Nonrandom behavior in the projection of random bipartite networks

Izat B. Baybusinov,[1] Enrico Maria Fenoaltea,[1,*] Jungying Cui,[2,†] and Yi-Cheng Zhang[1]
[1]*Physics Department, University of Fribourg, Chemin du Musée 3, 1700 Fribourg, Switzerland*
[2]*Research Center for Intelligence Traditional Chinese Medicine, Chongqing College of Traditional Chinese Medicine, Chongqing 402760, China*

(Received 26 March 2023; accepted 24 January 2024; published 20 February 2024)

There are two main categories of networks studied in the complexity physics community: Monopartite and bipartite networks. In this paper, we present a general framework that provides insights into the connection between these two classes. When a random bipartite network is projected into a monopartite network, under quite general conditions, the result is a nonrandom monopartite network, the features of which can be studied analytically. Unlike previous studies in the physics literature on complex networks, which rely on sparse-network approximations, we provide a complete analysis, focusing on the degree distribution and the clustering coefficient. Our findings primarily offer a technical contribution, adding to the current body of literature by enhancing the understanding of bipartite networks within the community of physicists. In addition, our model emphasizes the substantial difference between the information that can be extracted from a network measuring its degree distribution, or using higher-order metrics such as the clustering coefficient. We believe that our results are general and have broad real-world implications.

DOI: 10.1103/PhysRevE.109.024308

I. INTRODUCTION

In the past three decades, networks have been a major research subject [1–3] for the study of the relationships between multiple agents. Among them, bipartite networks are one particular class that is widely studied in the literature [4–7]. We can present a bipartite network as a set of people who are connected to a set of events. For example, agents can participate in events and, if everyone randomly chooses a few of them, we obtain a typical random bipartite graph. Traditional networks, instead, belong to the so-called class of monopartite networks. We can present a random monopartite network as formed by people only. If each agent chooses randomly to connect with a few other agents, then we obtain a random monopartite network with nodes solely represented by people. This is the well-known Erdos-Renyi network [8].

Networks of the latter class can be constructed from those of the former class by projection [9,10]: A new monopartite network can be constructed in which two individuals are connected if, in the bipartite counterpart, they have at least one common event. In the literature, such people-events networks are called affiliation networks [11,12], and their projections are studied in many different real systems. In many applications, these projections are explicit, such as the process of connecting diseases associated with the same gene [13], countries exporting the same product [14], people belonging to the same group or team [15], scientists collaborating on the same project [16,17], or actors appearing in the same movie [18,19]. In other cases, projection is more implicit, e.g., when connections in a social network arise from shared interests [20]. In addition, the projection method has been applied to make predictions about the bipartite network itself [21] or to design recommendation systems [22,23].

One of the reasons for studying projections is that some structural properties of real monopartite networks, particularly social networks, can be explained if we assume that such networks result from a projection of a latent bipartite network. Empirical evidence indicates that in social networks, the clustering coefficient—the fraction of connected triplets of nodes—is significantly higher than that of a random network with the same degree distribution [24,25]; at the same time, in the projection of any bipartite network, notably high clustering is generally observed compared to a random network with the same degree distribution [4,26,27]. In addition to various nontrivial mechanisms influencing the formation of a social group [28,29], the observed high clustering coefficient could be partially explained as a consequence of a projection mechanism [24]. In short, projected networks manifest nontrivial properties.

For these reasons, many physicists have explored the properties of projected networks, specifically focusing on their degree distribution and clustering coefficient, starting from the seminal paper by Newman, Watts, and Strogatz [4]. However, in the physics literature on complex networks, analytical results have consistently been derived under the assumption of sparse networks [30]. This approximation has two motivations. First, in most real-world cases, social network density is low [31] (indeed, the generating function method, a popular tool to investigate network properties, is exact only for treelike networks [32]). Second, projecting a bipartite network gives rise to nontrivial topological issues. Unlike traditional monopartite networks, where each tie is formed

