Extending the theoretical framing for physics education research: An illustrative application of complexity science

Jonas Forsman[*]

Department of Physics and Astronomy, Uppsala University, Box 516, 751 20 Uppsala, Sweden

Rachel Moll

Faculty of Education, Vancouver Island University, Nanaimo, British Columbia V9R 5S5, Canada

Cedric Linder

Department of Physics and Astronomy, Uppsala University, Box 516, 751 20 Uppsala, Sweden
and Department of Physics, University of the Western Cape, Bellville 7535, Cape Town, South Africa
(Received 12 March 2014; published 24 September 2014)

The viability of using complexity science in physics education research (PER) is exemplified by (1) situating central tenets of student persistence research in complexity science and (2) drawing on the methods that become available from this to illustrate analyzing the structural aspects of students' networked interactions as an important dynamic in student persistence. By drawing on the most cited characterizations of student persistence, we theorize that university environments are made up of social and academic systems, which PER work on student persistence has largely ignored. These systems are interpreted as being constituted from rules of interaction that affect the structural aspects of students' social and academic network interactions from a complexity science perspective. To illustrate this empirically, an exploration of the nature of the social and academic networks of university-level physics students is undertaken. This is done by combining complexity science with social network analysis to characterize structural similarities and differences of the social and academic networks of students in two courses. It is posited that framing a social network analysis within a complexity science perspective offers a new and powerful applicability across a broad range of PER topics.

PACS numbers: 01.40.Fk, 89.75.Fb, 89.65.-s

I. INTRODUCTION

In his 2013 Oersted lecture, Redish pointed out that, "despite the clear value of theoretical frameworks in scientific research, DBER [Discipline Based Education Research] scientists are often reluctant to situate their work within a theoretical framework" (p. 3) and that "part of the challenge in building educational theory" is that "human behavior is extremely complex." He further argued that, "Just as we have done in many areas of science, we create a theoretical framework that allows us to build descriptive models and that can evolve and change as we learn more" ([1], p. 2). However, to create such theoretical framing for physics education research (PER) it is inevitable that the physics education community will draw from other scientific areas such as education, physics, sociology, psychology, neuroscience, biology, and mathematics. A common theoretical theme found across all these sciences is the modeling of complex systems—or more specifically, the application of complexity science or complexity theory. This powerful perspective, however, has to date seen very limited application in PER (see, e.g., Refs. [2,3]). Our aim is to illustrate an application of it for a long-known thorny issue—student persistence.

The reason that student persistence is chosen begins with the work of Tinto, long regarded as a leading figure in the seminal work on modeling student persistence. Tinto describes how the view of student persistence has shifted from relatively simplistic psychological aspects such as "individual attributes, skills, and motivation" to a "detailed longitudinal model that made explicit connections between the environment, in this case the academic and social systems of the institution and the individuals who shaped those systems, and student retention [persistence] over different periods of time" ([4], p. 2–3). Such a shift in perspective, however, has thus far generally been missing in PER studies dealing with student persistence issues.

Thus, we set out to illustrate the use of complexity science in PER by addressing two research aims: (1) how to situate central constructs from student persistence research within a framework of complexity science, and (2) illustrate the viability of using methods available from complexity

[*]Corresponding author.
jonas.forsman@physics.uu.se

science to analyze the structural aspects of students' networked interactions.

II. CONCEPTUAL FRAMEWORK

A. Student persistence research

Student persistence—students persevering with their studies—in university-level physics and related engineering education has been widely studied for many years, yet it continues to be a significant international problem [5,6]. For example, in Sweden, the percentage of students who graduate with a Bachelor of Engineering in Physics in the designed degree program time is 63% of what it was 15 years ago, while during the same time period the number of students has grown by 50% [7–11].

To date, student persistence research has drawn extensively upon the work of Spady [12,13] and Tinto [14–16]. In 1987 Tinto showed that only 20% of students dropping out failed academically, meaning that the remaining 80% left their university studies for other reasons [15]. This insight sparked a new thrust of study in the area that became known as student persistence research. The contemporary field of student persistence has identified social phenomena as critical attributes of student persistence, instead of continuing to simply equate student dropout with academic failure.

Since then, student persistence research has built its knowledge base from a variety of applications of the work of Spady and Tinto, and even today the research field still draws extensively on these two foundational frameworks [17–20]. According to these frameworks, the aspect of students' university life that has the most influence on student persistence is the degree of students' social and academic integration as a function of students being able to come to terms with the social and academic rules, norms, values, and expectations found within the university community. These sets of rules, norms, values, and expectations are framed as Durkheimian *social systems* [21] of the university, i.e., the *social* "rules" and the *academic* "rules" of being a part of the university community [15,22].

Thus, social and academic systems were brought to the fore [12–16] in the field of student persistence. Several successful theoretical and empirical models of student retention are built around the separation of a social and an academic system [12–18]. Thus, in this article, it is analytically appropriate to treat these systems as two distinct systems. These systems have been argued to be interrelated in a complicated way [16]. However, Tinto [16] made the case that in order for students to have a chance to continue their studies, they need to be well integrated into both of these systems. Tinto went on to propose that students who are less successfully integrated into their university's social and academic life—from the students' point of view—tend to be less likely to continue with their studies ([16], p. 16).

Previous research involving the constructs of academic and social integration (and thus social and academic systems) have been criticized by some as being fuzzy and ambiguous [18]. In this article, we propose that this critique can be fruitfully dealt with by analyzing students' social and academic integration using a network perspective (cf. [20]) to bring out the "well connected" or "knowing the right people" attributes within the system.

How do students become socially and academically integrated? In other words, how do they become "well connected" (or "getting to know the right people")? Is it possible to empirically explore if this process is random or if it is governed by a set of rules (social systems)? If yes, then it is of paramount importance to do this in relation to exploring the fundamental nature, i.e., social and academic integration, of student persistence. Then, physics students' course interaction network is not only of importance when it comes to physics learning [23], but also for their persistence. However, if these networks are created randomly and each student has a random position in the network, then a network analysis of students' course interactions in relation to learning, or persistence, is limited to being descriptive and does not offer any practical or pedagogical implications.

B. Social network theory

Visualizations and analyses of the structure and dynamics of the social and academic networks are possible through social network theory. The emphasis in social network theory is on "structural relations" [24]; i.e., social network theory offers a way to explore the relations between different structures within networks. The essential components of a network are the *nodes* and the *links* between nodes. In other words, social network theory is a powerful way to explore, come to understand, and characterize structure connectivity through network measurements.

