

Improving exponential-family random graph models for bipartite networks

Alex Stivala ^{1,*}, Peng Wang^{2,3}, Alessandro Lomi⁴

¹Institute of Computing, Università della Svizzera italiana, Via Buffi 13, Lugano, Ticino 6900, Switzerland

²Centre for Transformative Innovation, Swinburne University of Technology, John Street, Hawthorn VIC 3122, Australia

³Institute of Computing, Università della Svizzera italiana, Via Buffi 13, Lugano, Ticino 6900, Switzerland

⁴Università della Svizzera italiana, Via Buffi 13, Lugano, Ticino 6900, Switzerland

*Corresponding author. Institute of Computing, Università della Svizzera italiana, Via Buffi 13, Lugano 6900, Switzerland. E-mail: alexander.stivala@usi.ch.

ABSTRACT

Bipartite graphs, representing two-mode networks, arise in many research fields. These networks have two disjoint node sets representing distinct entity types, for example persons and groups, with edges representing associations between the two entity types. In bipartite graphs, the smallest possible cycle is a cycle of length four, and hence four-cycles are the smallest structure to model closure in such networks. Exponential-family random graph models (ERGMs) are a widely used model for social, and other, networks, including specifically bipartite networks. Existing ERGM terms to model four-cycles in bipartite networks, however, are relatively rarely used. In this work we demonstrate some problems with these existing terms to model four-cycles, and define new ERGM terms to help overcome these problems. The position of the new terms in the ERGM dependence hierarchy, and their interpretation, is discussed. The new terms are demonstrated in simulation experiments, and their application illustrated on a canonical example of an empirical two-mode network.

KEYWORDS: bipartite graph; two-mode network; exponential-family random graph model; ERGM; four-cycle.

1. INTRODUCTION

Bipartite graphs are graphs whose nodes can be partitioned into two disjoint sets, such that an edge exists only between nodes in different sets. Such graphs have important applications in representing two-mode networks, which are networks in which there are two types of nodes, with edges possible only between nodes of different types. An important example of a two-mode network is an affiliation network, in which one type of node represents a person, the other type of node represents a group, and an edge represents membership of a person in a group [1]. Two-mode networks have applications not only in sociology, but also biology, ecology, political science, psychology, finance, and economics; for a recent review of applications and methods for two-mode networks, see Neal *et al.* [2]. Bipartite networks also arise as representing the meso-level network in the conceptualization and analysis of multilevel networks [3].

Two-mode networks can be studied by means of their projections onto one-mode (unipartite) networks, thereby allowing the use of existing methods for one-mode networks, however this

Received: 9 February 2025. **Editorial decision:** 10 June 2025. **Accepted:** 12 June 2025

© The Author(s) 2025. Published by Oxford University Press. This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted reuse, distribution, and reproduction in any medium, provided the original work is properly cited.

can result in lost information, and properties of the one-mode networks (such as high clustering coefficients) that are due to the projection process rather than the original data [4]. Although the former problem can be ameliorated by using both projections in analyses [5, 6], it is still desirable to study the original two-mode network directly, for which specific methods are required [4].

In studying one-mode networks, a central concept is triadic closure, the tendency for a path of length two (a ‘two-path’; three nodes connected by two edges) to be ‘closed’ into a triangle by the addition of a third edge. In the context of social networks, this is the process of a friend of a friend becoming themselves a friend, and is perhaps most well known via the ‘strength of weak ties’ [7] argument, whereby an open two-path of strong ties is the ‘forbidden triad’, which is ‘forbidden’ because the two actors with strong ties to a common third actor must themselves have a strong tie. In a bipartite network, however, a closed triad (triangle) is impossible; indeed it is a defining feature of bipartite graphs that only cycles of even length are possible [8]. Therefore, the smallest possible cycle in a bipartite graph is a four-cycle, and hence four-cycles are frequently used to measure closure in bipartite networks [9, 10].

An example of a social process that creates four-cycles is peer referral in director interlock networks [9, 11]. If a director on the boards of two companies recruits a director from one of them to also sit on the board of the other, then an open three-path is closed, forming a four-cycle.

Four-cycles in bipartite networks also have the particular importance that, together with the degree distribution, they explain the degree assortativity in the one-mode projected network [12, 13].

Exponential random graph models (ERGMs) are a widely used model for social [14–17], and other [18–20] networks. Specific forms of the ERGM have been developed for two-mode networks [21–24], however, as we shall show in this work, existing ERG models for bipartite networks often have problems modelling four-cycles, and hence can frequently not adequately model closure in bipartite networks.

In this work, we will show that, despite the importance of four-cycles in two-mode networks, ERGM terms to model four-cycles in such networks are relatively rarely used, in contrast to the ubiquity of terms modelling triadic closure in one-mode networks. We will then describe some problems with existing configurations for modelling four-cycles that could explain their relatively infrequent use, and propose new ERGM configurations for modelling four-cycles to help overcome these problems. We will discuss the interpretation of these new parameters, and their position in the dependence hierarchy of Pattison and Snijders [25]. We will then demonstrate the new configurations using simulation experiments and demonstrate their use on a canonical example of an empirical two-mode network.

2. THE ERGM

An exponential-family random graph model (ERGM) is a probability distribution with the form

$$\Pr_{\theta}(X = x) = \frac{1}{\kappa(\theta)} \exp \left(\sum_C \theta_C g_C(x) \right) \quad (2.1)$$

where

- $X = [X_{ij}]$ is square binary matrix of random tie variables,
- x is a realization of X ,
- C is a ‘configuration’, a set of nodes and a subset of ties between them, designed in order to model a particular structure of interest,
- $g_C(x)$ is the network statistic for configuration C ,
- θ is a vector of model parameters, where each θ_C is the parameter corresponding to configuration C ,

- $\kappa(\theta) = \sum_{x \in G_n} \exp(\sum_C \theta_{CGC}(x))$, where G_n is the set of all square binary matrices of order n (graphs with n nodes), is the normalizing constant to ensure a proper distribution.

We will use the notation x_{ij} (where $1 \leq i \leq n$, $1 \leq j \leq n$) for elements of the binary adjacency matrix x . In this work we consider only the case of undirected networks, so $x_{ij} = x_{ji}$, and the cardinality of G_n is $|G_n| = 2^{n(n-1)/2}$. In the case of bipartite networks, the two disjoint node sets are denoted A and B , with sizes $N_A = |A|$ and $N_B = |B|$ respectively, and so $N_A + N_B = n$, and $x_{ij} = 0$ if both i and j are in node set A or both i and j are in node set B . In this bipartite case, the normalizing constant is $\kappa(\theta) = \sum_{x \in G_{N_A, N_B}^{\text{Bipartite}}} \exp(\sum_C \theta_{CGC}(x))$, where $G_{N_A, N_B}^{\text{Bipartite}}$ is the set of all bipartite graphs with node set sizes N_A and N_B . This set has cardinality $|G_{N_A, N_B}^{\text{Bipartite}}| = 2^{N_A N_B}$.

Estimating the value of the parameter vector θ which maximizes the probability of the observed graph, that is, the maximum likelihood estimator (MLE), enables inferences regarding the under-representation (negative and statistically significant estimate) or over-representation (positive and statistically significant estimate) of the corresponding configurations. These inferences are conditional on the other configurations included in the model, which need not be independent.

Estimating the MLE of (2.1) is computationally intractable due to the normalizing constant $\kappa(\theta)$ (specifically, the size of the set of graphs it sums over). Therefore, Markov chain Monte Carlo (MCMC) methods are usually used [26–28]. One such algorithm is the ‘Equilibrium Expectation’ (EE) algorithm [29–31], which was recently shown to converge to the MLE [20, 31].

ERG models with only simple configurations (such as the Markov edge plus triangle model) can be prone to problems with phase transitions or ‘near-degeneracy’ [17, 32–38]. Such problems are typically avoided by the use of ‘alternating’ [33, 39, 40] or ‘geometrically weighted’ [41, 42] configurations. The former are parameterized with a decay parameter λ controlling the rate at which the weight of contributions from additional terms in the statistic decay. The corresponding parameter for the geometrically weighed configurations can be estimated as part of the model, in which case it becomes a ‘curved ERGM’ [43], however in this work we will use fixed values of λ for the ‘alternating’ configurations.

In order to model two-mode networks, and in particular affiliation networks, Wang *et al.* [21] define the alternating k -two-path statistics $K\text{-C}_A$ and $K\text{-C}_B$, implemented as $XACA$ and $XACB$ in MPNet [44, 45], and $\text{BipartiteAltKCyclesA}$ and $\text{BipartiteAltKCyclesB}$ in *EstimNetDirected* (<https://github.com/stivalaa/EstimNetDirected>):

$$z_{K\text{-C}_A}(\lambda) = z_{XACA}(\lambda) = z_{\text{BipartiteAltKCyclesA}}(\lambda) = \lambda \sum_{i \in B} \sum_{\{l \in B : l < i\}} \left[1 - \left(1 - \frac{1}{\lambda} \right)^{L_2(i, l)} \right] \quad (2.2)$$

where $\lambda > 1$ is the decay parameter, and $L_2(i, l)$ is the number of two-paths connecting i to l :

$$L_2(i, l) = \sum_{h \neq i, l} x_{ih} x_{hl}. \quad (2.3)$$

The corresponding change statistic is [21]:

$$\delta_{K\text{-C}_A}(\lambda)(i, j) = \delta_{XACA}(\lambda)(i, j) = \delta_{\text{BipartiteAltKCyclesA}}(\lambda)(i, j) = \sum_{l \in B} \left[x_{il} \left(1 - \frac{1}{\lambda} \right)^{L_2(j, l)} \right] \quad (2.4)$$

$$= \sum_{l \in N(i)} \left(1 - \frac{1}{\lambda} \right)^{L_2(j, l)}. \quad (2.5)$$

where $N(i)$ denotes the neighbours of node i , that is, nodes $k \neq i$ such that $x_{ik} = 1$, or, equivalently, $d(i, k) = 1$, where $d(u, v)$ is the geodesic distance from u to v . The statistic and change statistic for BipartiteAltKCyclesB (K-C_P or XACB) are defined similarly.

3. LITERATURE SURVEY

In order to get an overview of which effects are used in modelling bipartite networks with ERGMs, and how well they fit four-cycles, we conducted a comprehensive survey of publications which included ERG models of bipartite networks. To be included, a publication must contain one or more ERG models of one or more empirical bipartite (two-mode) networks. There must be sufficient detail given to know at least the parameters included in the models, and their estimated signs and statistical significance. Models of one-mode projections of two-mode networks were excluded; we only consider ERG models of the bipartite network itself. Models of multilevel networks were excluded, although if there is a model of just the cross-level (bipartite) network, this is included. The papers included in the survey are listed in [Supplementary Table SA1](#) in [Supplementary Appendix SA](#), which also contains further details of the literature survey.

Most of the models in [Supplementary Table SA1](#) were estimated with the BPNet [46], MPNet [44, 45], or statnet [47–53] software, but a small number were estimated either by maximum pseudo-likelihood estimation (MPLE) or with Bayesian methods using the Bergm [54, 55] software. The one model included that was estimated with Bergm contains no terms to model four-cycles or goodness-of-fit tests including four-cycles [56]. Of the ten models (across three publications) estimated by MPLE, only one contains a term to model four-cycles, and this is found to be positive and significant [57].

[Table 1](#) summarizes the parameter estimates in models from the literature in [Supplementary Table SA1](#) that were estimated using BPNet or MPNet, and contain the four-cycles parameter C4, or the bipartite alternating k -two-path parameters (K-C_P and K-C_A) defined in Wang *et al.* [21]. Less than a third (20/63) include the four-cycles parameter, less than half (30/63) include either of the two alternating k -two-path parameters, and less than 30% (17/63) include both K-C_P and K-C_A.

[Table 2](#) summarizes the parameter estimates in models from the literature in [Supplementary Table SA1](#) that were estimated using statnet, and which contain the four-cycle term, the bipartite geometrically weighted dyadwise shared partner distribution (gwb1dsp or gwb2dsp) term, the statnet equivalent of the K-C_P and K-C_A parameters, or the geometrically weighted non-edgewise shared partner (gwnsp) term. Only one model estimated with statnet contained an explicit term for four-cycles [58]. Further, only 5/43 models contain a gwb1dsp or gwb2dsp term at all, and only two models contain both (Lubell and Robbins [59] have two models, both of which include both gwb1nsp and gwb2nsp).