*enricofenoaltea@hotmail.it
†Jungying.Cui@unifr.ch

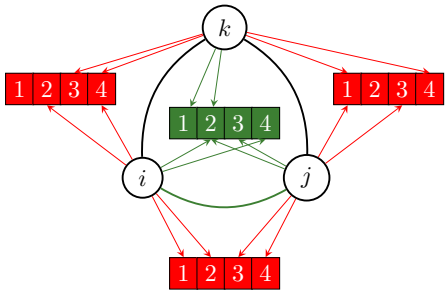

FIG. 1. Difference between a projected bipartite network and a random network (i.e., the projected network without considering correlations) with four events and three agents. In the bipartite configuration, there is a unique list of events (green squares). Meanwhile, in the random network case, there is an independent list of events (red squares) for each possible link, i.e., for each pair of individuals (white circles).

independently, this is not the case in projected bipartite networks (see Fig. 1) where an independent link in the bipartite system simultaneously generates multiple connections in the projected network [26]. However, with the increasing amount of data at our disposal and the availability of social media, the interconnectedness of most social systems is growing significantly [33]. In other words, the approximation of sparse networks may no longer be applicable. Therefore, it is essential to develop theoretical frameworks that can be validated or used as null models even for denser networks.

In this paper, we complement the results obtained in the community of physicists working in bipartite projection by exact analytical results on the degree distribution and clustering coefficient. In particular, we introduce a simple random bipartite network model, and we study the behavior of its projection when varying the bipartite network density. Apart from the technical contribution, our findings illustrate, within a simple and intuitive framework, the profound distinction of the information about a network when measured by its degree distribution and other higher-order moments such as the clustering coefficient.

The rest of the paper is organized as follows. In the second section, we introduce our random bipartite projection framework. The third section is divided into three parts: In the first, we compute exactly the degree distribution of the projected network generated by our model, and then we show that in the limit of sparse and large networks, it corresponds to that of an Erdos-Renyi network; in the second, we perform exact calculations also for the clustering coefficient, showing, however, that this metric does not reduce to that of a random network in the limit of sparse networks; in the third part, we show alternative interpretations to explain the emergence of correlations in the projection process and study the scaling properties of the clustering coefficient. Finally, in the last section, we conclude the paper.

II. THE FRAMEWORK

Keeping with the people and events metaphor for bipartite networks (although our framework could be applied to any bipartite network), we construct a random network consisting of K people and N events, and we assume that each person connects to any of the events independently with probability $\beta \in [0, 1]$. Thus, the elements of the adjacency matrix $\mathbf{n}$ of the bipartite network are written as

$$n_{i\alpha} = \begin{cases} 1 \text{ with probability } \beta, \\ 0 \text{ with probability } 1-\beta, \end{cases} \tag{1}$$

where individuals are labeled by $i = 1, \ldots, K$ and events by $\alpha = 1, \ldots, N$. From this adjacency matrix, we can compute all the network measures. For our purposes, it is useful to work with the degree distribution, i.e., the probability $B_N(m)$ that an individual is connected to m events out of N. Since the network is random, this probability is a binomial distribution:

$$B_N(m) = \binom{N}{m} \beta^m (1-\beta)^{N-m}, \tag{2}$$

as is the case in a random bipartite network. From this we construct the monopartite network, resulting from the projection of the bipartite network. Since two agents are connected when they share at least one event, the elements of the adjacency matrix $\mathbf{A}$ in the projected monopartite network can be written as

$$A_{ij} = \theta\left(\sum_{\alpha=1}^{N} n_{i\alpha} n_{j\alpha}\right), \tag{3}$$

where $\theta(\cdot)$ is the Heaviside function, and $i, j = 1, \ldots, K$.

III. ANALYTIC DEVELOPMENTS

We want to study the properties of $\mathbf{A}$ averaged over all the realizations of the bipartite network. In particular, we are interested in the probability of the entire adjacency matrix, i.e., the probability $P(A_{11}, A_{12}, \ldots, A_{KK}) \equiv P(\mathbf{A})$ to observe a specific configuration of $\mathbf{A}$. This can be written as

$$P(\mathbf{A}) = \sum_{\mathbf{n}\in\mathbf{n}_{NK}} s(\mathbf{n}) \prod_{i=1}^{K} \prod_{j=1}^{K} \delta_{A_{ij}}^{\theta\left(\sum_\alpha n_{i\alpha} n_{j\alpha}\right)}, \tag{4}$$

where δ_x^0 is the Kronecker delta (equal to 1 if $x = 0$ and 0 otherwise), $\mathbf{n}_{NK}$ is the set of all possible adjacency matrices of a bipartite network with K people and N events, and

$$s(\mathbf{n}) = \beta^{\sum_{ij} n_{ij}} (1-\beta)^{N^2 - \sum_{ij} n_{ij}}$$

is the probability to have a specific configuration $\mathbf{n} \in \mathbf{n}_{NK}$.

