# Disassortativity of percolating clusters in random networks

Shogo Mizutaka[*] and Takehisa Hasegawa[†]

*Department of Mathematics and Informatics, Ibaraki University, 2-1-1 Bunkyo, Mito 310-8512, Japan*

We provide arguments for the property of the degree-degree correlations of giant components formed by the percolation process on uncorrelated random networks. Using the generating functions, we derive a general expression for the assortativity of a giant component, $r$, which is defined as Pearson's correlation coefficient for degrees of directly connected nodes. For uncorrelated random networks in which the third moment for the degree distribution is finite, we prove the following two points: (1) Assortativity $r$ satisfies the relation $r \leqslant 0$ for $p \geqslant p_c$. (2) The average degree of nodes adjacent to degree $k$ nodes at the percolation threshold is proportional to $k^{-1}$ independently of the degree distribution function. These results claim that disassortativity emerges in giant components near the percolation threshold. The accuracy of the analytical treatment is confirmed by extensive Monte Carlo simulations.

## I. INTRODUCTION

All systems are considered as networks if they consist of elements, and the relation between the elements can be defined. Owing to the generality of the definition of networks, various systems such as ecosystems, metabolic interactions, the World Wide Web, and social relationships are regarded as networks. Thus far, network science has extracted common properties from real networks [1,2]. A representative one is the correlation between degrees of directly connected nodes [3,4]. If similar (dissimilar) degree nodes are more likely to connect to each other in a network, the network has positive (negative) degree-degree correlation. We often call a network with positive (negative) degree-degree correlation an assortative (disassortative) network. Newman discovered that social networks possess positive degree correlations whereas biological and technological networks are disassortative [3]. Following the seminal work of Newman, the degree correlations of complex networks have been studied extensively. One of the reasons for this is that the degree correlations affect the behavior of dynamics on networks. Much effort has been devoted to examining the relation between the degree-degree correlation and phenomenological models on networks such as failures, spreading of diseases or information, and synchronization, to gain a deep understanding of the character of real-world networks [5–9].

There are networks in which no direct path along edges exists between two nodes. Such networks consist of several connected components. It is noticed that the degree correlation of a component is different from that of the whole network if the network is not singly connected. Recent works have formalized the joint probability of degrees in the giant component (GC) whose size is proportional to that of the whole network by the generating function method and obtained the average degree $\bar{k}_{nn}(k)$ of nodes adjacent to degree $k$ nodes [10] and the assortativity $r$ defined by Pearson's correlation coefficient for nearest degrees [11]. As demonstrated for some random networks [10,11], the GC can have the negative degree-degree correlation (*disassortativity*) in spite that the whole network is degree-uncorrelated. In addition, Tishby *et al.* have shown that the correlation between degrees for the GC in the Erdős-Rényi random graph is always negative if the network is not singly connected and the average degree is greater than unity or, equivalently, the GC exists [11].

The above generating function method can be generalized to the case of the percolation problem *on* given substrate networks. In the percolation problem on networks, each node is occupied (not removed) with a given probability and is unoccupied (removed) otherwise. It is known that the system undergoes the emergence of a percolating cluster, i.e., a GC of occupied nodes, at a certain value of occupation probability called as the percolation threshold. It is, however, unknown what correlation percolating clusters on uncorrelated random networks exhibit, especially at and around the percolation threshold where the system exhibits critical behavior [12].

In this study, we analyze the degree correlation of the GC generated by the site percolation process on uncorrelated random networks with arbitrary degree distribution. It is already known that the site percolation process on uncorrelated networks does not induce any degree-degree correlation as long as we focus on the degree-degree correlation of the whole network consisting of occupied nodes [13]. We extract the GC from the whole network and examine what degree-degree correlation is observed from the GC. By formulating the generating function for the joint probability of degrees of the GC, we prove that the GC in random networks with arbitrary degree distribution $P(k)$ always shows disassortativity in terms of assortativity $r$ if the third moment $\langle k^3 \rangle$ of $P(k)$ is finite and the networks are not singly connected. In addition, by analyzing the average degree $\bar{k}_{nn}(k)$ of nodes adjacent to

