

RESEARCH

Open Access



Properties of the connected components in projections of random bipartite networks: effects of clique size fluctuations

Yuka Fujiki^{1,2} and Shogo Mizutaka^{3*}

*Correspondence:
shogo.mizutaka.kh50@vc.ibaraki.ac.jp

¹ Frontier Research Institute for Interdisciplinary Sciences (FRIS), Tohoku University, 6-3, Aramaki aza Aoba, Aoba-ku, Sendai, Miyagi 9808578, Japan

² Advanced Institute for Materials Research (AIMR), Tohoku University, 2-1-1, Katahira, Aoba-ku, Sendai, Miyagi 9808577, Japan

³ Graduate School of Science and Engineering, Ibaraki University, 4-12-1, Nakanarusawa, Hitachi, Ibaraki 3168511, Japan

Abstract

Projection is a helpful description for treating bipartite networks as (monopartite) networks with pairwise interactions. Projections induce correlation spontaneously, avoiding negative degree correlation, even if bipartite networks are entirely random. In this study, we examined the structure of projections of random bipartite networks characterized by the degree distribution of individual and group nodes through the generating function method. We decomposed a projection into two subgraphs, the giant component, and finite components and analyzed their degree correlation. We demonstrate that positive degree correlations in projections originating from the clique size fluctuation remain after the decomposition at the set of finite components, although the values of their clustering coefficient are still finite. The giant component can exhibit either positive or negative degree correlations based on the structure of the projection. However, they are positively correlated in most cases. In addition, we found that a projection removed the giant component coincides with one in the subcritical phase, i.e., the discrete duality relation, when the degree distributions for group and individual are of Poisson.

Keywords: Projections, Bipartite networks, Degree correlation, Clique clustering

Introduction

The structural complexity of networks consisting of nodes and edges emerges from the pairwise relations between two nodes and the higher-order structure induced by group relations among more than two nodes. Several researchers have collaborated to produce scientific results. Many molecules are involved in intravitreal chemical reactions. Some species compete or coexist in the same ecosystems. Such group relationships are widely observed in empirical networks ranging from nature to society, including in neural, biological, ecological, social, and technological networks (Newman et al 2002; Milo et al 2002; Ugander et al 2012; Petri et al 2014; Benson et al 2016; Levine et al 2017; Grilli et al 2017; Benson et al 2018; Sizemore et al 2018). Several studies focused on higher-order structures using the language of pairwise networks. The first treatment was based on bipartite networks and their projections. Bipartite networks consist of two disjoint sets of nodes and a set of edges between the nodes of different sets, which are translated into

projections by replacing a node set with cliques (Fig. 1). One seminal study corresponds to the generating function approach for bipartite networks (Newman et al 2001). Specifically, a random bipartite network characterized by two-degree distributions of two node sets was utilized as the null model of a network with a group structure. The generating function method analytically provides clustering and assortativity coefficients for projections, and conjectures high levels of clustering and assortative mixing in social networks (Newman et al 2001, 2002; Newman and Park 2003). Detailed structural properties have been investigated for projections of random bipartite networks and empirical ones (Fisher et al 2017; Vasques Filho and O’Neale 2018, 2020).

Another important property is the emergence of a giant (bipartite) component, that is, the percolation transition. A natural extension of the Molly–Reed criterion for a configuration model provides the transition condition (Newman et al 2001). With respect to a random (monopartite) network generated from Erdős–Rényi and configuration models, various structural properties of the giant component (GC) and finite components have been clarified. Hereafter, we denote a subgraph composed of all finite components, which is the set difference between the entire network and GC, as the FC. A broad class of random networks satisfies a discrete duality relation: the FC in the supercritical phase possessing GC behaves like a random network in the subcritical phase possessing no GC (Bollobás et al 2007; Durrett 2007; Janson and Riordan 2011). In FC, the degree distribution is characterized by that of the entire network with an exponential cutoff. In addition, similar to the entire network, the degree correlation of the FC is absent (Tishby et al 2018). The structural properties of the GC are different from those of the entire network and FC. Specifically, the degree distribution of GC coincides with the neighbor degree distribution of the entire network at the percolation threshold (Dorogovtsev et al 2008). In terms of degree correlation, GC is negatively correlated irrespective of the degree distribution of the entire network (Engel et al 2004; Bialas and Oleś 2008; Tishby et al 2018; Mizutaka and Hasegawa 2018), and the negative degree correlation extends within the percolation correlation length (Mizutaka and Hasegawa 2020). Several studies have examined the properties of connected components in networks with group structures in the context of clustered network models. The random clustered network model generates random networks in which independent distributions of single edges and triangles specify the degree of the nodes (Newman 2009; Miller 2009). The GC of a random clustered network can exhibit a positive or negative degree–degree correlation based on the clustering details (Hasegawa and Mizutaka 2020). The generalized configuration model is a generalization of a random clustered network model, in which the generated networks are specified as arbitrary distributions of subgraphs (Karrer and Newman 2010). Mann et al. used the

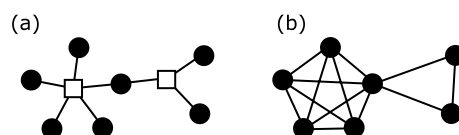


Fig. 1 A schematic view of a bipartite network consisting of seven individual nodes (solid spheres) and two group nodes (open squares) (a), and its projection (b). In the projection, two individual nodes sharing at least one adjacent group node in the bipartite networks are connected by an edge to each other

generating function method for the generalized configuration model when examining the structure of GC in a random network with arbitrary clique clustering (Mann et al 2022). They observed that GC possesses negative degree correlations for a single-size clique network. They also investigated networks comprising single edges and m -cliques and noted that the average degree of adjacent nodes increased with the node degree with an oscillation, implying a positive degree correlation. Component statistics in the random (monopartite) networks were collected to a certain extent. However, there is a paucity of understanding regarding networks with group structures. In particular, what determines the degree correlation of components and how the structures in the subcritical and supercritical phases are related is not yet completely understood.

In this study, we consider networks with clique clustering, which are projections of random bipartite networks (Newman et al 2001). In the projection procedure, each group node with degree m is replaced with a clique structure of size m (Fig. 1). The generating function method is used to analyze the statistical properties of GC and FC of the projections of random bipartite networks, and to evaluate the assortativity (Newman 2002) and global clustering coefficient (Newman 2001; Barrat and Weigt 2000). Our findings reveal that when all groups in a bipartite network have the same size, the projection and FC are degree uncorrelated, whereas the GC consistently exhibits a negative degree correlation, irrespective of the individual node degree distribution. Conversely, when employing the Poisson distribution for both the degree distributions of individual and group nodes, the GCs are positively correlated, except when the average degree of the group nodes is sufficiently small. For two examples of projections of random bipartite networks, we find that a projection that removes the GC corresponds to one in the subcritical phase, that is, the discrete duality relation. We also investigate the relationship between the fluctuation of clique size and degree correlation in a network with tunable clique size fluctuation and show that a small amount of clique size fluctuation results in the degree correlations of GC being positive. Finally, we discuss the real-world implications of our findings.

The remainder of this paper is organized as follows: We formulate the structural properties of the projections and their subgraphs using the generating function methods shown in Section [Formulation](#). We consider two examples to determine some properties of the projections of bipartite networks in Section [Examples](#). Section [Summary & discussion](#) presents a summary and discussion.