Less than a quarter (26/117) of the models include an explicit assessment of goodness-of-fit to four-cycles, and of those, the majority (18/26) are good. Only seven of these are for models that explicitly include C4 as a model parameter, and, as expected of any converged model containing this term, these models fit four-cycles well. Conversely, all of the models which are described as having a poor fit to four-cycles in the goodness-of-fit procedure are models that do not contain the

Table 1. Counts of parameters in models estimated by BPNet or MPNet (total 63) in the reviewed literature

	C4	K-C _P	K-C _A
Total estimated	20	30	25
Negative	14	20	15
Negative and signif.	5	14	9
Positive	6	10	10
Positive and signif.	4	7	2

Table 2. Counts of parameters in models estimated by statnet (total 43) in the reviewed literature^a

	cycle(4)	gwb1dsp	gwb2dsp	gwnsp
Total estimated	1	2	5	4
Negative	1	1	2	1
Negative and signif.	1	1	0	0
Positive	0	1	3	3
Positive and signif.	0	1	2	3

^aThe counts for gwb1dsp and gwb2dsp include those for the equivalent parameters gwb1nsp and gwb2nsp, respectively.

C4 parameter. However, of these eight models, five contain either $K-C_A$ or $K-C_P$, and one contains both.

This relative rarity of models containing terms to model closure (four-cycles) in bipartite networks, or assess goodness-of-fit to four-cycles, is in stark contrast to ERG modelling for one-mode networks, where terms modelling triadic closure, such as triangles, alternating- k -triangles, or geometrically weighted edgewise shared partners (gwesp in statnet) are almost always included in models, since triadic closure (as evidenced by clustering, or transitivity, in the network), is a well-known feature of social networks [7, 60, 61]. For example, Clark and Handcock [62] use the latent order logistic (LOLOG) model [63] to reproduce ERG models for 13 networks from peer reviewed papers, and all of these models contain alternating- k -triangles [39, 64–66], gwesp [67–72], or three-cycles [64, 73] terms, and in the majority of cases the results are able to be replicated with a triangle term in the LOLOG model [62]. And yet, despite higher than expected numbers of four-cycles (or bipartite clustering coefficient) being a notable feature of some bipartite networks, such as director interlock networks [9, 10], and collaboration and communication networks [10], the literature survey presented here shows that terms to model four-cycles are relatively rarely used in published ERG models of bipartite networks. This could be due to researchers choosing not to model bipartite closure, however, as shown by some examples in the following section, it can also be due to difficulties in obtaining converged model estimates when using existing model terms.

4. PROBLEMS WITH EXISTING BIPARTITE ERGM STATISTICS

It is notable that in the original paper proposing $K-C_A$ and $K-C_P$ configurations [21], they are explicitly described as k -two-path statistics, and in fact they are just bipartite versions of the one-mode alternating- k -two-path statistic, representing multiple shared partners (gwdsp in statnet, and then later gwb1dsp and gwb2dsp for bipartite networks). There is, however, already some ‘semantic slippage’ into interpreting them as ‘cycles’ or ‘closure’—even though they only actually include cycles (closure) when $k > 1$ ($k = 1$ is a two-path, $k = 2$ is a four-cycle; see Fig. 3 in Wang *et al.* [21]). For example, ‘... a better chance of achieving model convergence when closure effects ($K - C_P$ and $K - C_A$) are included in the model’ [21]. Even the name $K-C_A$ (or $K-C_P$) suggests ‘cycle’ or ‘closure’ by the use of the ‘C’ (C_4 is used for four-cycles in the paper). Wang *et al.* [22] goes back to naming the $K-C_P$ and $K-C_A$ statistics as A2P-A and A2P-B and describing them as ‘shared affiliations (alternating two-paths)’ [22]. However the MPNet terminology is XACA and XACB [44, 45], again with the ‘C’ suggestive of cycles or closure, and the EstimNetDirected [74] software (<https://github.com/stivalaa/EstimNetDirected>) refers to these effects as BipartiteAltKCyclesA and BipartiteAltKCyclesB.

In the influential book edited by Lusher *et al.* [14], Wang [23] describes $K-C_A$ and $K-C_P$ explicitly as ‘alternating A cycles’ and ‘alternating P cycles’ [23], with ([23], Fig. 10.11) captioned ‘Alternating 2-paths’ but with the figure panels labelled ‘A cycles (KCA)’ and ‘P cycles (KCP)’ respectively [23]. This description or interpretation is carried over into the empirical part of the book, with Harrigan and Bond [75], applying bipartite ERGM to a director interlock network, describing $K-C_P$ and $K-C_A$ as ‘alternating k -cycles’ for directors and corporations respectively [75], and in the ERGM

results tables as ‘Director 4-cycles’ and ‘Corporation 4-cycles’ [75]. It is also notable that Harrigan and Bond [75] describes the difficulty of fitting models with $K-C_A$ and $K-C_P$ and the resulting poor goodness-of-fit for the four-cycle statistic:

We found for this particular network that we cannot have both $K-C_P$ and $K-C_A$ in the same model due to convergence issues. In line with Wang, Sharpe, Robins, and Pattison (2009), we present two alternative models, one with each possible k -cycle parameter. . . . However, in line with Wang, Sharpe, Robins, and Pattison (2009), we had a poor fit on the classic 4-cycle parameter [C4], which suggests that improving these structural effects is a substantial area of future research [75].

Given the results of the literature review described in Section 3 it would seem, however, that no such research improving these structural effects has been published yet, and this work may be the first attempt to do so.

We may conclude from this that the $K-C_A$ and $K-C_P$ statistics count too many things other than four-cycles to be usefully used and interpreted as bipartite closure in many cases. Specifically, they count simple two-paths (k -two-paths with $k = 1$) as their first (highest weighted) term. This results in situations where long paths or cycles contribute to the $K-C_P$ and $K-C_A$ statistics, despite having exactly zero four-cycles (see Table 3). Perhaps even more problematically, high-degree nodes, or stars, for example the ‘Nine-star’ structure in Table 3, result in large values of $K-C_A$ or $K-C_P$ (XACA or XACB in MPNet terminology), depending on which node set the hub node is in, but also have exactly zero four-cycles. It seems clear that large networks are likely to contain many stars, and many paths (of length two or more; and note that a two-star is just a two-path), as well as large cycles [76, 77], and these will contribute to large values of the $K-C_P$ and/or $K-C_A$ statistics, but nothing to the number of four-cycles (the C4 statistic).

In addition, Robins *et al.* [39] report that interpretation of the alternating k -two-paths parameter is problematic: ‘In this article, we do not concentrate on alternating k -two-path parameters. For some data, we have found it important to include them in models but further work is needed to understand better their effect when included with other parameters’ [39]. Martin [78] described this as the authors ‘being somewhat mystified by this statistic’ [78], and that most ERG modellers would not be able to describe a ‘clear behavioral-process analogue to the once-canonical alternating two paths statistic’ [78].

Therefore, we propose new statistics that count four-cycles (and not two-paths). In the following two sections we propose two different new statistics. In Section 5, we propose a simple new statistic based on $K-C_A$ and $K-C_P$. Unfortunately, however, this statistic is shown to be more, rather than less, prone to degeneracy than $K-C_A$ and $K-C_P$. Therefore, in Section 6 we propose a new statistic based on counting four-cycles, but which is less prone to problems with near-degeneracy than the simple four-cycles parameter or the $K-C_A$ and $K-C_P$ parameters.

5. A SIMPLE, BUT UNSUCCESSFUL, NEW STATISTIC

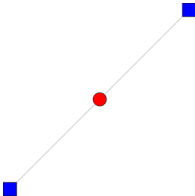
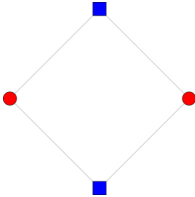
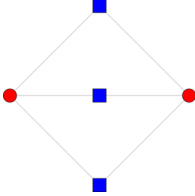
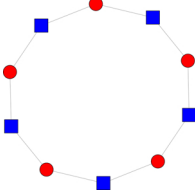
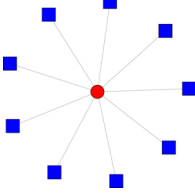
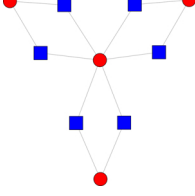
A simple solution to the problem, described in Section 4, that the $K-C_A$ (2.2) and $K-C_P$ statistics count simple two-paths as their first term (k -two-paths, $k = 1$), is to remove the first term and reverse the signs. In this way, the first, positive, term no longer counts open two-paths, but rather counts four-cycles. We define a new statistic $\text{BipartiteAltK4CyclesA}$ as

$$z_{\text{BipartiteAltK4CyclesA}}(\lambda) = - (z_{\text{BipartiteAltKCyclesA}}(\lambda) - z_{\text{TwoPathsA}}), \quad (5.1)$$

and its change statistic

$$\delta_{\text{BipartiteAltK4CyclesA}}(\lambda)(i, j) = - (\delta_{\text{BipartiteAltKCyclesA}}(\lambda) - \deg(i)) \quad (5.2)$$

Table 3. Statistics of some example bipartite networks^a

Name	Visualization	N _A	N _B	L	C4	XACA	XACB	BpNP4CA	BpNP4CB
Two-path		1	2	2	0	1	0	0	0
Four-cycle		2	2	4	1	1.5	1.5	2	2
Four-cycles-3		2	3	6	3	4.5	1.75	3.4641	4.24264
Ten-cycle		5	5	10	0	5	5	0	0
Nine-star		1	9	9	0	36	0	0	0
Four-fan-3		4	6	12	3	16.5	4.5	4.73205	6

^aL is the number of edges, C4 the number of four-cycles, and BpNP4CA and BpNP4CB are the new statistics BipartiteFourCyclesNodePowerA and BipartiteFourCyclesNodePowerB. Nodes in node set A are represented as red circles, and nodes in node set B as blue squares.

where

$$z_{\text{TwoPathsA}} = \sum_{i \in B} \sum_{\{l \in B : l < i\}} L_2(i, l) \quad (5.3)$$

is the number of two-paths connecting nodes in node set B (and which therefore go through a node in node set A), $L_2(i, l)$, defined by (2.3), is the number of two-paths connecting i to l , and $\deg(i)$, the degree of node i , is the change statistic for the number of two-paths through node i . BipartiteAltK4CyclesB and its change statistic are defined similarly.

Table 4 is a copy of Table 3, but with the new BipartiteAltK4CyclesA and BipartiteAltK4CyclesB statistics (labelled BpAK4CA and BpAK4CB respectively) included (and the N_A , N_B , L , BpNP4CA, and BpNP4CB columns removed to make space). This table shows the value of the new statistics on some small example networks, demonstrating, that, by design, they are zero for networks in which there are no four-cycles.

Unfortunately, however, simulation experiments indicate that the new BipartiteAltK4CyclesA and BipartiteAltK4CyclesB parameters are actually *more* problematic with respect to near-degeneracy than the original $K-C_A$ and $K-C_P$ parameters. Figure 1 shows the results of simulation experiments similar to those described by Wang *et al.* [21]. Bipartite networks were simulated with 30 nodes in node set A and 20 nodes in node set B with the Edge parameter set to -3.0 . In three different sets of simulations, for each of the BipartiteAltK4CyclesB, BipartiteAltK4CyclesB, and BipartiteFourCyclesNodePowerB (defined in Section 6) parameters, the parameter in question is varied from -1.00 to 10.0 in increments of 0.01 for each of two values of λ ($\lambda = 2$ and $\lambda = 5$) for BipartiteAltK4CyclesB and BipartiteAltK4CyclesB, and for each of two values of α ($\alpha = 1/2$ and $\alpha = 1/5$) for BipartiteFourCyclesNodePowerB. The networks were simulated using the SimulateERGM program from the EstimNetDirected software package, using the basic ERGM sampler, with a burn-in of 10^5 iterations and an interval of 10^4 iterations between each of 100 samples, to ensure that samples are drawn from the equilibrium ERGM distribution, and are not too autocorrelated.