Note that, according to Eq. (4), there are K^2 conditions on the probability of the adjacency matrix of the projected network. Consequently, computing it becomes a challenging task. Instead, we calculate a simpler quantity—the probability p that an element A_{ij} is 1. In a first-order approximation, p is given by the extremal distribution of sampling at least one element out of N (two people are connected if they share at least one event). Given that in our model the probability of not sharing a specific event is $(1-\beta^2)$, the probability of sharing at least one event is

$$p = 1 - (1-\beta^2)^N. \tag{5}$$

By symmetry, p is the same for any i and j, so the expected value of A_{ij} is $\langle A_{ij} \rangle \equiv \sum_{\mathbf{A}} P(\mathbf{A}) A_{ij} = p$. Hence, the link probability p provides no information about the bipartite structure

in the monopartite version. Neglecting the correlation arising from the transition from a bipartite to a monopartite network, one might mistakenly assume that the network is random with connection probability p, akin to Erdos-Renyi. Nevertheless, for finite N, correlations play an important role for the degree distribution, and the projected network cannot be approximated as a random network. In Fig. 1, we provide a simple example with $N = 4$ and $K = 3$ that illustrates the effect of neglecting correlations. To show it formally, in the following we compute the degree distribution $P(d)$ of the projected network.

A. Degree distribution

The degree d_i of the individual i in the monopartite network is given by $d_i = \sum_j A_{ij}$. If an individual with degree d participates in m events, there must be d individuals out of $K-1$ (we do not allow self-link) meeting her at one of those m events. Summing over m, we have

$$P(d) = \sum_{m=0}^{N} B_N(m) \binom{K-1}{d} [1 - B_m(0)]^d B_m(0)^{K-1-d}. \quad (6)$$

As mentioned in the physics literature of complex networks, the degree distribution of a projected network was derived in the sparse network limit, i.e., when $N \to \infty$ and βN is fixed. In this limit the sum in Eq. (6) is dominated by the maximum of $B_N(m)$, i.e., when $m = \beta N$. In this case, $P(d)$ is a binomial distribution with the probability parameter given by $1 - e^{-\beta^2 N} \approx p$. Thus, one retrieves the degree distribution of an Erdos-Renyi network with connection probability p.

In general, one can calculate the moment-generating function to show that $P(d)$ is binomial in the large-N limit:

$$\langle e^{\lambda d} \rangle = e^{\lambda(K-1)} \sum_{l=0}^{K-1} \Big[(\beta(1-\beta)^l + (1-\beta))^N \times \binom{K-1}{l} (e^{-\lambda} - 1)^l \Big], \quad (7)$$

and at large N we have $[\beta(1-\beta)^l + (1-\beta)]^N \approx e^{-l\beta^2 N}$. So we obtain

$$\langle e^{\lambda d_i} \rangle \approx e^{\lambda(K-1)} [1 - e^{-\beta^2 N}(1 - e^{-\lambda})]^{K-1}, \quad (8)$$

which corresponds to the moment-generating function of a binomial distribution with probability parameter $1 - e^{-\beta^2 N}$. Hence, with many events, the projected network and a random monopartite network cannot be distinguished when measuring their degree distributions. Note that the above argument would fail if K scaled with N.