Formulation

We treat the projections of random bipartite networks consisting of individual and group nodes. In random bipartite networks, the individual and group nodes are connected randomly. The random bipartite networks used in this study are locally tree-like and sparse. In the next subsection, we briefly review the structural properties of an entire network of random bipartite network projections (Newman et al 2001; Newman and Park 2003). In Sec. [Properties of finite components and the giant component](#), we develop generating functions for certain structural properties of the GC and FC in the projections.

Properties of the projections

We consider a random bipartite network of N individual and M group nodes. Let $p(k)$ and $\tilde{p}(m)$ be the degree distributions of individual and group nodes in the bipartite network, respectively. We define the two functions that generate $p(k)$ and $\tilde{p}(m)$ as $f_0(x) = \sum_k p(k)x^k$ and $g_0(x) = \sum_m \tilde{p}(m)x^m$, respectively. Let $q(k)$ ($\tilde{q}(m)$) be the excess degree distribution that an individual (group) node reached by following the edge outgoing from a group node (individual node) has k (m) edges apart from the one we followed. Using the derivative of $f_0(x)$ ($g_0(x)$), the generating function $f_1(x)$ ($g_1(x)$) of $q(k)$ ($\tilde{q}(m)$) are given as

$$f_1(x) = \frac{f'_0(x)}{f'_0(1)} = \sum_k \frac{(k+1)p(k+1)}{z_1} x^k, \tag{1}$$

$$g_1(x) = \frac{g'_0(x)}{g'_0(1)} = \sum_m \frac{(m+1)\tilde{p}(m+1)}{\tilde{z}_1} x^m. \tag{2}$$

Here, $q(k) = (k+1)p(k+1)/z_1$ ($\tilde{q}(m) = (m+1)\tilde{p}(m+1)/\tilde{z}_1$) and $f'_0(1)$ ($g'_0(1)$) is a normalizing constant. A bipartite network holds the relation $z_1N = \tilde{z}_1M$.

Next, we consider the projection of a random bipartite network. The number of cliques and the distribution of the size m cliques in the projection correspond to the number M of group nodes and the group-node degree distribution $\tilde{p}(m)$ in the bipartite network, respectively. The degree distribution $P(n)$ of the projection is equal to the distribution of the number of second-nearest neighbors of individual nodes in the bipartite network (see Fig. 1). As in the case of random networks (Newman et al 2001), a composition of the generating functions $f_0(x)$ and $g_1(x)$ provides a generating function for the distribution of such numbers. Thus, the generating function $G_0(x)$ of the degree distribution $P(n)$ is given as

$$G_0(x) = f_0(g_1(x)). \tag{3}$$

From the derivative of $G_0(x)$, the generating function for the excess degree distribution $Q(n)$ in the projection is

$$G_1(x) = \frac{G'_0(x)}{G'_0(1)} = g_2(x)f_1(g_1(x)), \tag{4}$$

where $G'_0(1)$ is a normalizing constant and $g_2(x) = g'_1(x)/g'_1(1)$.

The global clustering coefficient is defined as $C = 3N_{\Delta}/N_{\text{triplet}}$, where N_{triplet} and N_{Δ} denote the numbers of connected triplets and triangles, respectively. N_{triplet} and N_{Δ} are given by

$$N_{\text{triplet}} = \frac{1}{2}N \frac{\partial^2}{\partial x^2} G_0(x) \Big|_{x=1} = \frac{1}{2}N \frac{z_1}{\tilde{z}_1} \left(\tilde{z}_3 + \frac{z_2 \tilde{z}_2^2}{z_1 \tilde{z}_1} \right) \tag{5}$$

and

$$N_{\Delta} = \frac{1}{6}M \frac{\partial^3}{\partial x^3} g_0(x) \Big|_{x=1} = \frac{1}{6}M\tilde{z}_3, \tag{6}$$

respectively, where $z_n = \sum_k k(k-1)\dots(k-n+1)p(k)$ and $\tilde{z}_n = \sum_m m(m-1)\dots(m-n+1)\tilde{p}(m)$. We use the relation $Nz_1 = M\tilde{z}_1$ to obtain the global clustering coefficient C of the projection as (Newman et al 2001)

$$C = \frac{z_1\tilde{z}_1\tilde{z}_3}{z_1\tilde{z}_1\tilde{z}_3 + z_2\tilde{z}_2^2}. \tag{7}$$

Similar to $G_0(x)$, we obtain the generating function $E(x, y)$ of the joint probability $P(n, n')$, such that the two ends of a randomly chosen edge in the projection have excess degrees n and n' , as follows (Newman and Park 2003):

$$E(x, y) = g_2(xy)f_1(g_1(x))f_1(g_1(y)). \tag{8}$$

Notably, Eq. 4 is obtained from Eq. 8 with $y = 1$. The inequality $E(x, y) \neq G_1(x)G_1(y)$ indicates that projections can possess degree correlations even though individual and group nodes are generally randomly connected in bipartite networks. Using $E(x, y)$ and $G_1(x)$, we can calculate the assortativity r as

$$r = \frac{\partial_x \partial_y E(x, y) - (\partial_x G_1(x))^2}{(x \partial_x)^2 G_1(x) - (\partial_x G_1(x))^2} \Big|_{x=1, y=1}. \tag{9}$$

Inserting Eqs. 4 and 8 into Eq. 9, we obtain (Newman and Park 2003)

$$r = \frac{1 - \frac{\tilde{z}_3}{z_2} + \frac{\tilde{z}_4}{z_3}}{1 - \frac{\tilde{z}_3}{z_2} + \frac{\tilde{z}_4}{z_3} + \frac{z_2\tilde{z}_2}{z_1\tilde{z}_1} \left(1 + \frac{\tilde{z}_2}{z_3} + \frac{z_2^2}{z_1z_3} \left(\frac{z_3}{z_2} - \frac{z_2}{z_1} \right) \right)}. \tag{10}$$

The assortativity is a non-negative value $r \geq 0$ (see Appendix A). We assume that all moments appearing on the right-hand side of Eq. 10 are finite. That is, all derivatives at $x = 1$ and $y = 1$ are finite on the right-hand side of Eq. 9. When we employ a power law type for the individual and group node degree distributions, the high-order moments diverge in the thermodynamic limit. Note that careful treatment is required to determine whether the assortativity converges to a value or does not make sense, as in (Litvak and van der Hofstad 2013).