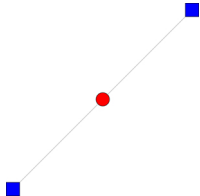
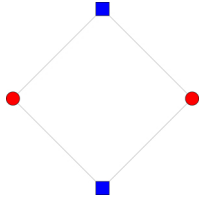
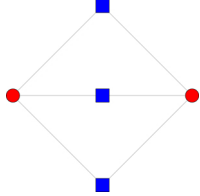
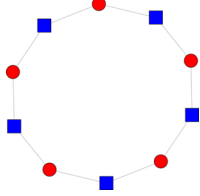
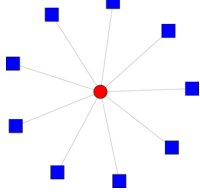
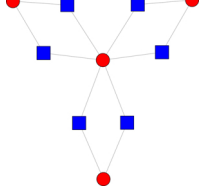
The right panel in Fig. 1a shows the results of a simulation similar to that shown in Fig. 9 of Wang *et al.* [21]: the number of edges increases smoothly, giving good coverage of the graph space (with respect to density, at least). The left panel shows the value of the statistic corresponding to the parameter BipartiteAltK4CyclesB itself. The behaviour of this curve is starting to look as if it could be prone to near-degeneracy, with a fairly steep increase at a critical value.

The graphs in Fig. 1b show the results for the new BipartiteAltK4CyclesB parameter. In this case, both the Edge statistic (right) and BipartiteAltK4CyclesB statistic itself (left) show a phase transition, where a critical value of the parameter separates an empty graph regime from a complete graph regime. This model is therefore near-degenerate, suggesting that this new parameter may in fact be less useful than the original $K-C_A$ and $K-C_P$ parameters.

For completeness, Fig. 1c shows the results for the new BipartiteFourCyclesNodePowerB parameter defined in the following section (Section 6). When $\alpha = 0.5$ an abrupt change from a near-empty to a full graph occurs, however, as shown by the result for $\alpha = 0.2$, this can be removed by decreasing the value of α .

In summary, we defined new BipartiteAltK4CyclesA and BipartiteAltK4CyclesB parameters as modified forms of the $K-C_A$ and $K-C_P$ parameters defined in Wang *et al.* [21], in order to count four-cycles but not open two-paths. However, near-degeneracy was exhibited in the simulation experiments (similar results occur with the larger simulated networks described in Section 7; data not shown). We conclude, therefore, that although these new parameters could potentially be useful in some cases, they are more prone to near-degeneracy and not as useful as the existing $K-C_A$ and $K-C_P$ parameters. Therefore, in the following section, we define new statistics that weight four-cycle counts in a different way.

Table 4. Statistics of some example bipartite networks^a

Name	Visualization	C4	XACA	XACB	BpAK4CA	BpAK4CB
Two-path		0	1	0	0	0
Four-cycle		1	1.5	1.5	0.5	0.5
Four-cycles-3		3	4.5	1.75	1.5	1.25
Ten-cycle		0	5	5	0	0
Nine-star		0	36	0	0	0
Four-fan-3		3	16.5	4.5	1.5	1.5

^aC4 is the number of four-cycles, and BpAK4CA and BpAK4CB are the new statistics BipartiteAltK4CyclesA and BipartiteAltK4CyclesB, respectively. Nodes in node set A are represented as red circles, and nodes in node set B as blue squares.

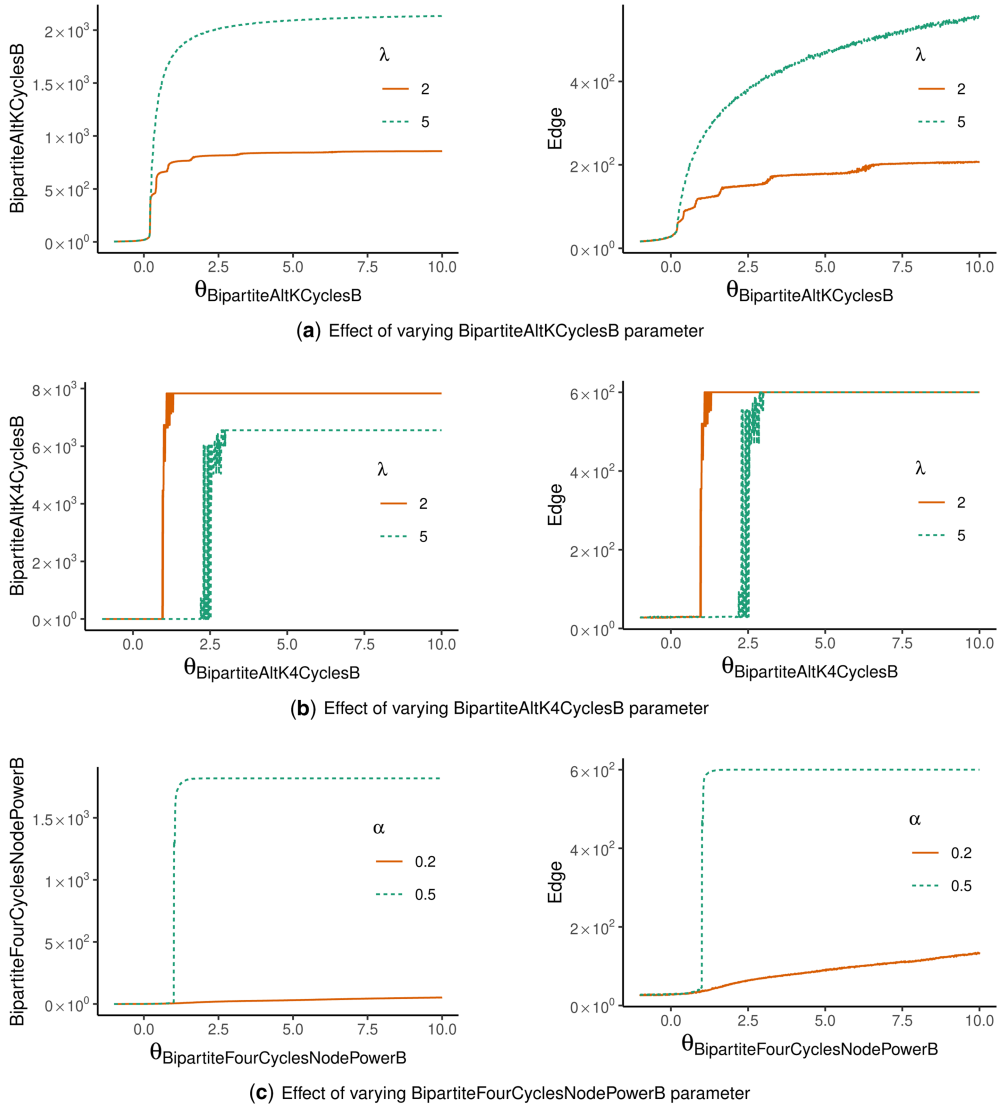


Figure 1. Effect of varying (a) the BipartiteAltKCyclesB parameter, (b) the BipartiteAltK4CyclesB parameter, and (c) the BipartiteFourCyclesNodePowerB parameter on the statistic corresponding to the parameter itself (left) and the Edge statistic (right). Each graph plots the mean value of the statistic (over 100 simulations) for two different values of the relevant λ or α parameter.

6. NEW STATISTICS FOR MODELLING FOUR-CYCLES IN BIPARTITE ERGMS

Since a four-cycle is a combination of two two-paths [33], the number of four-cycles is

$$C_4 = \frac{1}{2} \sum_{i < j} \binom{I_2(i, j)}{2}. \quad (6.1)$$

The sum in Equation (6.1) is over the $\binom{n}{2}$ pairs of nodes in the graph, with the factor of $\frac{1}{2}$ to account for the double-counting due to the symmetry of each four-cycle containing two distinct

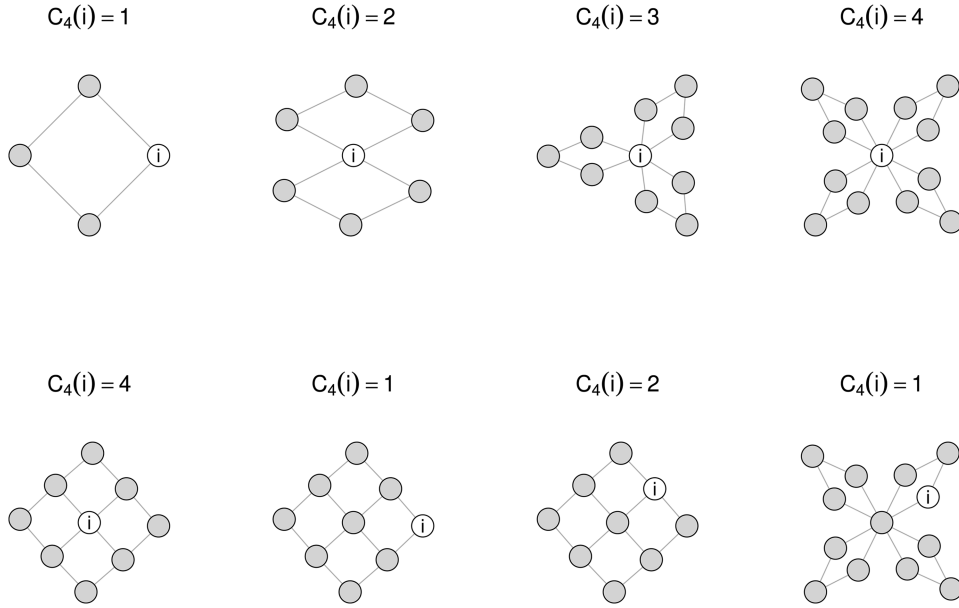


Figure 2. $C_4(i)$ is the number of four-cycles involving a node i .

pairs of nodes, each connected by two two-paths. The number of four-cycles containing a particular node i is

$$C_4(i) = \sum_{j \neq i} \binom{L_2(i, j)}{2} \quad (6.2)$$

$$= \sum_{\{j : d(i, j) = 2\}} \binom{L_2(i, j)}{2} \quad (6.3)$$

Some illustrative examples of the value of $C_4(i)$ for different nodes in some small graphs are shown in Fig. 2.

The total number of four-cycles (6.1) can also be expressed in terms of the number of four-cycles at each node (6.2) as

$$C_4 = \frac{1}{4} \sum_i C_4(i) \quad (6.4)$$

where the factor of $\frac{1}{4}$ accounts for the fact that each four-cycle is counted four times, once for each node it contains. The FourCyclesNodePower statistic is then defined as

$$z_{\text{FourCyclesNodePower}(\alpha)} = \sum_i [C_4(i)^\alpha] \quad (6.5)$$

where $0 < \alpha \leq 1$ is the exponent for, in the terminology of Wilson *et al.* [79], the ‘ α -inside’ weighting, since the subgraph counts ($C_4(i)$ here) are exponentiated before summing over all subgraphs. The ‘ α -outside’ weighting would be to exponentiate the statistic after summing over all

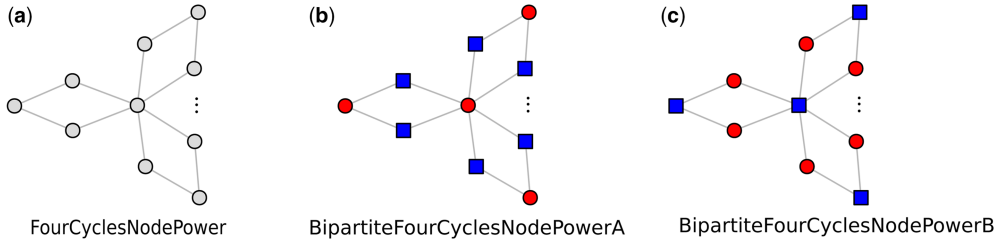


Figure 3. Representations of the new configurations (a) FourCyclesNodePower, (b) BipartiteFourCyclesNodePowerA, and (c) BipartiteFourCyclesNodePowerB. Nodes in node set A are shown as red circles, and nodes in node set B are shown as blue squares.

subgraphs, that is, in this case it would be $[\sum_i C_4(i)]^\alpha$. As discussed in Wilson *et al.* [79], the α -inside weighting leads to local dependence as usually used in ERGMs, while the α -outside weighting leads to global dependence, in which all ties are dependent on each other to some degree [79]. The nature of the local dependencies induced by the new change statistics defined here using the α -inside weighting are discussed in Section 6.3 below.

