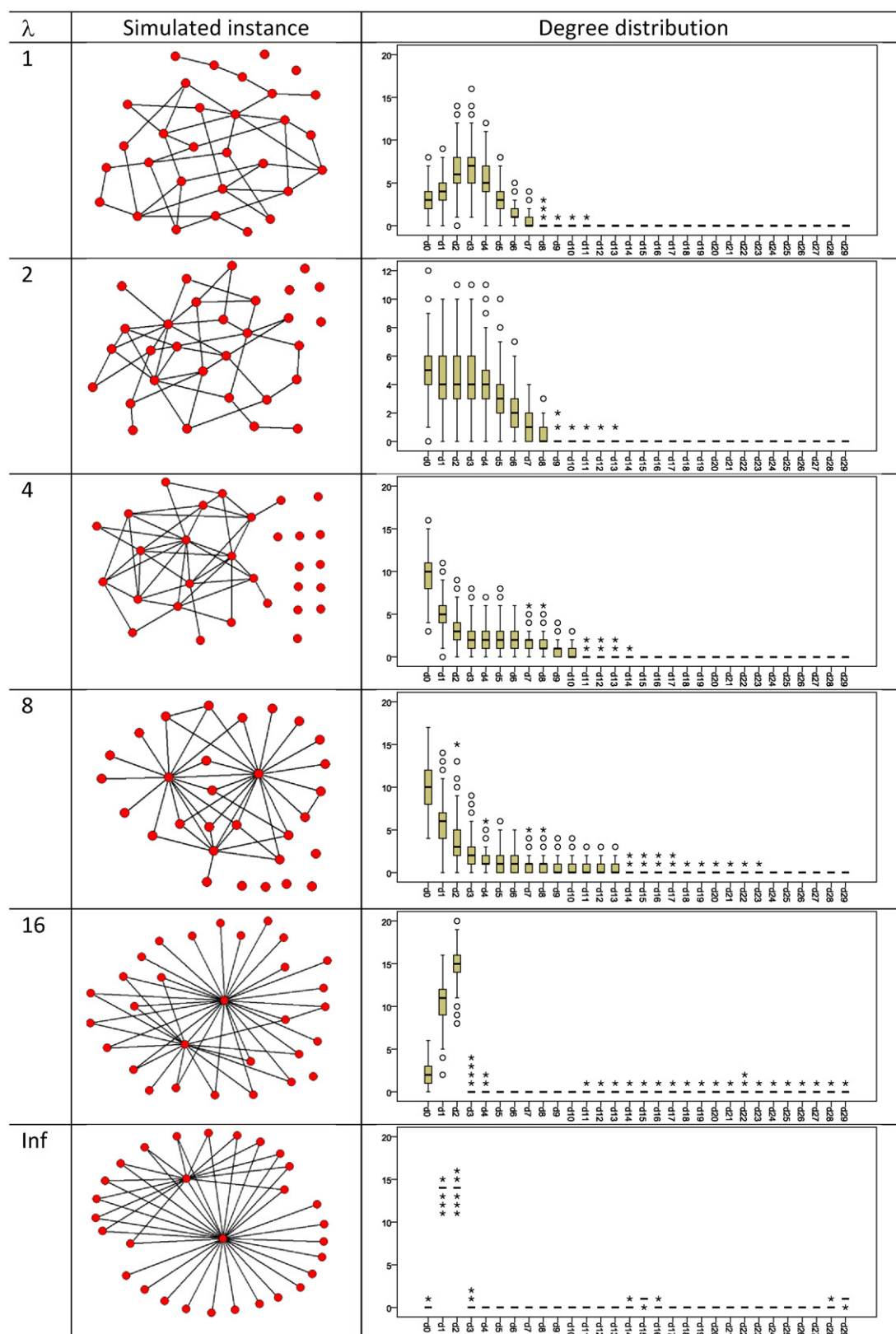
Fig. 16. Simulation using alternating-stars with different λ-values.