B. Clustering coefficient

In addition to the degree distribution, we also consider a higher-order measure, namely the clustering coefficient C of the projected network [19]. It is defined as the ratio of the number of closed triplets (three nodes connected to each other) to the number of open triplets (two unconnected nodes with a common neighbor). Formally,

$$C = 3 \frac{\langle \text{triplets} \rangle}{\langle \text{open triplets} \rangle} = \frac{\sum_{i,j,k} \langle A_{ij} A_{jk} A_{ki} \rangle}{\sum_i \langle (\sum_j A_{ij})(\sum_j A_{ij} - 1) \rangle}. \quad (9)$$

This quantity represents the probability that, when selecting two links with one end in common, the other two ends are also connected. In other words, it measures the level of connectivity between the neighbors of a node. While there exist different definitions of the clustering coefficient in the literature [34], we follow the most widely accepted definition. Understanding the properties of this metric is important not only in the study of social networks where, as mentioned in the Introduction, a high clustering coefficient has been empirically observed, but also in the study of nonsocial networks such as the brain or information network. Indeed, the clustering coefficient is clearly related to the concept of triadic motifs or 3-cliques, i.e., triangular patterns that play an important role in information processing within networks [35].

Note that, in the case of a random network generated with probability p, the clustering coefficient is exactly p [36]. To compute the numerator of Eq. (9) in our projected network, we must consider the probability that three agents i, j, k are connected. There are two contributions to this probability. The first is the probability P_2 that in the bipartite network i, j, k participate in the same event at least once [Fig. 2(a)]. Similarly to the derivation of p in Eq. (5), P_2 is given by the extremal distribution of sampling at least one element out of N. Since the probability of three people not sharing a specific event is $(1 - \beta^3)$, the probability that they share at least one is $P_2 = 1 - (1 - \beta^3)^N$. The second contribution is the probability P_3 that the three individuals participate pairwise in three different events [Fig. 2(b)]. This is computed as follows:

$$P_3 = \sum_{m,n,l} B_N(m) B_m(n) [B_n(0) B_{m-n}(l)][1 - B_{N-m}(0)]. \quad (10)$$

The first factor in this sum is the probability that the first individual, say i, participates in m events out of N; the second and third factors are the conditions that j participates in n of these m events and that k participates in l of the remaining $m - n$ events, respectively. In this way, individuals j and k both meet i, but they do not meet each other in any of the m events chosen by i. The last factor is the condition that j and k meet in the other $N - m$ events. Note that in the physics literature

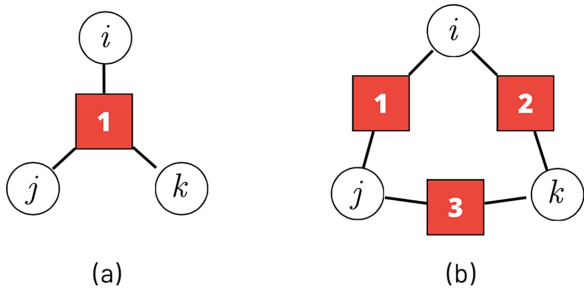

FIG. 2. Illustration of the two configurations of the bipartite network leading to a triangle in the projected network. Individuals are represented by white circles, and events by red squares. The configuration in (a) appears with probability P_2; the configuration in (b) appears with probability P_3.

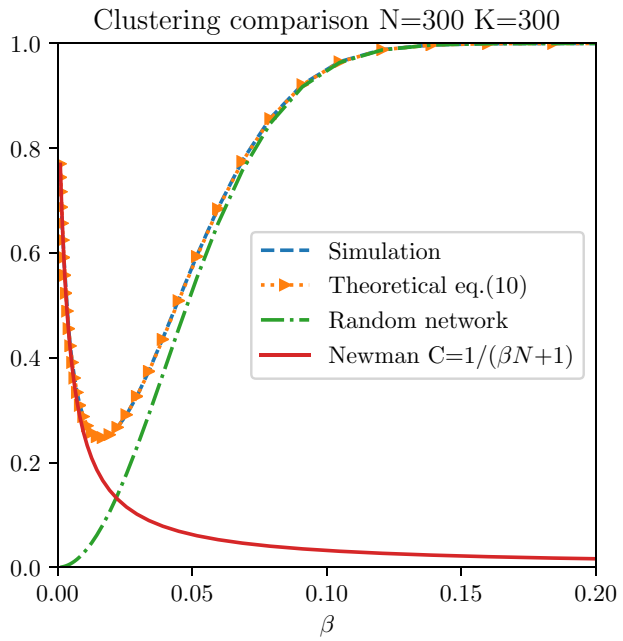

FIG. 3. Clustering coefficient vs β. The clustering of our projected network is compared with that of a random network with connection probability p and that of the Newman model in [4].