This study differentiates between the social and academic networks of the participating students' courses. The nodes in the networks represent individual students within the network, and the links represent the type of relationships (social or academic) between students. Networks are characterized by quantifying links between nodes. The nodes in a *directed network* have a *total degree* made up of the sum of *out-degree* (number of outgoing links) and *in-degree* (number of incoming links) measurements. A route through a sequence of nodes, which begins with the starting node and follows outgoing links to other nodes through the network to the end node, is denoted as a *path*. When there is a path between every node, then the network is considered *connected*. If there are many paths between two nodes, the distance and number of links in the different paths are used to find the shortest path between nodes [25].

In this study, the networks were formed as *directed networks* using the responses from a network survey (see Sec. III), which asked the students with whom they

interacted and how (socially or academically). Their responses were used to create models of their social, academic, and combined (social and academic) networks. The distribution of in-degree nodes of the different networks were analyzed to investigate whether the different networks show similarity. The power of this approach is that it may indicate the nature of the underlying processes in the creation of the two networks [26].

Using the average path length, it is possible to measure how far apart two nodes in general are in the network and to see how large or small the networks tend to be. Clustering coefficients [26] were measured. A clustering coefficient describes, in short, the likelihood that a node's two neighbors are also connected to each other (in this case, that two students who are interacting with one student are also interacting with each other). Clustering coefficients were used to illustrate how grouped the network tended to be and to see how tightly grouped the students were in their groups.

While student persistence has been shown to depend on students' social and academic networks, the conceptual and theoretical underpinning needed to apply a network analysis to explore them is currently lacking in the field of student persistence. Thus, the following question arises: How can the relationship between social integration and Durkheimian [21] social systems be described? This question arises because the central constructs of student persistence research are for students to be integrated into the social systems of the university. In order to be able to bring together social integration and network theory in the case of student persistence, we set out to draw on the field of complexity science.

C. Using complexity science in education research

This section provides an overview of the field of complexity science and introduces the conceptual thinking associated with complexity science. In doing so, we provide a theoretical framework through which it is possible to interpret the collection of tools provided by social network analysis. Complexity science aims to describe and understand complex systems and their capacity to show order, patterns, and structure [27]. Especially important is how these orders, patterns, and structures seemingly emerge spontaneously from local interactions (neighbor interactions) between the system's components or agents. Complexity science is currently used as a research framing across many disciplines [28] generating a theory that essentially "transcends disciplines" [29]. Complexity science is not characterized by any particular method, but it is a methodological perspective that employs a range of methods to study complex phenomena, and where the evolution of a complex system is understood to be largely unpredictable and uncontrollable. For more details on the historical development of complexity science, see Waldrop [29], and for an overview of current

applications of complexity science in a wide array of fields, see Mitchell [28]. For pioneering use of complexity science in education, see, for example, Refs. [27,30,31].

Complex systems are typically described as learning systems because they are adaptive and self-organizing [32]. Learning in this sense is taken to be characterized by "ongoing, recursively elaborative adaptations through which systems maintain their coherences within dynamic circumstances" ([33], p. 151). As a guiding perspective in education research, complexity science uses the characteristics of complex systems and conditions of emergence to both understand and prompt learning. This is done through, for example, complex network simulations of the effects of the rules of choosing what to learn. Here, the effects of the rules of choosing what to learn takes place amongst interacting agents [34]. Further studies also explored these effects in respect to different knowledge structures [35]. Although the behavior of complex systems such as society, organisms, or the Internet can only be conceptually discussed as somewhere in between complete order and complete disorder, any attempt to measure or distinguish one system as "more" complex than another typically breaks down [28]. Complexity, as a single construct, is hard to measure. But it is possible to examine structures and dynamics that are shared across complex systems.

Complex systems have been found to have a decentralized network structure where the connectedness of system nodes is characterized by a power-law distribution [27]. This power-law degree distribution is frequently found in real-world dynamic networks, such as the Internet [36–38], the World Wide Web [39,40], and citation networks [41]. The widely accepted interpretation [26] of finding a power-law degree distribution is that the system has evolved (and may continue to evolve) through preferential attachment, which is a central process for exploring how complex systems evolve and adapt. This characterization of decentralized networks means that there are very few components or nodes that are much more connected than others—i.e., a small subset becomes more "important" than others, and thus "attracts" more connections. However, other complex social systems have recently been found to follow other distributions such as the Gamma, Weibull, and Zipf distributions [42]. Accordingly, from these findings, the underlying processes of network formation could be different, but the process of preferential attachment can still play a role in these distributions, just not as a single process.

Decentralized network structures can be contrasted with two other types of network structures: (1) centralized networks with one central node where every other node is only connected to that central node and (2) distributed networks where all nodes have the same connectivity in the network. Information is shared efficiently in centralized networks, but they are vulnerable to breakdown due to dependency on the central node. On the other hand, distributed networks are robust to breakdowns, but inefficient in sharing information.

In the case of decentralized networks, when an "important" component is removed or breaks down, then the whole system will suffer considerable damage. However, the system will remain stable with the removal of any of the many less important, or less connected, nodes. This brings us to an important observation: since a one-time measurement was used in this study, the dynamic nature of the system was difficult to immediately appreciate. It was only after an in-depth consideration of the distribution of connectivity of the nodes (students) that it was possible to confidently infer the complex nature of students' social and academic interactions. In other words, it was through this framing—complexity science—that it became possible to interpret the complex nature of students' social and academic interactions in relation to the stochastic processes which can give rise to the most likely family of distributions these networks resembles.

Because of their decentralized structure, complex systems are made up of components, which are themselves complex systems. Thus, complex systems can be described as nested. Nested systems are "self-similar," meaning that they share the same type of distribution of the connectedness of nodes. Each level of complex organization exhibits similar structures and dynamics but operates within different time scales and in different units of analysis. For example, mathematics *learning-for-teaching* has been modeled as several nested complex systems: subjective understanding, classroom collectivity, curriculum structure, and mathematical objects [43]. By using complexity science, it is possible to consider individual students to be nested within social and academic networks, which, in turn, are nested within the larger network of students within the course.