Properties of finite components and the giant component

We introduce the probability u (\tilde{u}) of reaching a finite component through an edge outgoing from a group node (an individual node). To obtain the probability u , let us consider a situation in which an edge outgoing from a group node leads to an individual node with k edges other than the one we followed (excess degree k). By definition of the probability \tilde{u} , the probability that all k edges lead to finite components is \tilde{u}^k . Because we find an individual node with an excess degree k by following an edge outgoing from a group node with probability $q(k) = (k+1)p(k+1)/z_1$, the probability u is the sum

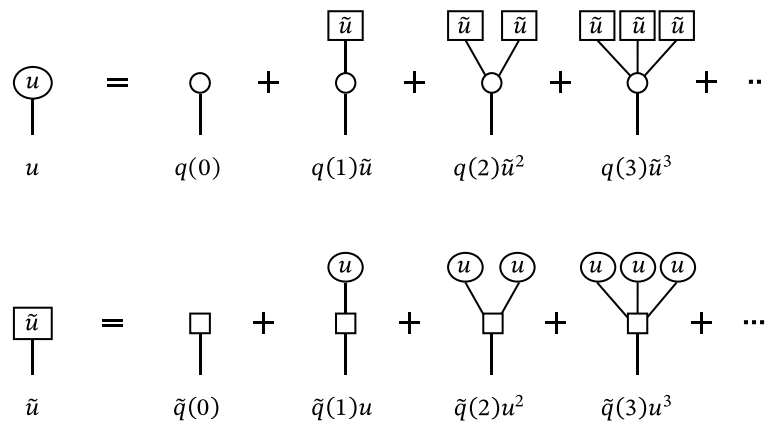


Fig. 2 Schematic representation of the sum rule in Eqs. 11 and 12. Open spheres (squares) denote individual (group) nodes. The left-hand side represents that an edge outgoing from a group (individual) node leads to a finite component with probability u (\tilde{u}). The right-hand side represents the summation for probabilities that an individual (group) node reached by following an edge outgoing from a group (individual) node has excess degree k (m) and the k (m) edges lead to finite components

of $q(k)\tilde{u}^k$ over k (see Fig. 2). Therefore, using Eqs. 1 and 2, we have self-consistent equations for u and \tilde{u}

$$u = f_1(\tilde{u}), \tag{11}$$

and

$$\tilde{u} = g_1(u), \tag{12}$$

respectively. Note that the equation for \tilde{u} is obtained in a manner similar to that of u . Expanding the above equations near $\epsilon = 1 - u$ and $\tilde{\epsilon} = 1 - \tilde{u}$, where ϵ ($\tilde{\epsilon}$) $\ll 1$, we obtain the percolation threshold above which GC exists (Newman et al 2001)

$$\frac{z_2 \tilde{z}_2}{z_1 \tilde{z}_1} = 1. \tag{13}$$

From the probability $P(n, FC) = P(n)u^n$ that a node chosen randomly from the projection has degree n and resides in a finite component, we obtain the generating function $G_0(x, FC)$, which generates $P(n, FC)$.

$$G_0(x, FC) = f_0(g_1(ux)). \tag{14}$$

This leads to the generating function $G_0(x|FC)$ of the degree distribution of the FC, as follows:

$$G_0(x|FC) = \frac{f_0(g_1(ux))}{1 - S}, \tag{15}$$

where

$$S = 1 - f_0(g_1(u)) \tag{16}$$

is the fraction of the GC size. The relation $G_0(x) = G_0(x, FC) + G_0(x, GC)$ gives the related generating functions for the GC, as follows:

$$G_0(x, GC) = f_0(g_1(x)) - f_0(g_1(ux)) \tag{17}$$

and

$$G_0(x|GC) = \frac{f_0(g_1(x)) - f_0(g_1(ux))}{S}. \tag{18}$$

Here, $G_0(x, GC)$ is the generating function of the probability that a randomly chosen node has degree n and resides in GC. $G_0(x|GC)$ is the generating function for the degree distribution of GC. Similar to the derivation of Eq. 4, we obtain the generating functions for the excess degree distributions of FC and GC from Eqs. 15 and 18, as follows:

$$G_1(x|FC) = \frac{g_2(ux)f_1(g_1(ux))}{g_2(u)u} \tag{19}$$

and

$$G_1(x|GC) = \frac{g_2(x)f_1(g_1(x)) - ug_2(ux)f_1(g_1(ux))}{1 - g_2(u)u^2}, \tag{20}$$

respectively.

From the probability $\tilde{p}(m, FC) = \tilde{p}(m)u^m$ that a randomly chosen group node has degree m and resides in a finite component, we obtain its generating function $g_0(x, FC)$, as follows:

$$g_0(x, FC) = g_0(ux). \tag{21}$$

We replace $G_0(x)$ and $g_0(x)$ in Eqs. 5 and 6 with $G_0(x, FC)$ and $g_0(x, FC)$ and obtain the number of triplets and triangles in the FC as

$$N_{\text{triplet}}^{\text{FC}} = \frac{1}{2}Nz_1u^2 \left(ug_1''(u) + (g_1'(u))^2 f_1(g_1'(u)) \right) \tag{22}$$

and

$$N_{\Delta}^{\text{FC}} = \frac{1}{6}M\tilde{z}_1u^3 g_1''(u), \tag{23}$$

respectively. GC is the set difference between the entire network and the FC. Thus, the numbers for the GC are given by

$$N_{\text{triplet}}^{\text{GC}} = N_{\text{triplet}} - N_{\text{triplet}}^{\text{FC}} \tag{24}$$

and

$$3N_{\Delta}^{\text{GC}} = 3N_{\Delta} - 3N_{\Delta}^{\text{FC}}. \tag{25}$$

The clustering coefficients of FC and GC are

$$C^{FC} = \frac{ug_1''(u)}{ug_1''(u) + (g_1'(u))^2 f_1'(g_1(u))}, \tag{26}$$

and

$$C^{GC} = \frac{\frac{\tilde{z}_3}{\tilde{z}_1} - u^3 g_1''(u)}{\frac{\tilde{z}_3}{\tilde{z}_1} + \frac{z_2 \tilde{z}_2^2}{z_1 \tilde{z}_1^2} - u^2 (ug_1''(u) + (g_1'(u))^2 f_1'(g_1(u)))}, \tag{27}$$

respectively.

Next, we obtain the generating function $E(x, y, FC)$ ($E(x, y, GC)$) for the joint probability $P(n, n', FC)$ ($P(n, n', GC)$) that two ends of a randomly chosen edge have excess degrees n and n' and that the edge resides in a finite component (the GC). $E(x, y, FC)$ and $E(x, y, GC)$ are expressed as follows:

$$E(x, y, FC) = g_2(uxy) f_1(g_1(ux)) f_1(g_1(uy)) \tag{28}$$

and

$$E(x, y, GC) = g_2(xy) f_1(g_1(x)) f_1(g_1(y)) - g_2(uxy) f_1(g_1(ux)) f_1(g_1(uy)), \tag{29}$$

respectively. We have generating functions $E(x, y|FC)$ and $E(x, y|GC)$ for the joint probabilities of FC and GC as follows:

$$E(x, y|FC) = \frac{g_2(uxy) f_1(g_1(ux)) f_1(g_1(uy))}{g_2(u) u^2} \tag{30}$$

and

$$E(x, y|GC) = \frac{g_2(xy) f_1(g_1(x)) f_1(g_1(y)) - g_2(uxy) f_1(g_1(ux)) f_1(g_1(uy))}{1 - g_2(u) u^2}, \tag{31}$$

respectively. We insert Eqs. 19, 20, 30, and 31 into Eq. 9 to evaluate the assortativity of FC and GC. Equation 8 is not equivalent to Eqs. 30 and 31. This implies that the degree correlation of the projections differs from their GC and FC.