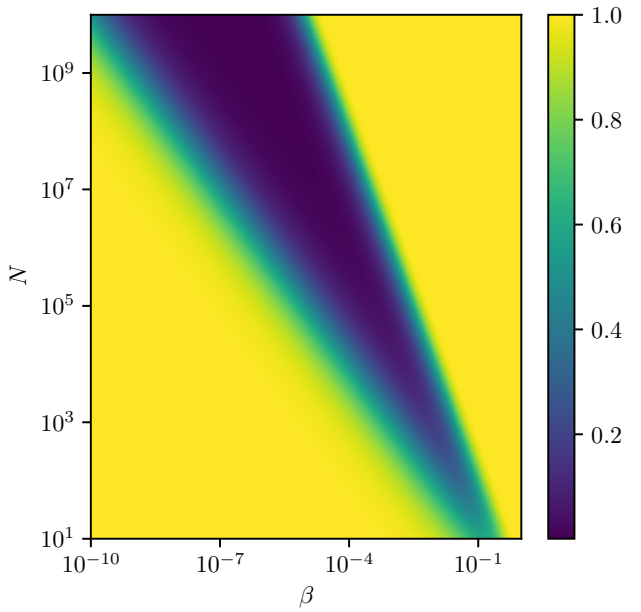

FIG. 4. Phase diagram of the clustering coefficient of Eq. (11) in the $[\beta, N]$ parameter space. When N increases, the width of the minimum valley increases since the left border scales as $1/N$ and the right one as $1/\sqrt{N}$.

on bipartite networks, this term has always been overlooked in the analytical derivations of the clustering coefficient, as only the sparse network limit has been considered (see, in particular, the derivation of the clustering coefficient in a projected network in [30]). Indeed, in this limit, P_2 dominates, and the effect of P_3 is negligible.

Now, the denominator of Eq. (9) is straightforwardly written as $K\langle d_i(d_i-1)\rangle$, and, after some algebra, the clustering coefficient can be written as

$$C = 3 - \frac{2 + (1+2\beta)^N(1-\beta)^{2N} - 3(1-\beta^2)^N}{1 - 2(1-\beta^2)^N + (1-2\beta^2+\beta^3)^N}. \tag{11}$$

Figure 3 shows that our analytical result is consistent with numerical simulations. It is interesting to note that our model is equivalent to a special case of a model studied in mathematics, namely the random intersection graph, where the same functional form for the clustering coefficient can be derived using a different approach [37].

From Eq. (11), the clustering is 1 at $\beta \approx 1/N$. Here the network is separated into disconnected communities composed of individuals that share all the events they participate in (i.e., within a community, each individual participates in all the events shared with their neighbors and does not participate in all others). Increasing β, spurious links start to appear between communities, reducing the clustering until $\beta \sim N^{-1/2}$ (see Fig. 4), where links between two randomly sampled individuals are highly frequent. From this point on, the increase in link density generates a higher clustering coefficient. Actually, the clustering reaches its minimum value when the network seems to move from a phase of isolated communities to one with a unique connected component. To evaluate this minimum, we compute C by expanding (11) at small β:

$$C = \frac{1 + N^2\beta^3}{N\beta + 1} + o(\beta^3). \tag{12}$$

The minimum is reached when β satisfies the following equation:

$$2\beta^3N^3 + 3\beta^2N^2 - N = 0. \tag{13}$$

For large N, the solution β_c scales as $N^{-2/3}$. Instead, the clustering at this minimum decreases quite slowly with N:

$$C(\beta_c) \approx N^{-1/3}. \tag{14}$$

Note that, when looking at the clustering coefficient, our projected network is fundamentally different from a random network, as shown in Fig. 3. Unlike the case of the degree distribution, this is also true for large N (sparse network). This is a consequence of the aforementioned correlations arising from dimensionality reduction due to the transition from a bipartite to a monopartite network. While this was already known from the results obtained in the sparse network limit (without the term P_3), our findings emphasize even more the special properties of projected networks. We not only show that they generally have a higher clustering coefficient compared to that of classic monopartite networks with the same degree distribution, but also that their clusterization with density is nonmonotonic. Therefore, an important message to draw is that when measuring the properties of a network (real or artificial), the degree distribution is not a sufficient measure to extract all the information. The formation process of such a network (e.g., whether it is the result of a projection or not) is crucial for higher-order measures such as the clustering coefficient and their dynamics when other parameters, such as density, change.