Another key aspect of the dynamics of complex systems is that they are recursively adaptive to their own internal and external interactions; in this case, adaptations continuously occur in response to the students' relationships with other students and teachers within the university and other networks outside the university. Through recursive adaptations, complex systems self-organize; properties, behavior, and structures all emerge without an external system or an internal "leader system" that controls the complex system [27]. For the systems analyzed in this study, the interpretation of the interactions taking place is a coadaptation between the rules of interactions and the students within the systems. In the study, only a "snapshot" of the social and academic systems was taken. However, given their complex systems properties, it is analytically appropriate to infer the processes that led to their emergence [26].

III. METHOD

A. Network survey

To explore the structures of a social and an academic network, a network survey was developed and tested. Appropriate ethical approval was obtained for this study.

A final version was given to students in two courses in an engineering program at a Swedish university (a four-and-a-half-year degree program made up of core courses in physics, mathematics, and computer science, and a research project requiring at least six months of full-time work) at the end of the Spring term 2010. The data were collected from the participating students who were attending introductory physics courses. Typically, the courses have a lecture format with separate problem-solving and student laboratory sessions. The teaching approaches range from traditional (chalk and talk) to being highly interactive. The design of the learning environment specifically provides for many accessible venues to accommodate both instructor- and student-initiated group activities. Following the methodology of Morrison [44], which is a widely accepted methodology used in social network analysis [45], the students were asked to list the names of the people in their courses with whom they interacted. The students were then asked to characterize the nature of their interactions with each of these people on a scale from *only social* (1) to *only academic* (5) with *both social and academic* being (3).

In Fig. 1 we provide the cover page of the network survey.

Students' interpretations of what constitutes social interaction and academic interaction were explored in focus-group discussions [46] with four groups (12 randomly selected students). Discussions produced eight examples of social interactions and nine examples of academic interaction. These examples were member checked [46]. Examples of social interaction involved doing the following with other students: pausing while studying, going to student pubs, participation in "exam parties," going out for lunch or dinner, doing sports or working out, playing board games, engaging in student organizations, and participating in student activities. Examples of academic interaction involved doing the following with other students: studying, discussing problems, discussing solutions to problems, doing laboratory work, doing hand-in exercises, study visits, going to the lectures, studying for exams, and going to labor-market events. The focus group descriptions were very similar to the descriptions by Spady [12,13] and Tinto [14–16] of social and academic systems.

The two engineering program courses where the survey was administered were a Mechanics II course (Course One) and a computing science course (Course Two). The Course One survey was answered by 68 students, which resulted in a network of 122 students (54 more students in the course were mentioned in the survey responses of others). The Course Two survey was answered by 66 students, which resulted in a network of 107 students (41 more students in the course were mentioned in the responses of others). Only the students who answered the network survey were used in analysis of degree distribution, path length, and clustering coefficient. All students mentioned were used in the visualizations of the networks.

In this survey, if you choose to participate, we ask you to write the names of other students within your course with whom you have been interacting. Further, you are asked to classify your interactions as either social or academic on a five degree scale. A marking to the far left would mean that you mainly have a social relationship. A marking to the far right would indicate that your relationship to that person is mainly academic – you are mainly studying in groups, discuss the courses and/or content of courses.

Name:		Social				Academic
X		O	◉	O	O	O

Indicate that you have a relationship to student X, which is mostly social (we sometimes study together and do coursework, but on the broader scale of things we do social things together).

Name:		Social				Academic
X		O	O	◉	O	O

Indicate that your relationship with X is evenly balanced between doing social or academic activities together.

Name:		Social				Academic
X		O	O	O	◉	O

Indicate that your relationship with student X is mostly academic, and you have – at some point – done something social together.

FIG. 1.   Cover page of the network survey.

B. Network visualizations and measurements

From the lists of names and relationship characterizations provided by the students in the network survey, three separate networks were created using Gephi 0.7 beta [47] for each course surveyed: a social network, an academic network, and a combined (both social and academic) network. The social network contained students' social interactions that corresponded with students' social bonds as noted by them in the network survey. A social bond was interpreted to exist when students characterized their relationship as "only social" or "both social and academic." The academic network corresponded to students' academic interactions as characterized by the students on the survey with responses of "both social and academic" or "only academic." The combined network contained all student interactions, both social and academic, that were noted by the students in the network survey. Students' responses on the network survey were interpreted in terms of directed social and/or academic bonds; if a student mentioned another student, a directed link was present. If both students mentioned each other, then it was interpreted that a bidirectional link existed between the two of them.

To explain the measurement procedures used in the analysis—average path length, degree, clustering coefficient, and degree distribution—the terms are now defined mathematically. These measurements are the most commonly used in the study of complex (networked) systems as the means of describing the systems and system creation and to elucidate system properties [26]. As described in the conceptual framework, the first construct of network theory is the nodes. The nodes represent components of a network. Secondly, the links represent the relationships or connections between nodes.

A network can be mathematically described through an adjacency matrix A with elements A_{ij} such that

$$A_{ij} = \begin{cases} w & \text{if there is a link between nodes } i \text{ and } j \\ 0 & \text{otherwise,} \end{cases}$$

where w is 1 if the network is unweighted or it can take any value when the network is weighted. The elements $i = j$ are zero when the nodes in the network do not have self-links (links to itself).

1. Average path length

The distance between adjacent nodes is defined as 1. Thus, the average path length is calculated as the average distance between any two nodes in the network.

2. Degree

A node's total degree is the number *of adjacent nodes.* The total degree (k_i) of node i in the example network is

then the sum of each row, and it can be calculated through the adjacency matrix A_{ij} by using

$$k_i = \sum_{j=1}^{n} A_{ij}. \tag{1}$$

In a directed network, there is a possibility of differentiating between incoming nodes and outgoing nodes. The sum of each row is then the in degree, and the sum of each column is the out degree. This can be calculated through the adjacency matrix, A_{ij} by using

$$k_i^{\text{in}} = \sum_{j=1}^{n} A_{ij}, \tag{2}$$

$$k_i^{\text{out}} = \sum_{i=1}^{n} A_{ij}. \tag{3}$$

3. Clustering coefficient

When two nodes are directly connected, these two nodes are adjacent. The clustering coefficient C_c (or transitivity), in short, is the likelihood that a node's two adjacent nodes are also adjacent (i.e., are directly connected) to each other. This is calculated as per

$$C_c = \frac{3 \times \text{number of triangles in the network}}{\text{number of connected triplets of nodes in the network}}, \tag{4}$$

where $0 \le C_c \le 1$, and 0 corresponds to no links in the network, while 1 indicates that all nodes are adjacent to each other.