The statistics described in this section so far are equally applicable to one-mode and two-mode (bipartite) graphs. When dealing with bipartite graphs, however, it is often useful to consider statistics of the two node sets separately. Hence we also define

$$z_{\text{BipartiteFourCyclesNodePowerA}}(\alpha) = \sum_{i \in A} [C_4(i)^\alpha] \quad (6.6)$$

and

$$z_{\text{BipartiteFourCyclesNodePowerB}}(\alpha) = \sum_{j \in B} [C_4(j)^\alpha] \quad (6.7)$$

for the two node sets A and B respectively. Because the sets A and B are disjoint, we have

$$z_{\text{FourCyclesNodePower}}(\alpha) = z_{\text{BipartiteFourCyclesNodePowerA}}(\alpha) + z_{\text{BipartiteFourCyclesNodePowerB}}(\alpha). \quad (6.8)$$

Representations of the new configurations are shown in Fig. 3. The vertical ellipsis in the figures is to indicate that the configuration includes any number [up to $\lfloor (n-1)/3 \rfloor$, since, apart from the one shared node, each four-cycle must include at most three distinct nodes] of four-cycles all involving a shared node (the central node in the figures).

Table 3 shows the values of the BipartiteFourCyclesNodePowerA (BpNP4CA) and BipartiteFourCyclesNodePowerB (BpNP4CB) statistics for some small example bipartite networks, with the parameter $\alpha = 0.5$ (the alternating k -two-path statistics XACA and XACB are shown with their parameter $\lambda = 2.0$). Note that, unlike the XACA and XACB ($K-C_A$ and $K-C_B$) statistics, these new statistics have the value zero for structures that contain no four-cycles (such as the two-path, ten-cycle, and nine-star structures in Table 3).

The ‘four-fan-3’ graph in Table 3 is the same graph as the representation of the BipartiteFourCyclesNodePowerA configuration in Fig. 3b (ignoring the vertical ellipsis so that exactly three four-cycles are present). Note, however, that this has the counterintuitive property that, although it is a representation of the BipartiteFourCyclesNodePowerA configuration, in fact the value of the BipartiteFourCyclesNodePowerB statistic is greater than that of the BipartiteFourCyclesNodePowerA statistic for this graph. This is because the value of the statistic [(6.6) or (6.7)] is the sum over all

nodes in the relevant node set (A or B , respectively) of the four-cycle count at each node (6.2) raised to the power α (α -inside' weighting). Therefore, in this graph, the nodes in mode B contribute more to the total as each one (of the six) is involved in exactly one four-cycle (and hence raising to the power of α still contributes one to the sum), while of the four nodes in mode A , three are involved in only one four-cycle, while the fourth is involved in three four-cycles and hence contributes only $3^\alpha \approx 1.73205$ (when $\alpha = 0.5$). Generalizing this four-fan-3 graph to four-fan- k ($k \geq 1$, and if $k = 1$ the graph is just a four-cycle) with the central high-degree node in node set A , we have:

$$N_A = k + 1 \quad (6.9)$$

$$N_B = 2k \quad (6.10)$$

$$n = 3k + 1 \quad (6.11)$$

$$L = 4k \quad (6.12)$$

$$C_4 = k \quad (6.13)$$

$$z_{\text{BipartiteFourCyclesNodePowerA}}(\alpha) = k^\alpha + k \quad (6.14)$$

$$z_{\text{BipartiteFourCyclesNodePowerB}}(\alpha) = 2k \quad (6.15)$$

Because $0 < \alpha \leq 1$, the BipartiteFourCyclesNodePowerA statistic (6.14) will always be less than (or equal to, if $\alpha = 1$ or $k = 1$) the BipartiteFourCyclesNodePowerB statistic (6.15) for this family of graphs.

6.1. Interpretation of the new parameters

Interpretation of the FourCyclesNodePower parameter is that a positive value increases the number of four-cycles and a negative value decreases the number of four-cycles, relative to a value of zero. A smaller value of the exponent α means that additional four-cycles including the same node contribute less than if those cycles involved distinct nodes.

Interpretation of the BipartiteFourCyclesNodePowerA and BipartiteFourCyclesNodePowerB parameters is rather more complicated and is illustrated in Figs 4 and 5. These figures show statistics (Fig. 4) and network visualizations (Fig. 5) of simulated bipartite networks with different combinations of positive, zero and negative BipartiteFourCyclesNodePowerA and BipartiteFourCyclesNodePowerB parameters. The simulated networks have 100 nodes in node set A and 50 nodes in node set B and are simulated with common ERGM parameters Edge, BipartiteAltStarsA [$\lambda = 2$], and BipartiteAltStarsB [$\lambda = 2$] set to -6.0 , -0.4 , and 1.0 , respectively. For BipartiteFourCyclesNodePowerA [$\alpha = 1/5$] and BipartiteFourCyclesNodePowerB [$\alpha = 1/5$] the negative ('neg') parameter value is -1.5 and the positive ('pos') parameter value is 6.5 .

Note that in a bipartite network, any four-cycle must contain two nodes in node set A and two nodes in node set B . So how can we get more four-cycles in one mode than the other? The answer is that the four-cycle counts for the two must be equal, but the weighted node-oriented four-cycle counts (6.6) and (6.7) can differ. As discussed in Section 6, the statistic BipartiteFourCyclesNodePowerA (6.6) is maximized by having four-cycles involving distinct pairs of nodes in node set A (rather than many four-cycles involving the same node in node set A). If the BipartiteFourCyclesNodePower parameter for A is positive and for B is zero (or negative) then we tend to get more mode A nodes involved in four-cycles, with the same mode B nodes participating in many four-cycles, since the statistic is higher by having different nodes in the four-cycles, than for having the same node involved in many four-cycles. In the examples illustrated in Fig. 5, this results in the mode A nodes being part of a denser core with lots of four-cycles with a smaller number of B nodes, resulting in isolated B nodes. And vice versa for A zero (or negative) and B positive ('zero.pos' and 'neg.pos'; these are perhaps clearer as there are more A nodes than B nodes in the network). Particularly in the 'neg.pos' case, we can see a core of mode B nodes connected to a small number of

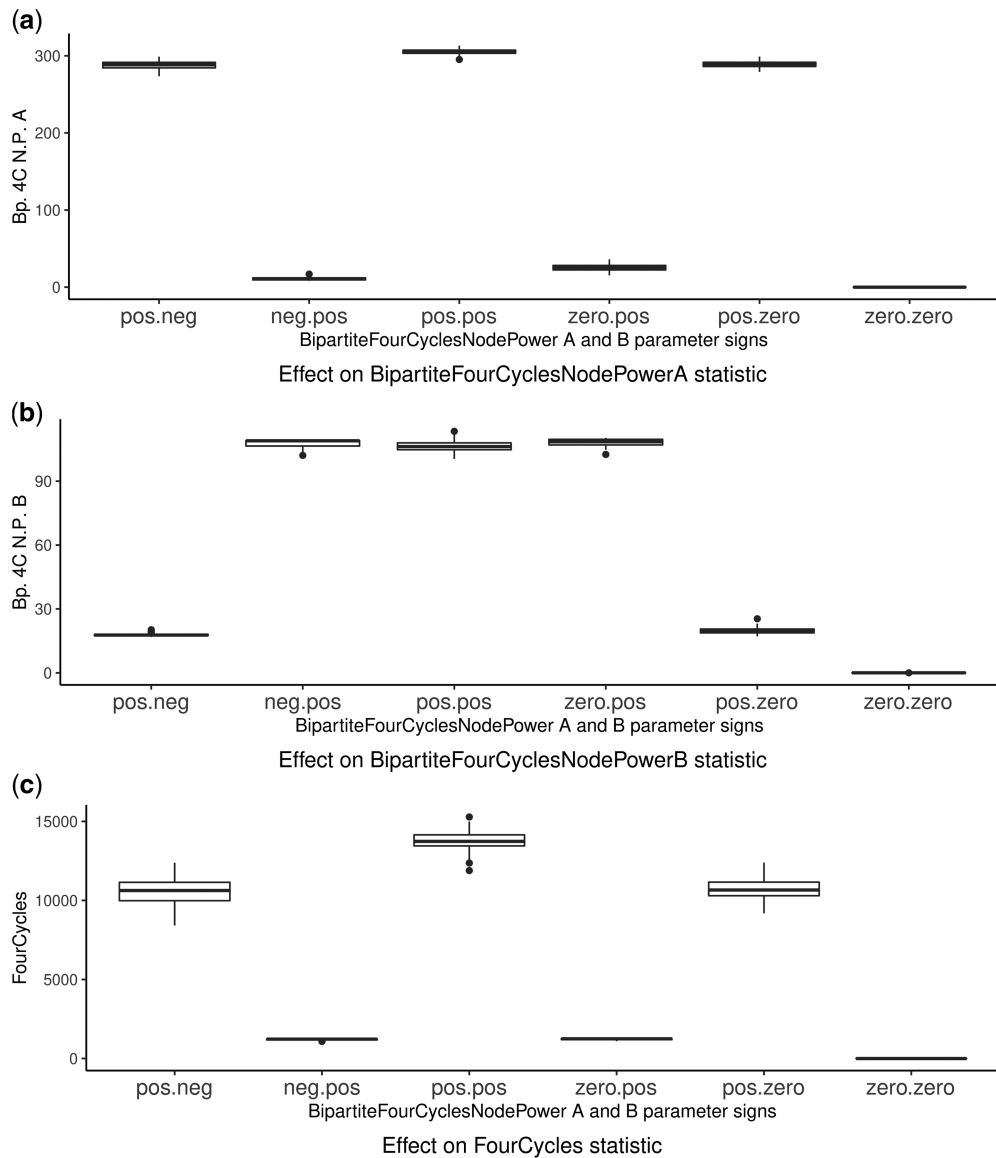


Figure 4. Effect of different combinations of negative, zero, and positive values of the BipartiteFourCyclesNodePowerA and BipartiteFourCyclesNodePowerB parameters on (a) the BipartiteFourCyclesNodePowerA statistic, (b) the BipartiteFourCyclesNodePowerB statistic and (c) the FourCycles statistic. Box plots show the statistics of 100 simulated networks.

central mode A nodes in four-cycles (as well as others not involved in four-cycles) and many mode A isolates. If BipartiteFourCyclesNodePower for both A and B are positive then there are even more four-cycles, but they are more evenly distributed between the A and B nodes.

To try to make this interpretation clearer, consider Fig. 6. The box plots in this figure show the number of unique nodes in each node set (A or B) that are involved in four-cycles. When the BipartiteFourCyclesNodePowerA parameter is positive and the BipartiteFourCyclesNodePowerB parameter is zero or negative, then a large number of node set A nodes are involved in four-cycles, but only a small number of node set B nodes are. So the same node set B nodes are involved in

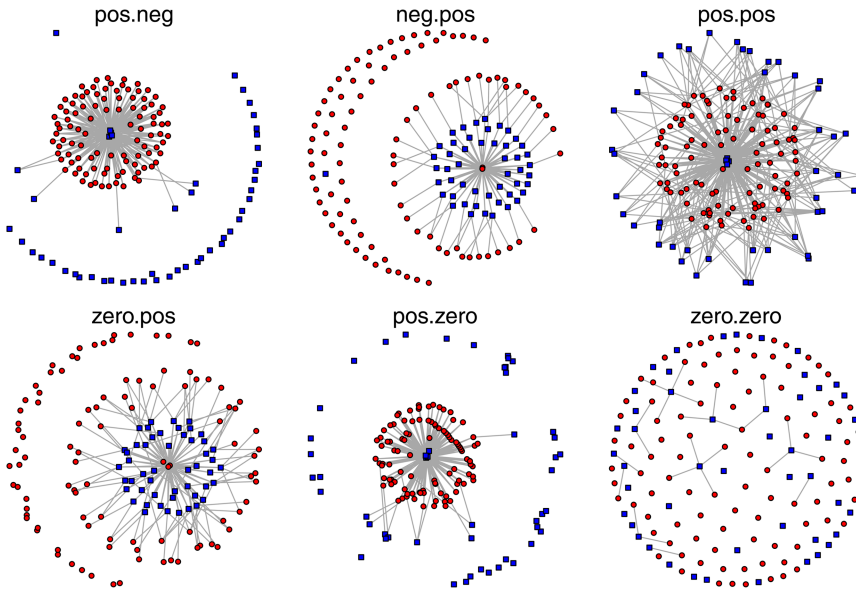


Figure 5. Examples of networks simulated with different combinations of negative, zero, and positive values of the BipartiteFourCyclesNodePowerA and BipartiteFourCyclesNodePowerB parameters, drawn from the simulations shown in Fig. 4. Nodes in node set A are shown as red circles, and nodes in node set B are shown as blue squares.

multiple four-cycles with many different node set A nodes. If BipartiteFourCyclesNodePowerB is positive and BipartiteFourCyclesNodePowerA is negative or zero, then the same nodes in mode set A are involved in multiple four-cycles with many different nodes from node set B.