C. Other properties of the model

In this section, we discuss additional properties and implications of our framework.

1. The geometrical argument

First we want to further elucidate why, in light of our model, the clustering coefficient of a projected network is higher than that of a standard monopartite network with the same degree distribution. We have already mentioned that this is due to the presence of correlations arising from the dimensionality reduction (bipartite to monopartite network). The presence of correlations can also be explained by a geometrical argument. For any individual i, let us define $\mathbf{v}_i = \{n_{i\alpha}\}_{\alpha=1,\dots,N}$ as the vector whose elements are 1 if i participates in event α, and zero otherwise. We can interpret this vector also as a vertex of an N-dimensional hypercube. Thus, each vector (or individual) is a randomly drawn vertex on the hypercube with an average distance from the origin equal to βN. On this hypercube, two individuals are connected whenever their inner product is positive, that is, when $\mathbf{v}_i \cdot \mathbf{v}_j > 0$. From this perspective, correlations in the projected network naturally appear as dimensional constraints on the hypercube.

2. The Hamiltonian approach

To generalize our results from a statistical mechanics perspective, one can introduce a Hamiltonian for the matrix $\mathbf{A}'$:

$$H[\mathbf{A}'] = \sum_{i=1}^{K}\sum_{j=1}^{K} A'_{ij}(1 - \mathbf{v}_i \cdot \mathbf{v}_j). \tag{15}$$

This describes a model in which individuals have a nonzero probability—determined by the Boltzmann weights and the associated temperature T—of being connected in the projected network even though they have no common events in the bipartite network. The minimum of the Hamiltonian is when $\mathbf{A}'$ is identically the adjacency matrix in Eq. (3). So, by studying this Hamiltonian in the zero-temperature limit, one can study the properties of $\mathbf{A}$. Computing the partition function Z, we find

$$Z = \prod_{ij}\sum_{A'_{ij}=0}^{1} e^{-\frac{1}{T}H[\mathbf{A}']} = \prod_{ij}\left(1 + e^{-\frac{1}{T}(1-\mathbf{v}_i\cdot\mathbf{v}_j)}\right), \tag{16}$$

and the link probability is

$$n_{ij} := \sum_{A_{ij}=0}^{1} A_{ij}P(\mathbf{A}_{ij}) = \frac{1}{e^{\frac{1}{T}(1-\mathbf{v}_i\cdot\mathbf{v}_j)} + 1}, \tag{17}$$

which is a Fermi distribution with chemical potential $\mu_{ij} = \mathbf{v}_i \cdot \mathbf{v}_j$. Note that these chemical potentials are not independent, so when averaging, correlations appear when computing quantities involving three or more individuals.

3. The scaling properties

Finally, it is interesting to study the properties of our system as a function of N as the scaling gives different behaviors. For example, in [30], Newman, Watts, and Strogatz describe a random bipartite network model in which, differently from above, they fix the average degree $z = O(1)$ in the bipartite network so that $\beta = z/N$, and they study the large-N limit. Since β scales as $1/N$, for large N the bipartite network is sparse. In this case, it is possible to neglect the term P_3, and the resulting monopartite network of people is clustered into different communities, as we have shown above (when $\beta < 1/N$). Thus, with the approach of [30], the effects arising at higher values of β cannot be observed (see Fig. 3), and the network always remains fragmented. In this context, our findings complement and extend this and the subsequent traditional physicists' models on bipartite projections, which are extensively employed as null models and benchmarks for various real systems. In fact, we show exact results that hold for any network density and not just in the (analytically) simpler limit of large N, which is appropriate only for sparse networks. Examining a naturally dense real network, or increasing the density artificially by reducing the number of events (for example, through aggregating similar events or limiting the data set), can lead to a richer scenario in which links between different communities facilitate communication and information diffusion. One can argue that interesting social phenomena can be detected by observing the system at a fixed scale (low N).