Examples

Projections of z-uniform bipartite networks

We consider a random bipartite network in which the degree distribution of the group nodes is $\tilde{p}(m) = \delta_{mz}$. Here, $\delta_{mz} = 1$ for $m = z$ and $\delta_{mz} = 0$ otherwise. Random bipartite networks with $\tilde{p}(m) = \delta_{mz}$ are referred to as z -uniform bipartite networks. The projections of z -uniform bipartite networks consist of cliques of single size z . Notably, they are reduced to ordinary random networks with the degree distribution $p(k)$ when 2 uniform bipartite networks are chosen. The generating functions for the joint degree distributions of the projections of z -uniform bipartite networks, their FC, and their GC are given as:

$$E(x, y) = (xy)^{z-2} f_1(x^{z-1}) f_1(y^{z-1}), \tag{32}$$

$$E(x, y|FC) = \frac{(xy)^{z-2} f_1((ux)^{z-1}) f_1((uy)^{z-1})}{u^2}, \tag{33}$$

and

$$E(x, y|GC) = \frac{(xy)^{z-2} f_1(x^{z-1}) f_1(y^{z-1}) - (uxy)^{z-2} f_1((ux)^{z-1}) f_1((uy)^{z-1})}{1 - u^z} \tag{34}$$

respectively. We insert $y = 1$ into the above expressions to obtain the generating functions for the excess degree distributions as follows:

$$G_1(x) = x^{z-2} f_1(x^{z-1}), \tag{35}$$

$$G_1(x|FC) = \frac{x^{z-2} f_1((ux)^{z-1})}{u}, \tag{36}$$

$$G_1(x|GC) = \frac{x^{z-2} f_1(x^{z-1}) - u(ux)^{z-2} f_1((ux)^{z-1})}{1 - u^z}. \tag{37}$$

The equations for $z = 2$ coincide with the known results for random networks (Bialas and Oleś 2008; Tishby et al 2018; Mizutaka and Hasegawa 2018). We find $E(x, y) = G_1(x)G_1(y)$ and $E(x, y|FC) = G_1(x|FC)G_1(y|FC)$ from Eqs. 32, 33, 35, and 36, without depending on z , which implies that the projections of z -uniform bipartite networks and their FC are degree-uncorrelated. By contrast, GC is degree-correlated ($E(x, y|GC) \neq G_1(x|GC)G_1(y|GC)$). The numerator of Eq. 9, using Eqs. 34 and 37 is as follows:

$$-\frac{u^z}{(1 - u^z)^2} (\partial_x G_1(x) - \partial_x G_1(x|FC))^2 \Big|_{x=1} \leq 0. \tag{38}$$

The equality in the inequality above holds if, and only if $u = 0$. The denominator of Eq. 9 has a non-negative value; thus, the assortativity of the GC is always negative, irrespective of z and $p(k)$. This implies that when we consider a network with cliques of a single size, the GC is negatively correlated with an arbitrary individual node degree distribution $p(k)$.

We assume that the individual node degree distribution is Poisson, $p(k) = e^{-\lambda} \lambda^k / k!$ with an average λ . In this case, the assortativity, global clustering coefficient, and percolation threshold for the projections are given by

$$r = 0, \tag{39}$$

$$C = \frac{1}{1 + \lambda^{\frac{z-1}{z-2}}}, \tag{40}$$

and

$$\lambda(z - 1) = 1, \tag{41}$$

respectively. From the probabilities (11) and (12), we obtain a self-consistent equation:

$$u = \exp\left(\lambda\left(u^{z-1} - 1\right)\right). \tag{42}$$

We evaluate the root u in Eq. 42 and calculate Eq. 9, using Eqs. 33 and 36 to obtain the assortativity of the FC. Similarly, the assortativity of GC is derived from Eqs. 34 and 37. In addition, we obtain the global clustering coefficients of FC and GC using Eqs. 26 and 27, respectively. If we focus only on the networks at the percolation threshold, the assortativity and global clustering coefficient of GC can be analyzed. To calculate the assortativity at the percolation threshold (41), we determine the generating function (31) for the joint probability of the GC near/at the percolation threshold ($u = 1 - \epsilon$ and $\epsilon \rightarrow 0$). To this end, we expand components of Eq. 31 as Eqs. B5–B7, in the appendix B. From Eqs. B5–B7, we obtain the generating functions $E(x, y|GC)$ and $G_1(x|GC)$ near the percolation threshold as Eqs. B8 and B9, respectively. For projections of z -uniform bipartite networks with $p(k) = e^{-\lambda}\lambda^k/k!$, we obtain

$$r_c^{GC} = -\frac{z-1}{1+2z(z-1)}. \tag{43}$$

Similarly, inserting $u = 1 - \epsilon$ to Eq. 27 and taking $\epsilon \rightarrow 0$, we obtain

$$C_c^{GC} = \frac{z(z-2)}{z^2-1}. \tag{44}$$

Fig. 3a and b show the assortativity and global clustering coefficient of GC as a function of λ for $z = 2, 3, 4$, and 5. The lines correspond to the results of Eq. 9, which are evaluated using generating function methods. The vertical thin lines represent the positions of the percolation thresholds, and the symbols (stars) represent r_c^{GC} obtained from Eq. 43. The solid symbols denote the numerical results of the Monte Carlo simulations. To generate a bipartite network with $p(k) = e^{-\lambda}\lambda^k/k!$ and $\tilde{p}(m) = \delta_{mz}$, we predetermine the number

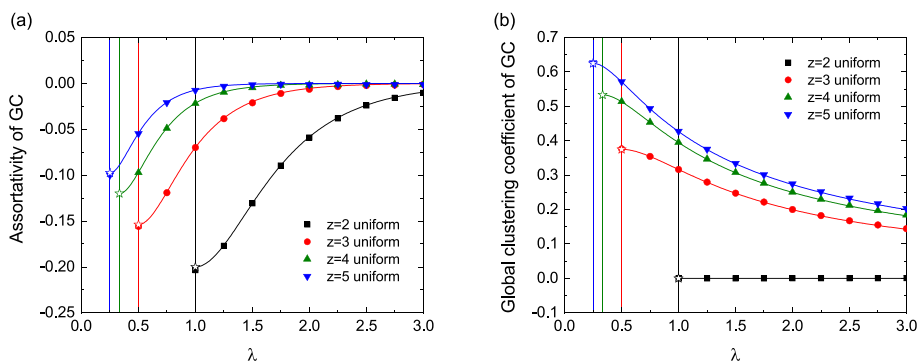


Fig. 3 Structural properties of the GCs decomposed from the projections of z -uniform bipartite networks. The vertical axis indicates **a** assortativity and **b** global clustering coefficients. The horizontal axis denotes the average of individual-node degrees distributed in a Poisson $p(k) = e^{-\lambda}\lambda^k/k!$. Each vertical line is the percolation threshold above which the GC exists and a star on the line represents the corresponding **a** r_c^{GC} and **b** C_c^{GC} given by Eqs. 43 and 44, respectively. The solid symbols denote numerical results from Monte Carlo simulations averaged over 1,000 configurations. We set the number of edges between two types of nodes as $W = 12,000$. The error bars for standard errors on the numerical results are smaller than the data points

W of edges between two types of nodes. Next we prepare $N (= \lfloor W/\lambda \rfloor)$ isolated individual nodes and $M (= \lfloor W/z \rfloor)$ isolated group nodes with z stubs. We connect each stub of group nodes with a randomly chosen individual node and form an edge until all stubs are consumed. Because all individual nodes are selected uniformly with probability $1/N$, the probability that an individual node has k edges is $p(k) = \binom{W}{k} (1/N)^k (1 - 1/N)^{W-k}$. The distribution $p(k) = \binom{W}{k} (1/N)^k (1 - 1/N)^{W-k}$ coincides with the Poisson distribution $p(k) = e^{-\lambda} \lambda^k / k!$ in the thermodynamic limit ($W \rightarrow \infty$) under a fixed $\lambda = W/N$. We have validated the correspondence of two distributions in our simulations with $W = 12,000$ (not shown). As shown in Fig. 3, the analytical treatments agree with the corresponding simulations. Figure 3a shows that the GC is negatively correlated, as expected from Eq. 38. Notably, the result of Erdős–Rényi random networks (Tishby et al 2018; Mizutaka and Hasegawa 2018) was recovered for $z = 2$. In Fig. 3a, the assortativity $r^{\text{GC}} (\leq 0)$ increases with z . In general, the correlation coefficient is invariant if two variables are multiplied by a constant. A k -degree individual node in a bipartite network becomes a node with degree $(z - 1)k$ in the projection. Therefore, it is essentially identical for degree correlations in the projections of z -uniform bipartite networks ($z > 2$) and for ordinary random graphs without clustering ($z = 2$). However, the assortativity for different values of z is discrepant because the degrees of the nodes involved in GC depend on k and z . Conversely, the clustering coefficient C^{GC} increases with decreases in λ , as shown in Fig. 3b. The average λ required for forming the GC decreases as the clique size increases, thus reducing the number of triplets $N_{\text{triplet}}^{\text{GC}}$ in the GC.

