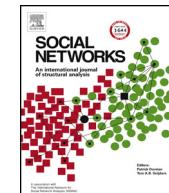




Contents lists available at ScienceDirect

Social Networks  
journal homepage: [www.elsevier.com/locate/socnet](http://www.elsevier.com/locate/socnet)



## Social selection models for multilevel networks

Peng Wang<sup>a,\*</sup>, Garry Robins<sup>b</sup>, Philippa Pattison<sup>c</sup>, Emmanuel Lazega<sup>d</sup>

<sup>a</sup> Centre for Transformative Innovation, Faculty of Business and Law, Swinburne University of Technology, Australia

<sup>b</sup> Melbourne School of Psychological Sciences, The University of Melbourne, Australia

<sup>c</sup> The University of Sydney, Australia

<sup>d</sup> Department of Sociology, CSO-CNRS, Science Po, Paris, France

### ARTICLE INFO

#### Keywords:

Social selection  
Exponential random graph models  
Multilevel networks  
Attribute effects

### ABSTRACT

Social selection models (SSMs) incorporate nodal attributes as explanatory covariates for modelling network ties (Robins et al., 2001). The underlying assumption is that the social processes represented by the graph configurations without attributes are not homogenous, and the network heterogeneity maybe captured by nodal level exogenous covariates. In this article, we propose SSMs for multilevel networks as extensions to exponential random graph models (ERGMs) for multilevel networks (Wang et al., 2013). We categorize the proposed model configurations by their similarities in interpretations arising from complex dependencies among ties within and across levels as well as the different types of nodal attributes. The features of the proposed models are illustrated using a network data set collected among French elite cancer researchers and their affiliated laboratories with attribute information about both researchers and laboratories (Lazega et al., 2006, 2008). Comparisons between the models with and without nodal attributes highlight the importance of attribute effects across levels, where the attributes of nodes at one level affect the network structure at the other level.

© 2015 Elsevier B.V. All rights reserved.

### 1. Introduction

Social network data are usually defined on a given number of nodes, their relational ties among each other, and information in regards to various measurements or attributes for the nodes. Network ties may form based on the interdependent nature of human social networks, that is, the presence of network ties may be dependent on the existence of other network ties (Frank and Strauss, 1986; Pattison and Robins, 2002, 2004; Snijders, 2011). For example, friends of friends may form a new friendship. Social network ties may also arise based on nodal attributes, that is, people may form social relationships based on certain individual characteristics. Research on how nodal attributes can impact on network structure is well established, and the idea of homophily, for instance, is well known in the literature (McPherson et al., 2001). For example, friendship may arise from common interests. The processes of tie formation based on nodal attributes are known as social selection processes (Robins et al., 2001), distinguishing them from social influence processes where there can be contagion or diffusion of attributes across network ties. Statistical models capable of capturing social selection processes

include the co-evolution stochastic actor-oriented models for network dynamics (Snijders et al., 2010a,b; Snijders, 2005, 2006), and Exponential Random Graph Models (ERGMs) for cross-sectional network analysis (Holland and Leinhardt, 1981; Frank and Strauss, 1986; Wasserman and Pattison, 1996; Snijders et al., 2006; Robins et al., 2007a).

ERGMs treat network ties as endogenous and model network structures using local network configurations that can be interpreted as the outcome of social structural processes. The underlying assumption is that network ties are interdependent, and the global network structure is a collective result of local tie interactions. Although cross-sectional social network data are defined only at one time point, interdependent ties imply self-organizing network processes, which can be interpreted as the result of ongoing social processes (Robins et al., 2007a; Snijders et al., 2010a,b; Lusher and Robins, 2013). Incorporating nodal attributes as explanatory covariates for modelling network ties, Robins et al. (2001) proposed Social Selection Models (SSMs) as extensions to ERGMs. As we often find the homogeneity assumption under ERGMs may be too strong, especially for large empirical networks, SSMs may be able to take network heterogeneity into account by using nodal attributes as exogenous covariates.

In the cases of statistical analysis of one-mode networks, various studies have shown that exogenous nodal attributes can shape the network structure together with the network tie

\* Corresponding author. Tel.: +61 392148230.  
E-mail address: [pengwang@swin.edu.au](mailto:pengwang@swin.edu.au) (P. Wang).

self-organizing processes (Lazega and Van Duijn, 1997; Robins et al., 2001; Goodreau et al., 2009; Lusher and Robins, 2010). For longitudinal network studies, the co-evolution models presented by Snijders et al. (2007a,b, 2010a,b) can investigate how network structure and actor attributes may affect each other when processes of selection and influence are both operating simultaneously.

The importance of nodal attributes in networks is further highlighted in the literature of two-mode network analysis. For two-mode or bipartite networks, ties are only defined between nodes of two different types, and the research focus is on the properties of the affiliation network structures and their inferred social processes. In the cases of affiliations between people and groups, Breiger (1974) presented the notion of *duality* reflected by bipartite network structures. Nodal attributes of both sets of nodes again may play important roles in the affiliation structure. For instance, in the case of interlocking directorate networks, companies with similar products or operating in similar fields (e.g. IT, or Finance) may be more likely to share directors; while directors from the same business schools may be more likely to sit on the same or share multiple company boards (Useem, 1984; Harrigan and Bond, 2013). Homophily in affiliation networks can then be seen as the tendency for nodes with similar attributes to have common affiliations. SSMs for bipartite networks were proposed by Agneessens and Roose (2008) and Wang (2013). For co-evolution models where both network structure and attribute values are modelled over time, Koskinen and Edling (2012) proposed models for bipartite networks. These models demonstrated the importance of attributes in affiliation network structures. Snijders et al. (2013) further proposed co-evolution models for one- and two-mode combined networks which share some similar model configurations proposed in this paper for ERGMs.

In the context of multilevel networks, Wang et al. (2013), following Wasserman and Iacobucci (1991), proposed a general data structure for multilevel networks, where one-mode network ties within levels form the micro- and macro-networks, and two-mode affiliation ties define the cross-level network. Note that we use the term *level* here because that is its common usage in empirical network context; the general data structure need not be construed as at different levels, but it does require two separate sets of nodes.

Lazega et al. (2008, 2013) developed theories and mechanisms of multilevel networks in organizational contexts. Using two-level networks of elite<sup>1</sup> French cancer researchers and their affiliated research laboratories, Lazega et al. (2008, 2013) proposed hypotheses on how multilevel network structure and various nodal attributes may affect or contribute to an individual's performance. These analyses used within-level network centralities and various nodal attributes to classify nodes. Actor performance was regressed against centralities and nodal attributes, suggesting possible network strategies individuals with different attributes may have used in the multilevel network.

Treating all of the within- and between-level networks as endogenous, Wang et al. (2013) proposed ERGMs for multilevel network structure. Using the same data set as in Lazega et al. (2008, 2013), Wang et al. (2013) demonstrated the interdependent nature of ties among the three networks, and highlighted how network ties at one level may affect the structure of the network at the other level. However, attributes of the researchers and of the laboratories were not taken into account. It is plausible that attributes such as geographical location of the laboratories, specialties, seniority and performance levels (as partial signals of status) of these research elites play important roles in structuring this complex system of individual and collective action, including both the advice network

among the researchers as well as the collaboration ties among laboratories.

In this paper, we propose social selection parameters to extend the original Wang et al. (2013) ERGM for multilevel networks. One of the most popular selection effects examined in one-mode networks is that of homophily, where an actor selects another based on a similar attribute. Homophily in a multilevel context may be understood in different perspectives. Firstly, similarities in nodal attribute values may increase the tendency to form within-level ties as well as between-level common affiliations, for example, researchers of the same specialties may be more active in seeking advice from one another and introduce each other to work in the same laboratory. Secondly, homophily in one level may encourage tie formations in the other level through between level ties, for example, advice seeking activities among researchers of the same specialties may encourage collaborations between their affiliated laboratories. The proposed SSMs in this paper may help to determine the extent homophily and other attribute processes may affect the within-level, between-level affiliation, as well as the overall multilevel network structure.

We start by reviewing the general forms of SSMs for one- and two-mode networks in Sections 2 and 3, and then introduce the extension to multilevel networks in Section 4. After reviewing some simple dyadic attribute interaction effects, Section 5 introduces various multilevel model specifications based on the types of nodal attributes, followed by their possible empirical interpretations. In Section 6 we make some comments about the model selection strategy we applied in the modelling example presented in Section 7. The modelling example uses the network data on the French elite cancer researchers and their affiliated laboratories (Lazega et al., 2008, 2013) with attribute information about both researchers and laboratories. We compare the fitted SSMs with the multilevel ERGMs without attributes presented in Wang et al. (2013), and show how different attributes affect the overall multilevel structure.

The proposed models are implemented in the MPNet software package as an extension to PNet (Wang et al., 2006) which uses Markov Chain Monte Carlo maximum likelihood estimation algorithm for parameter estimation, and simulation based strategies for model goodness of fit tests (Snijders, 2001).

## 2. Social selection models for one-mode networks

Let  $\mathbf{X}$  denote a network variable, and its realization as  $\mathbf{x}$ . For a one-mode directed network with  $n$  nodes, the graph sample space  $\Omega$  consists of  $2^{n(n-1)}$  possible graphs. The graph variable can be seen as a collection of network tie variables, that is, an adjacency matrix  $\mathbf{X} = \{X_{ij}\}$  where  $X_{ij} = 1$  if there is a tie between node  $i$  and node  $j$ , otherwise  $X_{ij} = 0$ . Let  $\mathbf{Y} = \{Y_i\}$  be a set of variables representing the attribute value for node  $i$ , where  $0 < i \leq n$ ; and  $\mathbf{y} = \{y_i\}$  be its realization. A general form of social selection models (SSMs) for one-mode networks can be expressed as

$$Pr(\mathbf{X} = \mathbf{x} | \mathbf{Y} = \mathbf{y}) = \frac{1}{k(\theta)} \exp \sum_{Q, \Lambda} \{\theta_Q Z_Q(\mathbf{x}) + \theta_\Lambda Z_\Lambda(\mathbf{x}, \mathbf{y})\}$$

where

- $Q$  is a network configuration of type  $Q$  comprising tie variables that are conditionally dependent given the rest of the network.
- $\Lambda$  is a joint attribute-network configuration comprising tie variables as well as nodal attribute variables.
- $Z_Q(\mathbf{x})$  are the sufficient statistics representing the network endogenous effects. They can be defined based on dependence assumptions among network tie variables. At their simplest they

<sup>1</sup> Having published eight articles per year in international journals for three years in a row (1996–1998).

have the following form but other functional forms are often used as described below.

$$Z_Q(\mathbf{x}) = \sum_{\mathbf{x}} \prod_{X_{ij} \in Q} X_{ij}$$

- $\theta_Q$  is the vector of parameters corresponding to the graph statistics without nodal attributes; we refer to them as *structural effects*, or the *structural part* of a SSM.
- $Z_A(\mathbf{x}, \mathbf{y})$  are the sufficient statistics for the interactions between the network and nodal attributes.
- $\theta_A$  is the vector of parameters corresponding to the graph statistics representing the interaction between network tie variables ( $X_{ij}$ ) and nodal attributes ( $Y_i$ ). We refer to them as *attribute effects*, or the *attribute part* of a SSM.
- $k(\theta)_Q$  is a normalizing constant defined based on the graph space and the actual model specification, that is,

$$k(\theta) = \sum_{\mathbf{x} \in \Omega} \exp \sum_{Q, A} \{\theta_Q Z_Q(\mathbf{x}) + \theta_A Z_A(\mathbf{x}, \mathbf{y})\}$$

For  $Z_Q(\mathbf{x})$ , dependence hypotheses include the Bernoulli (Erdős and Rényi, 1960; Holland and Leinhardt, 1981), the Markov (Frank and Strauss, 1986) and the social circuit assumption (Pattison and Robins, 2002, 2004; Snijders et al., 2006). Pattison et al. (2009) and Pattison and Snijders (2013) introduced a hierarchy of tie dependence summarizing the various assumptions in the ERGM literature, providing guidance for future ERGM specification development.

Nodal attribute variables are treated as fixed covariates in SSMs, and they can be binary, continuous or categorical. Binary attributes assign vectors of 1s and 0s to the nodes where “1” indicate nodes “have, or belong to” a particular attribute such as gender, or membership of a particular organization, etc. For binary attributes,  $Z_A(\mathbf{x}, \mathbf{y})$  often have the general form of

$$Z_A(\mathbf{x}, \mathbf{y}) = \sum_{\mathbf{x}} \prod_{X_{ij} \in A} X_{ij} \prod_{Y_k \in A} Y_k$$

Continuous attributes provide continuous measurements on the nodes, such as age, tenure, or performance measures, etc. Categorical attributes categorize nodes into a countable number of sets or groups by assigning integer IDs to categories, and labelling nodes using the corresponding IDs. The statistics to implement these effects may have different expressions depending on the type of attributes and the configurations involved, as instance below.

The derivation of graph statistics involving attributes is based on both the tie-variable dependence assumptions, as well as a variant of the Hammersley–Clifford theorem and moral graphs introduced by Robins et al. (2001). Robins et al. (2001) proposed SSM specifications for directed one-mode networks. After introducing the general form of multilevel SSMs, we review the Robins et al. (2001) specifications at the dyadic level, that is, interaction terms between attribute values and a single tie variable, as they are the baseline SSMs for the proposed extensions and are used in the modelling example in Section 7.