(a) AS(2.0)=-1 and Two-star = 0.25
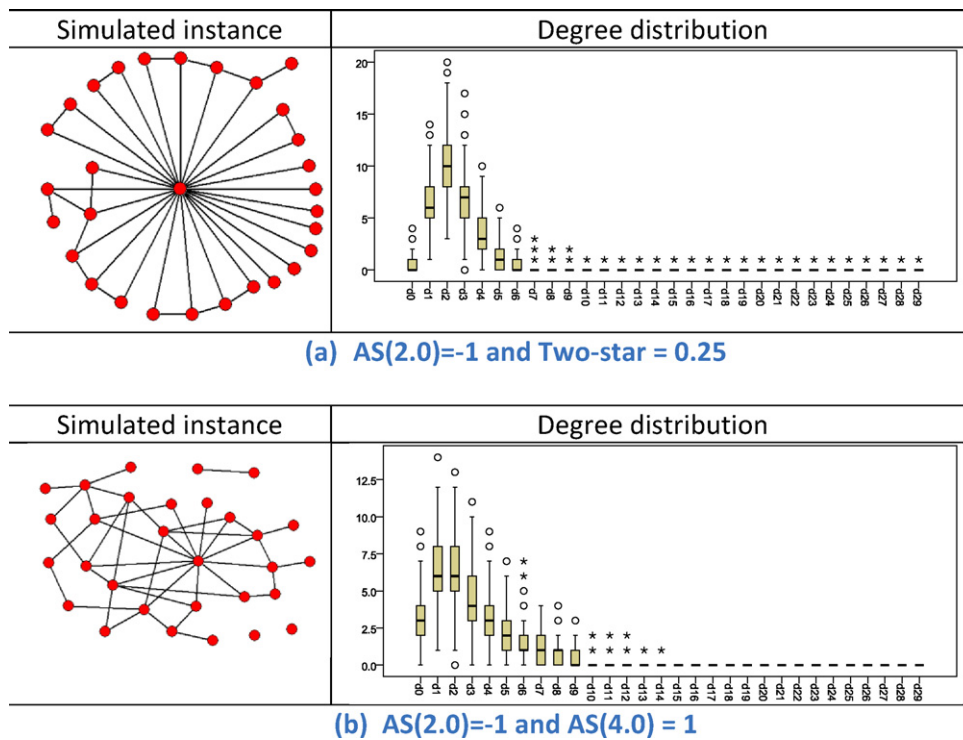

(b) AS(2.0)=-1 and AS(4.0) = 1

Fig. 17. Simulation using models with multiple (alternating-)star parameters.

alternating-star (AS) statistic uses geometric weighting parameter (λ)[5] on the graph degree-distribution. For a network with n nodes, it is defined as

$$z_{AS}(x, \lambda) = \sum_{k=2}^{n-1} (-1)^k \frac{z_{Sk}(x)}{\lambda^{k-2}}$$

where $z_{Sk}(x)$ is the number of stars of size k, and $i+$ is the degree of node i. Different λ-values may yield different degree distributions. From experience, the alternating statistics with $\lambda = 2.0$ (including AS, AT and A2P, etc.) generally provide robust model convergence. However, the converged model may still have difficulty in capturing the actual degree- or two-path-distribution for the cases where the distributions are heavy tailed. Hunter and Handcock (2006) proposed methods for estimating λ as a ratio parameter, although they are computationally intensive in practice. Robins and Lusher (2013) conducted some simulation studies to demonstrate how different values of lambda may result in different triangle and two-path distributions in the cases of alternating triangles and alternating two-paths. Robins and Lusher (2013) also demonstrated how we may improve model GOF by including both the alternating star (AS) and the two-star parameters in the model. In our modelling examples presented in this paper, obtaining excellent fit to the degree distributions required the two-star parameter as well as two or more AS parameters with different λ-values.

To illustrate model interpretation and the properties of AS with different λ values, we present simulations on networks with 30 nodes and a fixed density of 0.1. The AS parameter was set constant at 1.0, but the λ-values were changed from 2 to 16 in factors of 2s. We collected every 10,000th sample from a simulation of 10,000,000 iterations which gave a distribution of 1000 graphs. For each different λ-value, Fig. 16 presents an example graph from

the simulated graph distribution along with a box plot of the degree-distributions of the 1000 graphs. As the λ-value increased from 1 to 8, the range of the degree distributions increased from the highest maximum degree of 11 ($\lambda = 1$) to the highest maximum degree of 23 ($\lambda = 8$). The number of isolated nodes also increased from around 3 to about 10. The degree distributions look continuous (not multi-modal), and the tails become longer. With $\lambda = 16$, the degree distribution is more discontinuous with two modes, where most graphs had two or three very high-degree nodes, with the rest of the nodes having degrees below 5 with some variations. When λ equals infinity, which makes the alternating-star statistic equivalent to the two-star statistic, the graph distribution is almost frozen across the simulation, with two very high degree nodes. Based on these observations, we may conclude that a positive AS parameter with a greater λ value is likely to create high-degree nodes or hubs in a network. A greater λ may capture long-tailed degree distributions with greater variances. However, when λ is too big, the degree distribution may become multi-modal and frozen.