4. Degree distribution

The degree distribution [the distribution of in degrees (k^{in})] is a fundamental network property [26]. The distribution of out degree (k^{out}) is not commonly analyzed when processes of network creation are sought. The study of different distributions is critical to explore the possible processes that drive the construction of the network [26]. The degree distribution p_k [Eq. (5)] can be thought of as a probability for a randomly chosen node to have degree k, i.e., a probability density function:

$$p_k = \frac{\text{number of nodes with degree} k^{\text{in}}}{\text{total number of nodes}},$$
$$k^{\text{in}} = 1, 2, \ldots, \infty. \tag{5}$$

The commonly found power-law distribution in degree distributions of networks has been found to likely depend on "preferential attachment," which is where a node with high degree "attracts" more nodes to create a link to that node

[26]. However, other distributions suggest other driving processes such as removal or joining of nodes that have been modeled by a Weibull model [48]. Price's creation model [26] has been shown to follow a power-law behavior though the beta and gamma functions. As it stands now in the field, there are other processes used to explain different distributions. However, studies using these different distributions all suggest that the process of creating the network and the degree distribution of a particular network are closely tied together; see, for example, Refs. [26,48].

In order to estimate the similarity of the networks' degree distribution to known degree distributions, we used AIC_c [49] [Eq. (6)], which is a sample-size corrected version of the Akaike information criterion (AIC) [50]. The AIC_c was used because the networks in this study can be argued to be small:

$$\text{AIC}_c = 2k - 2\ln(L) + \frac{2k(k+1)}{n-k-1}, \tag{6}$$

where k is the number of parameters, n is the sample size, and L is the maximized value of the estimated likelihood function.

Eleven common distributions (power law, normal, log-normal, exponent, Poisson, Cauchy, Gamma, logistic, binomial, geometric, and Weibull) were tested and evaluated through the use of the AIC_c criterion.

In studying the relative fit between the degree distribution found in this study and other known distributions, it becomes possible to compare the process of network creation between two networks, but not to draw any conclusions about pinpointing the exact process through which each network is created.

IV. RESULTS

This section is divided into two parts. The first part addresses the first research aim, which is to situate social systems, i.e., *the social and the academic system* (and thus social integration), within a framework of complexity science. This is done in order to have a theoretical framing for the tools used to analyze students' networked interactions. The second part is further divided into three subsections that are used to illustrate the analysis of the students' interaction networks that addresses the second research aim. This aim involves illustrating the viability of using methods available from complexity science to analyze the structural aspects of students' networked interactions. The path lengths of these networks are investigated in the second section. In the third section, the distributions of the in degrees of the networks are analyzed, in order to evaluate their macroscopic structural similarities and differences. The second part establishes if the students' interaction networks are random, and if they are not, how differences in the social and academic interaction networks (and thus the rules of interaction) can be analyzed.

A. Situating core concepts from student persistence in a complexity science framework

Here, we show how the new theoretical framing is used to address the first research aim. Throughout the history of student persistence research that has included learning [17] and academic achievement [12–17], the idea that student persistence is situated within a complex system has been tacitly taken into account. For example, Spady ([12], p. 40) has argued that if research on student persistence is to be more fruitful, it needs to take a "… particular social system as the result of a complex social process" into account. More recently, Bean ([18], p. 238) argued that "students' experiences are complex, and their reasons for departure are complex." However, empirical research on student persistence that explicitly draws on complexity science is both rare and new [2,3]. This is probably because the use of complexity science is still new to educational research [27]. By drawing on results from complexity science in the natural sciences, which has identified many characteristics of complex systems, it becomes possible to apply the general concepts from this framework for an educational context. This makes it possible to identify complex systems as learning systems. As discussed in Sec. IIC, these systems are adaptive and self-organizing systems, whose global behavior is an unpredictable and emergent property from the interactions of component agents [27]. Complex systems can be characterized by specific structures and dynamics. For example, complex systems are made up of independent agents that are connected in a decentralized network of local interactions (cf., neighbor interaction [27]).

Like complexity, social networks have been present, but also in the background, in the development of theoretical models used to understand student persistence, especially in the work of Tinto [14–16] who recognized that the field of student persistence research needs to employ "network analysis and/or social mapping of student interaction… [to] …better illuminate the complexity of student involvement" ([17], p. 619). Recently, research has indicated that the number of possible social links to other students is of significance when predicting academic fit [51] (i.e., students' congruency with the academic system). Further exploration of peer effects in random housing allocation for college has shown that the person with whom you share a room in the college years tends to have an effect on grade-point average (GPA) scores [19]. However, these effects diminish over time [19]. When examining effects on previously identified critical constructs from student persistence research, the structure of the social networks can even have significant negative effects on persistence and related constructs [20]. At the same time, it has also been recognized that students' social networks can be a source of both support and stress for students [52]. These effects are clearly not linear, however; hitherto, complexity science together with social network analysis has not been employed to examine the structure of students' social network and academic networks as resulting from two sets of "social systems."

Through interpreting the foundational work of Spady and Tinto from a complexity science perspective, it is the conceptualization of the nature of adaptation that differentiates Spady, Tinto, and complexity science. While Spady ([12], p. 41) accepts that the outcome of the adaptation from his model, "normative congruence," is a multifaceted concept, he depicts the student as an active agent within a static environment. For example, he argues that students' interactions within the university "provide the student with the opportunity of assimilating successfully into both the academic and social systems of the college" ([12], p. 77). The notion of a static environment is also what Tinto describes when he claims that in the process of student interactions a student will "… continually modify his goal and institutional commitments in ways which lead to persistence and/or to varying forms of dropout" ([17], p. 94). Drawing on the notion of adaptation from complexity science [27], however, adaptation is then a codependent construct between both the individuals and the systems where the individuals are situated. Hence, it is not only the students' adaptation to the institutional environment that is critical, but also the institutional environment's adaptation to the students.

As the institutional environment is made up of the social system and the academic system of the university [12–17], it is proposed here that these systems can be analytically viewed as complex systems [27]. By considering Spady's and Tinto's characterization of the social system and the academic system and comparing these with the complex system perspective, the social system and academic system can then be characterized as consisting of what Sawyer, in his reading of Durkheim [21], called *social facts* [22]. Examples of these could be agency, intention, discourse patterns, collaborations, subcultures, norms, beliefs, and expectations (for readability, these social facts are taken to be equivalent to *rules of interaction*).

