Modeling student collaborations using valued ERGMs

James E. Wells

Science Division, College of the Sequoias, 915 S. Mooney Blvd ,Visalia, CA 93277 and
Department of Physics, University of Connecticut, 2152 Hillside Road, unit 3046, Storrs, CT 06269-3046

Network analytic techniques are particularly well suited to studying how students form groups, since interactions between people are affected by other interactions within the same community. Collaboration between students in class and out of class during a calculus-based, introductory physics course at a liberal arts college is described using networks. Students are represented by nodes, which are connected by edges, representing interactions between pairs of students. Both the nodes and the edges are associated with various covariates representing the characteristics of the student and the intensity of their collaboration. Exponential family random graph models (ERGMs), a network analytic technique analogous to logistic regression, are used to estimate the probability of the existence of a particular edge, based on the various covariates and the overall structure of the network. An extension to ERGMs, valued ERGMs, model the strength of the edges in addition to their existence. Both the binary and valued ERGMs found that reciprocal interactions, hierarchical interactions, and interactions within assigned groups are more likely to occur. The valued ERGM also found that students with higher course grades correlate with the strength of interactions that students report. There is some evidence that instructors may affect who students collaborate with outside of class.

I. INTRODUCTION

Networks model complex patterns of interactions within a system. Networks consist of nodes, which represent the elements of the system, connected by edges, which represent the interactions between the elements. Nodes and edges may have characteristics associated with them. Edges can also be directed [1–3]. Social network analysis is grounded in relational realism: the idea that relationships are the fundamental unit of society and they both reflect and shape individuals' actions [4].

Physics education research (PER) has long shown working in groups improves student learning [5]. Recently, several groups [6–10] have used social network analysis to explore how students' positions within the community of their classmates affect student outcomes and persistence. In each case, whether the networks were in class, out of class, or online, and whether the students were in an introductory or upper-division course, more centrally positioned students had better outcomes: either higher grades or a higher rate of persistence in the physics sequence.

However, none of these works examined factors affecting the formation of edges between students. Due to the interconnected nature of networks, the assumption of traditional statistics that the independent variables, in this case the edges, represent independent events is violated. An exponential family random graph model (ERGM) [11] is a network analytic technique analogous to logistic regression, but without the independence assumptions of traditional statistics. It gives the probability of an edge existing between two nodes, considering the existing structure in the network and the attributes of the nodes. This preliminary work seeks to use ERGMs to determine what factors affect the likelihood of student collaboration and to find if instructors may influence who students work with outside of class.

II. SAMPLE

The data for this study were taken during the Fall 2017 semester in a calculus-based, introductory physics class for majors and engineers at a liberal arts institution. The course was taught as an integrated lecture-lab style, meeting three times a week for 110 minutes per session. Students were assigned to groups of three to work together on classwork and laboratory activities. These groups also worked collaboratively on the group portion of exams. After each exam, the groups were reassigned; students worked with three different groups over the course of the semester. The groups were formed with the goal of mixing high-, medium-, and low-scoring students together, following [12]. The initial groups were based on students' FCI pre-test scores; the groups in latter periods were based on students' exam grades.

At the end of each week, students filled out a survey in class which asked them to list the people with whom they had a meaningful physics interaction either in class or out of

class that week. The survey included a list of their classmates names and students were invited to write in names not on the list if needed. If a student recorded an interaction, a directed edge was created from that student to the student named. Each week is represented by an in-class network and an out-of-class network. Students also indicated the frequency of these interactions: one time; more than once, but not every day; or every day as in [8]. These categories were given edge strengths of one, two, and three, respectively. The strengths assume that frequent interactions have high value to the students. Demographic data were recorded at the beginning of the semester when the goals of the project were explained to students and informed consent forms were completed.

The semester was split into three periods: before the first exam, between the first and second exams, and between the second and third exams. At the beginning of each of these three periods, the assigned in-class groups were changed. The results of the weekly surveys were aggregated to create an in-class and an out-of-class network for each period by simply adding up the edge strengths from each week in that period.