Model specifications with two or more star parameters with different λ values may be able to capture degree distributions that are long-tailed, continuous (not two- or multi-modal), and having some variations in both very high-degree nodes as well as low degree nodes. While a positive two-star parameter, or a positive AS parameter with a larger λ value, creates high-degree hubs, negative alternating star parameters with smaller λ values captures nodes with relatively low degrees and prevents the simulation from becoming frozen. Fig. 17 illustrates the effects of combining two or more star parameters in a single model. As in previous simulations, for networks with 30 nodes and a fixed density at 0.1, we fixed a positive two-star parameter at 0.25 and a negative AS ($\lambda = 2$) at -1. The positive two-star generated graphs with a very high-degree node connected to most of the other nodes in the network, while the negative AS created some variability in the low degree nodes (as compared to the frozen graph distribution generated only by the positive two-star).

To reduce the degree of the hubs, we replace the positive two-star parameter with an AS parameter with $\lambda = 4$. The simulation

---

[5] The λ-values for alternating-k-stars described here are equally applicable to the *geometrically weighted degree distribution* as described in Hunter and Handcock (2006) which uses a *ratio parameter* $\alpha = \ln(\lambda)$.

generated graph distributions with hubs having maximum degree of 14. Compare the degree distribution to the simulation with only one positive AS parameter with $\lambda = 4$, as shown in Fig. 16, the negative AS with $\lambda = 2$ made the majority of nodes having low degrees and with little variation, while the positive AS with $\lambda = 4$ made some high degree nodes possible. The decision about how many AS parameters with different λ-values should be included in a model depends on the observed degree distribution. In some cases (as demonstrated in our modelling example section), non-significant AS parameters are important for adequate fit to the data.

References

Agneessens, F., Roose, H., 2008. Local structural patterns and attribute characteristics in 2-mode networks: p* models to map choices of theatre events. Journal of Mathematical Sociology 32, 204–237.

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

Bellotti, E., 2011. Getting funded. Multilevel network of physicists in Italy. Social Networks, http://dx.doi.org/10.1016/j.socnet.2011.12.002.

Borgatti, S.P., Foster, P.C., 2003. The network paradigm in organizational research: a review and typology. Journal of Management 29, 991–1013.

Brass, D.J., Galaskiewicz, J., Greve, H.R., Tsai, W., 2004. Taking stock of networks and organizations: a multilevel perspective. Academy of Management Journal 47, 795–817.

Dansereau, F., Alutto, J., Yammarino, F., Dumas, M., 1984. Theory Testing in Organizational Behavior: The Variant Approach. Prentice-Hall, Englewood Cliffs, NJ.

Daraganova, G., Robins, G., 2013. Autologistic actor attribute models. In: Lusher, D., Koskinen, J., Robins, G.L. (Eds.), Exponential Random Graph Models for Social Networks: Theories, Methods and Applications. Cambridge University Press, New York, pp. 102–114.

Erdös, P., Rényi, A., 1960. On the evolution of random graphs. Publications of the Mathematical Institute of the Hungarian Academy of Science 5, 17–61.

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

Friedkin, N., 2006. A Structural Theory of Social Influence, vol. 13. Cambridge University Press, New York.

Handcock, M.S., 2003. Assessing Degeneracy in Statistical Models of Social Networks. Working Paper No. 39. Centre for Statistics and the Social Sciences, University of Washington.

Handcock, M.S., Gile, K.J., 2010. Modeling social networks from sampled data. The Annals of Applied Statistics, 5–25.

Hedström, P., Sandell, R., Stern, C., 2000. Mesolevel networks and the diffusion of social movements: the case of the Swedish Social Democratic Party. American Journal of Sociology 106, 145–172.

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

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

Holland, P.W., Leinhardt, S., 1981. An exponential family of probability distributions for directed graphs. Journal of the American Statistical Association 76 (373), 33–50.

Iacobucci, D., Wasserman, S., 1990. Social networks with two sets of actors. Psychometrika, 707–720.

Koehly, L.M., Peterson, S.K., Watts, B.G., Kempf, K.K.G., Vernon, S.W., Gritz, E.R., 2003. A social network analysis of communication about hereditary nonpolyposis colorectal cancer genetic testing and family functioning. Cancer Epidemiology Biomarkers and Prevention 12, 304–313.