### 3. Social selection models for two-mode networks

For bipartite networks, ties are defined between two different set of nodes. We denote one set as  $A$  with  $u$  nodes, and the other as  $B$  with  $v$  nodes, the bipartite network variable  $\mathbf{X}$  can be represented by a rectangular adjacency matrices with  $u$  by  $v$  tie variable entries ( $X_{ij}$ 's). As there are two sets of nodes, we label the attributes of the two sets as  $\mathbf{Y} = \{\mathbf{Y}^A, \mathbf{Y}^B\}$  with their realizations  $\mathbf{y} = \{y^A, y^B\}$ . Bipartite

SSMs has the same general expression as the one-mode SSMs where the structural part of the model ( $Z_Q(\mathbf{x})$ ) can be specified based on the hierarchy of dependence assumptions (Pattison and Snijders, 2013; Wang et al., 2013). The attribute part of the model ( $Z_A(\mathbf{x}, \mathbf{y})$ ) can be seen as a collection of different interaction terms, that is,

$$Z_A(\mathbf{x}, \mathbf{y}) = \left\{ \begin{array}{l} Z_A(\mathbf{x}, \mathbf{y}^A) \\ Z_A(\mathbf{x}, \mathbf{y}^B) \\ Z_A(\mathbf{x}, \mathbf{y}^A, \mathbf{y}^B) \end{array} \right\}$$

$Z_A(\mathbf{x}, \mathbf{y}^A)$  and  $Z_A(\mathbf{x}, \mathbf{y}^B)$  are statistics representing the interactions between the bipartite network structure and the nodal attributes from one set of the nodes; whereas  $Z_A(\mathbf{x}, \mathbf{y}^A, \mathbf{y}^B)$  represents interactions between the network structure and attributes from both sets of the nodes. Agneessens and Roose (2008) and Wang (2013) proposed some bipartite SSM specifications. Similar configurations are also used in the bipartite co-evolution models of Koskinen and Edling (2012). After introducing SSMs for multilevel networks in the next section, we review some simple bipartite SSM specifications representing the interaction between attributes and dyads which we applied in our modelling example.

### 4. Social selection models for two-level networks

Wang et al. (2013) proposed ERGMs for two-level networks. The network variable ( $\mathbf{M}$ ) consists of the within-level one-mode network  $\mathbf{A}$  with  $u$  nodes, the within level one-mode network  $\mathbf{B}$  with  $v$  nodes, and the cross-level ( $u, v$ ) bipartite network  $\mathbf{X}$ , that is,  $\mathbf{M} = \{\mathbf{A}, \mathbf{X}, \mathbf{B}\}$  with the corresponding realizations denoted as  $\mathbf{m} = \{\mathbf{a}, \mathbf{x}, \mathbf{b}\}$ . Each of the network variables is a collection of tie variables either within level ( $\mathbf{A} = \{A_{ij}\}$  and  $\mathbf{B} = \{B_{ij}\}$ ) or cross-level ( $\mathbf{X} = \{X_{ij}\}$ ).<sup>2</sup> Let  $\mathbf{Y}$  denote the set of attributes for nodes of levels A and B, that is,  $\mathbf{Y} = \{\mathbf{Y}^A, \mathbf{Y}^B\}$  with their realizations  $\mathbf{y} = \{y^A, y^B\}$ . Similar to SSMs for one- and two-mode networks, we can express the general form of SSMs for two-level networks as

$$\Pr(\mathbf{M} = \mathbf{m} | \mathbf{Y} = \mathbf{y}) = \frac{1}{k(\theta)} \exp \sum_{Q, A} \{\theta_Q Z_Q(\mathbf{m}) + \theta_A Z_A(\mathbf{m}, \mathbf{y})\}$$

where

$$Z_Q(\mathbf{m}) = \left\{ \begin{array}{l} Z_Q(\mathbf{a}) \\ Z_Q(\mathbf{b}) \\ Z_Q(\mathbf{x}) \\ Z_Q(\mathbf{a}, \mathbf{x}) \\ Z_Q(\mathbf{b}, \mathbf{x}) \\ Z_Q(\mathbf{a}, \mathbf{x}, \mathbf{b}) \end{array} \right\} \quad Z_A(\mathbf{m}) = \left\{ \begin{array}{l} Z_A(\mathbf{a}, \mathbf{y}^A) \\ Z_A(\mathbf{b}, \mathbf{y}^A) \\ Z_A(\mathbf{x}, \mathbf{y}^A, \mathbf{y}^B) \\ Z_A(\mathbf{a}, \mathbf{x}, \mathbf{y}^A, \mathbf{y}^B) \\ Z_A(\mathbf{b}, \mathbf{x}, \mathbf{y}^A, \mathbf{y}^B) \\ Z_A(\mathbf{a}, \mathbf{x}, \mathbf{b}, \mathbf{y}^A, \mathbf{y}^B) \end{array} \right\}$$

$Z_Q(\mathbf{m})$  is the vector of graph statistics involving network tie variables where  $Z_Q(\mathbf{a})$  and  $Z_Q(\mathbf{b})$  may follow standard ERGM specifications for one-mode networks (e.g. Snijders et al., 2006);  $Z_Q(\mathbf{x})$  may follow standard bipartite ERGM specifications (e.g. Wang et al., 2009); and  $Z_Q(\mathbf{a}, \mathbf{x})$ ,  $Z_Q(\mathbf{b}, \mathbf{x})$ ,  $Z_Q(\mathbf{a}, \mathbf{x}, \mathbf{b})$  may follow the two-level ERGMs as in Wang et al. (2013).

Our goal in this paper is to propose additional model specifications that incorporate attribute variables.  $Z_Q(\mathbf{m}, \mathbf{y})$  is the vector of interaction effects between network tie variables and nodal

<sup>2</sup> As noted in the introduction, the use of the term “level” here is conventional, although in many applications, including our example, it will be appropriate to the context to construe the nodes as at different levels, akin to standard multilevel modelling. More generally, however, the data structure refers to two distinct node sets, but with ties within and between node sets that are modelled as three distinct types of tie (Wasserman and Iacobucci, 1991).

attribute covariates. The specifications for  $Z_{\Lambda}(\mathbf{a}, \mathbf{y}^A)$  and  $Z_{\Lambda}(\mathbf{b}, \mathbf{y}^B)$  may follow the one-mode SSMs proposed by Robins et al. (2001);  $Z_{\Lambda}(\mathbf{x}, \mathbf{y}^A, \mathbf{y}^B)$  may follow the proposed SSMs for bipartite networks listed in Wang (2013).

We begin by reviewing some dyadic attribute interaction effects from these papers because they are used in the modelling example. Our more important focus is on the specifications of the remaining interaction terms which have not been previously specified for these models:  $Z_{\Lambda}(\mathbf{a}, \mathbf{x}, \mathbf{y}^A, \mathbf{y}^B)$ ,  $Z_{\Lambda}(\mathbf{b}, \mathbf{x}, \mathbf{y}^A, \mathbf{y}^B)$  and  $Z_{\Lambda}(\mathbf{a}, \mathbf{x}, \mathbf{b}, \mathbf{y}^A, \mathbf{y}^B)$ .

Note that the terms involving both  $\mathbf{y}^A$  and  $\mathbf{y}^B$  may have one of them absent depending on the availability of the relevant data (for instance, attribute data may only be observed on one set of nodes). In fact, we can see  $Z_{\Lambda}(\mathbf{a}, \mathbf{x}, \mathbf{y}^A, \mathbf{y}^B)$  as a collection of  $Z_{\Lambda}(\mathbf{a}, \mathbf{x}, \mathbf{y}^A)$ ,  $Z_{\Lambda}(\mathbf{a}, \mathbf{x}, \mathbf{y}^B)$  and  $Z_{\Lambda}(\mathbf{a}, \mathbf{x}, \mathbf{y}^A, \mathbf{y}^B)$ . The same considerations apply to  $Z_{\Lambda}(\mathbf{b}, \mathbf{x}, \mathbf{y}^A, \mathbf{y}^B)$  and  $Z_{\Lambda}(\mathbf{a}, \mathbf{x}, \mathbf{b}, \mathbf{y}^A, \mathbf{y}^B)$ .

## 5. Social selection model specifications

We review and propose SSM specifications for both non-directed and directed one-mode and bipartite networks, and categorize the proposed model configurations based on similarities in their interpretations. The configurations representing different interaction effects are listed in the order of complexity, including

- Attribute activity,
- Dyadic interaction,
- Attribute based centralization effects,
- Degree assortativity and
- Cross-level entrainment and exchange.

After describing each configuration, we provide its possible interpretations based on our subsequent empirical example. Our goal with this paper is to provide a general approach to SSMs for multilevel networks, but we feel it may be helpful to motivate particular specifications with a concrete example, to illustrate how particular effects may be interpreted. Hence we draw on the empirical context that is analyzed in more depth later in the article. When we are using our empirical example to illustrate interpretations in that specific context, as distinct from a more general description of interpretation, we use *italic* font to avoid confusion.

For our empirical example, we use the same two-level network data as in Lazega et al. (2008, 2013) and Wang et al. (2013) about French cancer researchers and their laboratories. The data comprises a directed researcher advice seeking network at level  $B$ , and a directed laboratory collaboration network at level  $A$ . The researcher-laboratory cross-level network is labelled as  $X$ . We refer to this network as the *RL-network* (Researcher-Lab network).

In this section, possible interpretations of the proposed configurations are illustrated with a couple of selected attributes. For researchers, we use director status as an example binary attribute where “1” indicates the researcher is a director of a laboratory; research performance, as an example continuous attribute; and specialty as an example categorical attribute. For laboratories, we use location as an example binary attribute where “1” indicates whether the laboratory is based in the capital city; laboratory size as an example continuous attribute; and the research specialty of the laboratory as an example categorical attribute. We label the attribute name as “[Attr]” when introducing a configuration more generally, and replace it with the actual names of the attribute when it comes to precise examples. We use the following conventions in all figures (Figs. 1–10) in this section

- For binary attributes, the coloured nodes have attribute values as “1s”;

- For continuous attribute, the coloured nodes represent the attribute values (and where appropriate a larger size represents a greater value).
- For categorical attributes where effects are defined for nodes with matching or mismatching category indices, the comparisons are based on the categories of coloured nodes only.
- Functions defined on attribute values of dyads are labelled as “ $f$ ” whose functional forms depend on the types of attributes and the corresponding network effects as described below “- $f$ ” is also used as postfix for labels of proposed configurations.

Selected sufficient statistics for the configurations described below are listed in Appendix A.

### 5.1. Attribute activity effects

Network activity effects based on nodal attributes can be applied to test whether nodes with a particular binary attribute, or having a greater value in a continuous attribute, are more “active” (tend to have higher degrees) within a given one-mode network (i.e. within level), or bipartite network (i.e. between-level). In a directed network, out-degree is typically interpreted as *network activity* or *expansiveness*; with in-degree interpreted as *popularity*.

Fig. 1 presents some possible [Attr]-activity effects for both binary and continuous attributes. Such effects have been proposed earlier (e.g. Robins et al., 2001; Wang, 2013) but are included here because they are fundamental effects in this class of models, and are important in our empirical example. For the reason of simplicity, we only present configurations at one level ( $B$ ) represented by circles for within level networks; and attributes for nodes of type  $B$  for between-level affiliation networks. The sufficient statistics for these effects are presented in Wang (2013).

The interpretation of these effects is quite straightforward: a positive effect suggests that actors with (more of) the attribute tend to have more of the given ties (i.e. outwards, inwards, or mutual in the directed unipartite case).

*In the RL-network context:*

- a positive Director-sender effect (Fig. 1(b)) suggests directors seek more advice from other researchers.
- a positive Performance-receiver effect (Fig. 1(c)) suggests researchers tend to seek more advice from researchers with higher performance scores.
- a positive Capital City-reciprocity (Fig. 1(d)) effect suggests that researchers work in the capital tend to exchange advice more than researchers from other locations.

For the cross-level bipartite networks ( $x$ ), similar interpretations apply, except that they relate specifically to cross-level ties.

*In the RL-network context:*

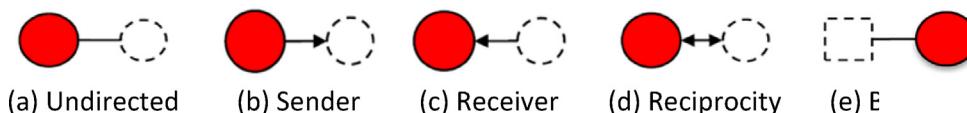
- a positive cross-level Director-activity effect (Fig. 1(e)) suggests directors are more likely to affiliate with more laboratories;

### 5.2. Dyadic interaction effects

Dyadic interaction effects involve attribute values from both nodes in a dyad. They have the following general form (using an undirected network  $B$  as an example)

$$Z_{[Attr]-f}(\mathbf{b}, \mathbf{y}^B) = \sum_{i=1}^v \sum_{j=1}^v B_{ij} f(Y_i^B, Y_j^B)$$

where  $f(Y_i^B, Y_j^B)$  may have different functional forms depending on the types of attributes.