Projections of double Poisson bipartite networks

Another simple case is the projection from a random bipartite network with $p(k) = e^{-\lambda} \lambda^k / k!$ and $\tilde{p}(m) = e^{-\tilde{\lambda}} \tilde{\lambda}^m / m!$. We refer to such bipartite networks as double Poisson bipartite networks. The assortativity and clustering coefficients of the projections of the double Poisson bipartite networks are given by

$$r = \frac{1}{1 + \lambda \tilde{\lambda} + \lambda} \tag{45}$$

and

$$C = \frac{1}{\lambda + 1}, \tag{46}$$

respectively (Newman and Park 2003). From Eq. 13, the percolation threshold is given by

$$\lambda \tilde{\lambda} = 1. \tag{47}$$

Equations 11 and 12 can be rewritten as

$$u = \exp(-\lambda(1 - \tilde{u})) \tag{48}$$

and

$$\tilde{u} = \exp(-\tilde{\lambda}(1 - u)), \tag{49}$$

respectively.

Using Eqs. 48 and 49, we determine the following relations between the generating functions:

$$G_0(x|FC) = G_0(x; \lambda^*, \tilde{\lambda}^*), \tag{50}$$

$$G_1(x|FC) = G_1(x; \lambda^*, \tilde{\lambda}^*), \tag{51}$$

and

$$E(x, y|FC) = E(x, y; \lambda^*, \tilde{\lambda}^*), \tag{52}$$

where $\lambda^* = \tilde{u}\lambda$ and $\tilde{\lambda}^* = u\tilde{\lambda}$. The relations indicate that the distributions $P(n|FC)$, $Q(n|FC)$, and $P(n, n'|FC)$ for projections with λ and $\tilde{\lambda}$ are identical to the distributions $P(n)$, $Q(n)$, and $P(n, n')$ for them with λ^* and $\tilde{\lambda}^*$, respectively. Hence, the FC in the supercritical phase ($\lambda\tilde{\lambda} > 1$) can be mapped onto the entire network using the different parameter sets λ^* and $\tilde{\lambda}^*$. The mapped network is in the subcritical phase ($\lambda^*\tilde{\lambda}^* < 1$), because it does not possess a GC. This is referred to as a discrete duality relation in studies on random graphs (Bollobás et al 2007; Durrett 2007). We can also observe the same relation in the projections of z -uniform bipartite networks with a Poisson individual degree distribution, for example, $E(x, y|FC) = E(x, y; \tilde{u}\lambda, z)$. When the projections of random bipartite networks satisfy the discrete duality relation, the FC of the projections is positively correlated because Eq. 9 is positive.

We determine the roots u and \tilde{u} in Eqs. 48 and 49 for the given parameters λ and $\tilde{\lambda}$ and Eq. 9 for the FC and GC to obtain the assortativity of the FC and GC. Similar to the derivation of Eq. 43, we obtain the assortativity r_c^{GC} of GC at the percolation threshold ($\lambda\tilde{\lambda} = 1$) as

$$r_c^{GC} = \frac{\tilde{\lambda}^3 + \tilde{\lambda}^2 + 2\tilde{\lambda} - 1}{3\tilde{\lambda}^3 + 11\tilde{\lambda}^2 + 17\tilde{\lambda} + 5}, \tag{53}$$

where the numerator denotes an increasing function of $\tilde{\lambda}$. Specifically, r_c^{GC} becomes negative when $\tilde{\lambda} \lesssim 0.393$. We also calculate the global clustering coefficients for FC and GC using Eqs. 26 and 27, respectively. At the percolation threshold, the global clustering coefficient of the GC is

$$C_c^{GC}(\lambda, \tilde{\lambda}) = \frac{\tilde{\lambda}^2 + 3\tilde{\lambda}}{\tilde{\lambda}^2 + 5\tilde{\lambda} + 3}. \tag{54}$$

Figure 4 shows the λ dependence of the assortativity (a) and the global clustering coefficient (b) for the entire network (dashed lines) and GC (solid lines) of the projections of double Poisson bipartite networks. The lines represent the results obtained from the generating functions. The symbols represent the results of Monte Carlo simulations. In our Monte Carlo simulations, to generate double Poisson bipartite

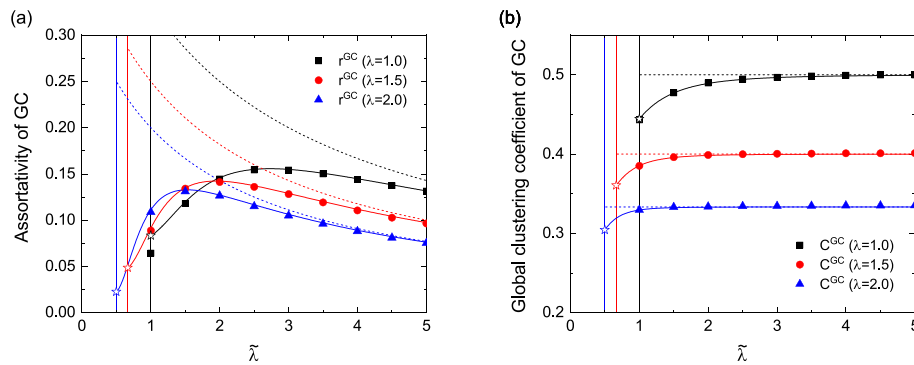


Fig. 4 Structural properties of projections of double Poisson bipartite networks. The lines represent **a** assortativity and **b** the global clustering coefficient calculated by the generating function method for the whole network (dashed lines) and GC (solid lines). Each vertical line is the percolation threshold and a star on the line represents the corresponding r_c^{GC} and C_c^{GC} given by Eqs. 53 and 54, respectively. The solid symbols represent the numerical results of Monte Carlo simulations averaged over 1,000 configurations of the GC, which is decomposed from the whole network with $W = 12,000$. The error bars for standard errors on the numerical results are smaller than the data points

networks with λ and $\tilde{\lambda}$, first, we predetermine the number of edges W between the two types of nodes. Second, we prepare $N (= \lfloor W/\lambda \rfloor)$ isolated individual nodes and $M (= \lfloor W/\tilde{\lambda} \rfloor)$ isolated group nodes. Next, we form an edge between a randomly chosen individual node and a randomly chosen group node repeatedly until the number of edges reaches the predetermined W . All individual and group nodes are selected uniformly with probabilities $1/N$ and $1/M$ per a process forming an edge, respectively. Thus, the individual and group degree node distributions $p(k)$ and $\tilde{p}(m)$ are given as $p(k) = \binom{W}{k} (1/N)^k (1 - 1/N)^{W-k}$ and $\tilde{p}(m) = \binom{W}{m} (1/M)^m (1 - 1/M)^{W-m}$, respectively. In our simulations, we have confirmed that two binomial distributions, $p(k)$ and $\tilde{p}(m)$, approximate Poisson distributions with $\lambda = W/N$ and $\tilde{\lambda} = W/M$, respectively (not shown). As shown in Fig. 4, the analytical treatments agree with the corresponding simulations. The assortativity r^{GC} and the clustering coefficient C^{GC} decrease with $\tilde{\lambda}$ near the percolation threshold. As $\tilde{\lambda}$ decreases, the size and number of cliques in the network decrease, implying that the network structure of the projections approaches a local tree-like structure. Therefore, GC tends to lose its positive correlation near the percolation threshold.