Equation (6.8) implies that there are two degrees of freedom for the three parameters; for example we can include both BipartiteFourCyclesNodePowerA and BipartiteFourCyclesNodePowerB in a model, but not also FourCyclesNodePower since it is the sum of the other two.

6.2. Change statistics for the new statistics

The change statistic [28, 33, 43], that is, the difference in the statistic caused by adding a new edge (i, j) , for the four-cycles statistic (6.1) is

$$\delta_{C_4}(i, j) = \sum_{k \in N(i)} L_2(j, k) \quad (6.16)$$

$$= \sum_{k \in N(j)} L_2(i, k). \quad (6.17)$$

The change statistic for the FourCyclesNodePower statistic (6.5) is then:

$$\begin{aligned} \delta_{\text{FourCyclesNodePower}(\alpha)}(i, j) &= [C_4(i) + \delta_{C_4}(i, j)]^\alpha - C_4(i)^\alpha \\ &\quad + [C_4(j) + \delta_{C_4}(i, j)]^\alpha - C_4(j)^\alpha \end{aligned}$$

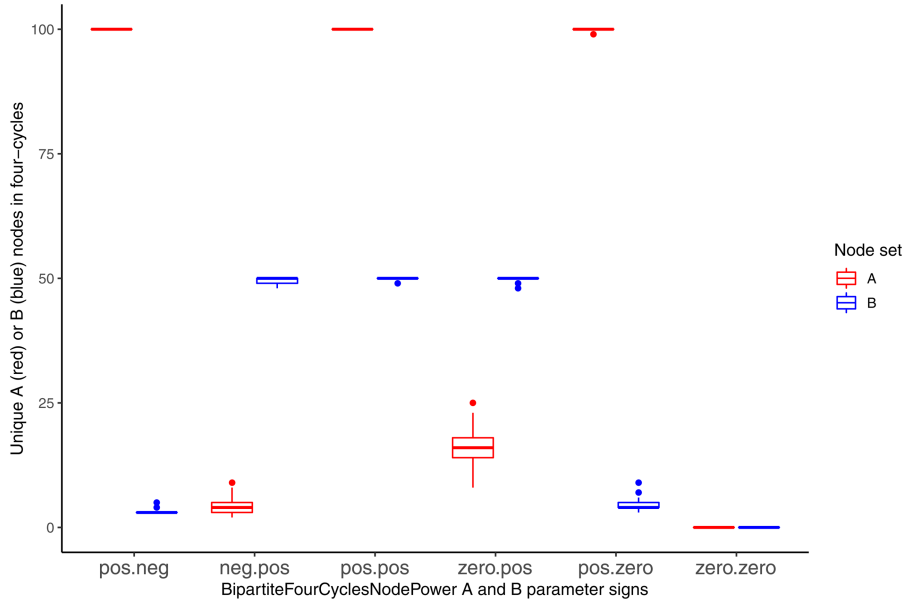


Figure 6. For each combination of negative, zero, and positive values of the BipartiteFourCyclesNodePowerA and BipartiteFourCyclesNodePowerB parameters, there are two box plots. They show the number of unique nodes in node set A (left, red) and node set B (right, blue) involved in four-cycles. The simulations are the same as those used in Fig. 4, and as in that figure, combinations of only negative and zero values are not shown as they result in extremely low density graphs with no four-cycles (similar to the zero.zero case).

$$\begin{aligned}
 & + \sum_{k \in N(i)} [(C_4(k) + L_2(k, j) + x_{kj}L_2(k, i))^\alpha - C_4(k)^\alpha] \\
 & + \sum_{\{k : k \in N(j) \wedge k \notin N(i)\}} [(C_4(k) + L_2(k, i))^\alpha - C_4(k)^\alpha]. \quad (6.18)
 \end{aligned}$$

The four terms in Equation (6.18) count the contributions from, respectively, node i , node j , the neighbours of node i , and the neighbours of node j which are not also neighbours of node i . Note that in the third term (the contribution from neighbours of node i), a node k can only be a neighbour of both node i and node j [that is, $k \in N(i) \wedge x_{kj} = 1$] if the network is not bipartite.

The change statistic for the bipartite four-cycles statistic for the node set A (6.6) is simpler than the general case (6.18), as we only count the contributions from the nodes in node set A. Specifically, we have:

$$\begin{aligned}
 \delta_{\text{BipartiteFourCyclesNodePowerA}(\alpha)}(i, j) &= [C_4(i) + \delta_{C_4}(i, j)]^\alpha - C_4(i)^\alpha \\
 & + \sum_{k \in N(j)} [(C_4(k) + L_2(k, i))^\alpha - C_4(k)^\alpha] \quad (6.19)
 \end{aligned}$$

where $i \in A$, $j \in B$, and $k \in A$. The change statistic for BipartiteFourCyclesNodePowerB (6.7) is defined analogously.

6.3. Position of the new statistics in the dependence hierarchy

The configurations allowed in a model are determined by the assumptions as to which ties are allowed to depend on which other ties. Pattison and Snijders [25] (subsequently elucidated by

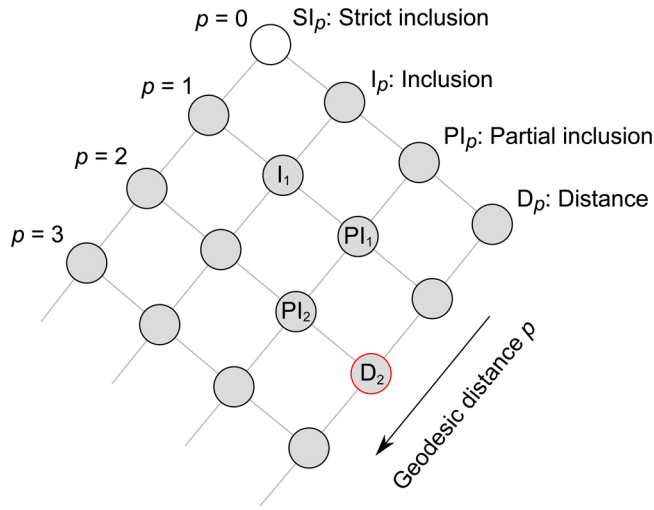


Figure 7. Dependence hierarchy, adapted from Wang *et al.* ([22], Fig. 3). SI_0 is not defined.

Wang *et al.* [22] for bipartite networks, and more recently by Pattison *et al.* [80]) created a two-dimensional hierarchy of dependence assumptions, where the two dimensions are two facets of proximity: the form of the proximity condition, and the maximum distance between dependent ties. This two-dimensional hierarchy of dependence assumptions is illustrated in Fig. 7, showing a partial order structure, in which, if one dependence condition can be implied by another, it can be reached by a downwards path from the first to the second [22].

In order to describe the proximity conditions, it is useful to define some notation for neighbourhoods in a graph. In Section 2 we defined $N(i)$ as the neighbours of node i , that is, nodes $k \neq i$ such that $d(i, k) = 1$. Following Pattison *et al.* [80], we now define $N_q(u)$, the q -neighbourhood of node u , as the set of nodes within geodesic distance q of u , that is, nodes v such that $d(u, v) \leq q$. Note that $N(u)$ as earlier defined in Section 2 is distinct from $N_1(u)$, the former being defined as nodes with a geodesic distance of exactly 1 from u (and hence excluding u itself), while the latter includes u itself as $d(u, u) = 0$. Further, we define $N_q(U)$, the q -neighbourhood of node set U , as $N_q(U) = \{v : v \in N_q(u) \text{ for some } u \in U\}$, that is, the set of nodes whose distance to some node in U is no more than q .

The four forms of proximity conditions, describing the nature of the proximity between the neighbourhoods of two pairs of nodes whose respective tie variables are hypothesized to be conditionally dependent only if the proximity condition holds, can be summarized as follows [80], in order of increasing generality:

1. Strict p -inclusion. SI_p ($p \geq 1$) holds if the p -neighbourhood of each node in a pair includes both of the nodes in the other pair.
2. p -inclusion. I_p ($p \geq 0$) holds if the p -neighbourhood of each pair of nodes includes the other pair.
3. Partial p -inclusion. PI_p ($p \geq 0$) holds if the p -neighbourhood of one pair of nodes includes the other pair.
4. p -proximity. D_p ($p \geq 0$) holds if the p -neighbourhood of one pair of nodes has a non-empty intersection with the other pair.

The dependence condition I_1 is equivalent to the widely-used ‘social circuit’ dependence assumption [22], in which two ties are conditionally dependent if they form a four-cycle if both present [33,

40, 81, 82], and PI_1 allows the ‘alternating pendant-triangle’ statistics recently described by Pattison et al. [80].

The simple but unsuccessful new statistics described in Section 5 are in the same dependence class (I_1 , ‘social circuit’) as the original $K-C_A$ and $K-C_P$ statistics from which they are derived.

The change statistic for FourCyclesNodePower (6.18) in computing the probability of a new edge (i, j) , depends only on edges between nodes in the two-neighbourhoods of i and of j , since any two nodes in a four-cycle must be at a geodesic distance of at most two from each other (a four-cycle is a pair of nodes with two two-paths between them). Hence any edges on which the probability of edge (i, j) depend must be in the two-neighbourhood of $\{i, j\}$. This puts the FourCyclesNodePower configuration in the dependence class two-proximity (D_2): the p -neighbourhood (with $p = 2$) of one pair of nodes has a non-empty intersection with the other pair [25, 80]. D_2 is labelled with a red outline in the dependence hierarchy diagram shown in Fig. 7.

Because of the partial order structure of the dependence hierarchy, the stricter proximity forms for a fixed p imply the more general ones (and D_p is the most general), and for a fixed proximity condition, smaller p implies all the larger p (so D_1 implies D_2 for instance). Hence, to show that the FourCyclesNodePower configuration, which is in D_2 , is not also in any more specific dependence class, it is sufficient to show that it is not in D_1 and also not in PI_2 (see Fig. 7). To do so, we can use Proposition 3 of Pattison et al. [80], which gives the properties that must hold for configurations implied by the dependence structures associated with each proximity condition.

Proposition 3(d) of Pattison et al. [80] states that ‘For D_p , each configuration is a subgraph in which every pair of edges lies on a path of length $\leq (p + 2)$ ’ [80]. So for a configuration to be in D_1 , every pair of edges must lie on a path of length three (or shorter). The FourCyclesNodePower configuration (Fig. 3a) does not meet this requirement, since there are edges that do not lie on a path of length three or less: consider, for example, a pair of edges incident to the outermost node in the figure on two different four-cycles. These do not lie on a path of length three (but are on a path of length four, satisfying the requirement for D_2 but not D_1). Hence the FourCyclesNodePower configuration is not in dependence class D_1 .

Proposition 3(c) of Pattison et al. [80] states that ‘For PI_p , each configuration is a subgraph in which every pair of edges lies either on a cyclic walk of length $\leq (2p + 2)$ or on a cyclic walk of length $\leq 2(p - r) + 1$ with an additional path of length $\leq r + 1$ attached to a node lying on the cyclic walk, for $0 \leq r \leq p - 1$ ’ [80]. So, for PI_2 , each configuration is a subgraph in which every pair of edges is on a cyclic walk of length $\leq (2p + 2) = 6$, or on a cyclic walk of length $\leq 2(p - 0) + 1 = 5$ with an additional path of length ≤ 1 attached to a node on the cyclic walk, or on a cyclic walk of length $\leq 2(p - 1) + 1 = 3$ with an additional path of length ≤ 2 attached to a node on the cyclic walk. Again, we can see that the configuration for FourCyclesNodePower does not meet these conditions, considering a pair of maximally distant edges (those incident to the outermost nodes in two different four-cycles in Fig. 3a). Such a pair of edges is neither on a six-cycle, and nor is it on a five-cycle with an additional path of length at most one or a three-cycle with an additional path of length at most two. Hence the FourCyclesNodePower configuration is not in dependence class PI_2 .

In the case of the BipartiteFourCyclesNodePower (A and B) statistics for bipartite networks, the same reasoning applies (see Fig. 3b and c).