Basic descriptive network statistics are given in Table I. Nodes represent students and edges represent the interactions between students. The density of a network is the ratio of edges in the network to possible edges. Components are the number of disconnected groups within the network. A single component means that there is a series of edges that connect every student to every other student in the network. Isolates are the number of nodes that are not connected to any other node. The diameter is the longest direct path between any two nodes in the network or, if the network has more than one component, in the largest component. The diameter measures the separation between students. The clustering coefficient is the fraction of groups of three students who are connected by two interactions (e.g. $A \rightarrow B$ and $B \rightarrow C$) that close the triangle (e.g. $A \rightarrow C$). The clustering coefficient measures the propensity of students to bring their colleagues together to form larger groups. Median activity is the median number of students each person reported working with during that period.

TABLE I. Basic network statistics for the in- and out-of-class networks in each of the three time periods.

Model Terms	Exam 01		Exam 02		Exam 03	
	In	Out	In	Out	In	Out
nodes	23	23	19	19	19	19
edges	95	119	69	57	80	60
density	0.19	0.24	0.20	0.17	0.23	0.18
components	1	1	1	5	2	6
isolates	0	0	0	0	0	2
diameter	7	5	7	5	4	5
clustering	0.44	0.39	0.48	0.49	0.45	0.54
median activity	4	5	3	2	4	2

III. EXPONENTIAL FAMILY RANDOM GRAPH MODELS

An ERGM compares the observed network to all the possible networks that can be generated with that number of nodes [11, 13]. Each network structure has an associated parameter that is adjusted to make the observed value of each network structure the most probable in the distribution of all possible networks generated from those parameters [13].

ERGMs proceed from the idea that networks are built from local structures. The processes that drive the formation of these structures can be endogenous, via self-organization; they can be due to the attributes of the actors in the network; or they can by exogenous, due to factors outside of the network and the actors [13].

An ERGM represents a network X using k terms that each represent a network structure. The model gives the log odds of a particular edge x_{ij} existing, given the network X without edge x_{ij}. Each term is associated with a coefficient θ_k that represents how the change in the number of the that particular structure $\delta_{ij,k}(x)$ in the network when x_{ij} is added affects the probability of the existence of x_{ij}. The mathematical form of an ERGM [13] is given by

$$\rho = \log \frac{\Pr(x_{ij} = 1 \mid X_{-ij} = x_{-ij}, \theta)}{\Pr(x_{ij} = 0 \mid X_{-ij} = x_{-ij}, \theta)} = \sum_{k=1}^{n} \theta_k \delta_{ij,k}(x). \tag{1}$$

The result of Eqn. 1 can be converted into a probability using the inverse logit function $\exp(\rho)/(1 + \exp(\rho))$.

Binary ERGMs model the existence of edges; valued ERGMs model the strength of the edges. A valued ERGM gives the probability that a student will give a certain strength to an interaction and uses the observed edge strengths in the model. A binary ERGM only predicts whether an edge will exist or not depending on the existence of edges in the network; the strengths given to the interactions by the students are not used.

IV. MODEL

The binary and the valued ERGMs fit the same network structures to the data. The terms in each model are shown in Tables II and III. Technical information on these terms is found in refs. [14, 15].

The reciprocity term represents the tendency for both students in a pair to report working with one another. Since the edges are directed and the surveys are filled out independently, it is not required that an interaction be reciprocal.

The transitive and cyclical terms account for hierarchical and collaborative interactions. A positive transitive edges term indicates that if edges $A \rightarrow B$ and $B \rightarrow C$ exist, it is more likely that $A \rightarrow C$ exists. This type of interaction indicates that person C is on the top of the hierarchy and person A is at the bottom. If edges $A \rightarrow B$ and $B \rightarrow C$ made the edge $C \rightarrow A$ to exist, this would represent a cyclical relationship, indicating a collaborative interaction [16, 17].

Homophily is the tendency of people to associate with others who share traits in common. For instance, a person who works with someone in their group would result in a homophilic edge with respect to group membership. If those two people also identified as different genders, the same edge would be heterophilic with respect to gender. The gender terms are are measured relative to the probability of a female - female edge. Heller and Hollabaugh found that gender composition affected group functioning [12], which the gender homophily term may be able to measure.

The grade and FCI pre-test score difference terms measure homophily in continuous variables: whether students with similar FCI pre-test scores (a proxy for previous physics experience) or similar final course grades tend to work together.

Node covariate terms test the effect of the pre-test FCI score and the course grade on students' activity and popularity. Activity is the number of people a student reports working with and popularity is the number of people who report working with that student.

Edge covariate terms compare the edges in the network being modeled to those in other networks. This can answer the question, "if two students worked together in class, are they more likely to work together outside of class?"