Next, we plot the color maps of assortativity analytically computed for the entire network, FC, and GC in $(\lambda, \tilde{\lambda})$ plane in Fig. 5. The dashed line in panel (a) represents the percolation threshold, given by Eq. 47. The black region in panels (b) and (c) is subcritical, in which GC does not exist. We confirm from panel (a) that the assortativity of the entire network is a decreasing function of λ and $\tilde{\lambda}$, as expected from Eq. 45. A strong positive correlation induced by isolated cliques is observed in the subcritical phase (region below the dashed line). Similarly, a strong positive correlation is also observed in panel (b), which is consistent with the discrete duality relation (Eqs. 50–52). Indeed, the value of assortativity at a point in panel (b) is the same as at a point determined by the discrete duality relation in panel (a). For instance, the assortativity value of 0.821 is observed at the point $(\lambda, \tilde{\lambda}) = (2.00, 3.10)$ in panel (b) and at the point $(0.147, 0.485)$ in panel (a). A

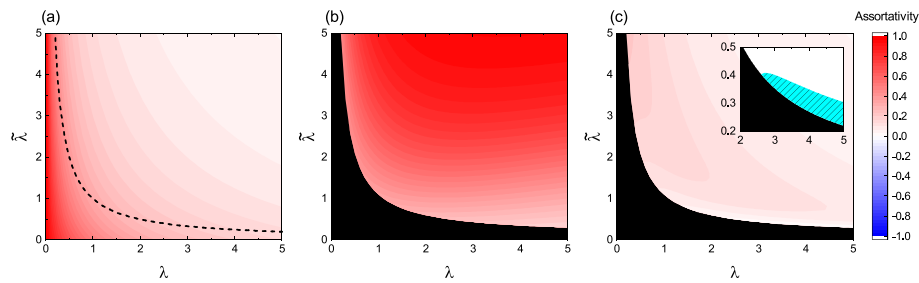


Fig. 5 Assortativities of the whole network **(a)**, FC **(b)**, and GC **(c)** of projection of double Poisson bipartite networks. The black dashed line in **(a)** represents the percolation threshold above which there exists the GCs. The area filled with black in **(b)** and **(c)** denotes subcritical. The hatched area in the inset of **(c)** is the region in which the assortativity of the GC is negative

negative correlation can be only seen in panel (b). The inset in (c) shows that r^{GC} near the percolation threshold can be negative with a small $\tilde{\lambda}$ (for example, $r^{GC} \approx -0.003$ at the point (5.00, 0.30), and see also Eq. 53).

Model connecting the two examples

The two examples in the previous sections exhibit different characteristics from the perspective of assortativity. Because we consider the same $p(k)$ (Poisson distribution) for both examples in Figs. 3 and 5, the difference in degree correlations arises from $\tilde{p}(m)$, which corresponds to the clique size distribution in the projections. To confirm the effect of the fluctuation in clique sizes, we consider the following bipartite network model, which connects z -uniform bipartite networks and double Poisson bipartite ones continuously. First, we prepare a bipartite network with W edges, whose degree distributions obey $p(k) = e^{-\lambda} \lambda^k / k!$ and $\tilde{p}(m) = \delta_{mz}$. Next, a randomly selected edge is removed from the group-node side and connected to a randomly selected group node. This randomization is repeated W' times. The generated bipartite networks are identical to z -uniform bipartite networks (double Poisson bipartite networks) if the fraction $p = W'/W$ of the replaced edges is zero (one). The variance of $\tilde{p}(m)$ is given as $\sigma_m^2 = pz(2 - p)$ when the number of group nodes is sufficiently large (see Appendix C). Thus, the fraction p tunes the clique size fluctuation. Figure 6 shows the assortativity for $z = 5$ as a function of the average λ of $p(k)$ for projections with different values of p . The solid and dashed lines represent the results for the GC and entire network, respectively. For $p \leq 0.05$, we observe that the assortativity of GC decreases as λ decreases and becomes negative, whereas for $p \geq 0.1$, the assortativity of GC never becomes negative even though λ approaches the percolation threshold. The inset of Fig. 6 presents the results for $z = 2$. For $p = 0$, the projections coincide with ordinary random graphs that must exhibit the strongest negative degree correlation in the present setting. Even in this case, GC does not show a negative correlation for $p \gtrsim 0.1$. Thus, a small fluctuation (small p) makes the degree correlation of GC positive, which implies that the projections of most bipartite networks display a positive degree correlation, including their GCs.

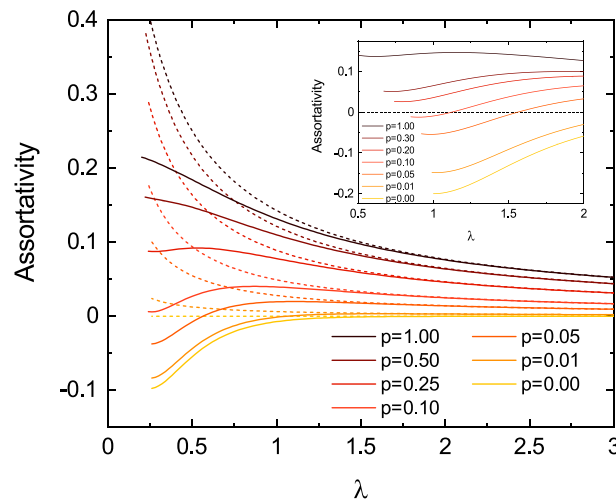


Fig. 6 Comparison of assortativity of the GC (solid lines) and whole network (dashed lines) for $z = 5$. The inset is for the GC of the projection for $z = 2$

Table 1 Assortativity of projections in the thermodynamic limit

Bipartite network	$p(k)$	$\tilde{p}(m)$	Whole	FC	GC
z -uniform	Arbitrary	δ_{mz}	0	0	≤ 0
Double Poisson	Poisson	Poisson	> 0	> 0	≥ 0

Summary & discussion

In this study, we have examined the statistical properties of the projections of random bipartite networks characterized by the individual node degree distribution $p(k)$ and the group node degree distribution $\tilde{p}(m)$. The projections can be decomposed into two subgraphs: a giant component (GC) and a subgraph composed of all finite components (FC). We have presented a method of using generating functions for evaluating the structural properties of both subgraphs and derived a general expression for the assortativity and global clustering coefficient of GC and FC. We have treated two examples of z -uniform bipartite networks, and double Poisson bipartite networks. We have validated the derived equations using numerical computations and examined the structural properties of the subgraphs.