_____

[*]shogo.mizutaka.sci@vc.ibaraki.ac.jp
[†]takehisa.hasegawa.sci@vc.ibaraki.ac.jp

degree $k$ nodes, we show that $\bar{k}_{nn}(k)$ at the percolation threshold is proportional to $k^{-1}$ as long as $\langle k^3 \rangle < \infty$ and is also a decreasing function of $k$ near $p = p_c$. These results mean that the GC possesses disassortativity near the percolation threshold. The validity of the analytical treatment is confirmed by extensive Monte Carlo simulations.

The rest of this paper is organized as follows. In Sec. II we formulate the assortativity $r$ for the GC created by site percolation using the generating functions. The comparison between analytical treatment proposed in Sec. II and simulations is shown in Sec. III. In addition, we show exact expressions of assortativity $r$ at the critical point for $z$-regular random networks and Erdős-Rényi random graphs. In Sec. IV we further show the disassortativity of the GC by showing that $\bar{k}_{nn}(k)$ is a decreasing function of degree $k$. Section V is devoted to the summary and discussion.

## II. ANALYTICAL TREATMENTS

Let us consider an uncorrelated random network with an arbitrary degree distribution $P(k)$. First, let $G_0(x)$ be the generating function for the probability, $P(k)$, of a randomly chosen node having degree $k$, as

$$G_0(x) = \sum_{k=0}^{\infty} P(k) x^k. \tag{1}$$

Using Eq. (1), the generating function $G_1(x)$ for the probability of an edge leading to a degree $k$ node is given by

$$G_1(x) = G_0'(x)/G_0'(1) = \sum_{k=1}^{\infty} \frac{k P(k)}{\langle k \rangle} x^{k-1}, \tag{2}$$

where $G_0'(x)$ is the derivative of $G_0(x)$ with respect to $x$ and $\langle k \rangle$ is the mean of the degree distribution $P(k)$, $\langle k \rangle = \sum_k k P(k)$. In this study we concentrate on the site percolation problem on a given substrate network with $P(k)$: each node is occupied with probability $p$ and is unoccupied otherwise. In general, there exists a threshold $p_c$ above which an infinitely large cluster, i.e., a GC, emerges in the thermodynamic limit, which means that the fraction $S$ of nodes belonging to the GC becomes $S > 0$ from $p > p_c$. We denote by $u$ the probability

that one end of an edge randomly chosen from the substrate network does not lead to a GC. The probability $u$ is given as the solution of the following self-consistent equation:

$$u = q + p G_1(u), \tag{3}$$

where $q = 1 - p$. Using the probability $u$, we have the fraction $S$ as

$$S = p[1 - G_0(u)]. \tag{4}$$

The percolation threshold $p_c$ is given with the condition that Eq. (3) has a nontrivial solution of $u < 1$, yielding $S > 0$. For uncorrelated random networks, it is known as $p_c = \langle k \rangle / \langle k(k-1) \rangle$ (see Refs. [14,15]).

Let us focus on only degree correlations of GCs formed by the site percolation in uncorrelated networks. First, we consider the conditional probability $P(\text{GC}, k, k'|l, m)$ that a randomly chosen edge has two ends with degree $k$ and $k'$ and belongs to the GC conditioned on the two ends having originally $l$ and $m$ neighbors in a substrate network. As $p G_1(u)$ is the probability that one end of an edge is occupied and does not lead to the GC, $p^2[1 - G_1^{k-1}(u) G_1^{k'-1}(u)]$ represents the probability that a randomly chosen edge leads to two occupied ends with degree $k$ and $k'$ and belongs to the GC. Therefore, we can write the probability $P(\text{GC}, k, k'|l, m)$ as

$$P(\text{GC}, k, k'|l, m) = p^2 \left[1 - G_1^{k-1}(u) G_1^{k'-1}(u)\right] \binom{l-1}{k-1}$$
$$\times p^{k-1} q^{l-k} \binom{m-1}{k'-1} p^{k'-1} q^{m-k'}. \tag{5}$$