Because of the larger sample space when values are considered, the valued ERGM models some terms to reduce the sample space. They are included in Table III, because they are essential for understanding the effect sizes, as discussed in Sec. V. The intercept term represents the log odds of an edge between two female students when every other term in the model were zero.

V. RESULTS

The model coefficients and uncertainties for the binary and the valued models may be found in Table II and Table III respectively. Edge probabilities are found from the models using the inverse logit function. The probability for an edge between two female students that does not change any another structure in the network in the binary exam 01 in-class network is $\exp(-3.3)/(1 + \exp(-3.3)) = 0.04$. If that edge were part of a reciprocal relationship, the probability rises to $\exp(-3.3 + 3.3)/(1 + \exp(-3.3 + 3.3)) = 0.5$. If both students are also in the same group, the probability is $\exp(-3.3 + 3.3 + 18)/(1 + \exp(-3.3 + 3.3 + 18)) \approx 1$. If adding a single edge added three new transitive edges to the network, the transitive edge coefficient would be multiplied by three when added to the log odds. For terms like grade difference, the coefficient is multiplied by the difference in grades of the two students. Because the effect size depends on how a particular edge affects all the network structures in the model, all the terms in Tables II and III must be included in the log odds calculation. To calculate probabilities in the valued model, the strength of the edge is multiplied by the coefficient.

TABLE II. Binary Network Fit Coefficients θ_k. Bold entries marked with a * are significant at the $\alpha < 0.05$ level. The uncertainty is the standard error of the coefficient.

Model Terms	Exam 01		Exam 02		Exam 03	
	In	Out	In	Out	In	Out
intercept	$-3.3 \pm 0.7^*$	$-2.3 \pm 0.6^*$	$-10 \pm 4^*$	-5.1 ± 3.1	$-19 \pm 6^*$	-5.2 ± 4.7
reciprocity	$3.3 \pm 0.7^*$	$2.0 \pm 0.4^*$	$2.2 \pm 0.7^*$	$2.0 \pm 0.6^*$	1.5 ± 0.8	$3.7 \pm 1.0^*$
transitive edges	$1.0 \pm 0.4^*$	$0.7 \pm 0.3^*$	$1.1 \pm 0.5^*$	$0.9 \pm 0.4^*$	$1.1 \pm 0.5^*$	0.4 ± 0.6
cyclical edges	-0.6 ± 0.3	-0.3 ± 0.2	0.3 ± 0.4	0.1 ± 0.3	-0.1 ± 0.4	0.8 ± 0.6
group 01 homophily	$18 \pm NA^*$	0.2 ± 0.4	0.4 ± 0.5	$1.1 \pm 0.4^*$	0.2 ± 0.8	-0.4 ± 0.6
group 02 homophily			$20 \pm NA^*$	$-5.1 \pm 1.3^*$	$-8.4 \pm 2.0^*$	$-4.1 \pm 1.8^*$
group 03 homophily					$6.1 \pm 1.3^*$	-2.4 ± 1.8
$M \rightarrow F$	$-1.4 \pm 0.7^*$	-0.2 ± 0.4	-0.7 ± 0.6	-0.7 ± 0.5	1.2 ± 0.8	-0.1 ± 0.7
$F \rightarrow M$	-0.4 ± 0.5	0.0 ± 0.4	-1.0 ± 0.6	$-1.1 \pm 0.4^*$	0.2 ± 0.8	$-2.3 \pm 0.8^*$
$M \rightarrow M$	0.0 ± 0.4	$-0.7 \pm 0.3^*$	-0.6 ± 0.4	$-0.9 \pm 0.4^*$	-0.9 ± 0.8	$-1.7 \pm 0.6^*$
grade difference	-0.2 ± 0.3	0.0 ± 0.2	0.0 ± 0.0	0.3 ± 0.4	$0.13 \pm 0.05^*$	-0.2 ± 0.5
FCI difference	0.1 ± 0.4	-0.2 ± 0.2	0.3 ± 0.5	0.0 ± 0.0	0.4 ± 0.5	0.0 ± 0.0
node cov activity (grade)	0.0 ± 0.0	0.0 ± 0.0	0.1 ± 0.1	0.0 ± 0.1	$0.2 \pm 0.1^*$	$0.2 \pm 0.1^*$
node cov popularity (grade)	0.0 ± 0.0	0.0 ± 0.0	0.0 ± 0.0	0.0 ± 0.1	0.0 ± 0.1	$-0.2 \pm 0.1^*$
node cov activity (FCI)	0.0 ± 0.0	0.0 ± 0.0	-0.1 ± 0.1	0.0 ± 0.0	$-0.15 \pm 0.07^*$	$-0.14 \pm 0.06^*$
node cov popularity (FCI)	0.0 ± 0.0	0.0 ± 0.0	0.0 ± 0.1	0.0 ± 0.0	-0.1 ± 0.1	$0.14 \pm 0.06^*$
edge cov in-class (e01)		$1.3 \pm 0.3^*$	0.0 ± 0.1	0.0 ± 0.0	$1.1 \pm 0.2^*$	0.4 ± 0.2
edge cov out-of-class (e01)			-0.1 ± 0.2	0.0 ± 0.1	0.2 ± 0.1	$0.7 \pm 0.2^*$
edge cov in-class (e02)				$0.7 \pm 0.2^*$	0.0 ± 0.1	-0.0 ± 0.1
edge cov out-of-class (e02)					0.1 ± 0.2	0.1 ± 0.2
edge cov in-class (e03)						0.3 ± 0.2