**Fig. 1.** Attribute activity effects.

- For binary attributes, the interaction represents a homophily effect for nodes having attribute values as "1"s, and  $f(Y_i^B, Y_j^B) = Y_i^B, Y_j^B$ .
- For categorical attributes, the graph statistics can be defined based on whether the nodes involved within the dyad have the matching or mismatching category indices representing homophily or heterophily respectively. In the case of homophily,  $f(Y_i^B, Y_j^B)$  is an indicator function where  $f(Y_i^B, Y_j^B) = 1$ , if  $Y_i^B = Y_j^B$ , 0 otherwise.
- For continuous attributes, the most common functions of  $f(Y_i^B, Y_j^B)$  include sum, product and absolute difference. A positive sum effect indicates that nodes with higher total attribute values tend to form more ties with each other. The product function represents the interactions of the two attribute values. The absolute difference represents the heterophily among nodes.

**Fig. 2** illustrates possible dyadic interaction configurations in non-directed and directed one-mode networks (within level B), where " $f$ " indicates different functional forms.

In the RL-network context,

- a positive Director-interaction effect suggests directors are more likely to seek advice from other directors;
- a positive Specialty-match effect indicates advice seeking is more likely to take place among researchers of the same specialty.

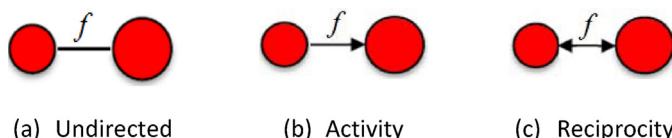
In the RL-network context,

- at the researcher level, a negative Performance-difference effect suggests researchers tend to seek advice from others with similar performance scores;
- at the laboratory level, a positive Size-sum effect suggests collaboration tends to happen more between larger laboratories.

These attribute activity and dyadic interaction configurations serve as important baseline effects in most SSMs in the literature (e.g. Robins et al., 2001; de la Haye et al., 2010; Wang, 2013; Koskinen and Edling, 2012) as well as in our modelling example. Now we move to attribute effects that are specific to the multilevel structure.

### 5.3. Attribute based centralization effects

Attribute based centralization effects can be represented as interactions between nodal attributes and star-like graph configurations, such as the Markov two-stars (Frank and Strauss, 1986), as the number of two-stars is associated with the standard deviation of the degree distribution. For one-mode networks, Robins et al. (2001) proposed some attribute based centrality one-mode configurations, and Agneessens and Roose (2008) and Wang (2013)



**Fig. 2.** Dyadic interaction effects (Note: " $f$ " may have different functional forms).

proposed attribute based popularity effects for bipartite networks. They can be used for within-level and cross-level SSM specifications.

These earlier effects can be included in models but here we propose cross-level effects that involve the interactions of nodal attributes and at least two different types of tie variables from the within-level and the cross-level networks.

**Fig. 3** shows some possible multilevel attribute centralization effects for binary and continuous attributes for both nondirected and directed networks. They can also be seen as representing an association between the within- and cross-level activities of actors.

Positive [Attr]-star2AX effects suggest nodes with attribute values "1"s for binary attributes, or higher continuous attribute values, are popular in both within and cross-level networks.

In the RL-network context,

- a positive CapitalCity-inStar2AX effect (**Fig. 3(c)**) suggests capital city laboratories with more researchers attract more collaboration. In short, a correlation between size (in terms of the number of researchers) and collaborations is stronger for capital city laboratories.
- a positive Performance-inStar2BX effect (**Fig. 3(f)**) indicates that the correlation between advice seeking in-degree and multiple laboratory affiliation is stronger for high performing researchers. For instance, high performing researchers affiliated with multiple laboratories may attract more advice-seeking ties.

### 5.4. Cross level triadic closures

Multilevel network closure can be represented by cross-level triangles that involve both within- and cross-level ties. Within the cross-level bipartite network (X), triangles are not possible by construction. However, the bipartite two-paths can be closed by within level ties, and we refer to the effects represented by such closures as cross level triadic closure effects. The corresponding configurations are labelled as TXAX or TXBX where T stands for Triangle, and XAX/XBX denote the types of ties involved in an upright triangle in the left to right order. This labelling convention follows Wang et al. (2013) and some bivariate ERGMs as in Huiszing et al. (2012) and Wang (2013).

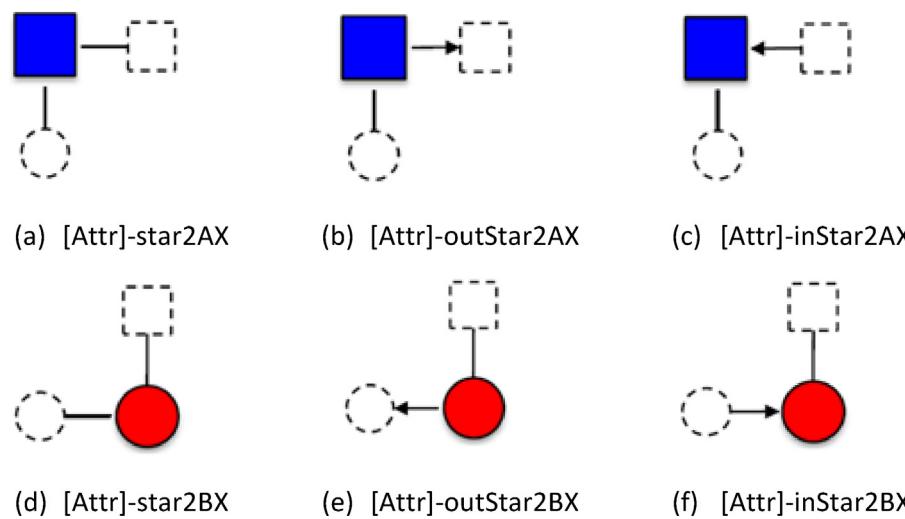
We can include effects that test whether nodes with a particular attribute are more likely to be involved in multilevel closure. There are different possibilities, but here we present configurations where the focal actor (i.e. with the attribute) is involved in two cross-level ties that are then closed at the other level.

We can define configurations such as [Attr<sup>A</sup>]-TXBX or [Attr<sup>B</sup>]-TXAX for both binary and continuous attributes as shown in **Fig. 4**.

The [Attr<sup>A</sup>]-TXBX represents the tendency for B-partners of a focal A-node (with the corresponding attribute) to be connected. Note that for these configurations ([Attr<sup>A</sup>]-TXBX and [Attr<sup>B</sup>]-TXAX), the attribute in question is at one level and the tie that closes the X-ties is at the other level.

In the RL-network context

- a positive Size-TXBX effect (**Fig. 4(b)**) indicates researchers within larger laboratories are more likely to seek advice from their colleagues within the same laboratory.



**Fig. 3.** Attribute based centralization effects (two-stars).

- a positive Director-TXAX effect (Fig. 4(d)) indicates laboratories with common directors tend to collaborate with one another.

Additionally, we can derive configurations representing various forms of interactions between the attributes of two nodes and the cross level triadic closures. These attributes could relate to dyads within the one level or across both levels.

Fig. 5 presents some possible configurations representing the interactions between the dyadic attributes within one level and the cross level triadic closure. As before, “-f” can have different functional forms depending on the types of attributes. For example, for categorical attributes “-f” can be “-match” or “-mismatch”; for continuous attributes “-f” can be “-sum”, “-product”, or “-difference”. Here, the attributes relate to a pair of connected nodes within the one level involved in the cross-level triangulation. In these cases, the attributes and the tie that completes the closure are all present within the one level.

Binary attribute closure configurations represent attribute homophily associated with cross-level closures.

In the RL-network context,

- a positive CapitalCity-TXAX effect (Fig. 5(b)) indicates that laboratories based in the capital city are more likely to collaborate with one another when a researcher is affiliated to both of them.

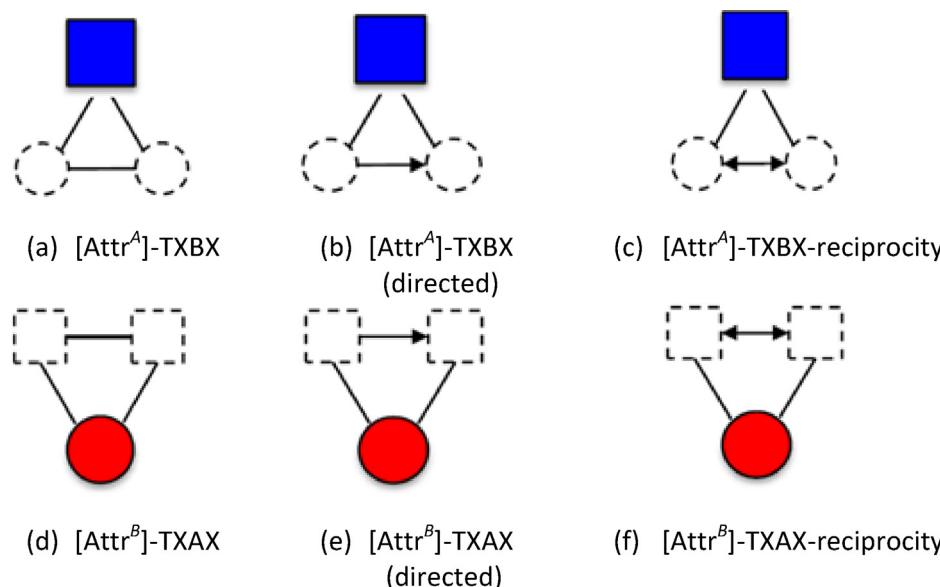
For categorical attributes, “f” is an indicator function on category indices, and we may have match and mismatch attributes for the two nodes in the within level dyads.

- A positive Specialty-TXBX-reciprocity-match (Fig. 5(f)) effect suggests researchers within the same laboratory and of the same specialty are more likely to seek advice from each other.

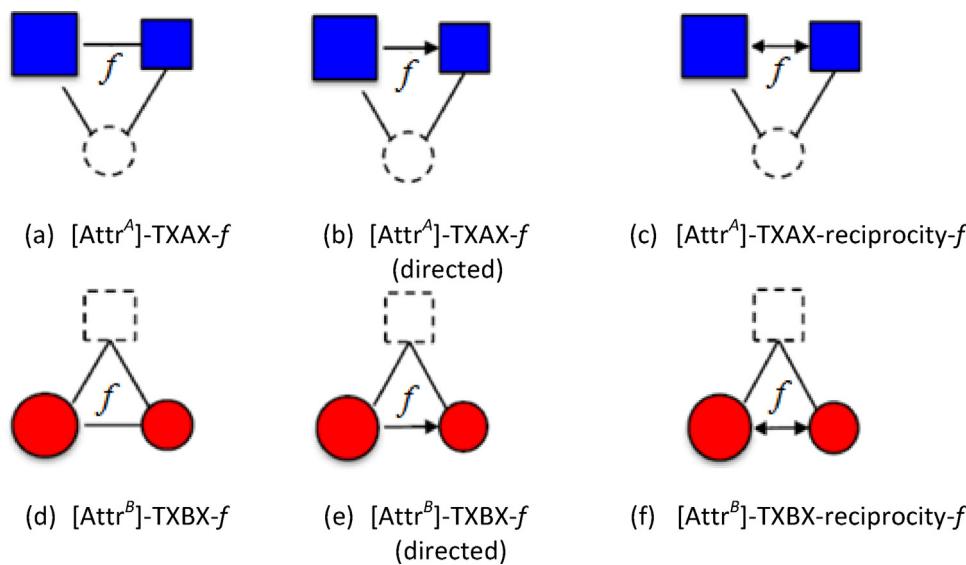
For continuous attributes, “f” may have the forms of sum, absolute difference or product.

In the RL-network context,

- a positive Performance-TXBX-sum effect (Fig. 5(e)) suggests that high performance researchers within the same laboratory are likely to seek advice from one another;



**Fig. 4.** Interaction between nodal attributes and cross level triadic closure ([Attr<sup>A</sup>]-TXBX and [Attr<sup>B</sup>]-TXAX).



**Fig. 5.** Interaction between dyadic attributes and cross level triadic closure (Note: “*f*” may have different functional forms).

- a negative Size-TXAX-difference effect (Fig. 5(b)) suggests that laboratories of similar sizes tend to form more collaboration when they share a researcher.

When the cross-level closure effects involve attributes of nodes at different levels, both the expressions of the statistics and the interpretations are very different from the previous effects. Fig. 6 shows configurations for the interactions of attributes of nodes across levels and the cross level triadic closure, we label them as  $[Attr^A]-[Attr^B]$ -TXAX/TAXB followed by the function on the attribute values “*f*”. We use  $[Attr^A]-[Attr^B]$ -TXAX-*f* (Fig. 6(a)) in all related examples below.

In some special cases, the attributes for nodes of types A and B can be the same depending on the context.

- For instance, in the cases of binary attributes, if both the laboratories and researchers are categorized by the same set of specialties, then we can test whether researchers and laboratories of a particular specialty represented by a binary dummy variable tend to form cross-level triadic closures.

The same applies to categorical attributes which represent a more generalized form of homophily (Match) or heterophily (Mismatch).

- For example, using the full categorization of specialties for both researchers and laboratories (instead of focusing on one category through a dummy variable), a positive Specialty-Specialty-mismatch-TXAX effect indicates researchers and laboratories of different specialties form more cross level triadic closures.

The attributes of the nodes from two different levels can also be different, and the interpretations will be more specific.

- For example, a positive CapitalCity-Director-TXAX effect suggests that directors in capital city laboratories are more likely to be affiliated with other laboratories that have existing collaborations with the capital city laboratories.