Let $P(k, k')$ and $P(\text{GC})$ be the joint distribution of degrees in the substrate network and the probability that an edge belongs to the GC, respectively. The relations $P(\text{GC}, k, k') = \sum_{l,m} P(\text{GC}, k, k'|l, m) P(l, m)$ and $P(k, k'|\text{GC}) = P(\text{GC}, k, k')/P(\text{GC})$ are satisfied. We also have $P(\text{GC}) = p^2[1 - G_1^2(u)]$ immediately. For convenience, we denote $P(k, k'|\text{GC})$ as $P_{\text{GC}}(k, k')$ and the subscript GC is used for conditional probabilities conditioned on the GC. Using these relations, we find the joint distribution $P_{\text{GC}}(k, k')$ of degrees on the GC as

$$P_{\text{GC}}(k, k') = \frac{1 - G_1^{k-1}(u) G_1^{k'-1}(u)}{1 - G_1^2(u)} \sum_{l \geqslant k, m \geqslant k'} \binom{l-1}{k-1} p^{k-1} q^{l-k} \binom{m-1}{k'-1} p^{k'-1} q^{m-k'} \frac{l P(l)}{\langle k \rangle} \frac{m P(m)}{\langle k \rangle}, \tag{6}$$

where we use the relation $P(k, k') = [k P(k)/\langle k \rangle][k' P(k')/\langle k \rangle]$ because the substrate network is uncorrelated. The generating function $B(x, y)$ for $P_{\text{GC}}(k, k')$ is obtained as follows (see the Appendix for details):

$$B(x, y) = \sum_{k \geqslant 1, k' \geqslant 1} P_{\text{GC}}(k, k') x^{k-1} y^{k'-1} = \frac{G_1(q + px) G_1(q + py) - G_1[q + p G_1(u) x] G_1[q + p G_1(u) y]}{1 - G_1^2(u)}. \tag{7}$$

From $B(x, y)$, the generating function $S(x)[= B(x, 1) = B(1, x)]$ for the marginal distribution $Q_{\text{GC}}(k)[= \sum_{k'} P_{\text{GC}}(k, k')]$, which is the probability of an edge reaching a node with degree $k$ conditioned on the edge in the GC, is

$$S(x) = \sum_{k=1}^{\infty} Q_{\text{GC}}(k) x^{k-1} = \frac{G_1(q + px) - G_1(u) G_1[q + p G_1(u) x]}{1 - G_1^2(u)}. \tag{8}$$

Obviously, these generating functions $B(x, y)$ and $S(x)$ are reduced to expressions for generating functions in Ref. [11] when $p = 1$. Thus, the present formalism is a generalization of the previous method, in which the site percolation process is

incorporated. In accordance with the argument in Ref. [11], assortativity $r$ of the GC is given by $B(x, y)$ and $S(x)$ as

$$r = \frac{\partial_x \partial_y B(x, y) - [\partial_x S(x)]^2}{(x\partial_x)^2 S(x) - [\partial_x S(x)]^2}\bigg|_{x=y=1}. \tag{9}$$

Substituting Eqs. (7) and (8) into Eq. (9), we can find the general result for assortativity of the GC as

$$r = \frac{-p\tilde{g}_0^2(g_1 - \tilde{g}_1)^2}{(1 - \tilde{g}_0^2)(g_1 - \tilde{g}_0^2\tilde{g}_1 + pg_2 - p\tilde{g}_0^3\tilde{g}_2) - p(g_1 - \tilde{g}_0^2\tilde{g}_1)^2}, \tag{10}$$

where

$$g_n = G_1^{(n)}(1) = \frac{\partial^n G_1(x)}{\partial x^n}\bigg|_{x=1} \tag{11}$$

and

$$\tilde{g}_n = G_1^{(n)}(u) = \frac{\partial^n G_1(x)}{\partial x^n}\bigg|_{x=u}. \tag{12}$$