A. In-class networks

The in-class binary models had three terms that were most important to edge formation: reciprocity, transitivity, and group homophily. Reciprocity is a structure commonly seen in social networks. The positive transitive-edges coefficient indicates a tendency for hierarchical relationships. It is not surprising that edges between students in the same group were more likely to be observed, since students sat with their groups and worked on activities and labs together. In the first two periods, group homophily causes the probability to be ≈ 1 and does not have a calculated standard error. However, including those terms greatly reduced the Akaike and Bayesian Information Criteria and improved agreement with the observed statistics in simulated networks.

Reciprocity, transitivity, and group homophily increase the likelihood of edges in the valued networks as well. Additionally, a student's activity is positively correlated with their final course grade. The coefficients are not large, but they are multiplied by the grade difference, when calculating the log odds, which can make a meaningful contribution to the probability. The binary model did not find this, so it is the strength of the interactions and not their number that correlates with course grade.

Many of the node covariate terms in both the binary and valued models are nonzero only in the exam 03 period. This could reflect changing evaluations of their fellow students over time or could be an artifact of the structure of the network at that time. A temporal analysis would be needed to answer this question.

B. Out-of-class networks

In the binary model, reciprocal edges and intra-group edges have increased likelihood. In the valued model, reciprocal edges, transitive edges, and intra-group edges are also more likely to occur. In both the binary and valued model, several node covariates become nonzero in the exam 03 period, as in the in-class networks. Male students are less likely to be named by others in the exam 03 period in both the binary and valued models. A temporal analysis may help explain these results.

The group homophily term with the in-class groups is always an important negative contribution in both models. However, the edge covariate terms with the in-class networks are positive. One interpretation of these results is being in the same group increases the probability of two students working together outside of class only if they gave their in-class interactions high strength. The probability is not as high as

TABLE III. Valued Network Fit Coefficients θ_k. Bold entries marked with a * are significant at the $\alpha < 0.05$ level. The uncertainty is the standard error of the coefficient.