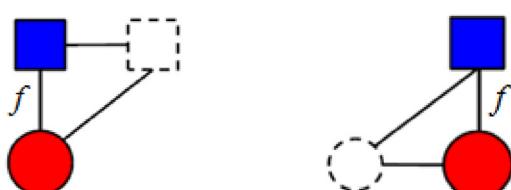
For categorical and continuous attributes, such configurations usually require a common or compatible attribute defined on nodes from both levels. For example, it might not make much sense to include a Size-Performance-TXAX-sum parameter in a model when size and performance are measured on different scales. However, one can consider the interaction among a binary attribute from one level, a continuous attribute from the other level and the cross level triadic closure.

- For example, a positive CapitalCity-Performance-TXAX-product effect suggests high performance researchers in capital city laboratories are more likely to be affiliated with other collaborating laboratories.

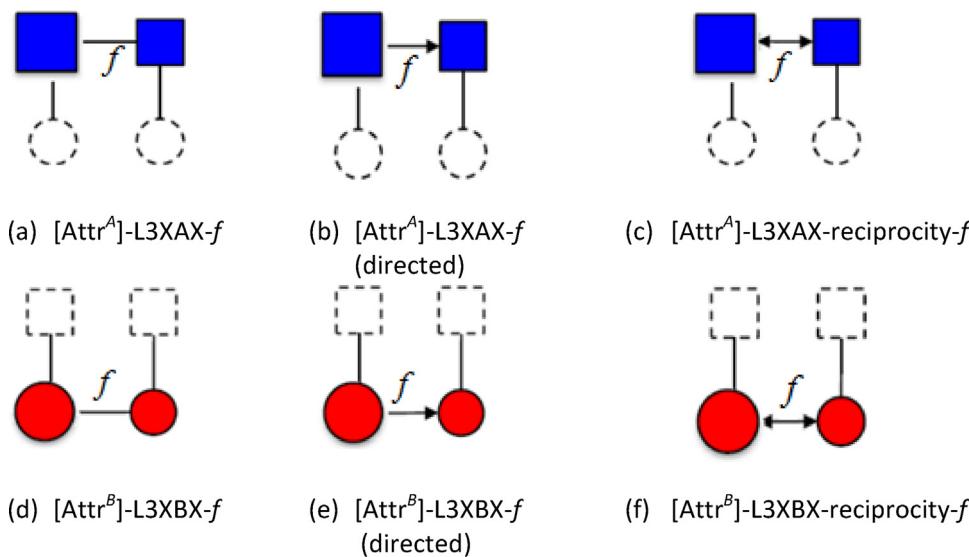
### 5.5. Degree assortativity effects

Degree assortativity effects are represented by cross-level three-path configurations of various forms. They can be interpreted as the tendency for popular or active nodes to themselves be connected. They also serve as preconditions for four-cycle closures. There are two forms of interactions among the within and cross-level networks depending on the number of different types of tie variables involved. Fig. 7 presents some possible three-paths involving the interaction between one of the within-level networks and the cross-level network, whereas Fig. 8 shows possible three-paths involve both within-level networks, as well as the cross-level network.

When the interactions only involve tie variables from the cross-level and one of the within level networks, the binary  $[Attr^A]$ -L3XAX or  $[Attr^B]$ -L3XBX parameters test the tendency for nodes with the same attribute and are active in the cross-network to connect to each other. We use  $[Attr^B]$ -L3XBX-*f* (Fig. 7 (e)) in the following examples



**Fig. 6.** Interaction between nodal attributes and cross level triadic closure (Note: “*f*” may have different functional forms).



**Fig. 7.** Degree assortativity effects (Note: "f" may have different functional forms).

In the RL-network context,

- a positive Director-L3XBX effect may indicate that directors that are affiliated with more laboratories are more likely to seek advice from one another.

For interactions between categorical attributes and the three-paths, we can test whether active affiliating nodes with matching or mismatching categories tend to connect to each other.

- For example, a positive Specialty-L3XBX-match effect suggests researchers that are active in the affiliation network and with the same specialty tend to seek advice from each other (in this context, active in the affiliation network means researchers affiliated to more laboratories).

For continuous attributes, using absolute difference between attribute values as an example,

- a negative Performance-L3XBX-difference effect suggests that researchers active in the affiliation network are more likely to seek advice from each other if they have similar performance scores.

When the three-path configuration involve ties from the cross-level network as well as both within level networks, we can define [Attr<sup>A</sup>]-[Attr<sup>B</sup>]-L3AXB-f configurations representing the tendency for popular within level nodes to form affiliation ties. Fig. 8 shows some possible such configurations for both undirected and directed within level networks.

In directed within level networks, L3AXBs represent assortativities of various combinations of nodal in- and out-degrees. As shown in Fig. 8(b)–(e), different effect labels are used to reflect the differences. For example, L3AXBout (Fig. 8(b)) represents the out-degree assortativity; L3BXApath (Fig. 8(e)) represents the assortativity between the out-degree of A nodes and in-degree of B nodes following the path in the order of B-X-A.

As previously discussed, we can also have the same or different types of attributes for nodes at the two different levels.

In the RL-network context,

- when a common set of specialties are defined on both researchers and laboratories, with a binary dummy variable for a given specialty, a

positive Specialty-Specialty-L3AXBin effect (Fig. 8(c)) suggests that for that specialty, laboratories popular in the collaboration network also have popular researchers in the advice seeking network.

- When treating the specialty attribute as categorical attribute, a positive Specialty-Specialty-L3AXBmatch effect gives a similar interpretation but this time across all specialties.
- In the case where the attributes [Attr<sup>A</sup>] and [Attr<sup>B</sup>] are of different types, a positive CapitalCity-Performance-L3AXBpath effect (Fig. 8(d)) suggests popular capital city laboratories have high performing researchers who seek more advice from other researchers.

### 5.6. Cross-level alignment

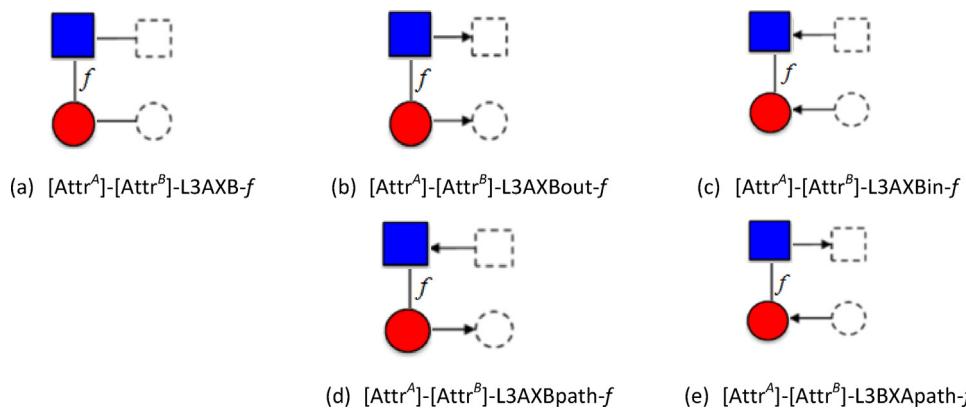
The degree assortativity effects define the preconditions for cross level four-cycle closures. The cross level four-cycle configurations generally represent a cross-level alignment effect where ties from the two different levels are aligned based on the cross-level affiliations of the nodes involved (see "tetradic" configurations in Lazega et al., 2013). By adding a tie (either within-level or cross-level) to the degree assortativity configuration, we can derive the four-cycle configurations of various forms. For nondirected networks, we label the cross level four-cycles as C4AXB; and for directed networks, we can define C4AXB-entrainment, -exchange, or -reciprocity, based on the directions of within level ties. Configurations representing their interactions with within-level attributes are shown in Fig. 9.

For the interactions between the binary within-level attributes and the cross-level four-cycles, they represent the tendency for nodes with a particular attribute to form alignment ties given that their affiliated nodes at the other level are also connected.

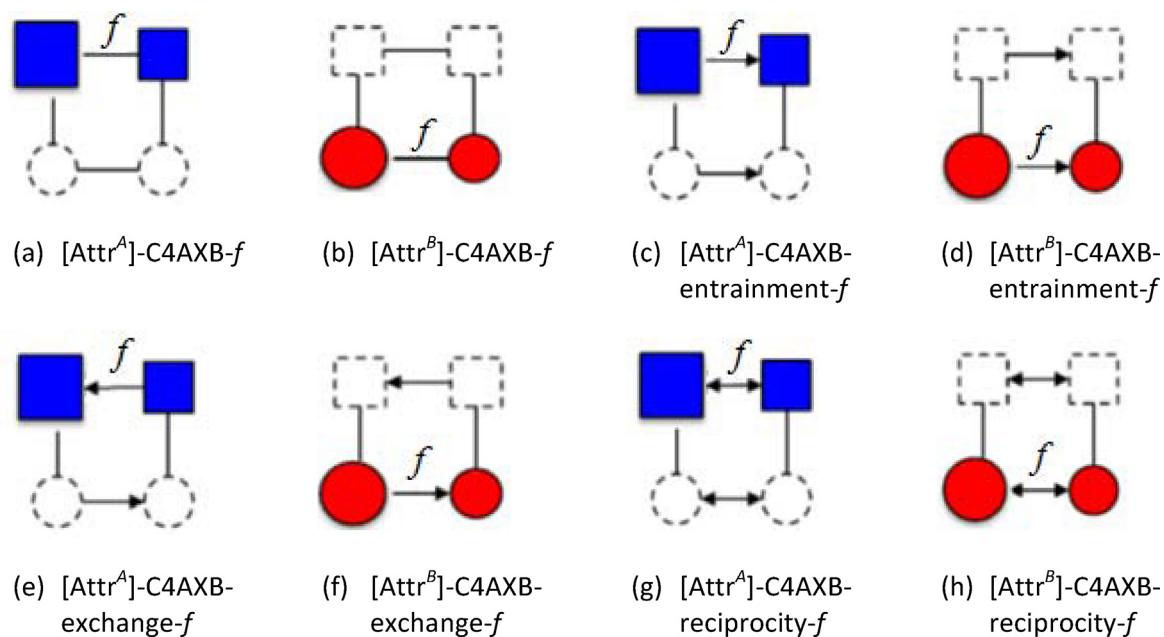
In the RL-network context,

- a positive Director-C4AXB-entrainment effect (Fig. 9(d)) suggests that directors are more likely to seek advice from directors in collaborating laboratories;
- a positive Director-C4AXB-exchange effect (Fig. 9(f)) on the other hand suggests that directors are more likely to ask advice from directors whose affiliated laboratory is seeking collaboration with the advice seeker's laboratory.

Categorical cross-level alignment effects represent the interaction between cross-level four-cycles and within-level homophily



**Fig. 8.** Cross level degree assortativity effects (Note: "f" may have different functional forms).



**Fig. 9.** Binary Cross level alignment effect. (Note: "f" may have different functional forms).

or heterophily. The expressions for these statistics are similar to the binary configurations except we need to replace the attribute functions with indicator functions.

In the RL-network context,

- a positive Specialty-C4AXB-entrainment-mismatch effect (Fig. 9(d)) suggests that researchers are more likely to seek advice from researchers in collaborating laboratories with different specialties.

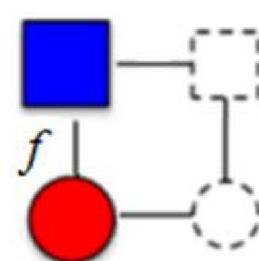
For continuous attributes, different functional forms may apply to the attribute values of the pair of nodes involved. Using [Attr<sup>A</sup>]-C4AXB-reciprocity-difference as an example in the RL-network context,

- a negative Performance-C4AXB-reciprocity-difference effect (Fig. 9(h)) indicates researchers with similar performance scores are more likely to seek advice from each other if they are affiliated with collaborating laboratories.

The interaction effects may involve attributes from both levels as shown in Fig. 10 for non-directed networks.

In the RL-network context,

- when we treat a particular specialty as a dummy variable, then a positive binary Specialty-Specialty-C4AXB effect (Fig. 10) may indicate laboratories and their affiliated researchers with that specialty are more likely to collaborate with other laboratories and seek advice from their researchers. When we treat the specialty as a categorical attribute, the same interpretation is generalized to all specialties.



**Fig. 10.** Cross level alignment effect ([Attr<sup>A</sup>]-[Attr<sup>B</sup>]-C4AXBs-f) (Note: "f" may have different functional forms).

## 6. Model specification

The complexity involved in the dependencies among ties within and across levels as well as the different types of nodal attributes implies a rather large number of possible configurations. Although we consider that the configurations listed in section 5 are the most important, they are only a subset of possible combinations of cross-level configurations and nodal attributes. This presents a challenge for model fitting and selecting a good specification in practice. For ERGMs, formal model selection criteria have not been developed, as model convergence is subject to the empirical network structure and the model specification, that is, not all model specifications will converge for a given network dataset. The use of indices common in other areas of statistics (AIC, BIC, Mahalanobis distance) is possible but somewhat uncertain as the correct degrees of freedom are not known. More work needs to be done in this area for ERGMs in general.

In this section we want to describe the stepwise model selection strategies we used, as a practical guide to other researchers. The aim is to find parsimonious model specifications that provide adequate fit to the observed network data.

First of all, as for all statistical models, the research question, theory and a good understanding of how they may apply to the empirical network context can provide good guidance on what should be included in a model. For example, in the RL-network context, it is reasonable to hypothesize that researchers and laboratories based in the same locations are more likely to seek advice or collaborate with one another, so the Location-Match parameter may be included in the model.