Koskinen, J., Edling, C., 2010. Modelling the evolution of a bipartite network – peer referral in interlocking directorates. Social Networks, http://dx.doi.org/10.1016/j.socnet.2010.03.001.

Kozlowski, S.W.J., Klein, K.J., 2000. A multilevel approach to theory and research in organizations: contextual, temporal, and emergent processes. In: Klein, K.J., Kozlowski, S.W.J. (Eds.), Multilevel Theory, Research, and Methods in Organizations: Foundations, Extensions, and New Directions. Josse-bass, San Francisco, pp. 3–90.

Latapy, M., Magnien, C., Vecchio, N.D., 2008. Basic notions for the analysis of large two-mode networks. Social Networks 30, 31–48.

Lazega, E., Jourda, M.-Th., Lazega, D., Mounier, L., 2011. Network lift from dual alters: extended opportunity structures from a multilevel and structural perspective. In: Sunbelt XXXI, The Annual Meeting of the International Network for Social Network Analysis, Tampa, Florida, February 12, 2011.

Lazega, E., Jourda, M.-Th., Lazega, D., Mounier, L., 2010. Network parachutes from tetradic substructures: "organizational expansion" and the measurement of the contribution of laboratories to their scientists' relational capital and performance, Cahiers de l'ORIO, N° 1. Working paper. University of Paris-Dauphine.

Lazega, E., Jourda, M.-Th., Mounier, L., Stofer, R., 2008. Catching up with big fish in the big pond? Multilevel network analysis through linked design. Social Networks 30, 157–176.

Lazega, E., Mounier, L., Jourda, M.-Th., Stofer, R., 2006. Organizational vs. personal social capital in scientists' performance: a multilevel network study of elite French cancer researchers (1996–1998). Scientometrics 67 (1), 27–44.

Lazega, E., Mounier, L., Stofer, R., Tripier, A., 2004. Discipline scientifique et discipline sociale: Réseaux de conseil, apprentissage collectif et innovation dans la recherche française sur le cancer (1997–1999). Recherches Sociologiques 35, 3–27.

Lomi, A., 2011. Social settings and social networks: models and ideas. Presentation at Applications of Social Network Analysis (ASNA), Zurich, Switzerland, September, 2011.

Lubbers, M.J., 2003. Group composition and network structure in school classes: a multilevel application of the p* model. Social Networks 25 (4), 309–332.

Lubbers, M.J., Snijders, T.A.B., 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.

Mason, W., Conrey, F., Smith, E., 2007. Situating social influence processes: dynamic, multidirectional flows of influence within social networks. Personality and Social Psychology Review 11 (3), 279–300.

Pattison, P.E., Robins, G.L., 2002. Neighborhood-based models for social networks. Social Methodology 32, 301L 337.

Pattison, P.E., Robins, G.L., 2004. Building models for social space: neighbourhood-based models for social networks and affiliation structures? Mathematics and Social Sciences 42 (168), 11–29.

Pattison, P.E., Robins, G.L., Snijders, T.A.B., 2009. A Hierarchy of Dependence Assumptions for Exponential Random Graph Models for Social Networks. Unpublished Manuscript. Melbourne School of Psychological Sciences, The University of Melbourne, Australia.

Pattison, P.E., Robins, G.L., Snijders, T.A.B., Wang, P., 2011. Conditional Estimation of Exponential Random Graph Models from Snowball Sampling Designs. Unpublished Manuscript. Melbourne School of Psychological Sciences, The University of Melbourne, Australia.

Pattison, P.E., Snijders, T.A.B., 2013. Modeling social networks: next steps. In: Lusher, D., Koskinen, J., Robins, G.L. (Eds.), Exponential Random Graph Models for Social Networks: Theories, Methods and Applications. Cambridge University Press, New York, pp. 287–301.

Pattison, P., Wasserman, S., 1999. Logit models and logistic regressions for social networks. II. Multivariate relationships. British Journal of Mathematical and Statistical Psychology 52, 169–193.

Rinaldo, A., Fienberg, S.E., Zhou, Y., 2009. On the geometry of discrete exponential random families with application to exponential random graph models. Electronic Journal of Statistics 3, 446–484.

Robins, G., Pattison, P., Kalish, Y., Lusher, D., 2007a. An introduction to exponential random graph (p*) models for social networks. Social Networks 29 (2), 173–191.