These rules of interaction, divided into the realms of social life and academic life within an institution, form and are formed by social interactions and academic interactions between students, staff, and faculty, thus creating two systems of rules: the social and the academic. Furthermore, due to these interaction types being ubiquitous, the two types of interactions are continuously creating two entangled and ever-present systems. From a complexity science perspective, the emergence of a social and an academic system is dependent on interactions within the system while at the same time the interactions within the system are dependent on the two component systems, and thus there are recursive adaptations between student interaction and system structure. In other words, the networked interactions, the social system, and the academic system are all coadapting. Thus, differences in the structure and dynamics of the social system and the academic system
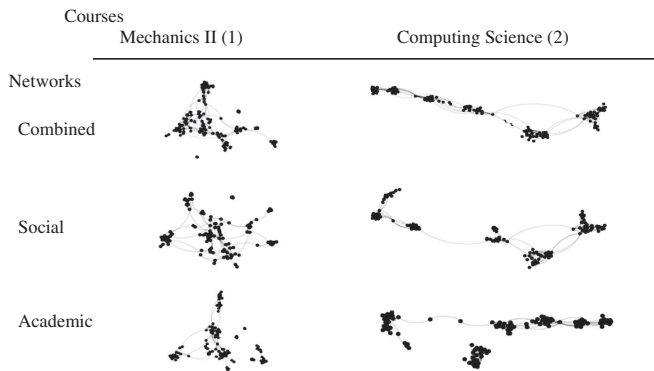
Courses
Mechanics II (1)   Computing Science (2)

Networks

Combined

Social

Academic

FIG. 2.   Network visualizations of the combined, social, and academic networks of Course One and Course Two students.

TABLE I.   Clustering coefficient and average path length for all networks of each course. C1 is Course One, C2 is Course Two, $S + A$ denotes the combined network, S denotes the social network, A denotes the academic network, and the asterisk denotes standard error of 1000 random simulated networks.

Course networks	Simulation of random network–mean clustering coefficient	Clustering coefficient	Average path length
C1 $S + A$	$0.027 \pm 0.004^*$	0.393	3.802
C1 S	$0.027 \pm 0.005^*$	0.364	3.572
C1 A	$0.026 \pm 0.005^*$	0.247	4.479
C2 $S + A$	$0.050 \pm 0.004^*$	0.519	5.424
C2 S	$0.041 \pm 0.004^*$	0.424	4.252
C2 A	$0.034 \pm 0.005^*$	0.374	2.844

should influence structures and dynamics of the networked interactions of individuals who are participating. This leads to the formation of a social network and an academic network (which are nested within a larger combined network), and the coformation of the social and academic systems.

B. Analyzing students' interaction networks

1. Visualizations of the network structure

Using Gephi 0.7 beta [47], network visualizations were generated for the Course One and Course Two data (Fig. 2) in order to make qualitative observations and visual interpretations of the network measurements. To attain clearer network visualizations, all students mentioned in the survey are represented in the network visualizations. Visual representations of the combined, social, and academic networks for each course were seen to be sharing similar structural characteristics.

For each course, the academic network displayed a group of students who were disconnected (or more disentangled) from the larger network. For example, for Course One (Fig. 2) the academic network had one large cluster and two small ones, which were not connected to the larger network. Course Two (Fig. 2) had one very large cluster that was disconnected. In contrast, most of the students appeared connected in the combined and social networks for each course.

The algorithm used to create the visualizations emphasizes structural clusters more than how each student is connected. Still, the positions of clusters and individual students are related to how the students within each cluster are connected. Thus, the visualizations of Course One portray networks where more students are interacting between groups than the visualizations of Course Two.

2. Average path length and clustering coefficients

Average path length is calculated to globally measure how "far away" students tended to be from one another in each network (Table I). The measurement is the average number of steps from a particular node to all other nodes, following the shortest paths [26].

A calculation of the clustering coefficient [26] was carried out to compare how tightly grouped the networks were (see Table I). The measurement is for the whole network, which results in a unique numerical value for the network, as per Eq. (4). The clustering coefficient of the analyzed networks was compared to the mean and standard error of the clustering coefficient in 1000 random network simulations with the same number of nodes and links.

In Course One students were closest to everyone else (i.e., had the lowest average path length) in the social network when compared to the other networks. Some students in this course tended to be farther apart in the combined network, or even unreachable (as no paths existed between them) in their academic network. This trend was different in Course Two, where the academic network had a much shorter average path length than the social or combined network. This could be seen visually where the academic network divided into two separate academic networks with no connections between them (Fig. 2). Therefore, the students had no direct academic connection to roughly one-third of the course. The reason being that the calculation of the average path length was smaller because there were fewer long paths in the network.

For all of the networks of both courses the measured clustering coefficient was a factor of 4–15 times larger than the corresponding simulated random network. This result shows that the structure of the network is far from random, and implies that students' interactions follow sets of rules [26]. The analysis illustrates how, for the two courses, the tendency of tightly connected groups peaked in the combined network of students and dropped to its lowest in the academic network. This was corroborated by qualitative observations of the visualizations (Fig. 2).