Model Terms	Exam 01		Exam 02		Exam 03	
	In	**Out**	**In**	**Out**	**In**	**Out**
intercept	$-\mathbf{1.1 \pm 0.2}^*$	$-\mathbf{1.0 \pm 0.2}^*$	-1.0 ± 0.6	-2.0 ± 1.0	$-\mathbf{1.7 \pm 0.4}^*$	-1.1 ± 0.3
reciprocity	$\mathbf{0.9 \pm 0.2}^*$	$\mathbf{0.7 \pm 0.1}^*$	$\mathbf{0.8 \pm 0.2}^*$	$\mathbf{1.0 \pm 0.2}^*$	0.1 ± 0.2	0.4 ± 0.2
transitive edges	$\mathbf{0.4 \pm 0.1}^*$	$\mathbf{0.4 \pm 0.1}^*$	$\mathbf{0.5 \pm 0.1}^*$	0.2 ± 0.1	$\mathbf{0.5 \pm 0.1}^*$	$\mathbf{0.4 \pm 0.2}^*$
cyclical edges	-0.1 ± 0.1	-0.1 ± 0.1	0.0 ± 0.1	0.1 ± 0.1	0.1 ± 0.1	0.2 ± 0.1
group 01 homophily	$\mathbf{0.5 \pm 0.1}^*$	$-\mathbf{0.6 \pm 0.2}^*$	0.6 ± 0.4	$\mathbf{0.4 \pm 0.1}^*$	0.3 ± 0.1	-0.4 ± 0.2
group 02 homophily			$\mathbf{44 \pm NA}^*$	$-\mathbf{0.8 \pm 0.2}^*$	$-\mathbf{1.0 \pm 0.2}^*$	-0.5 ± 0.3
group 03 homophily					$\mathbf{1.0 \pm 0.2}^*$	$-\mathbf{1.3 \pm 0.4}^*$
M $\rightarrow$ F	$-\mathbf{0.2 \pm 0.1}^*$	0.0 ± 0.1	0.1 ± 0.1	-0.1 ± 0.2	$\mathbf{0.4 \pm 0.2}^*$	-0.3 ± 0.2
F $\rightarrow$ M	-0.1 ± 0.1	0.0 ± 0.1	-0.1 ± 0.1	-0.3 ± 0.2	0.3 ± 0.2	$-\mathbf{0.6 \pm 0.2}^*$
M $\rightarrow$ M	-0.1 ± 0.1	0.0 ± 0.1	0.0 ± 0.0	-0.1 ± 0.1	0.1 ± 0.1	$-\mathbf{0.3 \pm 0.1}^*$
grade difference	0.0 ± 0.1	0.0 ± 0.1	0.0 ± 0.0	0.1 ± 0.3	0.0 ± 0.0	$\mathbf{0.05 \pm 0.01}^*$
FCI difference	-0.1 ± 0.1	0.0 ± 0.1	0.1 ± 0.3	0.0 ± 0.0	0.0 ± 0.1	0.2 ± 0.1
node cov activity (grade)	$\mathbf{0.003 \pm 0.002}^*$	0.0 ± 0.0	$\mathbf{0.051 \pm 0.004}^*$	0.0 ± 0.0	$\mathbf{0.07 \pm 0.04}^*$	0.0 ± 0.0
node cov popularity (grade)	0.0 ± 0.0	0.0 ± 0.0	0.0 ± 0.0	0.0 ± 0.0	0.0 ± 0.0	0.0 ± 0.0
node cov activity (FCI)	0.0 ± 0.0	0.0 ± 0.0	-0.02 ± 0.01	0.0 ± 0.0	0.0 ± 0.0	$-\mathbf{0.06 \pm 0.02}^*$
node cov popularity (FCI)	0.0 ± 0.0	0.0 ± 0.0	0.0 ± 0.0	0.0 ± 0.0	0.0 ± 0.0	$\mathbf{0.05 \pm 0.02}^*$
edge cov in-class (e01)		$\mathbf{0.07 \pm 0.02}^*$	0.0 ± 0.0	0.0 ± 0.0	$\mathbf{0.17 \pm 0.03}^*$	0.0 ± 0.0
edge cov out-of-class (e01)			0.0 ± 0.0	0.0 ± 0.0	0.0 ± 0.0	$\mathbf{0.11 \pm 0.03}^*$
edge cov in-class (e02)				$\mathbf{0.15 \pm 0.03}^*$	0.0 ± 0.0	0.0 ± 0.0
edge cov out-of-class (e02)					0.0 ± 0.0	0.0 ± 0.0
edge cov in-class (e03)						$\mathbf{0.15 \pm 0.05}^*$
non-zero edges	$-\mathbf{2.8 \pm 0.3}^*$	$-\mathbf{1.8 \pm 0.3}^*$	$-\mathbf{7.0 \pm 0.2}^*$	$-\mathbf{3.2 \pm 0.7}^*$	$-\mathbf{8 \pm 3}^*$	-6 ± 4
interval	$\mathbf{0.8 \pm 0.2}^*$					
at least			$\mathbf{1.3 \pm 0.4}^*$			
at most				$-\mathbf{1.2 \pm 0.4}^*$		$\mathbf{1.7 \pm 0.7}^*$

two students in different groups who interacted strongly in class. If two students are not in the same group, but interacting nearly every day in class, they are probably friends. It is no surprise that they would work together out of class. Groupmates who did not interact frequently in class when they were assigned shared tasks were not likely to freely chose to work together outside of class.