The denominator of the right-hand side in Eq. (10) is equal to $(1 - \tilde{g}_0^2)\sigma_{Q_{GC}}/p$, where $\sigma_{Q_{GC}}$ is the variance of $Q_{GC}(k)$, and is a positive real number. Then, the sign of assortativity is determined by the numerator. Therefore, the assortativity satisfies an inequality $r \leqslant 0$ for $p_c \lesssim p \leqslant 1$. The factor $(g_1 - \tilde{g}_1)^2$ in Eq. (10) claims that if a GC exists, it always exhibits disassortativity independently of the degree distribution because $(g_1 - \tilde{g}_1)^2$ becomes a nonzero positive value for $p \gtrsim p_c$. The result is persistent even at $p = 1$ when the substrate network is not singly connected, which is consistent with previous results in Refs. [10,11]. The zero assortativity is observed only when the network is singly connected at $p = 1$ because then the factor $\tilde{g}_0$ becomes zero. It is noted here that assortativity $r$ cannot be negative in infinitely large networks with $\langle k^3 \rangle = \infty$ (see Ref. [16]). The factor $g_2$ appearing in the denominator contains $\langle k^3 \rangle$ and reflects the feature.

## III. NUMERICAL CHECK

To evaluate the validity of our analytical treatment for uncorrelated random networks, we compare analytical estimates of the assortativity $r$ with corresponding simulation results. In our simulations, we utilize the configuration model which realizes uncorrelated random networks according to a predefined degree distribution. In the following subsections, we concentrate on typical examples, i.e., $z$-regular random graphs, Erdős-Rényi random graphs, and scale-free networks.

### A. $z$-regular random graphs

First, let us consider $z$-regular random graphs as a simple illustrative example. The degree distribution $P(k)$ of the $z$-regular random graph is

$$P(k) = \delta_{kz}, \tag{13}$$

whose percolation threshold is given as

$$p_c = \frac{1}{z-1}. \tag{14}$$

Figure 1 shows the $p$ dependence of assortativity $r$. Grayscale tube lines represent analytical estimates obtained
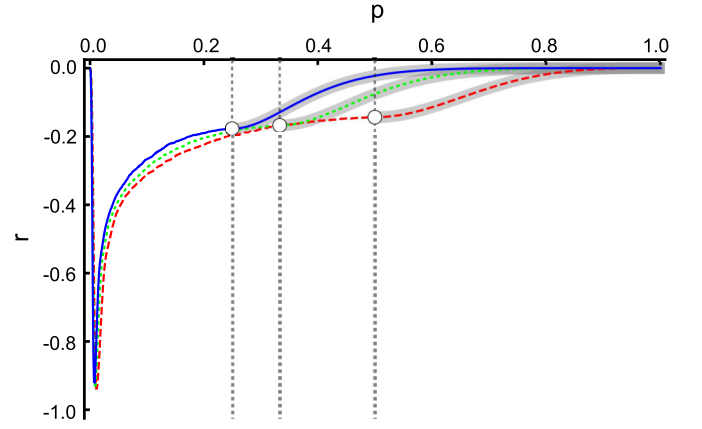


FIG. 1. Comparison with the analytical treatment and simulation results for $p$ dependence of assortativity $r$. The $z$-regular random graph is utilized as the substrate network. The grayscale tube lines represent the results obtained by Eq. (10) with the aid of Eqs. (2) and (3). The simulation results are for $z$-regular random graphs with $z = 5$ (solid blue line), 4 (dotted green line), and 3 (dashed red line). The number of nodes used for simulations is $N = 10^6$.