3. Degree distributions

Estimation of the degree distribution (distribution of nodes in degree) was carried out for each of the networks
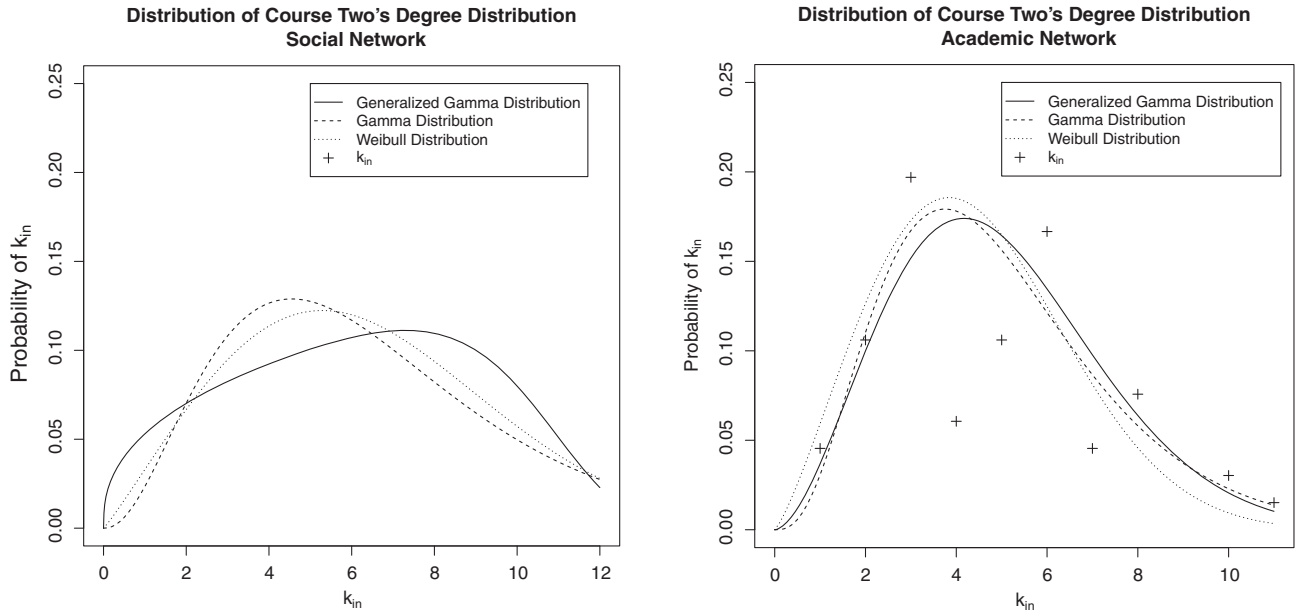
FIG. 3.   Degree distribution for Course Two for the social (left) and the academic (right) network, and the estimations for the generalized Gamma, Gamma, and Weibull distributions.

because it becomes possible to compare the process of network creation between two networks, as discussed in Sec. III B 4. Two classes of distributions were found to be likely candidates for describing the distributions of the networks found in each course: the Gamma distribution and the Weibull distribution. However, these distributions cannot easily be compared, as they can differ significantly in form.

In order to evaluate the parameters for these two distributions by using the same model, the generalized gamma function was used [53]:

$$f(x; a, v, p) = \frac{|p| x^{pv-1} e^{-(x/a)^p}}{a^{pv} \Gamma(v)}, \qquad (7)$$

where a, v, and p are the parameters of the generalized Gamma function.

These three distributions (Gamma, Weibull, and generalized Gamma) are examined together with the degree distributions of the networks to show how similar these distributions can be. The similarities of the estimated degree distributions are shown in Fig. 3 for the social and the academic networks of Course Two.

However, both the Gamma and the Weibull distributions are covered by the generalized Gamma function's special cases [Gamma distribution $f(x; a, v, 1)$, Eq. (7), with $p = 1$; Weibull distribution $f(x; a, 1, p)$, Eq. (7), with $v = 1$), which allows one to evaluate if parameter a is significantly different for these types of networks. The estimations of a were undertaken using nonparametric bootstrap methods based on maximum likelihood estimations, which is implemented by the package

FITDISTRPLUS [54] in the statistical environment r [55]. The estimation of the a parameter using this methodology also gave an estimation of the 95% confidence intervals for the a parameter. These confidence intervals are used to identify the differences between the networks (Table II).

The estimated parameter (a) for the generalized Gamma distribution between Course One and Course Two showed that the networks between these two courses are significantly different. The social network of Course Two fell outside the 95% confidence interval in comparison to the academic network of the same course; the implications being that the rules of interaction differ between the social and the academic interaction networks. Furthermore, both classes of distribution—as estimated from the degree distribution—are found in other systems that are identified as being complex ([26], p. 494; [42]). In other words, they also follow a set of interaction rules. From here it becomes possible to empirically describe such course interactions as complex systems.

TABLE II.   Estimation of parameter a in the generalized Gamma function with 95% confidence intervals.

	a	2.5%	97.5%
Course One, combined network	1.57	1.18	1.99
Course One, social network	1.38	1.04	1.73
Course One, academic network	1.26	0.92	1.61
Course Two, combined network	8.44	7.64	9.25
Course Two, social network	7.29	6.39	8.17
Course Two, academic network	5.59	5.00	6.32

V. DISCUSSION

This article has addressed the viability of using complexity science as a frame to extend methodological applications for physics education research. In doing so, we set out to address two research aims: (1) how to situate central constructs from student persistence research within a framework of complexity science and (2) illustrate the viability of using methods available from complexity science to analyze the structural aspects of students' networked interactions.

The structures of the social, academic, and combined networks of the participating students were described by examining the clustering coefficients, path length, and degree distribution of their nodes. The interpretations of these measurements were also informed by the visualizations of the networks that were produced. The clustering coefficient was used to show that the networks are far from random. The path length was used to show that the different courses had different interaction networks. By analyzing the degree distribution, it became possible to illustrate that these introductory-level courses had similar degree distributions to those of other complex systems [40] and could analytically be viewed as complex systems. These results support claims that students' social and academic networks are complex in nature. Realizing this complex nature of student interaction networks cannot provide a dynamic understanding of how these networks are evolving. However, the findings, together with the theoretical framework described above, offer the first steps toward building such understanding.

The exclusion of external actors (for example, other students, instructors, and parents) that interact with the students within a course was a limitation that was introduced by choice because the research was focused on discussion of students' course interactions and how that could be framed as a complex system. This choice was made to make it possible to study only one level of complexity—student interactions, in order to minimize mixing of multiple nested levels together under one analytic method. As a result of this choice, the study also has a somewhat limited sample size. At the same time, the sample size is of "normal" numerical value for its context for analyzing such students' course interactions.

Measurements in this article did not include the strength of links (edge weights) between students. This is because it is not necessary to do so when studying average path length, degree (in this case), clustering coefficient, and degree distribution. However, future studies of students' social and academic interaction need to take this property of links into account as it could prove to be vital in describing the function of students' course interaction networks.

The study reports on data obtained from a one-time measurement. A limitation arising from this is that it is not possible to identify dynamic changes in the social and academic network that can provide definitive interpretations about how the time-dependent evolution of complex systems can impact student persistence. However, the study embeds student persistence research into a complexity science framework using network theory and in this way can be argued to provide sufficient empirical rationale for analyzing students' course interactions as complex systems.

Differences in the social and academic networks were found that suggest that there could be different underlying constructs within Tinto's [14–16] social and academic systems, which are dependent on the agent interactions within the two different networks. Tinto [16] posited that the social and academic systems were nested within each other. Complexity science defines nested systems as complex systems that exist within a greater complex system and that have similar structures and dynamics, yet operating on different scales. Our analysis supports these claims and suggests that the social and academic networks are *definitely* entangled in each other and together are nested to create a larger network of interaction for the students. It is important to point out that the study implies that both the social and academic networks' structures and dynamics will affect the larger systems' structure and dynamics (i.e., the combined networks, which previous studies have examined).