However, some systems are inherently sparse, and it would not make sense to reduce the number of events to increase their density. For instance, in real offline social networks, there is a cognitive constraint on the number of connections each individual can handle [38]. Our exact results can also offer additional insights into these systems. For example, we can fix, in the manner of Newman, the average degree in the projected network to be of the order of 1:

$$\langle d_i\rangle = Kp = K[1 - (1 - \beta^2)^N] \sim O(1) \tag{18}$$

for any i. In this way, we obtain a new scaling:

$$1 - (1 - \beta^2)^N \sim 1/K \;\rightarrow\; \beta \sim \sqrt{\frac{1}{NK}}. \tag{19}$$

So the probability to participate in an event depends on both N and K. If $K \approx \sqrt{N}$, the population is clustered in different communities. Instead, if $K < \sqrt{N}$, the system is more cohesive. This mechanism can help to understand and manage social phenomena such as the observed social fragmentation and polarization [33,39].

IV. CONCLUSIONS

To summarize, in this paper we have examined the projection of a random bipartite network of people connected to events. Averaging over all the realizations, it is possible to extract information about the bipartite structure only with measures involving three or more agents. Indeed, lower-order quantities (e.g., the degree distribution) are not distinguishable from those of a random network if the number of events is large enough. This is evident by mapping the bipartite structure into a hypercube and noting that this imposes geometrical constraints. We have shown analytically how these constraints affect the properties of the projected network. In particular, we have investigated the scaling properties of the clustering coefficient in the monopartite network, showing, for example, that when there are many events, the system is fragmented into smaller communities.

We have also demonstrated that our model adds value over classical works by physicists on bipartite projection. In fact, we have obtained exact analytical results outside the limit of sparse networks. Consequently, our model can be validated for a wider range of real bipartite systems, or used as a null model. For example, one can analyze an evolving bipartite network system, where the number of links increases over time, and compare the results of our model with the observed clustering coefficient. It might also be interesting to analyze real monopartite evolving networks with the hypothesis that their topology is a consequence of an underlying bipartite network. Indeed, in most data sets of social systems, we can only detect the monopartite network, that is, the connections between individuals, and the above approach could help reveal the hidden reasons (such as the coparticipation in different events). This could be a driving force for the apparent person-to-person relationships.

Moreover, a natural extension of our framework, which needs to be explored, is the study, using the same analytical approach, of the properties of the projected network when weights are included. This occurs when the links between two individuals in the projected network are weighted, corresponding to the number of events the two individuals share in the bipartite network. In this paper, we chose not to include weights not only for the sake of simplicity but also because our primary focus was on studying the standard projected network. Indeed, the intriguing effects we observed stem precisely from the fact that, by excluding weights, a significant amount of information from the bipartite network is lost during the projection process. However, our model can be extended to include weights, and the results could, for example, give benchmarks on the effectiveness of similarity measures in real data sets.

In conclusion, our work offers an alternative framework for understanding the differences between monopartite and bipartite networks, and although we do not claim that our model accurately represents any specific real-world system, we believe it can provide valuable insights as a benchmark model.