Secondly, simulation studies of the various model specifications can provide good insight on how simulated graph structure changes by changing parameter values, and how we may interpret the corresponding effects. Robins et al. (2005), Robins and Lusher (2013) and Wang et al. (2009, 2013) illustrate simulation studies on ERGMs for one-mode, two-mode and multilevel networks. Simulation studies on the proposed attribute configurations in this paper may also help interpret SSMs for multilevel network data.

Thirdly, as model degeneracy remains an issue for ERGMs especially for models with more complex configurations, some standard “baseline” model specifications in the ERGM literature may be good starting points for model development. These include the dyadic independent models (Erdős and Rényi, 1960; Holland and Leinhardt, 1981) with attributes at the dyadic level; and model specifications following the social circuit assumptions (Snijders et al., 2006; Robins et al., 2009; Wang et al., 2009) without attribute covariates. The reasons and features of these two models as good starting models are described below.

1. The dyadic independent models assume all tie-variables are independent; hence model convergence is guaranteed (except for some very special cases which are unlikely in empirical practice). Fitting a dyadic independent model with nodal attributes can reveal important covariates at the dyadic level at least. Non-significant attribute effects in dyadic models are less likely to be important in models with structural parameters, hence may be eliminated from the model specification. This is particularly helpful when a lot of attributes are available. Once we determine which attributes to include in a model, endogenous structural configurations can then be added, including the social circuit model specifications. Even so, the discarded attributes may still be important in higher-order interactions. Model goodness of fit (GOF) tests, as discussed below, may help decide whether they should be included in models with structural configurations.
2. The social circuit models include effects that are often observed in human social networks, including effects for density

(Edge or Arc), reciprocity if directed, degree centralization (Alternating-stars), network closure (Alternating-triangles or Alternating-two-paths for two-mode networks) and its precondition (Alternating-two-paths). See Robins et al. (2009) and Wang et al. (2013) for some example social circuit model configurations that apply to our modelling example. Social circuit models are known to alleviate model degeneracy issues while capturing important network features. Attribute covariates and across level network interactions may then be added to the model based on theory and model GOF tests.

We start modelling with both of these baseline models, and run parallel estimation sessions. The dyadic independent models help simplify covariate selection, while the social circuit models capture some essential social processes with good chance of obtaining model convergence. Once we obtain converged baseline models, we may add configurations representing interactions between one- and two-mode networks (A and X, or B and X) and the interaction effects between the attributes and the two networks first, then move on to effects representing interactions across the two levels (A, X, and B) and their higher-order interactions with attributes. We follow model GOF test based model selection strategies described below to add these higher-order configurations to the model specification.

Fourthly, model goodness of fit (GOF) tests may assist in fine tuning the model specification. Once a starting model is converged, a GOF test may reveal observed graph statistics with or without attributes that may be considered as extreme compare to the graph distribution simulated from the starting model. These statistics have t-ratios greater than 2.0 in absolute values (Snijders, 2001; Hunter et al., 2008; Robins and Lusher, 2013).

For extreme statistics related to nodal degrees or degree distributions (e.g. stars of various sizes, standard deviation and skewness of the degree distributions), we recommend adding related Markov two-stars (two-in-stars, two-out-stars, or two-paths) to the model, or fine tuning the lambda values used in the Alternating-star parameters (see Wang et al. (2013) for some simulation studies on the impact of varying lambda values on Alternating-stars). Interaction effects between stars and attributes such as the ones described in Robins et al. (2001), Agneessens and Roose (2008) and Wang (2013), and the ones shown in Fig. 3, may also help model fits on degree distributions, especially when activities of nodes may be strongly related to their attribute values, and the degree distributions have long-tails. For example, in a research advice seeking context, high performance researchers may be considered as good sources of advice, and we may consider adding an interaction effect between performance and a Markov two-in-star to models having difficulty fitting the in-degree distribution.

For extreme statistics relate to closure (e.g. triangles, triad census or clustering coefficients), we recommend adding related forms of Alternating-triangles and Alternating-two-paths (see Robins et al. (2009) for possible forms of Alternating-triangles/-two-paths and their interpretations), or fine tuning the lambda values applied to the Alternating-triangles/-two-paths parameters that are already included in the model. (see Robins and Lusher (2013) for a simulation study of alternating-triangles with different lambda values). Adding higher-order interactions with attributes (such as configurations shown in Figs. 4–10) may also provide better fits if structural parameters alone are not sufficient.

Note that direct inclusions of all extreme configurations based on a GOF test in a subsequent model is not recommended, as network statistics are highly correlated, and some configurations are strongly related to model degeneracy (e.g. the Markov triangle parameters for large networks). Based on experience, inclusion of relatively lower-order configurations (e.g. a Markov two-path) may help model fit to higher-order configurations (e.g. different forms

of directed Markov triangles). We also find inclusion of lower-order configurations with the current GOF t-ratios between 1 and 2 better candidates than higher-order configurations with t-ratios greater than 2. We can apply this strategy iteratively until the converged model provides adequate fit to most of the graph statistics included in the GOF test. Of course, it is not necessary for a model to fit every graph feature (although the more the better), just as it is not necessary for a regression to explain close to 100% of the variance. It is a theoretical decision for the researcher to decide the important structures that must be fit well and those that may be left to one side. In our case, our final model does in fact plausibly reproduce all of the graph statistics we examined, as will be seen below.

Finally, GOF can also assist model reduction through a backward selection strategy. Once a converged model provides adequate fit to the graph statistics in the GOF test including attribute configurations, we may consider removing some of the non-significant or close to zero effects provided that the reduced model can also converge and provide similar fit in the GOF test. We applied this strategy in reaching our final model in the modelling example (parameter reduction from Model 3 to Model 4).

The next section uses some of the proposed two-level SSM specifications on the data set of French cancer research elites and their affiliated laboratories (Lazega et al., 2004, 2006, 2008) as extensions to the multilevel ERGMs proposed in Wang et al. (2013). We compare and discuss four different model speculations that demonstrate the model selection strategies described above.

## 7. Modelling example

### 7.1. Attribute data

As described in Wang et al. (2013), the two level data set of French cancer research elites and their affiliated laboratories contains formal collaboration ties among 82 laboratories, advice seeking ties among 97 researchers, and the researcher-laboratory affiliation ties (Lazega et al., 2006, 2008). Younger and lower level researchers more or less adapt their behaviours and relational choices to the official structure but also manage to cut across constraining formal boundaries to create their own networks of opportunities and exchanges for the development of their own career. In addition to the information about their within level connections and affiliations, the data set also has several attributes about both laboratories and researchers.

Attributes for laboratories:

#### **Paris (Location of laboratories)**

A binary attribute where 1 indicates the laboratory is located in the Paris region and 0 for other parts of France, the so-called Province. Thirty-six out of the 82 laboratories are in Paris, often in well-known institutions that are quite close to each other geographically.

#### **Size**

A continuous attribute for the number of researchers within each of the laboratories. The laboratory size attribute has a range of 4–100, an average size of 29 and a standard deviation of 22.16.

Attributes for researchers:

#### **Paris (Location of researchers)**

A binary attribute where 1 indicates the researcher is based in the Paris region, and 0 for other parts of France (the same attribute applies for laboratories). There are 45 out of the 97 researchers located in Paris.

#### **Director status**

A binary attribute where 1 indicates the researcher is director of the laboratory in which s/he is affiliated (and as such also the respondent to the questionnaire on inter-organizational networks in this system), 0 otherwise.

### **Age of researchers**

A continuous attribute for the age of researchers with range 31–67, average age 48.21 and standard deviation 7.76. Especially for this population, this indicator of seniority is also considered an indicator of status.

### **Performance levels**

A continuous attribute measuring research performance based on the average impact factors associated with publications over the period of five years from 2000 to 2004. It has a range of 1.06–212.0, an average score of 39.0, and a standard deviation of 28.5.

### **Specialties**

A categorical attribute describes each researcher's specialty. There are seven specialties: Diagnostic-prevention-epidemiology, Surgery/radiology, Haematology, Solid tumours/chimio, Pharmacology, Molecular/Cellular and Molecular/Genetic.

Wang et al. (2013) presented visualizations of the network as well as some statistics for the degree distributions. We add selected attribute information and present the networks in Fig. 11. Fig. 11(a) shows the laboratory collaborations where red squares are laboratories based in pairs, and blue squares are based on other locations. The sizes of the squares represent sizes of laboratories. Fig. 11(b) presents researcher advice network where red circles are researchers based in Paris. The sizes of the circles represent their performance over the 5 years (2000–2004). Fig. 11(c) shows the researcher-laboratory affiliations, and Fig. 11(d) shows the overall multilevel network where laboratories are shown as blue squares, director researchers are shown as black circles, and other researchers are shown as red circles. These visualizations are plotted by Pajek software (de Nooy et al., 2005). Besides the rather simple affiliation structure, it is not obvious how the nodal attributes may affect the network structure. The interpretations from the multilevel ERGMs presented in the next section provide more detailed descriptions of the network structure and enable inferences about underlying network processes.

### 7.2. Model comparisons and interpretations

We have fitted models following various model specifications. Here we compare four models that reflect the model selection processes to reach a final model that fits most graph statistics adequately. These models include:

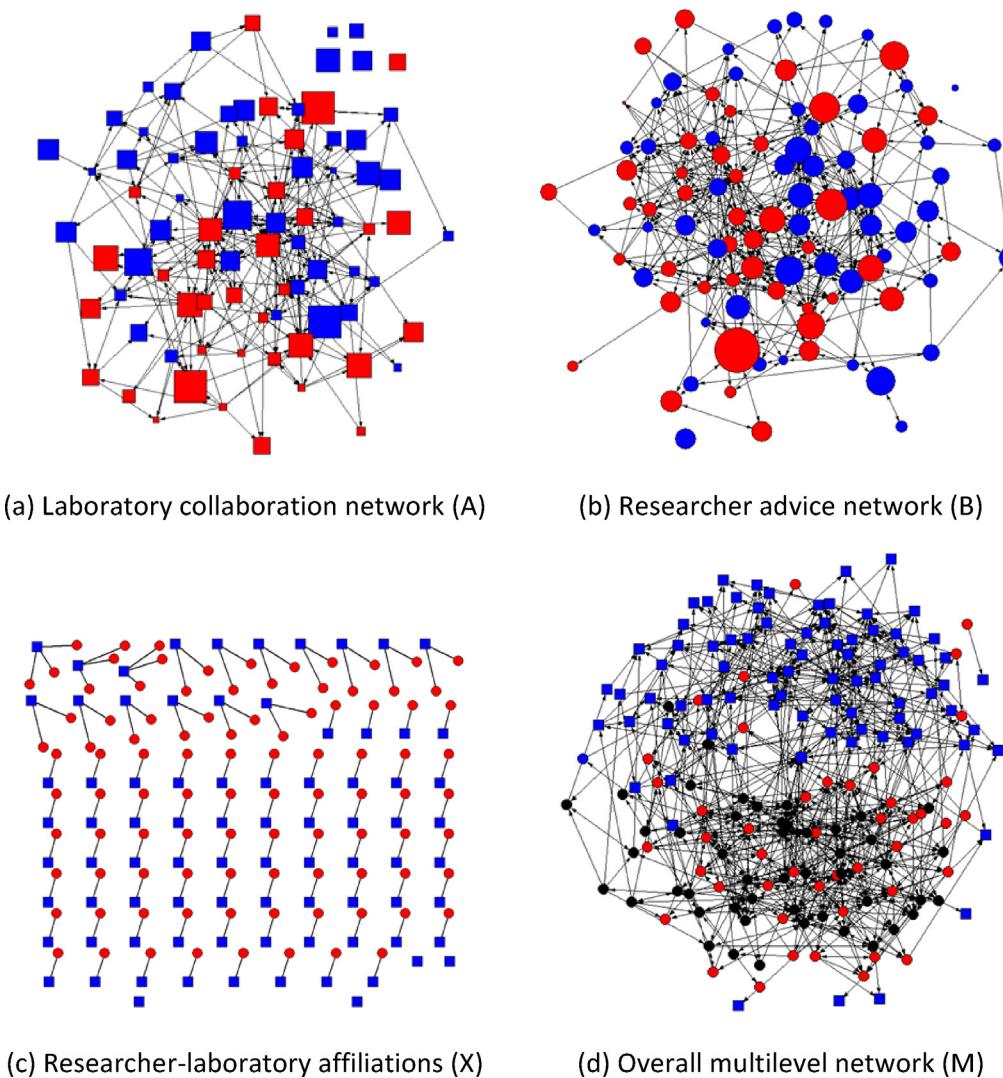
Model 1: A multilevel ERGM without attribute interactions. This is the same model as in the modelling example of Wang et al. (2013). We use Model 1 as a reference model to test whether the addition of attribute interaction effects will replace some of the structural effects.

Model 2: A multilevel SSM with only dyadic independent effects for all available attributes, that is, the interaction effects between nodal attributes and arc or reciprocal ties only. We can test whether attributes are important at a dyadic level, assuming that endogenous structural processes do not exist. This is analogous to a logistic regression model predicting ties only from attributes, assuming independent dyads.