6.4. Implementation

The new ERGM effects FourCyclesNodePower, BipartiteFourCyclesNodePowerA, and BipartiteFourCyclesNodePowerB are implemented in the EstimNetDirected [74] software, available from <https://github.com/stivalaa/EstimNetDirected>. BipartiteFourCyclesNodePowerA and BipartiteFourCyclesNodePowerB are also implemented, as b1np4c and b2np4c, as user-contributed statnet model terms [83, 84], available from <https://github.com/stivalaa/ergm.terms.contrib>. An example is described in Section 8.

To count the number of unique type *A* and *B* nodes that are involved in four-cycles in a two-mode network, the CYPATH software (<http://research.nii.ac.jp/~uno/code/cypath.html>) [85] was used to enumerate all of the four-cycles (which are necessarily chordless in a bipartite network).

Scripts for data conversion, statistical analysis, and generating plots were written in R [86] using the igraph [87, 88] and ggplot2 [89] packages.

7. SIMULATION EXPERIMENTS

In order to investigate the effect of the `BipartiteFourCyclesNodePowerA` parameter, and compare it to that of the `K-CA` [21] (known as `XACA` in `MPNet` and `BipartiteAltKCyclesA` in `EstimNet-Directed`) parameter, we conducted some simulation experiments. In these experiments, bipartite networks with 750 nodes in node set *A* and 250 nodes in set *B* were simulated with the `Edge`, `BipartiteAltStarsA` [$\lambda = 2$], and `BipartiteAltStarsB` [$\lambda = 2$] parameters set to -8.50 , -0.20 , and 2.00 , respectively. In one set of experiments, the `BipartiteAltKCyclesA` parameter was varied from -1.00 to 1.00 in increments of 0.01 , for each of three values of λ : 2 , 5 , and 10 . In another set of experiments, the `BipartiteFourCyclesNodePowerA` parameter was varied from -1.00 to 2.00 in increments of 0.01 , for each of three values of α : $1/10$, $1/5$, and $1/2$. The networks were simulated using the `SimulateERGM` program from the `EstimNetDirected` software package, using the tie/no-tie (TNT) sampler [90], with a burn-in of 10^7 iterations and an interval of 10^5 iterations between each of 100 samples, to ensure that samples are drawn from the equilibrium ERGM distribution, and are not too autocorrelated.

The results of these simulations are shown in Fig. 8. Using the `BipartiteAltKCyclesA` parameter (left column) results in phase transition or near-degeneracy behaviour, with the statistic showing a sudden sharp increase at a critical value of the parameter. At this critical value, the graph density (Fig. 8a left plot) also sharply increases, as does the number of four-cycles (Fig. 8c left plot), which, at parameter values less than the critical value, hardly increased at all. This behaviour is similar to that of the simple Markov (edge-triangle) model described in Koskinen and Daraganova [40], and is characteristic of near-degeneracy in ERGMs. This can prevent estimation of models which contain parameters that cause this behaviour, and yet occurs in this case even when using an ‘alternating’ statistic, designed to try to avoid such behaviour [21]. Note that changing the λ parameter appears merely to change the maximum value of the statistic; it does not remove or ‘smooth out’ the phase transition. In contrast, when using the new `BipartiteFourCyclesNodePowerA` parameter (right column) the phase transition behaviour is less apparent even for the highest value of α ($1/2$), and can be smoothed out further as we decrease α . So, by appropriately setting α , we can use the `BipartiteFourCyclesNodePowerA` parameter to generate smoothly varying numbers of four-cycles, without suddenly tipping from a low-density low-clustering regime to a high-density high-clustering regime with nothing in between, which did not seem to be possible on this example with the `BipartiteAltKCyclesA` parameter.

8. EMPIRICAL EXAMPLE

The new statistics `BipartiteFourCyclesNodePowerA` and `BipartiteFourCyclesNodePowerB` were implemented in `statnet` as `b1np4c` and `b2np4c`, using the facility to define custom `ergm` model terms [83, 84]. They are available at <https://github.com/stivalaa/ergm.terms.contrib>.

Here we demonstrate this implementation by using the `statnet` `ergm` package [51] with the new user terms to estimate a model for the Davis ‘Southern Women’ network [91], obtained via the `latentnet` R package [92, 93]. The network represents the participation of 18 women (first mode) in 14 social events (second mode). This is a well-known affiliation network [94], having been used as an example by Breiger [1] and many papers since, including two [21, 95] in the literature survey (Supplementary Table SA1). Although the original publication [91] contains information on the event times, with some exceptions [94, 96] this is not usually used, and we do not use this information here.

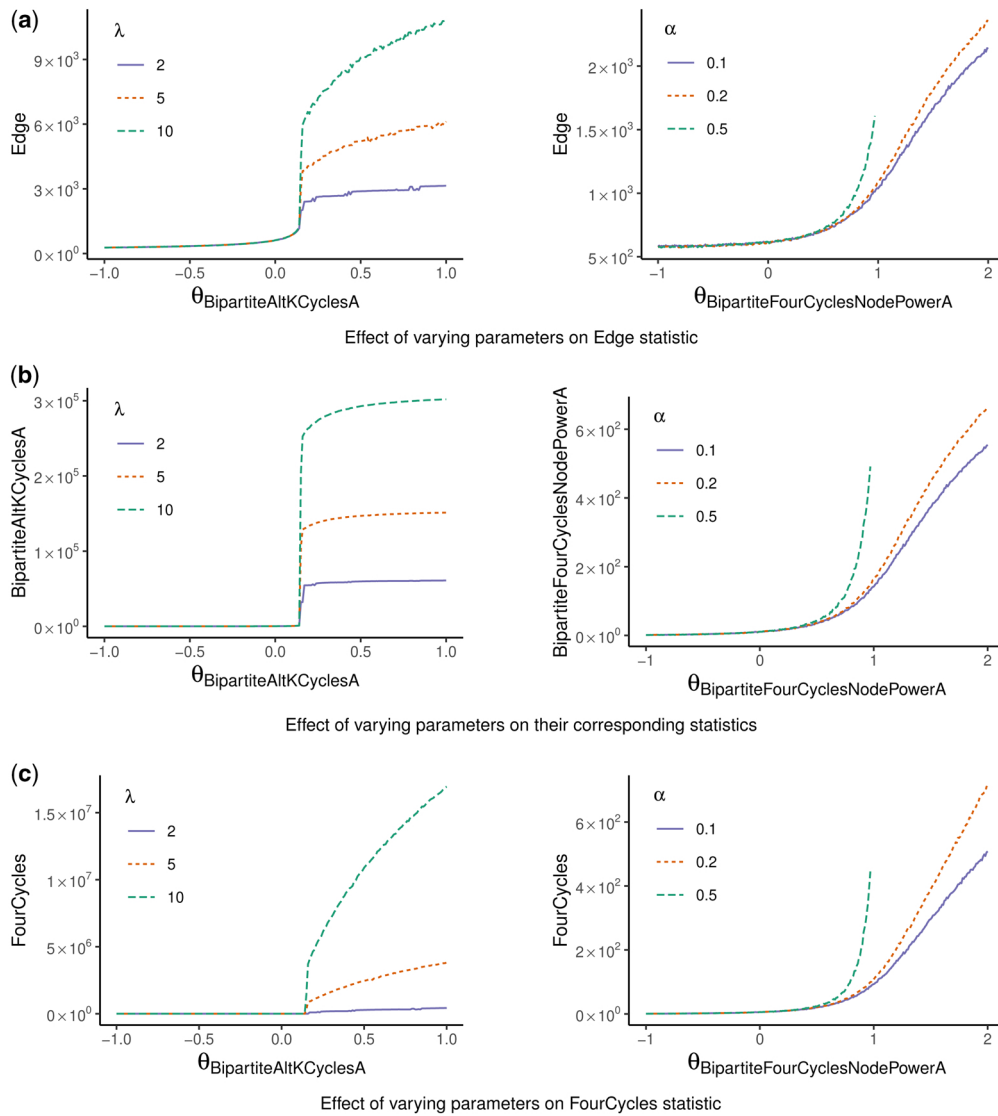


Figure 8. Effect of varying the BipartiteAltKCyclesA parameter (left) and BipartiteFourCyclesNodePowerA parameter (right) on (a) the Edge statistic, (b) the statistic corresponding to the parameter itself and (c) the FourCycles statistic. Each graph plots the mean value of the statistic (over 100 simulations) for three different values of the λ or α parameter for BipartiteAltKCyclesA and BipartiteFourCyclesNodePowerA, respectively.

The b1np4c and b2np4c model terms take the α ($0 < \alpha \leq 1$) value as a parameter, with a default value of $\alpha = 0.5$ if omitted. For example, to estimate a model with the b2np4c (BipartiteFourCyclesNodePowerB) term with $\alpha = 1/5$ (Model 4 in Table 5):

```
davis_model4 <- ergm(davis ~ edges + gwb1degree(1, TRUE) +
                     gwb2degree(1, TRUE) + b2np4c(1/5),
                     control = control.ergm(main.method = "Stochastic-Approximation"))
```

Table 5. ERGM parameter estimates for the Southern Women network, estimated with statnet. The table was generated directly from the statnet models with the texreg R package [97].

	Model 1	Model 2	Model 3	Model 4
edges	-2.07*** (0.34)	-0.20 (0.24)	0.47 (0.31)	-5.90*** (0.58)
b1star2	0.07 (0.07)			
b2star2	0.18*** (0.04)			
gwb1deg.fixed.1		-0.83 (1.03)	-7.04*** (1.51)	10.60*** (1.80)
gwb2deg.fixed.1		-2.26** (0.85)	7.76** (2.48)	-6.95*** (1.57)
gwb1dsp.fixed.0.5			0.45*** (0.12)	
gwb2dsp.fixed.0.5			-1.33*** (0.30)	
b2np4c.fixed.0.2				17.30*** (1.86)
AIC	319.34	328.85	308.77	285.41
BIC	329.93	339.44	326.41	299.53
Log Likelihood	-156.67	-161.43	-149.38	-138.71

*** $P < 0.001$; ** $p < 0.01$; * $p < 0.05$.

Table 5 shows four models for the Southern Women network, estimated with the stochastic approximation algorithm [27]. Model 1, with only the edges and two-star terms for each mode, is the same as ([21], Model (8.3)). Model 2, using the geometrically weighted degree terms rather than two-stars, is similar to ([21], Model (8.5)), but using the statnet gwb1degree and gwb2degree terms rather than the BPNet alternating k -star terms $K\text{-}S_p$ and $K\text{-}S_A$. Note the reversal of interpretation of signs between $K\text{-}S_p/K\text{-}S_A$ and gwb1degree/gwb2degree [98, 99]. Model 3, adding the geometrically weighted dyadwise shared partner terms, is similar to ([21], Model (8.6)). Model 4 uses the new b2np4c term rather than the geometrically weighted dyadwise shared partner terms gwb1dsp and gwb2dsp (models with b1np4c did not converge).

Cycle length distribution goodness-of-fit plots for the four models are shown in Fig. 9, and statnet goodness-of-fit plots in Supplementary Fig. SB1 (Supplementary Appendix SB). Note that all four models fit acceptably well on all the statistics included in the goodness-of-fit tests (degree distributions for each mode, dyadwise shared partners, and geodesic distance distribution), as well as the cycle length distributions. In particular, the fit to four-cycle counts is good for all models, including Model 1 and Model 2, which do not contain any terms to model four-cycles. It therefore appears that Model 1 is the most parsimonious explanation of this data, just as discussed in Wang *et al.* [21]. This model has a positive and statistically significant event two-star parameter (b2star2), indicating ‘greater discrepancies in the popularity of events than expected in a random network’ [21], taking into account the density (edges) and actor two-star (b1star2) effects.

That ERG models containing only terms to model density and degree distributions (and specifically not any terms to model dyadwise shared partner distributions or four-cycles) also fit the four-cycle count well indicates that the observed number of four-cycles could have occurred simply by chance [21]. Note that the same applies also to six-cycles. This is consistent with the results for this network described by Opsahl [10], where the observed value of the two-mode global clustering coefficient defined in that paper is not extreme in the distribution of that coefficient in random networks.