Measurements of the clustering coefficients for each type of network and for each course showed similar trends between the courses; all the networks were much more clustered than a random network would be, and the social networks were more tightly linked than the academic networks. Hence, the formation of such social and academic networks is not random and should be seen as dependent on "rules of interaction" that can be understood as the social and academic norms that exist in the institution (i.e., the social and academic systems of the university). In other words, the formation of the structures of these networks is dependent on neighbor interaction [27], where the neighbor interaction is dependent on the social and academic systems. Furthermore, due to the clustering coefficient being lower in the academic network, it is possible to argue that the rules of interaction [22] (for instance, agency, intention, discourse patterns, collaborations, subcultures, norms, beliefs, and expectations) in the academic system do not influence the "clustering" on a system scale as much as the rules of interaction in the social system—they may even be isolating students from each other academically. An example of a "rule of interaction" in the realm of the academic system (which could act as an isolator) is being widely recognized as being the most scholarly student in the course. Such a rule of interaction either potentially helps a student form new and wider academic ties or presents a potential for the student to become isolated both socially and academically.

Average path length measurements illustrated differences between social and academic networks, where social networks tended to have shorter path lengths. Consequently, students' social networks were structured such that most

students were close to each other, whereas academic networks were more distinct—students work closely with some but not all of their fellow students. Visualizations of student networks illustrated similar trends, where the academic networks for each course were more likely to have a completely separate group of students who work together. Drawing on a complexity science perspective, further tentative characterization of these "rules of interaction" (that is, the social facts [22] in the social and academic system of the university) becomes possible through the analysis of the outcomes of the social system rules, as experienced by the social network of students in Course One and Course Two. The rules of interaction in the social system can then be described as creating social bridges between most students in the course, creating such a network where most students could eventually get ahold of any other student by asking around in the network. However, the academic system (i.e., the rules of interaction of the academic system) does not produce the same frequency of what could be called academic bridges (links that connect academic groups together) between all students, even creating islands of academic interaction that are not connected to the other students of the course at all.

This illustrative study found sufficient empirical evidence for the presence of distinct social and academic systems [12–16] that emerge from and can play a role of codependence on student-to-student interactions within the larger social and academic system of the university. This finding suggests that researchers, educators, and policy makers not only need to address critical aspects of the academic environment, but the same kind of research rigor needs to be used to address the social side of studying. The analyses illustrate a method and conceptual framing that could be used to better understand the structure and dynamics of these networks and exemplify that it makes sense to separate the interaction network of courses into two separate networks.

The empirical implications provide researchers, practitioners, and policy makers with a characterization of the complexities of participating student interactions that opens the way for the development of complexity science as an important framing for PER. This research perspective offers new ways to incorporate observations, measurements, and recognition of complex structures and dynamics that are inherent to a given PER problem. In terms of codependency of network structures and "rules of interaction" by using the constructs of social facts [22] together with complexity science [27], student persistence research is framed from a complexity science perspective where network structures are brought to the fore. The measurements of network structures can thus be interpreted as the outcomes of different rules of interaction, both social and academic. If measurements of network structure are to be used to predict educationally significant dynamics such as student persistence, a complexity science framing facilitates the interpretation of network measurements.

This article began by proposing the use of complexity science as a suitable research framework for PER and has provided an illustration of what kind of research becomes possible by framing a social network analysis within a complexity science perspective. For this to be possible there is a need to incorporate and develop new, effective methodologies that are aligned with complexity science. Such methodologies could encompass, for example, agent simulation within complex networks [35,56] to delve into how differences are formed from an agent interaction perspective, how they develop over time. Renowned researchers in the field of complexity in education [21,27] have posited implications for research and policy issues when these kinds of issues are considered. Our illustrative analysis complements well the future directions of research of physics students' networked interactions as described by Brewe, Kramer, and Sawtelle [23], and we envisage many aspects of the experience of teaching and learning physics taking on new visibility through guiding applications of complexity science.

[1] E. F. Redish, arXiv:1308.3911.

[2] J. Forsman, C. Linder, R. Moll, D. Fraser, and S. Andersson, A new approach to modelling student retention through an application of complexity thinking, Stud. Higher Educ. **39**, 68 (2014).

[3] J. Forsman, Licentiate thesis, Uppsala University, Sweden, 2011.

[4] V. Tinto, Research and practice of student retention: What is next?, J. Coll. Student Retention **8**, 1 (2006).

[5] Organization for Economic Cooperation, and Development, *Education at a Glance 2009: OECD Indicators* (OECD Publishing, Paris, 2009).

[6] Committee on Science, Engineering, and Public Policy, *Rising above the Gathering Storm: Energizing and Employing America for a Brighter Economic Future* (The National Academies Press, Washington, DC, 2007).

[7] Statistics Sweden, and National Agency for Higher Education (2003). Higher education. Students and graduated

students in undergraduate education 2001/02. Series: Utbildning och forskning [education and research] UF20 SM0301.

[8] Statistics Sweden, and National Agency for Higher Education (2005). Higher Education. Students and graduated students in undergraduate education 2003/04. Series: Utbildning och forskning [education and research] UF20 SM0501.

[9] Statistics Sweden, and National Agency for Higher Education (2007). Higher education. Students and graduated students in undergraduate education 2005/2006. Series: Utbildning och forskning [education and research] UF20 SM0701.

[10] Statistics Sweden, and National Agency for Higher Education (2009), Students and graduated students at first and second cycle studies 2007/08. Series: Utbildning och forskning [education and research] UF20 SM0902.

[11] Statistics Sweden, and National Agency for Higher Education (2010). Higher education. Students and graduated students at first and second cycle studies 2008/09. Series: Utbildning och forskning [education and research] UF20 SM1002.

[12] W. Spady, Dropouts from higher education: An interdisciplinary review and synthesis, Interchange **1,** 64 (1970).

[13] W. Spady, Dropouts from higher education: Toward an empirical model, Interchange **2,** 38 (1971).

[14] V. Tinto, Dropout from higher education: A theoretical synthesis of recent research, Rev. Educ. Res. **45,** 89 (1975).

[15] V. Tinto, *Leaving College: Rethinking the Causes and Cures of Student Attrition* (University of Chicago Press, Chicago, 1987).