Model 3: A multilevel SSM with both structural effects and attribute interaction effects. The structural part of the model follows the same specification as for Model 1; the specification for the attribute part of the model follows as much as Model 2 with the addition of some cross-level attribute interaction configurations based on a model selection strategy by model goodness of fit (GOF) tests as discussed in section 6. See detailed model comparisons below.

Model 4: A simplified version of Model 3 where most of the non-significant attribute parameters are removed while providing adequate fit to all graph statistics.

The within level endogenous effects follow the Robins et al. (2009) ERGM specifications for one-mode directed networks,



**Fig. 11.** The multilevel network among French cancer research elites (Lazega et al., 2008).

including the baseline density effect (Arc); reciprocity effect; degree centralization effects (AinS, AoutS); correlations between in- and out-degrees (Two-path); Transitive and cyclic closures (AT-T, AT-C); and shared activity effect (A2P-U).

The cross-level network endogenous effects included in our models follow Wang et al. (2013) multilevel ERGM specifications, including effects representing affiliation based popularity (AXS\_1A.in); affiliation based activity (AXS\_1A.out); affiliation based closure (TXBX.arc); in- and out-degree assortativity through affiliations (various three-path configurations); and cross-level entrainment and exchange (various four-cycle configurations).

To highlight the importance of network endogenous effects or the structural part of SSMs, we first compare Model 2 with Model 3 and 4. Model 2 is a dyadic independent model which assumes the existences of network ties are independent from one another (Erdős and Rényi, 1960; Holland and Leinhardt, 1981). Model 2 does not have any higher-order structural effects besides the baseline density and reciprocity effects and their interactions with nodal attributes, whereas Model 3 and 4 include within as well as cross-level higher-order effects.

Table 1 lists the parameter estimates (est.) and the estimated standard errors (s.e.) for the dyadic independent model (Model 2) and the attribute part of Model 3 and 4 labelled as Model 3 (a)

and Model 4 (a). Table 2 compares Model 1, the multilevel ERGM without attributes as in Wang et al. (2013), with the structural part of Model 3 and 4 labelled as Model 3 (b) and Model 4 (b). The significant parameter estimates are listed in bold font where the ratio between the estimate and its estimated standard error is greater than 2.0. Note that direct comparisons between the multilevel ERGM (Model 1) and the dyadic independent SSM (Model 2) are not shown in the same table, as the only common component between the two models are the density (Arc) and Reciprocity effects, therefore we list the Arc and Reciprocity effects of Model 2 in Table 1; and such effects for Model 3 and 4 are listed together with Model 1 in Table 2.

We compare the models in both the differences in the significances of the parameter estimates as well as model goodness of fits. For model goodness of fit (GOF) tests, 291 graph statistics are involved for each of the fitted model. These graph statistics without nodal attributes include within- and cross-level structure effects in the literature (e.g. Holland and Leinhardt, 1981; Frank and Strauss, 1986; Snijders et al., 2006; Wang et al., 2009, etc.), as well as the ones introduced in Wang et al. (2013). The statistics involve nodal attributes include statistics follow the within level one-mode SSMs (Robins et al., 2001) and statistics proposed in this paper. The full detail of the GOF analysis can be obtained from the authors on request.

**Table 1**

Comparison between a dyadic independent SSM (Model 2) and the attribute part of two SSMs (Model 3 and 4).

	[Attr]	Effects	Model 2		Model 3 (a)		Model 4 (a)	
			est.	(s.e.)	est.	(s.e.)	est.	(s.e.)
Laboratory collaboration network	Paris	Arc	<b>-3.574</b>	<b>0.181</b>				
		Reciprocity	<b>1.719</b>	<b>0.316</b>				
		Sender	0.158	0.188	0.069	0.168		
		Receiver	-0.325	0.212	-0.194	0.208		
		Interaction	<b>0.689</b>	<b>0.300</b>	0.375	0.293		
	Size	Sender	<b>-0.009</b>	<b>0.004</b>	-0.004	0.003		
		Receiver	0.001	0.004	0.000	0.004		
		Interaction	0.007	0.004	0.006	0.004		
Researcher advice network	Director	Arc	<b>-3.381</b>	<b>0.403</b>				
		Reciprocity	<b>3.582</b>	<b>0.186</b>				
		Sender	0.227	0.156	0.094	0.136		
		Receiver	0.174	0.148	0.183	0.157		
		Interaction	-0.126	0.161	-0.064	0.172		
	Paris	Sender	<b>-0.416</b>	<b>0.156</b>	<b>-0.350</b>	<b>0.105</b>	<b>-0.386</b>	<b>0.101</b>
		Receiver	-0.274	0.142	-0.132	0.131		
		Interaction	<b>1.147</b>	<b>0.174</b>	<b>0.618</b>	<b>0.127</b>	<b>0.569</b>	<b>0.094</b>
	Age	Sender	0.003	0.009	0.004	0.007		
		Receiver	-0.011	0.008	-0.008	0.008		
		Difference	<b>-0.038</b>	<b>0.007</b>	<b>-0.026</b>	<b>0.006</b>	<b>-0.023</b>	<b>0.006</b>
	Performance	Sender	0.003	0.002	0.001	0.002		
		Receiver	<b>0.009</b>	<b>0.002</b>	<b>0.005</b>	<b>0.002</b>	<b>0.005</b>	<b>0.002</b>
		Difference	<b>-0.011</b>	<b>0.002</b>	<b>-0.007</b>	<b>0.002</b>	<b>-0.007</b>	<b>0.002</b>
	Specialty	Match	<b>1.076</b>	<b>0.135</b>	<b>0.800</b>	<b>0.149</b>	<b>0.786</b>	<b>0.132</b>
		Match Reciprocity	<b>-0.678</b>	<b>0.312</b>	<b>-0.824</b>	<b>0.342</b>	<b>-0.795</b>	<b>0.321</b>
Cross level	Director	C4AXB Entrainment					<b>0.840</b>	<b>0.180</b>

Estimates in bold font indicate significant effects where the estimated effects exceed twice the standard errors in absolute values.

The parameter estimates for the dyadic independent model (Model 2) shown in Table 1 suggests both laboratory collaboration and researcher advice seeking tend to be reciprocated, as indicated by the positive Reciprocity parameter estimates.

For laboratories, without considering other higher order structural effects in the network, the positive Paris-Interaction parameter indicates laboratories within Paris region are more likely to collaborate with one another. The negative Size-Sender effect

indicates larger laboratories nominated fewer other laboratories as collaboration partners.

For researchers, the director status seems unimportant (at least at the dyadic level), perhaps reflecting the attitude of researchers towards formal hierarchical status. The negative Paris-sender effect indicates that researchers in Paris region are less likely to seek advice from others. However, the positive Paris-Interaction effect suggests researchers within the Paris region are more likely to

**Table 2**

Comparison between a model without attributes (Model 1) and the structural part of two SSMs (Model 3 and 4).

	Effects	Model 1		Model 3 (b)		Model 4 (b)	
		est.	(s.e.)	est.	(s.e.)	est.	(s.e.)
Laboratory collaboration network	Arc	<b>-3.831</b>	<b>0.556</b>	<b>-3.832</b>	<b>0.587</b>	<b>-3.815</b>	<b>0.574</b>
	Reciprocity	<b>1.679</b>	<b>0.381</b>	<b>1.610</b>	<b>0.399</b>	<b>1.525</b>	<b>0.413</b>
	2-path	<b>-0.079</b>	<b>0.029</b>	<b>-0.075</b>	<b>0.029</b>	<b>-0.090</b>	<b>0.029</b>
	Isolates	<b>2.017</b>	<b>0.760</b>	<b>2.024</b>	<b>0.786</b>	<b>2.057</b>	<b>0.769</b>
	AinS (4.00)	<b>0.640</b>	<b>0.268</b>	<b>0.686</b>	<b>0.288</b>	<b>0.737</b>	<b>0.267</b>
	AoutS (4.00)	<b>0.320</b>	<b>0.086</b>	<b>0.300</b>	<b>0.093</b>	<b>0.334</b>	<b>0.089</b>
	AinS (2.00)	-0.889	0.614	-0.949	0.653	-1.039	0.618
	AT-T (2.00)	<b>0.446</b>	<b>0.127</b>	<b>0.419</b>	<b>0.119</b>	<b>0.420</b>	<b>0.130</b>
	Researcher advice network	<b>-4.084</b>	<b>0.118</b>	<b>-3.847</b>	<b>0.321</b>	<b>-3.975</b>	<b>0.143</b>
	Reciprocity	<b>3.313</b>	<b>0.212</b>	<b>3.336</b>	<b>0.246</b>	<b>3.361</b>	<b>0.235</b>
Lab and affiliation	AT-T (2.00)	<b>1.085</b>	<b>0.072</b>	<b>1.042</b>	<b>0.079</b>	<b>1.046</b>	<b>0.074</b>
	AT-C (2.00)	<b>-0.384</b>	<b>0.068</b>	<b>-0.371</b>	<b>0.073</b>	<b>-0.360</b>	<b>0.073</b>
	A2P-U (2.00)	<b>-0.071</b>	<b>0.020</b>	<b>-0.073</b>	<b>0.021</b>	<b>-0.083</b>	<b>0.021</b>
	AXS1Ain (2.00)	0.24	0.131	0.161	0.134		
Researcher and affiliation Cross level interactions	AXS1Aout (2.00)	<b>-0.324</b>	<b>0.129</b>	-0.184	0.141		
	TXBX	<b>1.958</b>	<b>0.275</b>	<b>2.022</b>	<b>0.283</b>	<b>1.945</b>	<b>0.265</b>
	L3AXBin	-0.006	0.018	-0.009	0.018	0.004	0.017
	L3AXBout	-0.012	0.008	-0.015	0.010	-0.016	0.010
	L3AXBpath	<b>-0.051</b>	<b>0.015</b>	<b>-0.042</b>	<b>0.016</b>	<b>-0.043</b>	<b>0.016</b>
	L3BXApAth	-0.003	0.010	-0.008	0.011	-0.013	0.010
	C4AXB entrainment	<b>0.634</b>	<b>0.104</b>	<b>0.644</b>	<b>0.111</b>	<b>0.524</b>	<b>0.114</b>
	C4AXB exchange	<b>0.639</b>	<b>0.109</b>	<b>0.645</b>	<b>0.111</b>	<b>0.659</b>	<b>0.108</b>
	C4AXB exchange reciprocal A	<b>-0.293</b>	<b>0.065</b>	<b>-0.291</b>	<b>0.081</b>	<b>-0.256</b>	<b>0.096</b>
	C4AXB exchange reciprocal B	<b>-0.295</b>	<b>0.136</b>	<b>-0.277</b>	<b>0.144</b>	<b>-0.328</b>	<b>0.148</b>

Estimates in bold font indicate significant effects where the estimated effects exceed twice the standard errors in absolute values.

seek advice from each other than from researchers based in other parts of France. The combination of these two effects suggests that Parisian researchers seek advice from each other but less so from the provinces. The negative homophily effects on Age-Difference and Performance-Difference suggest researchers seek advice from others with similar age or performance scores. The positive Performance-Receiver effect indicates researchers with higher performance scores are more likely to receive advice seeking ties. From the estimates of the Specialty categorical attribute effects, we can see that advice seeking is more likely to happen within specialty. However, there is a tendency against reciprocation of advice seeking within specialty indicated by the negative Specialty-Match-Reciprocity effect, reflecting perhaps the tendency not to seek advice from colleagues “below” in the status hierarchy, even within the same specialty.

In a model GOF test, Model 2 fitted 163 out of the 291 graph statistic with t-ratios smaller than 2.0 in absolute values. Most of the adequately fitted statistics are attribute effects (i.e. interactions between attributes and the arc or the reciprocity configurations). As we would expect from a model without any higher order structural effects, the poorly fitted 128 statistics include most of the higher order configurations (with or without interactions with nodal attributes), such as in-/out-stars, triadic-closures of different forms, skewness/standard deviations of degree distributions, clustering coefficients and cross level interactions of various forms.

To capture the higher-order configurations, we use the ERGM specification obtained for the network without attributes (Model 1) as the specification for the structural part of the SSM. Model 3 is such an example whose structural parameters are shown in Table 2, and the attribute part is shown in Table 1. Before comparing Model 3 and 1, we briefly present the interpretation of Model 1 as recalled from Wang et al. (2013). At the laboratory level, collaboration tends to be reciprocal (positive reciprocity), centralized (positive A<sub>inS</sub> and A<sub>outS</sub>), clustered (positive AT-T), as well as following shorter paths (negative two-path). Within the researcher level, advice seeking tends to be reciprocal (positive reciprocity), clustered (positive AT-T) and hierarchical (negative AT-C); researchers also tend not to share multiple advisors (negative A<sub>2P-U</sub>). Laboratories with more researchers are less likely to nominate collaborations with other laboratories (negative AXS1A<sub>out</sub>). Researchers within the same laboratory tend to seek advice from one another. Researchers from popular laboratories tend to seek advice from researchers from collaborating laboratories (negative L3AXBpath and positive C4AXBexchange). There is a tendency for co-occurrence between laboratory collaborations and researcher advice seeking (positive C4AXBentrainment). However, stronger forms of cross-level entrainment are less likely (negative C4AXB exchange reciprocal A and B) possibly due to redundancy in knowledge or resource.