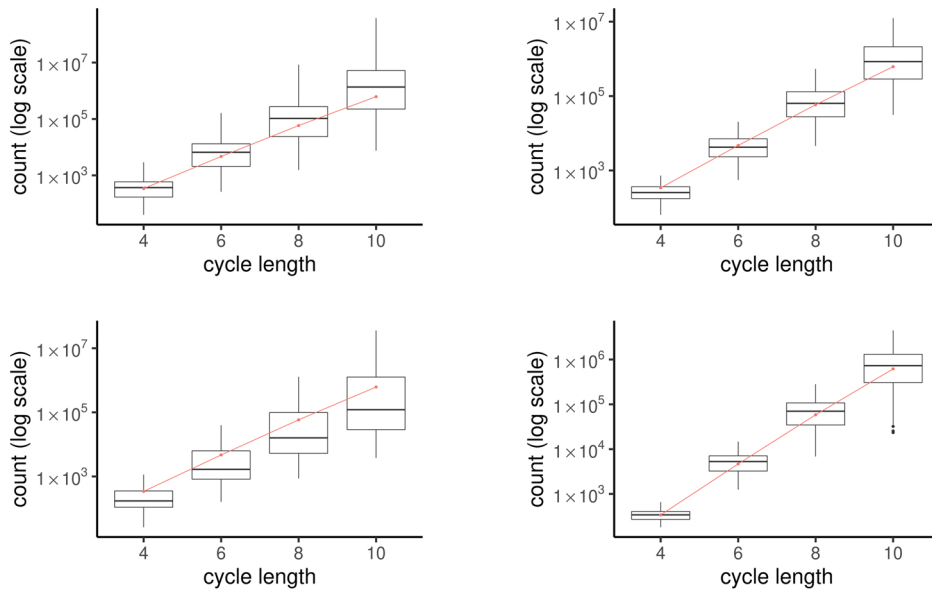


Figure 9. Cycle length distribution goodness-of-fit plots for the Southern Women network ERGM (Table 5) Model 1 (top left), Model 2 (top right), Model 3 (bottom left), and Model 4 (bottom right). Observed network statistics are plotted as red points (joined by red lines as a visual aid) with the statistics of 100 simulated networks plotted as black box plots.

9. CONCLUSION

The existing parameters for modelling shared partners or four-cycles in ERGMs for two-mode networks frequently lead to convergence problems, especially when parameters for both modes are included in the model. A literature survey shows that the majority of published ERG models of two-mode networks do not include these parameters. In addition, the majority of published models do not include an assessment goodness-of-fit to four-cycles.

In this work, we defined new ERGM effects to explicitly model four-cycles, and in the case of two-mode networks, four-cycle counts for the two modes separately. Simulation experiments show that ERGMs using the new parameters are able to generate bipartite networks with smoothly varying numbers of four-cycles, where the existing parametrisations cannot.

These new parameters come with both conceptual and computational costs. The conceptual cost is having to use a more general dependence class in the dependence hierarchy, without an underlying theoretical justification. The social circuit dependence assumption has theoretical justifications, as does partial inclusion (PI_1) for pendant-triangle configurations [80], but we are forced to use the more general class D_2 simply by the dependency induced by the new configuration, without any theoretical basis.

It is notable that our attempt to solve the problems with near-degeneracy observed with the existing $K-C_A$ and $K-C_P$ parameters with new parameters in the same dependence class (I_1 , ‘social circuit’) failed due to these parameters being even more prone to near-degeneracy (Section 5). These new parameters were motivated by the observation (Section 4) that the $K-C_A$ and $K-C_P$ statistics include not just four-cycles, but also simple two-paths (k -two-paths with $k = 1$), and that this could be a cause of near-degeneracy with these parameters. However, as shown in Section 5, modifying the $K-C_A$ and $K-C_P$ parameters so that they only include four-cycles (and remain in the same dependency class) did not solve the problem. The problem was only solved (Section 6) by abandoning the dyadwise shared partner structure of the $K-C_A$ and $K-C_P$ statistics, and moving instead to a node-oriented structure, counting the number of four-cycles in which each node is

involved (and using ‘ α -inside’ weighting). This has the consequence that these new statistics are no longer in the I_1 dependence class, but rather in the more general D_2 class. Of course, just because we were unable to solve the problems with the $K-C_A$ and $K-C_P$ parameters without creating a statistic in a more general dependence class does not mean it is impossible to do so, and an interesting direction for future work could be to either try to find a statistic for four-cycles that is not prone to near-degeneracy and is in the I_1 (or perhaps PI_1 , PI_2 , I_2 or D_1) dependence class, or to try to prove that it is not possible.

The computational cost is incurred by having to traverse the two-neighbourhood of a dyad in computing the change statistic, and even with precomputation of two-path counts [74], this can be prohibitively slow (or require impractical amounts of memory) on dense and/or high four-cycle count or very large networks, or those with very high degree nodes.

Another shortcoming is the counterintuitive and possibly confusing interpretation of the new parameters for bipartite networks (`BipartiteFourCyclesNodePowerA` and `BipartiteFourCyclesNodePowerB`; see Section 6.1). The new parameter for one-mode networks (`FourCyclesNodePower`) is also applicable to two-mode networks but does not treat the modes separately, and is relatively straightforward to interpret simply as a four-cycle closure parameter, so one solution is to use this, but this could omit important differences in the modes in a two mode network. Even the relatively straightforward `gwdegree` parameter used in `statnet` causes significant confusion due to the counterintuitive meaning of its sign [78, 98–100]. Having to use such relatively complex statistics, such as alternating, geometrically weighted, or the ‘ α -outside’ or ‘ α -inside’ [79] weightings often solves the problem of near-degeneracy, but comes at the cost of making the statistics difficult to interpret [38]. An alternative solution to the conceptual (and possibly computational) problem is to use instead the LOLOG model [62, 63] or tapered ERGM [34, 38, 101] where a simple four-cycle parameter is unlikely to cause near-degeneracy problems, much as these new models enable a simple triangle parameter in one-mode networks where alternating or geometrically weighted parameters are required in standard ERGMs [38, 62, 102]. LOLOG, however, does not (currently) handle bipartite networks [62, 103], while tapered ERGM has the advantage of being implemented in the `statnet` framework and can use any terms in the `statnet` ergm package.

One further shortcoming of the new parameters is that the weighting parameter α is fixed, and not estimated as part of the model. A potential avenue of future work is to explore the possibility of estimating this parameter in the context of a curved ERGM in the `statnet` ergm package.

One final issue is fitting cycles of length larger than four in bipartite networks. Of particular importance are six-cycles, which have been suggested as the basis (rather than four-cycles) for measuring closure in two-mode networks [10, 12, 13]. The dependence class required for the new statistics described in this work also admits six-cycles as configurations, however we did not attempt to fit models with six-cycles as a parameter. It seems likely that, like the simple four-cycles parameter, attempting to do so would lead to problems with near-degeneracy, necessitating the creation of another weighted configuration analogous to those defined here for four-cycles (which would then be in another, even more general, dependence class).

ACKNOWLEDGEMENTS

This work was performed on the OzSTAR national facility at Swinburne University of Technology. The OzSTAR program receives funding in part from the Astronomy National Collaborative Research Infrastructure Strategy (NCRIS) allocation provided by the Australian Government, and from the Victorian Higher Education State Investment Fund (VHESIF) provided by the Victorian Government.

SUPPLEMENTARY DATA

[Supplementary data](#) is available at *COMNET Journal* online.

FUNDING

This work was supported by the Swiss National Science Foundation [200778].

REFERENCES

1. Breiger RL. The duality of persons and groups. *Soc Forces* 1974;**53**:181–90.
2. Neal ZP, Cadieux A, Garlaschelli D et al. Pattern detection in bipartite networks: a review of terminology, applications, and methods. *PLOS Complex Syst* 2024;**1**:e0000010.
3. Wang P, Robins G, Pattison P et al. Exponential random graph models for multilevel networks. *Soc Netw* 2013;**35**:96–115.
4. Latapy M, Magnien C, Del Vecchio N. Basic notions for the analysis of large two-mode networks. *Soc Netw* 2008;**30**:31–48.
5. Everett MG, Borgatti SP. The dual-projection approach for two-mode networks. *Soc Netw* 2013;**35**:204–10.
6. Everett MG. Centrality and the dual-projection approach for two-mode social network data. *Method Innov* 2016;**9**:2059799116630662.
7. Granovetter MS. The strength of weak ties. *Am J Sociol* 1973;**78**:1360–80.
8. Weisstein EW. (2024) Bipartite Graph. From MathWorld—A Wolfram Web Resource. <https://mathworld.wolfram.com/BipartiteGraph.html>
9. Robins G, Alexander M. Small worlds among interlocking directors: network structure and distance in bipartite graphs. *Comput Math Organ Theory* 2004;**10**:69–94.
10. Opsahl T. Triadic closure in two-mode networks: redefining the global and local clustering coefficients. *Soc Netw* 2013;**35**:159–67.
11. Koskinen J, Edling C. Modelling the evolution of a bipartite network—Peer referral in interlocking directorates. *Soc Netw* 2012;**34**:309–22.
12. Vasques Filho D, O’Neale DRJ. Transitivity and degree assortativity explained: the bipartite structure of social networks. *Phys Rev E* 2020a;**101**:052305.
13. Vasques Filho D, O’Neale DR. The role of bipartite structure in R&D collaboration networks. *J Complex Netw* 2020b;**8**:cnaa016.
14. Lusher D, Koskinen J, Robins G. (eds.). *Exponential Random Graph Models for Social Networks: Theory, Methods, and Applications. Structural Analysis in the Social Sciences*. New York: Cambridge University Press, 2013.
15. Amati V, Lomi A, Mira A. Social network modeling. *Annu Rev Stat Appl* 2018;**5**:343–69.
16. Koskinen J. Exponential random graph modelling. In: Atkinson P, Delamont S, Cernat A et al. (eds.), *SAGE Research Methods Foundations*. London: SAGE, 2020. <https://doi.org/10.4135/9781526421036888175>
17. Koskinen J. Exponential random graph models. In: McLevey J, Scott J, Carrington PJ (eds.), *The Sage Handbook of Social Network Analysis*, 2nd edn., chapter 33. London: Sage, 2023.
18. Cimini G, Squartini T, Saracco F et al. The statistical physics of real-world networks. *Nat Rev Phys* 2019;**1**:58–71.
19. Ghafouri S, Khasteh SH. A survey on exponential random graph models: an application perspective. *PeerJ Comput Sci* 2020;**6**:e269.
20. Giacomarra F, Bet G, Zocca A. Generating synthetic power grids using exponential random graph models. *PRX Energy* 2024;**3**:023005.
21. Wang P, Sharpe K, Robins GL et al. Exponential random graph (p*) models for affiliation networks. *Soc Netw* 2009;**31**:12–25.
22. Wang P, Pattison P, Robins G. Exponential random graph model specifications for bipartite networks—a dependence hierarchy. *Soc Netw* 2013;**35**:211–22.
23. Wang P. (2013) Exponential random graph model extensions: models for multiple networks and bipartite networks. In: Lusher D, Koskinen J, Robins G (eds.), *Exponential Random Graph Models for Social Networks: Theory, Methods, and Applications*, chapter 10. New York: Cambridge University Press, 115–29.
24. Bomirya RP, Kuvelkar AR, Hunter DR, et al. Modeling Homophily in Exponential-Family Random Graph Models for Bipartite Networks. arXiv, arXiv:2312.05673v1, 2023, preprint: not peer reviewed.
25. Pattison PE, Snijders T. Modeling social networks: next steps. In: Lusher D, Koskinen J, Robins G, (eds.), *Exponential Random Graph Models for Social Networks: Theory, Methods, and Applications*, chapter 22. New York: Cambridge University Press, 2013, 287–301.
26. Geyer CJ, Thompson EA. Constrained Monte Carlo maximum likelihood for dependent data. *J R Stat Soc B* 1992;**54**:657–83.
27. Snijders TAB. Markov chain Monte Carlo estimation of exponential random graph models. *J Soc Struct* 2002;**3**:1–40.