[1] R. Albert and A.-L. Barabási, Rev. Mod. Phys. **74**, 47 (2002).
[2] S. Boccaletti, V. Latora, Y. Moreno, M. Chavez, and D.-U. Hwang, Phys. Rep. **424**, 175 (2006).
[3] A. S. d. Mata, Braz. J. Phys. **50**, 658 (2020).
[4] M. E. Newman, D. J. Watts, and S. H. Strogatz, Proc. Natl. Acad. Sci. (USA) **99**, 2566 (2002).
[5] M. J. Barber, Phys. Rev. E **76**, 066102 (2007).
[6] J. J. Ramasco, S. N. Dorogovtsev, and R. Pastor-Satorras, Phys. Rev. E **70**, 036106 (2004).
[7] P. Holme, F. Liljeros, C. R. Edling, and B. J. Kim, Phys. Rev. E **68**, 056107 (2003).
[8] E. N. Gilbert, Ann. Math. Stat. **30**, 1141 (1959).
[9] T. Zhou, J. Ren, M. Medo, and Y.-C. Zhang, Phys. Rev. E **76**, 046115 (2007).
[10] G. Cimini, A. Carra, L. Didomenicantonio, and A. Zaccaria, Commun. Phys. **5**, 76 (2022).
[11] K. Faust, Social Netw. **19**, 157 (1997).
[12] L. K. Gallos, D. Rybski, F. Liljeros, S. Havlin, and H. A. Makse, Phys. Rev. X **2**, 031014 (2012).
[13] K.-I. Goh, M. E. Cusick, D. Valle, B. Childs, M. Vidal, and A.-L. Barabási, Proc. Natl. Acad. Sci. (USA) **104**, 8685 (2007).
[14] A. Tacchella, M. Cristelli, G. Caldarelli, A. Gabrielli, and L. Pietronero, Sci. Rep. **2**, 723 (2012).
[15] M. A. Porter, P. J. Mucha, M. E. Newman, and C. M. Warmbrand, Proc. Natl. Acad. Sci. (USA) **102**, 7057 (2005).
[16] Y. Sun, R. Barber, M. Gupta, C. C. Aggarwal, and J. Han, in *Proceedings of the 2011 International Conference on Advances in Social Networks Analysis and Mining* (IEEE, Piscataway, NJ, 2011), pp. 121–128.
[17] S. Uddin, L. Hossain, and K. Rasmussen, PLoS ONE **8**, e57546 (2013).
[18] L. A. N. Amaral, A. Scala, M. Barthelemy, and H. E. Stanley, Proc. Natl. Acad. Sci. (USA) **97**, 11149 (2000).
[19] D. J. Watts and S. H. Strogatz, Nature (London) **393**, 440 (1998).
[20] R. L. Breiger, Soc. Forces **53**, 181 (1974).
[21] A. R. Benson, R. Abebe, M. T. Schaub, A. Jadbabaie, and J. Kleinberg, Proc. Natl. Acad. Sci. (USA) **115**, 11221 (2018).
[22] L. Lü, M. Medo, C. H. Yeung, Y.-C. Zhang, Z.-K. Zhang, and T. Zhou, Phys. Rep. **519**, 1 (2012).
[23] P. Resnick and H. R. Varian, Commun. ACM **40**, 56 (1997).
[24] M. E. J. Newman and J. Park, Phys. Rev. E **68**, 036122 (2003).
[25] A. Isfandyari-Moghaddam, M. K. Saberi, S. Tahmasebi-Limoni, S. Mohammadian, and F. Naderbeigi, J. Inf. Sci. **49**, 1126 (2023).
[26] M. Latapy, C. Magnien, and N. Del Vecchio, Social Netw. **30**, 31 (2008).
[27] J.-L. Guillaume and M. Latapy, Inf. Process. Lett. **90**, 215 (2004).
[28] E. M. Fenoaltea, F. Meng, R.-R. Liu, and M. Medo, Phys. Rev. Res. **5**, 013023 (2023).
[29] C. Stadtfeld, K. Takács, and A. Vörös, Social Netw. **60**, 129 (2020).
[30] M. E. J. Newman, S. H. Strogatz, and D. J. Watts, Phys. Rev. E **64**, 026118 (2001).
[31] J.-L. Guillaume and M. Latapy, Physica A: Stat. Mech. Appl. **371**, 795 (2006).
[32] M. Li and B.-H. Wang, in *Journal of Physics: Conference Series* (IOP, Bristol, UK, 2015), Vol. 604, p. 012013.
[33] T. Minh Pham, I. Kondor, R. Hanel, and S. Thurner, J. R. Soc. Interface. **17**, 20200752 (2020).
[34] M. Á. Serrano and M. Boguna, Phys. Rev. E **74**, 056114 (2006).
[35] R. Milo, S. Shen-Orr, S. Itzkovitz, N. Kashtan, D. Chklovskii, and U. Alon, Science **298**, 824 (2002).
[36] Y. Li, Y. Shang, and Y. Yang, Inf. Sci. **382-383**, 350 (2017).
[37] M. Deijfen and W. Kets, Probab. Eng. Inf. Sci. **23**, 661 (2009).
[38] R. I. Dunbar, J. Hum. Evol. **22**, 469 (1992).
[39] T. M. Pham, A. C. Alexander, J. Korbel, R. Hanel, and S. Thurner, Sci. Rep. **11**, 17188 (2021).