The structural part of the Model 3 is largely consistent with Model 1 except that the negative AXS1A<sub>out</sub> parameter estimate is no longer significant. Model 3 suggests the negative tendency for laboratories with more researchers to nominate collaborations indicated by Model 1 is subsumed into the attribute interaction effects. Besides this small difference, the structural part of Model 3 follows the same interpretations as Model 1. It is worth noting that all the significant attribute interaction effects in Model 3 are related to attributes of researchers rather than laboratories, despite the fact that the AXS1A<sub>out</sub> effect is oriented towards laboratories. This demonstrates that nodal attribute interaction effects at one level may be able to explain structural effects mostly related to nodes at a different level.

We saw from Table 1 that in the attribute part of Model 3, the previously significant location (Paris) interaction effect and the Size-sender effect of laboratories in Model 2 became non-significant. So these effects are explained by the higher order

structural features of the network, including the between and across level interactions. Other significant attribute parameter estimates in Model 2 remain significant in Model 3 despite the inclusion of structural effects. However, compared with Model 2, most of the significant attribute effects in Model 3 are smaller in scale relative to their estimated standard errors, which shows that the higher order structural effects also subsume part of the remaining significant attribute effects.

In terms of model GOF, Model 3 fitted most of the 291 graph statistics except the Director-C4AXBentrainment statistic (*t*-ratio = 2.46) and the skewness of researcher out degree distribution (*t*-ratio = 2.05). Given the already positive and significant structural C4AXB-entrainment effect, there are more cross level entrainment four-cycles involving directors than without directors. As well, the observed researcher out-degrees were more skewed than the model distribution. Notice that the Director-C4AXBentrainment configuration involves an outgoing tie from one of the directors (see Fig. 10 for the C4AXBentrainment configuration), so that a better control over the Director-C4AXBentrainment may also help improving the fit on the skewness of the researcher out-degree distribution.

Model 4 simplifies Model 3 by removing all non-significant attribute effects (as shown in Table 1) and the interaction between the lab collaboration and the cross-level affiliation, that is, AXS1A<sub>in</sub> and AXS1A<sub>out</sub> as shown in Table 2. Note that we keep the non-significant structural effects (L3AXB<sub>in</sub>, L3AXB<sub>out</sub> and L3BXApath) in Model 4, as without them Model 4 does not converge. The non-significant Laboratory-AinS helps model GOF on laboratory degree distributions as discussed in Wang et al. (2013). In addition, Model 4 includes the Director-C4AXB-entrainment parameter to capture the otherwise poorly fitted statistic. As a result, Model 4 provides adequate fit to all 291 graph statistics including the standard deviations and skewness of the degree distributions, so we conclude that Model 4 is the final model for this network data with attributes.

Similar to Model 3, the structural part of Model 4 is broadly consistent with Model 1 (Table 2), and the attribute part is qualitatively similar to Model 3 (Table 1), so the previous interpretations hold. The additional positive Director-C4AXB-entrainment parameter confirms that the directors are more likely to form entrainment advice seeking ties with directors from collaborating laboratories. It also improves the model fit on the researcher out-degree distribution.

It is an interesting observation that the Director attribute had no significant interactions with dyadic network structures (i.e. arc and reciprocity), but it is important for the cross-level four-cycle entrainment. This is an important message to learn from this example. We usually assume non-significant attribute effects at the dyadic level indicate that the attribute itself is not important for the network structure. This may be true for the within level network, that is, the Director attribute of the researchers has little effect on the advice seeking structure in this inter-organizational, semi-collegial context (see Piña-Stranger and Lazega (2011) for the differences between advice seeking at the intra- and inter-organizational levels). However, in a multilevel network context, an attribute effect that is non-significant at the dyadic level in a within level network may be still important in explaining network structures across levels. The laboratory collaboration ties are strongly aligned with advice seeking ties between directors. The researchers' director status does shape the laboratory collaboration network structure.

That directors' ties may affect laboratory collaborations is hardly empirically surprising in this particular context. However, it is not obvious in general that attributes at one level should shape ties at another. To our knowledge, such a proposition has never previously been put in the social networks theoretical literature, nor has it been empirically demonstrated.

From the positive Director-C4AXB-entrainment effect, we see that directors are important as invisible ‘embedded brokers’ in the creation of collaboration ties between laboratories and their affiliated researchers as suggested by Lazega et al. (2013). The descriptive analysis of Lazega et al. (2013) for this dataset suggests that the performance of a specific subset of researchers decreases when they become Directors of their lab and start working for the collective, which is consistent with the story that Directors become ‘embedded brokers’. The young researchers with lower than the median impact factor scores working in laboratories with higher than median in- and out-degree centrality scores, known as Little Fish in Big Pond or LFBP (Lazega et al., 2008), tend to catch up over time with the Big Fish in Big Pond (BFBP) in terms of performance. Connections between laboratories (created by these ‘embedded brokers’) increase connections with ‘dual alters’ (Lazega et al., 2013), that is, a potential collaborator at a different laboratory with whom the collaboration may be established by the directors or “embedded brokers”. This strategy of getting contact with dual alters through the laboratory collaboration channel has a leverage effect for performance when the dual alters who can be reached through this channel also have complementary resources.

## 8. Conclusions and discussions

As network data usually has information about nodes, incorporation of this extra information into our modelling framework will give us a more detailed and complete view of the network structure and the underlying network processes. Social selection models (SSMs) assume the existence of network ties depends not only on the presence of other ties but also on nodal level attributes. Building on the SSMs for one- and two-mode networks (Robins et al., 2001; Agneessens and Roose, 2008; Wang, 2013), we proposed SSM specifications for two-level networks as extensions to the multilevel ERGMs proposed in Wang et al. (2013).

The proposed configurations are based on the interactions between nodal attributes and well established social network theories such as homophily or heterophily, network popularity and network closure, etc. We discussed how these theoretical ideas may apply in multilevel network contexts. A detailed list of possible model configurations was presented together with example interpretations. They may serve as references for readers who wish to apply this class of models in different contexts.

Given that there are different types of nodal attributes and the relatively large numbers of possible configurations, how to find the right model specification becomes a non-trivial issue. As the

formal ERGM model selection procedures are yet to be developed due to the complexity of the network object and model specifications, we present an ad hoc model selection strategy based on experience. As demonstrated in the modelling example, the model goodness of fit test based selection strategy helps us eliminate non-significant parameters and add additional effects in order to find a converged parsimonious model that provides adequate fit to the network structure.

From the model comparisons in the modelling example, we can see several interesting and important implications about attribute information in a multilevel network context. First of all, a model with only attribute effects but not structural effects, such as Model 2, ignores the complex social structure, hence may produce biased interpretations, that is, some apparently significant attribute effects may indeed be explained by the network structure. Secondly, Model 1 ignores the attribute information which may make the structural part of the model overly complicated, as attribute effects may explain at least part of the network structure. Finally and most importantly, the cross level interactions in multilevel networks introduced another important aspect of dependencies between attributes and structure: attributes of nodes at one level may affect the structure of the network at the other level, regardless of whether the attribute is important or not at its own level.

The interpretations of our final model provide statistical confirmation of some of the interpretations of this network proposed by Lazega et al. (2008, 2013). This analysis of multi-level networks seems a sensible approach to key questions that sociologists ask when they try to combine, beyond ecological fallacies, both individual and contextual factors. Action and performance measured at the individual level strongly depend of the capacity of others to construct and to use organizations as “tools with a life of their own” (Selznick, 1949), and thus to manage interdependencies at different levels in a strategic manner. Without the models, we would not have statistical evidence for the rather detailed descriptions of both the multilevel networks and attributes as one picture.

## Appendix A.

For non-directed within-level networks, let  $A_{i+}$ ,  $X_{i+}$  and  $B_{i+}$  denote the degrees of node  $i$  in the corresponding networks;  $L_{2ij}^A(\mathbf{x}) = \sum_{k=1}^v X_{ik}X_{jk}$  denote the number of two-paths between nodes  $\{i, j\} \in A$  in the cross-level network  $X$ ; and  $I(Y_i^A, Y_j^A)$  be an indicator function such that  $I(Y_i^A, Y_j^A) = 1$  if  $Y_i^A = Y_j^A$ , 0 otherwise.

Effects	Figures	Statistics
[Attr]-star2AX(Binary or continuous)	Fig. 3(a)	$Z_{[\text{Attr}]-\text{star2AX}}(\mathbf{a}, \mathbf{x}, \mathbf{y}^A) = \sum_{i=1}^u A_i + X_i + Y_i^A$
[Attr <sup>A</sup> ]-TXBX (Binary or continuous)	Fig. 4(a)	$Z_{[\text{Attr}^A]-\text{TXBX}}(\mathbf{b}, \mathbf{x}, \mathbf{y}^A) = \sum_{i=1}^u \sum_{j=1}^u \sum_{k=j+1}^v X_{ij}X_{ik}B_{jk}Y_i^A$
[Attr <sup>A</sup> ]-TXAX (Binary or continuous)	Fig. 5(a)	$Z_{[\text{Attr}^A]-\text{TXAX}}(\mathbf{a}, \mathbf{x}, \mathbf{y}^A) = \sum_{i=1}^u \sum_{j=i+1}^u A_{ij}L_{2ij}^A Y_i^A Y_j^A$
[Attr <sup>A</sup> ]-TXAX-match (Categorical)	Fig. 5(a)	$Z_{[\text{Attr}^A]-\text{TXAX}}(\mathbf{a}, \mathbf{x}, \mathbf{y}^A) = \sum_{i=1}^u \sum_{j=i+1}^u A_{ij}L_{2ij}^A I(Y_i^A Y_j^A)$
[Attr <sup>A</sup> ]-TXAX-sum (Continuous)	Fig. 5(a)	$Z_{[\text{Attr}^A]-\text{TXAX-sum}}(\mathbf{a}, \mathbf{x}, \mathbf{y}^A) = \sum_{i=1}^u \sum_{j=i+1}^u A_{ij}L_{2ij}^A (Y_i^A + Y_j^A)$
[Attr <sup>A</sup> ]-[Attr <sup>B</sup> ]-TXAX (Binary)	Fig. 6(a)	$Z_{[\text{Attr}^A]-[\text{Attr}^B]-\text{TXAX}}(\mathbf{a}, \mathbf{x}, \mathbf{y}^A, \mathbf{y}^B) = \sum_{i=1}^u \sum_{j=1}^v \sum_{k=j+1}^u X_{ij}X_{ik}A_{jk}Y_j^A Y_i^B$

Effects	Figures	Statistics
[Attr <sup>A</sup> ]-[Attr <sup>B</sup> ]-TXAX-match(Categorical)	Fig. 6(a)	$Z_{[\text{Attr}^A]-[\text{Attr}^B]-\text{TXAX}-\text{match}}(\mathbf{a}, \mathbf{x}, \mathbf{y}^A, \mathbf{y}^B) = \sum_{i=1}^v \sum_{j=1}^u \sum_{k=j+1}^u X_{ij} X_{ik} A_{jk} I(Y_j^A, Y_i^B)$
[Attr <sup>A</sup> ]-[Attr <sup>B</sup> ]-TXAX-sum (Continuous)	Fig. 6(a)	$Z_{[\text{Attr}^A]-[\text{Attr}^B]-\text{TXAX}-\text{sum}}(\mathbf{a}, \mathbf{x}, \mathbf{y}^A, \mathbf{y}^B) = \sum_{i=1}^v \sum_{j=1}^u \sum_{k=j+1}^u X_{ij} X_{ik} A_{jk} (Y_j^A + Y_i^B)$
[Attr <sup>A</sup> ]-L3XAX (Binary)	Fig. 7(a)	$Z_{[\text{Attr}^A]-\text{L3XAX}}(\mathbf{a}, \mathbf{x}, \mathbf{y}^A) = \sum_{i=1}^u \sum_{j=i+1}^u A_{ij} X_i + X_j + Y_i^A Y_j^A$
[Attr <sup>A</sup> ]-L3XAX-match (Categorical)	Fig. 7(a)	$Z_{[\text{Attr}^A]-\text{L3XAX}-\text{match}}(\mathbf{a}, \mathbf{x}, \mathbf{y}^A) = \sum_{i=1}^u \sum_{j=i+1}^u A_{ij} X_i + X_j + I(Y_i^A Y_j^A)$
[Attr <sup>A</sup> ]-L3XAX-difference (Continuous)	Fig. 7(a)	$Z_{[\text{Attr}^A]-\text{L3XAX}-\text{difference}}(\mathbf{a}, \mathbf{x}, \mathbf{y}^A) = \sum_{i=1}^u \sum_{j=i+1}^u A_{ij} X_i + X_j +  Y_i^A - Y_j^A $
[Attr <sup>A</sup> ]-[Attr <sup>B</sup> ]-L3AXB (Binary)	Fig. 8(a)	$Z_{[\text{Attr}^A]-[\text{Attr}^B]-\text{L3AXB}}(\mathbf{a}, \mathbf{b}, \mathbf{x}, \mathbf{y}^A, \mathbf{y}^B) = \sum_{i=1}^u \sum_{j=1}^v X_{ij} A_i + B_j + Y_i^A + Y_j^B$
[Attr <sup>A</sup> ]-C4AXB (Binary)	Fig. 9(a)	$Z_{[\text{Attr}^A]-\text{C4AXB}}(\mathbf{a}, \mathbf{b}, \mathbf{x}, \mathbf{y}^A) = \frac{1}{2} \sum_{i=1}^u \sum_{j=i+1}^u \sum_{k=1}^v \sum_{l=1}^v A_{ij} X_{ik} X_{jl} B_{kl} Y_i^A Y_j^A$
[Attr <sup>A</sup> ]-[Attr <sup>B</sup> ]-C4AXB Binary	Fig. 10	$Z_{[\text{Attr}^A]-[\text{Attr}^B]-\text{C4AXB}}(\mathbf{a}, \mathbf{b}, \mathbf{x}, \mathbf{y}^A, \mathbf{y}^B) = \frac{1}{2} \sum_{i=1}^u \sum_{j=i+1}^u \sum_{k=1}^v \sum_{l=1}^v A_{ij} X_{ik} X_{jl} B_{kl} Y_i^A Y_k^B$

When within-level networks are directed, let  $A_{i+}$  and  $B_{i+}$  denote the out-degrees of node  $i$ , and  $A_{+i}$  and  $B_{+i}$  denote the in-degrees of node  $i$ .