Our results demonstrate that assortativity varies based on the subgraph and degree distributions, whereas nonzero clustering coefficients appear in both GC and FC of the two examples, with the exception of the 2-uniform bipartite network. The assortativity values of the networks are summarized in Table 1. For a z -uniform bipartite network, we have analytically demonstrated that projection and FC have no degree correlation. Simultaneously, the GC is negatively correlated irrespective of the value of z and the distribution $p(k)$ unless the entire network corresponds to the GC (Eq. 38) for a general case and Fig. 3a for a concrete example). For the double Poisson bipartite network, the

projection and FC display strong positive correlations (Fig. 5a and b). The GC shows a weak positive correlation over a wide range of average degrees λ and $\tilde{\lambda}$, and a weak negative correlation in a narrow region (Fig. 5c). In the region in which a negative correlation is observed, $\tilde{\lambda}$ is small and most of the cliques involved in the GC are small, including a simple edge, resulting in a tree-like structure. In Section [Model connecting the two examples](#), we have investigated the relation between the sign of assortativity in the GC and the fluctuation in the degrees of the group nodes. We have found that a small fluctuation results in the degree correlation of GC being positive.

In this study, we have found the discrete duality relation (Eqs. 50–52) claiming that the FC in the supercritical phase can be mapped onto the projection in the subcritical phase, that is they have the same structure. Thus, given that the assortativity r for the entire network is nonnegative (Eq. 10), the assortativity r^{FC} is also nonnegative. Whether this holds true for the general case is left for future studies. Clarifying this is essential for understanding bipartite networks and their projections.

Finally, we discuss the implications of our findings. The projections of random bipartite networks can explain the high assortativity of empirical social networks (Newman and Park 2003), whereas assortativity in the projections of empirical bipartite networks can be positive or negative (Fisher et al 2017). In the present study, we have demonstrated that the GCs of projections often show positive degree correlations but can also be negative when the fluctuation of the degrees of group nodes is slight. It would be helpful to analyze the degree distribution of group nodes of empirical data that show negative degree correlations, considering that empirical data are components, including a kind of GC obtained by sampling from the subject of interest. A random bipartite network can model data if the group-node degrees are not widely distributed. If a negative degree correlation is observed, although the degrees of the group nodes are widely distributed, this suggests a strong correlation in the bipartite structure. In the future, it will be necessary to characterize nontrivial structures that the random bipartite network model cannot capture if such data are observed.

Appendix A Assortativity for projections

Equation 10 can be rewritten as:

$$r = \frac{\tilde{\beta}_2}{\tilde{\beta}_2 + \frac{z_2 \tilde{z}_2^2}{z_1 \tilde{z}_1^2} \left(\tilde{\beta}_1 + \frac{\tilde{z}_2^2}{z_1 z_2} \beta_1 \right)}, \tag{A1}$$

where $\beta_n = z_n z_{n+1} - z_{n+1}^2 + z_n z_{n+2}$, and similarly $\tilde{\beta}_n = \tilde{z}_n \tilde{z}_{n+1} - \tilde{z}_{n+1}^2 + \tilde{z}_n \tilde{z}_{n+2}$. The sum of the terms in β_n is

$$z_{n_1} z_{n_2} = \sum_k k(k-1) \cdots (k-n_1+1) p(k) \sum_k k(k-1) \cdots (k-n_2+1) p(k) \tag{A2}$$

$$= \sum_{k_1, k_2} \frac{p(k_1) p(k_2) k_1! k_2!}{(k_1 - n_1)! (k_2 - n_2)!}, \tag{A3}$$

and thus,

$$\begin{aligned}
 \beta_n &= \sum_{k_1, k_2} \frac{p(k_1)p(k_2)k_1!k_2!}{(k_1 - n)!(k_2 - n)!} [(k_2 - n) - (k_1 - n)(k_2 - n) + (k_2 - n)(k_2 - n - 1)] \\
 &= \sum_{k_1, k_2} \frac{p(k_1)p(k_2)k_1!k_2!}{(k_1 - n)!(k_2 - n)!} (k_2 - k_1)(k_2 - n) \\
 &= \frac{1}{2} \sum_{k_1, k_2} \frac{p(k_1)p(k_2)k_1!k_2!}{(k_1 - n)!(k_2 - n)!} (k_2 - k_1)^2 \geq 0.
 \end{aligned}
 \tag{A4}$$

Here, we utilize $\sum_{k_1, k_2} f(k_1)f(k_2)(k_1 - k_2) = 0$ and $\sum_{k_1, k_2} f(k_1)f(k_2)k_1^2 = \sum_{k_1, k_2} f(k_1)f(k_2)k_2^2$ for an arbitrary function $f(x)$. We also verified that $\tilde{\beta}_n \geq 0$. Because all terms in the numerator and denominator in Eq. A1 are non-negative, the assortativity of the projection of a random bipartite network is always non-negative. If a clique-size fluctuation exists, the value of Eq. 10 is always positive, $r > 0$.

Appendix B Derivation of r_c^{GC}

To derive the assortativity r_c^{GC} of the GC at the percolation threshold, first we explore the generating function $E(x, y|GC)$ (Eq. 31) of the joint probability near the percolation threshold ($u = 1 - \epsilon$ and $\tilde{u} = 1 - \tilde{\epsilon}$). In the situation, the function $g_1(ux)$, which is a component of Eq. 31, is expanded by the order $O(\epsilon)$ as

$$\begin{aligned}
 g_1(ux) &= \frac{1}{\tilde{z}_1} \sum_m m\tilde{p}(m)x^{m-1}(1 - \epsilon)^{m-1} \\
 &\approx g_1(x) - \epsilon xg_1'(x).
 \end{aligned}
 \tag{B5}$$

Similarly, we have

$$\begin{aligned}
 g_2(uxy) &= \frac{1}{\tilde{z}_2} \sum_m m\tilde{p}(m)(xy)^{m-2}(1 - \epsilon)^{m-2} \\
 &\approx g_2(xy) - \epsilon xyg_2'(xy).
 \end{aligned}
 \tag{B6}$$

Using Eq. B5, we expand $f_1(g_1(ux))$ as

$$\begin{aligned}
 f_1(g_1(ux)) &\approx f_1(g_1(x) - \epsilon xg_1'(x)) \\
 &= \frac{1}{z_1} \sum_k kp(k)g_1(x)^{k-1} \left(1 - \epsilon x \frac{g_1'(x)}{g_1(x)}\right)^{k-1} \\
 &\approx f_1(g_1(x)) - \epsilon x f_1'(g_1(x))g_1'(x).
 \end{aligned}
 \tag{B7}$$

With noting that the leading order is $O(\epsilon)$ in the numerator and the denominator of Eq. 31 inserted above results (B5)–(B7), we obtain $E(x, y|GC)$ near the percolation threshold as

$$\begin{aligned}
 E(x, y|GC) &= (xyg_2'(xy)f_1(g_1(x))f_1(g_1(y)) + xg_2(xy)f_1'(g_1(x))g_1'(x)f_1(g_1(y)) \\
 &\quad + yg_2(xy)f_1(g_1(x))f_1'(g_1(y))g_1'(y)) / (\tilde{z}_3/\tilde{z}_2 + 2) + O(\epsilon).
 \end{aligned}
 \tag{B8}$$

Inserting $y = 1$ into Eq. B8, we have $G_1(x|GC)$ near the percolation threshold as

$$G_1(x|GC) = (xg'_2(x)f_1(g_1(x)) + xg_2(x)f'_1(g_1(x))g'_1(x) + z_2\tilde{z}_2g_2(x)f_1(g_1(x))/z_1\tilde{z}_1) / (\tilde{z}_3/\tilde{z}_2 + 2) + O(\epsilon). \tag{B9}$$

Taking $\epsilon \rightarrow 0$, we can obtain the generating functions $E(x, y|GC)$ and $G_1(x|GC)$ at the percolation threshold. Choosing the individual and group node degree distributions, which determine the functional forms of corresponding generating functions, we can obtain the assortativity r_c^{GC} of the GC at the percolation threshold from Eq. 9 for a concrete model such as Eqs. 43 and 53. Similarly, we can calculate the clustering coefficient C_c^{GC} of the GC at the percolation threshold.