from Eq. (10), and the other lines are drawn from Monte Carlo simulations. In our simulations, we generated 10 network realizations and performed site percolation $10^3$ times on each realization to take the average of $r$ at given values of $p$. On each run, we specify the largest component, which corresponds to the GC for $p > p_c$, based on the Newman-Ziff algorithm [17]. The assortativity of the largest component is evaluated and compared with the result obtained by analytical treatment. Our analytical estimates for $r$ match perfectly with the numerical data for $p > p_c$ in all cases. The vertical dashed lines from left to right indicate the percolation thresholds $p_c$ when $z = 5$, 4, and 3, respectively. Our numerical data assert that the assortativity $r$ does not show the singular behavior just at and around the percolation threshold $p_c$ even when the system size goes to infinitely large. This implies that the analytical expression for the assortativity $r$ at $p = p_c$ can be obtained. Approximating the probability $u$ at $p \gtrsim p_c$ as $u \sim 1 - \epsilon$ where $\epsilon$ is an infinitesimal value, we have the relation

$$G_1(u) \sim 1 - \frac{\langle k(k-1) \rangle}{\langle k \rangle}\epsilon. \tag{15}$$

The assortativity $r_c$ at $p = p_c$ is given by substituting Eq. (15) into Eq. (10) and taking $\epsilon \to 0$ as

$$r_c = -\frac{p_c g_2^2}{2g_1(2g_1^2 + g_2 + 3p_c g_1 g_2 + p_c g_3) - p_c(2g_1^2 + g_2)^2}. \tag{16}$$

Using Eqs. (13), (14), and (16), we have the assortativity $r_c$ at $p = p_c$ for a $z$-regular random graph:

$$r_c = -\frac{z-2}{5z-8}. \tag{17}$$

Large symbols on the edges of grayscale tube lines in Fig. 1 are $r_c$ given by Eq. (17) to confirm the accuracy of analytical treatment.
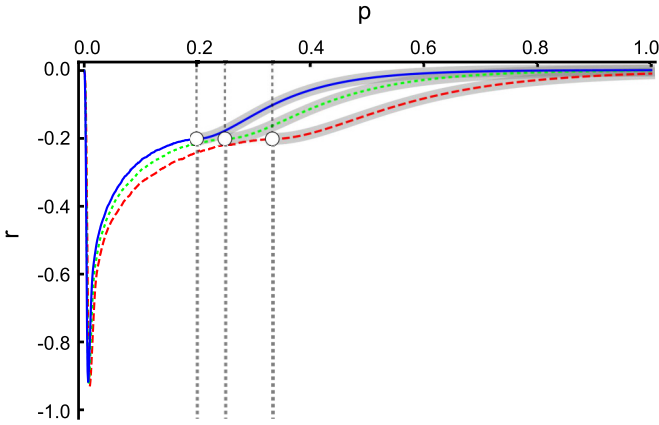
FIG. 2. Comparison with the analytical treatment and simulation results for $p$ dependence of assortativity $r$. The substrate network obeys the degree distribution with Eq. (18). The grayscale tube lines represent the results obtained by Eq. (10) with the aid of Eqs. (2) and (3). The simulation results are for Erdős-Rényi random graphs with $\langle k \rangle = 5$ (solid blue line), 4 (dotted green line), and 3 (dashed red line). The number of nodes used for simulations is $N = 10^6$. Ten realizations of networks were generated, and the site percolation process was performed $10^3$ times for each network to obtain simulation results.

### B. Erdős-Rényi random graphs

The degree distribution $P(k)$ and the percolation threshold $p_c$ for Erdős-Rényi random graphs are

$$P(k) = \frac{\langle k \rangle^k e^{-\langle k \rangle}}{k!} \quad (18)$$

and

$$p_c = \frac{1}{\langle k \rangle}, \quad (19)$$

respectively. Figure 2 shows the assortativity $r$ as a function of $p$. The analytical estimates for $r$ match perfectly with the numerical data for $p > p_c$ as is the case with $z$-regular random graphs. The assortativity $r_c$ at $p = p_c$ for Erdős-Rényi random graphs is given as

$$r_c = -\frac{1}{5}, \quad (20)$$

independently of the average degree $\langle k \rangle$ of original graphs.