VI. CONCLUSIONS AND FURTHER WORK

The valued and binary ERGMs broadly agree on the factors that affect the likelihood of students working together: transitive interactions, reciprocal interactions and interactions with assigned group members make collaborations more likely to occur. The valued model can find additional factors that depend on edge strength, such as the effect of course grade on activity. This suggests that future network investigations in PER should measure the strength as well as the number of interactions. These data also indicate that an instructor may be able to increase the likelihood of students working together outside of class if they assign those students to the same group in class.

There are many reasons to be wary of extrapolating these results beyond the sample described here. Residential liberal arts colleges are different than community colleges or research universities. There are other possible effects that are not represented in the data and model such as race, socio-economic status, or physics self-efficacy. External factors such as schedule conflicts or the proximity of students' living spaces may also influence the likelihood of collaboration. Also, no data were collected on existing relationships between the students before the course began.

Ongoing work is attempting to replicate these findings with other sections at the same institution with a larger variety of demographic data and to incorporate temporal information into the analysis.

[1] S. Wasserman and K. Faust, *Social Network Analysis: Methods and Applications*, Structural Analysis in the Social Sciences (Cambridge University Press, 1994).

[2] K. A. Zweig, *Network Analysis Literacy: A Practical Approach to the Analysis of Networks*, Lecture Notes in Social Networks (Springer Vienna, Vienna, 2016).

[3] M. E. J. Newman, *Networks*, second edition ed. (Oxford University Press, Oxford, United Kingdom ; New York, NY, United States of America, 2018).

[4] B. Carolan, *Social Network Analysis and Education: Theory, Methods & Applications* (Thousand Oaks, California, 2014).

[5] P. Heller, R. Keith, and S. Anderson, Teaching problem solving through cooperative grouping. part 1: Group versus individual problem solving, Am. J. Phys. **60**, 627 (1992).

[6] R. Dou, E. Brewe, J. P. Zwolak, G. Potvin, E. A. Williams, and L. H. Kramer, Beyond performance metrics: Examining a decrease in students' physics self-efficacy through a social networks lens, Phys. Rev. Phys. Educ. Res. **12**, 020124 (2016).

[7] J. P. Zwolak, R. Dou, E. A. Williams, and E. Brewe, Students' network integration as a predictor of persistence in introductory physics courses, Phys. Rev. Phys. Educ. Res. **13**, 010113 (2017).

[8] J. P. Zwolak, M. Zwolak, and E. Brewe, Educational commitment and social networking: The power of informal networks, Phys. Rev. Phys. Educ. Res. **14**, 010131 (2018).

[9] A. Traxler, A. Gavrin, and R. Lindell, Networks identify productive forum discussions, Phys. Rev. Phys. Educ. Res. **14**, 020107 (2018).

[10] D. L. Vargas, A. M. Bridgeman, D. R. Schmidt, P. B. Kohl, B. R. Wilcox, and L. D. Carr, Correlation between student collaboration network centrality and academic performance, Phys. Rev. Phys. Educ. Res. **14**, 020112 (2018).

[11] J. Harris, *An Introduction to Exponential Random Graph Modeling* (SAGE Publications, Inc., Thousand Oaks, California, 2013).

[12] P. Heller and M. Hollabaugh, Teaching problem solving through cooperative grouping. part 2: Designing problems and structuring groups, American Journal of Physics **60**, 637 (1992), https://doi.org/10.1119/1.17118.

[13] D. Lusher, J. Koskinen, and G. Robins, eds., *Exponential Random Graph Models for Social Networks: Theory, Methods, and Applications*, Structural Analysis in the Social Sciences (Cambridge University Press, 2012).

[14] M. Morris, M. S. Handcock, and D. R. Hunter, Specification of Exponential-Family Random Graph Models: Terms and Computational Aspects, J Stat Softw **24** (2008).

[15] D. R. Hunter, M. S. Handcock, C. T. Butts, S. M. Goodreau, and M. Morris, ergm: A package to fit, simulate and diagnose exponential-family models for networks, J Stat Softw **24**, 1 (2008).

[16] P. Zappa and A. Lomi, Knowledge sharing in organizations: A multilevel network analysis, in *Multilevel Network Analysis for the Social Sciences: Theory, Methods and Applications*, edited by E. Lazega and T. A. Snijders (Springer International Publishing, Cham, 2016) pp. 333–353.

[17] D. Gašević, S. Joksimović, B. R. Eagan, and D. W. Shaffer, SENS: network analytics to combine social and cognitive perspectives of collaborative learning, Computers in Human Behavior **92**, 562 (2019).