Appendix C Variance of clique size fluctuation

We derived the variance of the group node degrees using the generating function method. The probability that m edges are connected to a group node after random trimming of W' edges is

$$P_{\text{trim}}(m) = \binom{z}{m} p^{z-m} (1-p)^m. \tag{C10}$$

The probability that m edges will be added to a group node is

$$P_{\text{add}}(m) = \binom{W'}{m} q^m (1-q)^{W'-m}, \tag{C11}$$

where $q = 1/M$. The generating functions of the probabilities are

$$G_{\text{trim}}(x) = (p + (1-p)x)^z, \tag{C12}$$

$$G_{\text{add}}(x) = (qx + 1 - q)^{W'}. \tag{C13}$$

The degree distribution of group nodes $\tilde{p}(m)$ after trimming and adding W' edges can be represented as a convolution of $P_{\text{trim}}(m)$ and $P_{\text{add}}(m)$:

$$\tilde{p}(m) = \sum_{m'} P_{\text{trim}}(m') P_{\text{add}}(m - m'). \tag{C14}$$

The generating function of a convolution is the product of the generating functions of the indices of the convolution, and thus, the generating function of $\tilde{p}(m)$ is:

$$g_0(x) = (p + (1-p)x)^z (qx + 1 - q)^{W'}. \tag{C15}$$

The variance of the group-node degrees is

$$\sigma_m^2 = pz(2 - p - q). \tag{C16}$$

This value is an increasing function of the edge randomization probability p .

Author contributions

All authors contributed to the conception and design of the study.

Funding

Y. F. was supported by JSPS KAKENHI Grant Number 21K21302 and 23K13010. S. M. was supported by JSPS KAKENHI Grant Number 21K13853.

Availability of data and materials

All data generated or analyzed during this study are included in this published article.

Competing interests

The authors declare that they have no Conflict of interest.

Received: 5 April 2024 Accepted: 19 August 2024

Published online: 20 September 2024

References

- Barrat A, Weigt M (2000) On the properties of small-world network models. *Eur Phys J B* 13(3):547–560
- Benson AR, Gleich DF, Leskovec J (2016) Higher-order organization of complex networks. *Science* 353(6295):163–166
- Benson AR, Abebe R, Schaub MT et al (2018) Simplicial closure and higher-order link prediction. *Proc Natl Acad Sci USA* 115(48):E11221–E11230
- Bialas P, Oleś AK (2008) Correlations in connected random graphs. *Phys Rev E* 77(3):036124
- Bollobás B, Janson S, Riordan O (2007) The phase transition in inhomogeneous random graphs. *Random Struct Algor* 31(1):3–122
- Dorogovtsev SN, Goltsev AV, Mendes JF (2008) Critical phenomena in complex networks. *Rev Mod Phys* 80(4):1275
- Durrett R (2007) *Random Graph Dynamics*. Cambridge University Press
- Engel A, Monasson R, Hartmann AK (2004) On large deviation properties of erdős-rényi random graphs. *J Stat Phys* 117:387–426
- Fisher DN, Silk MJ, Franks DW (2017) The perceived assortativity of social networks: methodological problems and solutions. In: Missaoui R, Abdesslem T, Latapy M (eds) *Trends in social network analysis. Lecture notes in social networks*. Springer, Cham, pp 1–19
- Grilli J, Barabás G, Michalska-Smith MJ et al (2017) Higher-order interactions stabilize dynamics in competitive network models. *Nature* 548(7666):210–213
- Hasegawa T, Mizutaka S (2020) Structure of percolating clusters in random clustered networks. *Phys Rev E* 101(6):062310
- Janson S, Riordan O (2011) Duality in inhomogeneous random graphs, and the cut metric. *Random Struct Algor* 39(3):399–411
- Karrer B, Newman ME (2010) Random graphs containing arbitrary distributions of subgraphs. *Phys Rev E* 82(6):066118
- Levine JM, Bascompte J, Adler PB et al (2017) Beyond pairwise mechanisms of species coexistence in complex communities. *Nature* 546(7656):56–64
- Litvak N, van der Hofstad R (2013) Uncovering disassortativity in large scale-free networks. *Phys Rev E* 87(2):022801
- Mann P, Smith VA, Mitchell JB et al (2022) Degree correlations in graphs with clique clustering. *Phys Rev E* 105(4):044314
- Miller JC (2009) Percolation and epidemics in random clustered networks. *Phys Rev E* 80(2):020901
- Milo R, Shen-Orr S, Itzkovitz S et al (2002) Network motifs: simple building blocks of complex networks. *Science* 298(5594):824–827
- Mizutaka S, Hasegawa T (2018) Disassortativity of percolating clusters in random networks. *Phys Rev E* 98(6):062314
- Mizutaka S, Hasegawa T (2020) Emergence of long-range correlations in random networks. *J Phys: Complexity* 1(3):035007
- Newman ME (2009) Random graphs with clustering. *Phys Rev Lett* 103(5):058701
- Newman ME, Park J (2003) Why social networks are different from other types of networks. *Phys Rev E* 68(3):036122
- Newman ME, Strogatz SH, Watts DJ (2001) Random graphs with arbitrary degree distributions and their applications. *Phys Rev E* 64(2):026118
- Newman ME, Watts DJ, Strogatz SH (2002) Random graph models of social networks. *Proc Natl Acad Sci USA* 99:2566–2572
- Newman MEJ (2001) Scientific collaboration networks: network construction and fundamental results. *Phys Rev E* 64(1):016131
- Newman MEJ (2002) Assortative mixing in networks. *Phys Rev Lett* 89(20):208701
- Petri G, Expert P, Turkheimer F et al (2014) Homological scaffolds of brain functional networks. *J R Soc Interface* 11(101):20140873
- Sizemore AE, Giusti C, Kahn A et al (2018) Cliques and cavities in the human connectome. *J Comput Neurosci* 44:115–145
- Tishby I, Biham O, Katzav E et al (2018) Revealing the microstructure of the giant component in random graph ensembles. *Phys Rev E* 97(4):042318
- Ugander J, Backstrom L, Marlow C et al (2012) Structural diversity in social contagion. *Proc Natl Acad Sci USA* 109(16):5962–5966
- Vasques Filho D, O’Neale DR (2018) Degree distributions of bipartite networks and their projections. *Phys Rev E* 98(2):022307
- Vasques Filho D, O’Neale DR (2020) Transitivity and degree assortativity explained: The bipartite structure of social networks. *Phys Rev E* 101(5):052305

Publisher’s Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.