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

Robins, G., Lusher, D., 2013. Illustrations: simulation, estimation and goodness of fit. In: Lusher, D., Koskinen, J., Robins, G.L. (Eds.), Exponential Random Graph Models for Social Networks: Theories, Methods and Applications. Cambridge University Press, New York, pp. 167–185.

Robins, G., Pattison, P., 2006. Multiple Networks in Organizations. Technical Report. Melbourne School of Psychological Sciences, The University of Melbourne, Australia.

Robins, G., Pattison, P., Woolcock, J., 2005. Small and other worlds: global network structures from local processes. American Journal of Sociology 110 (4), 894–936.

Robins, G.L., Pattison, P.E., Wang, P., 2009. Closure, Connectivity and Degrees: New Specifications for Exponential Random Graph (p*) Models for Directed Social Networks, 31, pp. 105–117.

Robins, G.L., Alexander, M., 2004. Small worlds among interlocking directors: network structure and distance in bipartite graphs. Journal of Computational and Mathematical Organization Theory 10, 69–94, Invited paper.

Robins, G.L., Elliott, P., Pattison, P., 2001a. Network models for social selection processes. Social Networks 23, 1–30.

Robins, G.L., Pattison, P., Elliott, P., 2001b. Network models for social influence processes. Psychometrika 66, 161–190.

Skvoretz, J., Faust, K., 1999. Logit models for affiliation networks. In: Sobel, M., Becker, M. (Eds.), Sociological Methodology. Blackwell, New York, pp. 253–280.

Snijders, T.A., 2002. Markov chain Monte Carlo estimation of exponential random graph models. Journal of Social Structure 3, 2.

Snijders, T., Baerveldt, C., 2003. A multilevel network study of the effects of delinquent behavior on friendship evolution? Journal of Mathematical Sociology 27 (2–3), 123–151.

Snijders, T.A., 2009. Analyzing the joint dynamics of several networks. In: Sunbelt XXIX International Social Network Conference, USA, March.

Snijders, T., Van de Bunt, G., Steglich, C., 2010. Introduction to stochastic actor-based models for network dynamics? Social Networks 32 (1), 44–60.

Snijders., T.A., Bosker, R., 2012. Multilevel Analysis: An Introduction to Basic and Advanced Multilevel Modelling, 2nd ed. SAGE Publications, London.

Snijders, T.A., Lomi, A., Torlo, V., 2012. Multiplex dynamics of one-mode and two-mode networks, with an application to friendship and labor market preference. Social Networks, http://dx.doi.org/10.1016/j.socnet.2012.05.005.

Snijders, T.A., Pattison, P.E., Robins, G.L., Handcock, M., 2006. New specifications for exponential random graph models. Sociological Methodology 36, 99L 153.

Torlo, V., Lomi, A., Snijders, T., 2010. The co-evolution of social networks and students' affiliation preferences. In: Sunbelt XXX International Social Networks Conference, Italy, July 2010.

Wang, P., 2013. Exponential random graph model extensions: models for multiple networks and bipartite networks. In: Lusher, D., Koskinen, J., Robins, G.L. (Eds.), Exponential Random Graph Models for Social Networks: Theories, Methods and Applications. Cambridge University Press, New York, pp. 115–129.

Wang, P., Pattison, P.E., Robins, G.L., 2013. Exponential random graph model specifications for bipartite networks – a dependence hierarchy. Social Networks, http://dx.doi.org/10.1016/j.socnet.2011.12.004.

Wang, P., Robins, G.L., Pattison, P.E., 2006. PNet: a program for the simulation and estimation of exponential random graph models. Melbourne School of Psychological Sciences, The University of Melbourne, Australia.

Wang, P., Sharpe, K., Robins, G.L., Pattison, P.E., 2009. Exponential random graph (p*) models for affiliation networks. Social Networks 31, 12–25.

Wasserman, S., Pattison, P.E., 1996. Logit models and logistic regression for social networks. I. An introduction to Markov graphs and p*. Psychometrika 6 (3), 401–425.

Wasserman, S., Iacobucci, D., 1991. Statistical modelling of one-mode and two-mode networks: simultaneous analysis of graphs and bipartite graphs. British Journal of Mathematical and Statistical Psychology, 13–43.

Watts, D.J., 1999. Small Worlds: The Dynamics of Networks Between Order and Randomness. Princeton University Press, Princeton, NJ.