Effects	Figures	Statistics
[Attr]-outStar2AX Binary and continuous	Fig. 3(b)	$Z_{[\text{Attr}]-\text{outStar2AX}}(\mathbf{a}, \mathbf{x}, \mathbf{y}^A) = \sum_{i=1}^u A_{i+} X_{i+} Y_i^A$
[Attr]-inStar2AX Binary and continuous	Fig. 3(c)	$Z_{[\text{Attr}]-\text{inStar2AX}}(\mathbf{a}, \mathbf{x}, \mathbf{y}^A) = \sum_{i=1}^u A_{+i} X_{+i} Y_i^A$
[Attr <sup>A</sup> ]-TXBX	Fig. 4(b)	$Z_{[\text{Attr}^A]-\text{TXBX}}(\mathbf{b}, \mathbf{x}, \mathbf{y}^A) = \sum_{i=1}^u \sum_{j=1}^v \sum_{k=1}^v X_{ij} X_{ik} B_{jk} Y_i^A$
[Attr <sup>A</sup> ]-TXBX-reciprocity	Fig. 4(c)	$Z_{[\text{Attr}^A]-\text{TXBX}-\text{reciprocity}}(\mathbf{b}, \mathbf{x}, \mathbf{y}^A) = \sum_{i=1}^u \sum_{j=1}^v \sum_{k=1}^v \sum_{l=k+1}^v X_{ij} X_{ik} B_{jk} Y_i^A$
[Attr <sup>A</sup> ]-TXAX	Fig. 5(b)	$Z_{[\text{Attr}^A]-\text{TXAX}}(\mathbf{a}, \mathbf{x}, \mathbf{y}^A) = \sum_{i=1}^u \sum_{j=1}^v A_{ij} L_{2ij}^A Y_i^A Y_j^A$
[Attr <sup>A</sup> ]-TXAX-reciprocity	Fig. 5(c)	$Z_{[\text{Attr}^A]-\text{TXAX}-\text{reciprocity}}(\mathbf{a}, \mathbf{x}, \mathbf{y}^A) = \sum_{i=1}^u \sum_{j=1}^v A_{ij} A_{ji} L_{2ij}^A Y_i^A Y_j^A$
[Attr <sup>A</sup> ]-C4AXB-exchange (Binary)	Fig. 9(e)	$Z_{[\text{Attr}^A]-\text{C4AXB}}(\mathbf{a}, \mathbf{b}, \mathbf{x}, \mathbf{y}^A) = \frac{1}{2} \sum_{i=1}^u \sum_{j=i+1}^u \sum_{k=1}^v \sum_{l=1}^v A_{ij} X_{ik} X_{jl} B_{lk} Y_i^A Y_j^A$
[Attr <sup>A</sup> ]-C4AXB-entrainment-match (categorical)	Fig. 9(c)	$Z_{[\text{Attr}^A]-\text{C4AXB}-\text{entrainment}-\text{match}}(\mathbf{a}, \mathbf{b}, \mathbf{x}, \mathbf{y}^A) = \frac{1}{2} \sum_{i=1}^u \sum_{j=1}^u \sum_{k=1}^v \sum_{l=1}^v A_{ij} X_{ik} X_{jl} B_{kl} I(Y_i^A Y_j^A)$
[Attr <sup>A</sup> ]-C4AXB-reciprocity-difference (continuous)	Fig. 9(g)	$Z_{[\text{Attr}^A]-\text{C4AXB}-\text{reciprocity}-\text{difference}}(\mathbf{a}, \mathbf{b}, \mathbf{x}, \mathbf{y}^A) = \frac{1}{2} \sum_{i=1}^u \sum_{j=i+1}^u \sum_{k=1}^v \sum_{l=1}^v A_{ij} A_{ji} X_{ik} X_{jl} B_{kl} B_{lk}  Y_i^A - Y_j^A $

## References

- Agneessens, F., Roose, H., 2008. Local structural properties and attribute characteristics in 2-mode networks: p\*models to map choices of theater events 1. J. Math. Soc. 32 (3), 204–237.
- Breiger, R.L., 1974. The duality of persons and groups. Soc. Forces 53 (2), 181–190.
- de la Haye, K., Robins, G., Mohr, P., Wilson, C., 2010. Obesity-related behaviors in adolescent friendship networks. Soc. Netw. 32 (3), 161–167.
- de Nooy, W., Mrvar, A., Batagelj, V., 2005. Exploratory Social Network Analysis with Pajek, Structural Analysis in the Social Sciences 27. Cambridge University Press, ISBN: 0521602629.
- Erdős, P., Rényi, A., 1960. On the evolution of random graphs. Publ. Math. Inst. Hung. Acad. Sci. 5, 17–61.
- Frank, O., Strauss, D., 1986. Markov graphs. J. Am. Stat. Assoc. 81 (395), 832–842.
- Goodreau, S.M., Kitts, J.A., Morris, M., 2009. Birds of a feather, or friend of a friend? Using exponential random graph models to investigate adolescent social networks. Demography 46 (1), 103–125.
- Harrigan, N., Bond, M., 2013. Differential impact of directors' social and financial capital on corporate interlock formation'. In: Lusher, D., Koskeni, J., Robins, G. (Eds.), Exponential Random Graph Models for Social Networks: Theories, Methods and Applications. Cambridge University Press, Cambridge, UK.

- Holland, P.W., Leinhardt, S., 1981. An exponential family of probability distributions for directed graphs. *J. Am. Stat. Assoc.* 76 (373), 33–50.
- Huiszing, G., Van Duijn, M.A.J., Snijders, T.A.B., Wang, P., Sainio, M., Salmivalli, C., Veenstra, R., 2012. Univariate and multivariate models of positive and negative networks: liking, disliking, and bully–victim relationships. *Soc. Netw.* 34 (4), 645–657.
- Hunter, D.R., Goodreau, S.M., Handcock, M.S., 2008. Goodness of fit of social network models. *J. Am. Stat. Assoc.* 103, 248–258.
- Koskinen, J., Edling, C., 2012. Modelling the evolution of a bipartite network—peer referral in interlocking directorates. *Soc. Netw.* 34 (3), 309–322.
- Lazega, E., Van Duijn, M., 1997. Position in formal structure, personal characteristics and choices of advisors in a law firm: a logistic regression model for dyadic network data. *Soc. Netw.* 19, 375–397.
- 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), 3–28.
- Lazega, E., Mounier, L., Jourda, M.T.S., Stofer, R., 2006. Organizational vs. personal social capital in scientists' performance: a multi-level network study of elite French cancer researchers (1996–1998). *Scientometrics* 67 (1), 27–44.
- Lazega, E., Jourda, M.T., Mounier, L., Stofer, R., 2008. Catching up with big fish in the big pond? Multi-level network analysis through linked design. *Soc. Netw.* 30 (2), 159–176.
- Lazega, E., Jourda, M.-T., Mounier, L., 2013. Network lift from dual alters: extended opportunity structures from a multilevel and structural perspective. *Eur. Sociol. Rev.* 29, 1226–1238.
- Lusher, D., Robins, G., 2013. Formation of social network structure. In: Lusher, D., Koskinen, J., Robins, G. (Eds.), *Exponential Random Graph Models for Social Networks: Theory, Methods, and Applications*. Cambridge University Press, Cambridge, UK.
- Lusher, D., Robins, G., 2010. A social network analysis of hegemonic and other masculinities. *J. Mens Stud.* 18 (1), 22–44.
- McPherson, M., Smith-Lovin, L., Cook, J.M., 2001. Birds of a feather: homophily in social networks. *Annu. Rev. Soc.* 415–444.
- Pattison, P.E., Robins, G.L., 2002. Neighborhood-based models for social networks. *Soc. Method.* 32, 301–337.
- Pattison, P.E., Robins, G.L., 2004. Building models for social space: neighbourhood-based models for social networks and affiliation structures? *Math. Soc. Sci.* 42 (168), 11–29.
- Pattison, P., Snijders, T., 2013. Statistical models for social networks: future directions. In: *Exponential Random Graph Models for Social Networks: Theories, Methods and Applications*. Cambridge University Press, Cambridge, UK.
- Pattison, P., Robins, G., Snijders, T., Wang, P., Unpublished manuscript 2009. A hierarchy of dependence assumptions for exponential random graph models for social networks. Psychological Sciences, University of Melbourne.
- Piña-Stranger, Á., Lazega, E., 2011. Bringing personalized ties back in: their added value for biotech entrepreneurs and venture capitalists interorganizational networks. *Sociol. Q.* 52, 268–292.
- Robins, G., Elliott, P., Pattison, P., 2001. Network models for social selection processes. *Soc. Netw.* 23 (1), 1–30.
- Robins, G., Pattison, P., Woolcock, J., 2005. Small and other worlds: global network structures from local processes. *Am. J. Soc.* 110 (4), 894–936.
- Robins, G., Lusher, D., 2013. Illustrations: simulation, estimation and goodness of fit. In: Lusher, D., Koskinen, J., Robins, G. (Eds.), *Exponential Random Graph Models for Social Networks: Theory, Methods, and Applications*. Cambridge University Press, Cambridge, UK.
- Robins, G., Pattison, P., Wang, P., 2009. Closure, connectivity and degree distributions: exponential random graph ( $p^*$ ) models for directed social networks. *Soc. Netw.* 31, 105–117.
- Robins, G., Pattison, P., Kalish, Y., Lusher, D., 2007a. An introduction to exponential random graph ( $p^*$ ) models for social networks. *Soc. Netw.* 29 (2), 173–191.
- Selznick, P., 1949. *TVA and the Grass Roots: A Study of Politics and Organization*, vol. 3. University of California Press.
- Snijders, T.A.B., 2001. The statistical evaluation of social network dynamics. *Sociol. Methodol.* 40, 361–395.
- Snijders, T.A.B., Steglich, C.E.G., Schweinberger, M., 2007a. Modeling the co-evolution of networks and behavior. In: van Montfort, K., Oud, H., Satorra, A. (Eds.), *Longitudinal Models in the Behavioral and Related Sciences*. Lawrence Erlbaum, Mahwah, NJ, pp. 41–71.
- Snijders, T.A.B., van de Bunt, G.G., Steglich, C.E.G., 2010a. Introduction to stochastic actor-based models for network dynamics. *Soc. Netw.* 32, 44–60.
- Snijders, T.A., Lomi, A., Torló, V.J., 2013. A model for the multiplex dynamics of two-mode and one-mode networks, with an application to employment preference, friendship, and advice. *Soc. Netw.* 35 (2), 265–276.
- Snijders, T.A., Pattison, P.E., Robins, G.L., Handcock, M.S., 2006. New specifications for exponential random graph models. *Sociol. Methodol.* 36 (1), 99–153.
- Snijders, T.A., Steglich, C.E., Schweinberger, M., 2007b. Modeling the co-evolution of networks and behavior. *Longit. Models Behav. Relat. Sci.*, 41–71.
- Snijders, T.A., Van de Bunt, G.G., Steglich, C.E., 2010b. Introduction to stochastic actor-based models for network dynamics. *Soc. Netw.* 32 (1), 44–60.
- Snijder, T.A.B., 2011. Statistical models for social networks. *Annu. Rev. Sociol.* 37, 129–151.
- Useem, M., 1984. *The Inner Circle: Large Corporations and the Rise of Business Political Activity in the US and UK*. Oxford University Press, New York.
- Wang, P., Sharpe, K., Robins, G.L., Pattison, P.E., 2009. Exponential random graph ( $p^*$ ) models for affiliation networks. *Soc. Netw.* 31 (1), 12–25.
- Wang, P., 2013. Exponential random graph model extensions: models for multiple networks and bipartite networks. In: *Exponential Random Graph Models for Social Networks: Theories, Methods and Applications*. Cambridge University Press, Cambridge, UK.
- Wang, P., Robins, G., Pattison, P., 2006. PNet: A Program for the Simulation and Estimation of Exponential Random Graph Models. Melbourne School of Psychological Sciences, University of Melbourne.
- Wang, P., Robins, G., Pattison, P., Lazega, E., 2013. Exponential random graph models for multilevel networks. *Soc. Netw.* 35 (1), 96–115.
- Wasserman, S., Iacobucci, D., 1991. Statistical modelling of one-mode and two-mode networks: simultaneous analysis of graphs and bipartite graphs. *Br. J. Math. Stat. Psychol.* 44 (1), 13–43.
- Wasserman, S., Pattison, P., 1996. Logit models and logistic regressions for social networks: I. An introduction to Markov graphs and  $p^*$ . *Psychometrika* 61 (3), 401–425.