[16] V. Tinto, Classrooms as communities: Exploring the educational character of student persistence, J. Higher Educ. **68,** 599 (1997).

[17] P. T. Terenzini and E. T. Pascarella, Toward the validation of Tinto's model of college student attrition: A review of recent studies, Res. High. Educ. **12,** 271 (1980).

[18] J. P. Bean, in *College Student Retention: Formula for Student Success*, edited by A. Seidman (Praeger, Westport, 2005), pp. 215–244.

[19] B. Sacerdote, Peer effects with random assignment: Results for Dartmouth roommates, Q. J. Econ. **116,** 681 (2001).

[20] L. S. Thomas, Ties that bind: A social network approach to understanding student integration and persistence, J. Higher Educ. **71,** 591 (2000).

[21] E. Durkheim, *Suicide: A Study in Sociology*, translated by M. Johansson (Argos, Lund, 2004) (original work published in 1961).

[22] R. K. Sawyer, *Social Emergence Societies as Complex Systems* (Cambridge University Press, New York, 2009).

[23] E. Brewe, L. Kramer, and V. Sawtelle, Investigating student communities with network analysis of interactions in a physics learning center, Phys. Rev. ST Phys. Educ. Res. **8,** 010101 (2012).

[24] D. Knocke and S. Yang, *Social Network Analysis Quantitative Applications in the Social Sciences*, 2nd ed. (Sage Publications, Thousand Oaks, 2008).

[25] L. C. Freeman, Centrality in social networks conceptual clarification, Soc. Networks **1,** 215 (1978).

[26] M. E. J. Newman, The structure and function of complex networks, SIAM Rev. **45,** 167 (2003).

[27] B. Davis and D. Sumara, *Complexity and Education: Inquiries into Learning, Teaching and Research* (Lawrence Erlbaum Assoc., Hillsdale, NJ, 2006).

[28] M. Mitchell, *Complexity: A Guided Tour* (Oxford University Press, New York, 2009).

[29] M. M. Waldrop, *Complexity: The Emerging Science at the Edge of Order and Chaos* (Simon and Schuster, New York, 1992).

[30] B. Davis, D. Sumara, and R. Luce-Kapler, *Engaging Minds: Changing Teaching in Complex Times* (Routledge, New York, 2008).

[31] B. Davis and D. Sumara, Cognition, Complexity, and Teacher Education, Harv. Educ. Rev. **67,** 105 (1997).

[32] F. Capra, *The Hidden Connections: A Science for Sustainable Living* (Doubleday, New York, 2002).

[33] B. Davis and E. Simmt, Mathematics-for-teaching: An ongoing investigation of the mathematics that teachers (need to) know, Educ. Stud. Math. **61,** 293 (2006).

[34] J. B. Batista and L. da F. Costa, Knowledge acquisition by networks of interacting agents in the presence of observation error, Phys. Rev. E **82,** 016103 (2010).

[35] L. da F. Costa, Learning about knowledge: A complex network approach, Phys. Rev. E **74,** 026103 (2006).

[36] E. W. Zegura, K. L. Calvert, and M. J. Donahoo, A quantitative comparison of graph-based models for Internet topology, IEEE/ACM Trans. Netw. **5,** 770 (1997).

[37] M. Faloutsos, P. Faloutsos, and C. Faloutsos, On power-law relationships of the Internet topology, Comput. Commun. Rev. **29,** 251 (1999).

[38] R. Pastor-Satorras, A. Vázquez, and A. Vespignani, Dynamical and correlation properties of the Internet, Phys. Rev. Lett. **87,** 258701 (2001).

[39] R. Albert, H. Jeong, and A.-L. Barbási, Error and attack tolerance of complex networks, Nature (London) **406,** 378 (2000).

[40] D. Butler, Souped-up search engines, Nature (London) **405,** 112 (2000).

[41] S. Redner, How popular is your paper? An empirical study of the citation distribution, Eur. Phys. J. B **4,** 131 (1998).

[42] X. Cheng, C. Dale, and J. Liu, Statistics and social network of YouTube videos, in *Proceedings of the 16th International Workshop on Quality of Service, Enschede, The Netherlands, 2008* (Institute of Electrical and Electronics Engineers, Inc., Piscataway, NJ, 2008), pp. 229–238.

[43] B. Davis, *Inventions of Teaching: A Genealogy* (Lawrence Erlbaum Assoc., Mahwah, NJ, 2004).

[44] E. W. Morrison, Newcomers' relationships: The role of social network ties during socialization, Acad. Manag. J. **45,** 1149 (2002).

[45] P. V. Marsden, in *The Sage Handbook of Social Network Analysis*, edited by J. P. Scott and P. J. Carrington (Sage, London, 2011), pp. 370–388.

[46] C. Robson, *Real World Research: A Resource for Social Scientists and Practitioner-Researchers*, 2nd ed. (Blackwell Publishing, Oxford, 2002).

[47] http://www.gephi.org.

[48] J. M. McPherson, P. A. Popielarz, and S. Drobnic, Social networks and organizational dynamics, Am. Soc. Rev. **57,** 153 (1992).

[49] C. M. Hurvich and C.-L. Tsai, Regression and time series model selection in small samples, Biom. J. **76**, 297 (1988).

[50] H. Akaike, Statistical predictor identification, Ann. Inst. Stat. Math. **22**, 203 (1970).

[51] T. Rizzuto, J. LeDoux, and J. Hatala, It's not just what you know, it's who you know: Testing a model of the relative importance of social networks to academic performance, Soc. Psychol. Educ. **12**, 175 (2009).

[52] T. Maundeni, The role of social networks in the adjustment of African students to British society: Students' perceptions, Race Ethnicity Educ. **4**, 254 (2001).

[53] E. W. Stacy and G. A. Mihram, Parameter estimation for a generalized gamma distribution, Technometrics **7**, 349 (1965).

[54] M. L. Delignette-Muller, R. Pouillot, J.-B. Denis, and C. Dutang, FITDISTRPLUS: Help to fit of a parametric distribution to non-censored or censored data, *R Package version* 1.0–1 (2013).

[55] R Core Team, *R: A Language and Environment for Statistical Computing* (R Foundation for Statistical Computing, Vienna, 2013).

[56] S. Boccaletti, C. Latora, Y. Moreno, M. Chavez, and D. U. Hwang, Complex networks: Structure and dynamics, Phys. Rep. **424**, 175 (2006).