28. Hunter DR, Krivitsky PN, Schweinberger M. Computational statistical methods for social network models. *J Comput Graph Stat* 2012;**21**:856–82.
29. Byshkin M, Stivala A, Mira A *et al.* Auxiliary parameter MCMC for exponential random graph models. *J Stat Phys* 2016;**165**:740–54.
30. Byshkin M, Stivala A, Mira A *et al.* Fast maximum likelihood estimation via Equilibrium Expectation for large network data. *Sci Rep* 2018;**8**:11509.
31. Borisenko A, Byshkin M, Lomi A. A simple algorithm for scalable Monte Carlo inference. arXiv:1901.00533v4, 2020, preprint: not peer reviewed.
32. Handcock MS. (2003) Assessing degeneracy in statistical models of social networks. Working Paper no. 39, Center for Statistics and the Social Sciences, University of Washington. <https://csss.uw.edu/Papers/wp39.pdf>
33. Snijders TAB, Pattison PE, Robins GL *et al.* New specifications for exponential random graph models. *Sociol Methodol* 2006;**36**:99–153.
34. Fellows I, Handcock M. Removing phase transitions from Gibbs measures. In: Singh A, Zhu J (eds.), *Proceedings of the 20th International Conference on Artificial Intelligence and Statistics, Fort Lauderdale, FL, USA, volume 54 of Proceedings of Machine Learning Research*, 2017, 289–97.
35. Schweinberger M. Instability, sensitivity, and degeneracy of discrete exponential families. *J Am Stat Assoc* 2011;**106**:1361–70.
36. Chatterjee S, Diaconis P. Estimating and understanding exponential random graph models. *Ann Stat* 2013;**41**:2428–61.
37. Schweinberger M. Consistent structure estimation of exponential-family random graph models with block structure. *Bernoulli* 2020;**26**:1205–33.
38. Blackburn B, Handcock MS. Practical network modeling via tapered exponential-family random graph models. *J Comput Graph Stat* 2023;**32**:388–401.
39. Robins G, Snijders TAB, Wang P *et al.* Recent developments in exponential random graph (p^*) models for social networks. *Soc Netw* 2007;**29**:192–215.
40. Koskinen J, Daraganova G. Exponential random graph model fundamentals. In: Lusher D, Koskinen J, Robins G (eds.), *Exponential Random Graph Models for Social Networks: Theory, Methods, and Applications*, chapter 6. New York: Cambridge University Press, 2013, 49–76.
41. Hunter DR. Curved exponential family models for social networks. *Soc Netw* 2007;**29**:216–30.
42. Stivala A. Overcoming near-degeneracy in the autologistic actor attribute model. arXiv:2309.07338v2, 2023, preprint: not peer reviewed.
43. Hunter DR, Handcock MS. Inference in curved exponential family models for networks. *J Comput Graph Stat* 2006;**15**:565–83.
44. Wang P, Robins G, Pattison P, Koskinen J. *MPNet: Program for the Simulation and Estimation of (p) Exponential Random Graph Models for Multilevel Networks*. Melbourne School of Psychological Sciences, The University of Melbourne. 2014. <http://www.melnet.org.au/s/MPNetManual.pdf>.
45. Wang P, Stivala A, Robins G *et al.* *PNet: Program for the simulation and estimation of (p) exponential random graph models for multilevel networks*. 2022. <http://www.melnet.org.au/s/MPNetManual2022.pdf>
46. Wang P, Robins G, Pattison P. *PNet: Program for the Estimation and Simulation of p^* Exponential Random Graph Models*. Department of Psychology, The University of Melbourne. 2009. <https://www.melnet.org.au/s/PNetManual.pdf>
47. Handcock MS, Hunter DR, Butts CT *et al.* statnet: software tools for the representation, visualization, analysis and simulation of network data. *J Stat Softw* 2008;**24**:1548–7660.
48. Hunter DR, Handcock MS, Butts CT *et al.* ergm: a package to fit, simulate and diagnose exponential-family models for networks. *J Stat Softw* 2008;**24**:nihpa54860.
49. Hummel RM, Hunter DR, Handcock MS. Improving simulation-based algorithms for fitting ERGMs. *J Comput Graph Stat* 2012;**21**:920–39.
50. Handcock MS, Hunter DR, Butts CT *et al.* *statnet: Software Tools for the Statistical Analysis of Network Data*. The Statnet Project (<http://www.statnet.org>). R package version 2019.6. 2019. <https://CRAN.R-project.org/package=statnet>
51. Handcock MS, Hunter DR, Butts CT *et al.* *ergm: Fit, Simulate and Diagnose Exponential-Family Models for Networks*. The Statnet Project (<https://statnet.org>). R package version 4.7.5. 2024. <https://CRAN.R-project.org/package=ergm>.
52. Krivitsky PN, Hunter DR, Morris M *et al.* ergm 4: new features for analyzing exponential-family random graph models. *J Stat Soft* 2023;**105**:1–44.
53. Krivitsky PN, Hunter DR, Morris M, Klumb C. ergm 4: computational improvements. arXiv:2203.08198v1, 2022, preprint: not peer reviewed.
54. Caimo A, Friel N. BERGM: Bayesian exponential random graphs in R. *J Stat Soft* 2014;**61**:1–25.
55. Caimo A, Bouranis L, Krause R *et al.* Statistical network analysis with BERGM. *J Stat Soft* 2022;**104**:1–23.

56. Balest J, Secco L, Pisani E et al. Sustainable energy governance in South Tyrol (Italy): a probabilistic bipartite network model. *J Clean Prod* 2019;**221**:854–62.
57. Agneessens F, Roose H. Local structural properties and attribute characteristics in 2-mode networks: p^* models to map choices of theater events. *J Math Sociol* 2008;**32**:204–37.
58. Benton RA, You J. Endogenous dynamics in contentious fields: evidence from the shareholder activism network, 2006–2013. *Socius* 2017;**3**:2378023117705231.
59. Lubell M, Robbins M. Adapting to sea-level rise: centralization or decentralization in polycentric governance systems? *Policy Stud J* 2022;**50**:143–75.
60. Watts DJ, Strogatz SH. Collective dynamics of ‘small-world’ networks. *Nature* 1998;**393**:440–2.
61. Newman ME, Park J. Why social networks are different from other types of networks. *Phys Rev E Stat Nonlin Soft Matter Phys* 2003;**68**:036122.
62. Clark DA, Handcock MS. Comparing the real-world performance of exponential-family random graph models and latent order logistic models for social network analysis. *J R Stat Soc Ser A Stat Soc* 2022;**185**:566–87.
63. Fellows IE. A new generative statistical model for graphs: the latent order logistic (LOLOG) model. arXiv, arXiv:1804.04583v1, 2018, preprint: not peer reviewed.
64. Pauksztat B, Steglich C, Wittek R. Who speaks up to whom? A relational approach to employee voice. *Soc Netw* 2011;**33**:303–16.
65. Doreian P, Conti N. Social context, spatial structure and social network structure. *Soc Netw* 2012;**34**:32–46.
66. Wong LHH, Gygyax AF, Wang P. Board interlocking network and the design of executive compensation packages. *Soc Netw* 2015;**41**:85–100.
67. Goodreau SM. Advances in exponential random graph (p^*) models applied to a large social network. *Soc Netw* 2007;**29**:231–48.
68. Heidler R, Gamper M, Herz A et al. Relationship patterns in the 19th century: the friendship network in a German boys’ school class from 1880 to 1881 revisited. *Soc Netw* 2014;**37**:1–13.
69. Sailer K, McCulloh I. Social networks and spatial configuration—how office layouts drive social interaction. *Soc Netw* 2012;**34**:47–58.
70. Fischer M, Sciarini P. Unpacking reputational power: intended and unintended determinants of the assessment of actors’ power. *Soc Netw* 2015;**42**:60–71.
71. Toivonen R, Kovanen L, Kivela M et al. A comparative study of social network models: network evolution models and nodal attribute models. *Soc Netw* 2009;**31**:240–54.
72. Ackland R, O’Neil M. Online collective identity: the case of the environmental movement. *Soc Netw* 2011;**33**:177–90.
73. Anderson CJ, Wasserman S, Crouch B. A p^* primer: logit models for social networks. *Soc Netw* 1999;**21**:37–66.
74. Stivala A, Robins G, Lomi A. Exponential random graph model parameter estimation for very large directed networks. *PLoS One* 2020;**15**:e0227804.
75. Harrigan N, Bond M. Differential impact of directors’ social and financial capital on corporate interlock formation. In: Lusher D, Koskinen J, Robins G (eds.), *Exponential Random Graph Models for Social Networks: Theory, Methods, and Applications*, chapter 20. New York: Cambridge University Press, 2013, 260–71.
76. Stivala A. Geodesic cycle length distributions in delusional and other social networks. *J Soc Struct* 2020;**21**:35–76.
77. Stivala A. Geodesic cycle length distributions in fictional character networks. arXiv, arXiv:2303.11597v1, 2023, preprint: not peer reviewed.
78. Martin JL. Comment on geodesic cycle length distributions in delusional and other social networks. *J Soc Struct* 2020;**21**:77–93.
79. Wilson JD, Denny MJ, Bhamidi S et al. Stochastic weighted graphs: flexible model specification and simulation. *Soc Netw* 2017;**49**:37–47.
80. Pattison PE, Robins GL, Snijders TA et al. Exponential random graph models and pendant-triangle statistics. *Soc Netw* 2024;**79**:187–97.
81. Pattison P, Robins G. Neighborhood-based models for social networks. *Sociol Methodol* 2002;**32**:301–37.
82. Pattison P, Robins G. Building models for social space: neighbourhood-based models for social networks and affiliation structures. *Math Sci Hum* 2004;**42**:11–29.
83. Hunter DR, Goodreau SM, Handcock MS. ergm.userterms: a template package for extending statnet. *J Stat Soft* 2013;**52**:1–25.
84. Hunter DR, Goodreau SM. Extending ERGM functionality within statnet: building custom user terms. 2019. https://statnet.org/workshop-ergm-userterms/ergm.userterms_tutorial.pdf. Statnet Development Team

85. Uno T, Satoh H. An efficient algorithm for enumerating chordless cycles and chordless paths. In: Džeroski S, Panov P, Kocev D *et al.* (eds.), *International Conference on Discovery Science*, volume 8777 of LNAI. Springer, 2014, 313–24.
86. R CORE TEAM. *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing, 2022.
87. Csárdi G, Nepusz T. The igraph software package for complex network research. *InterJournal* 2006;Complex Systems:1695. <http://igraph.sf.net>
88. Antonov M, Csárdi G, Horvát S *et al.* igraph enables fast and robust network analysis across programming languages. arXiv, arXiv:2311.10260v1, 2023, preprint: not peer reviewed.
89. Wickham H. *ggplot2: Elegant Graphics for Data Analysis*. New York: Springer-Verlag, 2016.
90. Morris M, Handcock M, Hunter D. Specification of exponential-family random graph models: terms and computational aspects. *J Stat Softw* 2008;**24**:1548–7660.
91. Davis A, Gardner BB, Gardner MR, *et al.* *Deep South: A Social Anthropological Study of Caste and Class*. Chicago: University of Chicago Press. 1941.
92. Krivitsky PN, Handcock MS. Fitting position latent cluster models for social networks with latentnet. *J Stat Softw* 2008;**24**:1–23.
93. Krivitsky PN, Handcock MS. *latentnet: Latent Position and Cluster Models for Statistical Networks*. The Statnet Project (<https://statnet.org>). R package version 2.11.0. 2024. <https://CRAN.R-project.org/package=latentnet>
94. Lerner J, Lomi A. A dynamic model for the mutual constitution of individuals and events. *J Complex Netw* 2022;**10**:cnac004.
95. Kevork S, Kauermann G. Bipartite exponential random graph models with nodal random effects. *Soc Netw* 2022;**70**:90–9.
96. Everett MG, Broccatelli C, Borgatti SP *et al.* Measuring knowledge and experience in two mode temporal networks. *Soc Netw* 2018;**55**:63–73.
97. Leifeld P. texreg: conversion of statistical model output in R to LaTeX and HTML tables. *J Stat Soft* 2013;**55**:1–24.
98. Levy M, Lubell M, Leifeld P, Cranmer S. Interpretation of GW-Degree Estimates in ERGMs. 2016. [10.6084/m9.figshare.3465020.v1](https://doi.org/10.6084/m9.figshare.3465020.v1).
99. Levy M. gwdegree: improving interpretation of geometrically-weighted degree estimates in exponential random graph models. *J Open Source Softw* 2016;**1**:36.
100. Stivala A. Reply to “Comment on Geodesic Cycle Length Distributions in Delusional and Other Social Networks”. *J Soc Struct* 2020;**21**:94–106.
101. Handcock MS, Krivitsky PN, Fellows I. *ergm.tapered: Tapered Exponential-Family Models for Networks*. R package version 1.1-0. 2022. <https://github.com/statnet/ergm.tapered>.
102. Stivala A. New network models facilitate analysis of biological networks. arXiv, arXiv:2312.06047v1, 2023, preprint: not peer reviewed.
103. Fellows IE. *lolog: Latent Order Logistic Graph Models*. R package version 1.3.1. 2023. <https://CRAN.R-project.org/package=lolog>.