### C. Scale-free networks

Finally, we consider scale-free networks whose degree distribution obeys $P(k) \sim k^{-\gamma}$ for $2 \leqslant k \leqslant k_{cut}$. To argue the effect of network heterogeneity on the degree correlation of the GC, in scale-free networks with $k_{cut} \to \infty$, we start by comparing analytical and numerical results for the case with a finite cutoff degree, i.e., $k_{cut} < \infty$. Figure 3 shows the results for the scale-free networks with exponent $\gamma = 2.5$ and cutoff degree $k_{cut} = 10^3$. Monte Carlo data asymptotically reach the analytical line as increasing the system size $N$, which implies that the analytical treatment is valid for infinite networks with a finite cutoff degree. This also indicates the disassortativity of the GCs formed by occupied nodes on the scale-free networks with a finite cutoff degree analytically and numerically. The
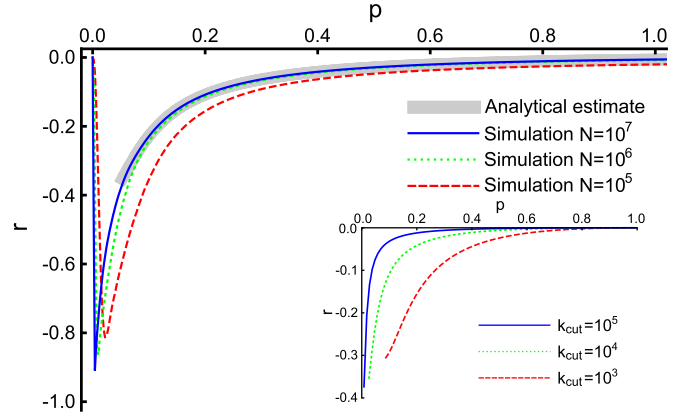


FIG. 3. Comparison with the analytical treatment and simulation results for $p$ dependence of assortativity $r$. Main panel: The results are for scale-free networks with exponent $\gamma = 2.5$ and cutoff degree $k_{cut} = 10^3$ of the degree distribution. The grayscale tube line represents the result obtained by Eq. (10) with the aid of Eqs. (2) and (3). The solid blue, dotted green, and dashed red lines are simulation results for $N = 10^7$, $10^6$, and $10^5$, respectively. Ten realizations of networks were generated, and the site percolation process was performed $10^3$ times for each network to obtain simulation results. Inset: The analytical estimates of $r$ for scale-free networks with $\gamma = 2.5$ and $k_{cut} = 10^5$ (solid blue), $10^4$ (dotted green), and $10^3$ (dashed red).

validity of our analytical treatment holds for different values of $\gamma$ and $k_{cut}$ (not shown). Based on the analytical treatment, we display the $p$ dependence of $r$ for scale-free networks with $\gamma = 2.5$ and different values of $k_{cut}$ (the inset of Fig. 3). It is known for $\gamma < 3$ that the percolation threshold approaches zero as $k_{cut}$ increases. With increasing $k_{cut}$, the trend that the assortativity $r$ goes to 0 rapidly above $p_c$ which is located on the left end of the line, is enhanced. This result indicates that when $k_{cut} \to \infty$, $p_c$ goes to 0 and the assortativity $r$ becomes 0 for $p > 0$.

In addition, the assortativity $r_c$ at $p = p_c$ as a function of the scale-free exponent $\gamma$ is shown in Fig. 4. The grayscale tube line and the symbols represent the analytical estimate of $r_c$ for networks without and with finite cutoff degrees, respectively. For $\gamma > 5$, $p_c > 0$ and $r_c < 0$ even for $k_{cut} \to \infty$. Most symbols are on the grayscale tube line, indicating that $r_c$ is not sensitive to $k_{cut}$ for $\gamma > 5$. For $3 < \gamma < 5$, $p_c > 0$ and $r_c = 0$ when $k_{cut} \to \infty$. The fashion that $r_c \to 0$ at $k_{cut} \to \infty$ is also reflected on the $k_{cut}$ dependence of $r_c$, i.e., the movement of symbols at fixed $\gamma$. The zero assortativity of the GC is because the right-hand side of Eq. (16) includes $\langle k^4 \rangle$, which diverges for $3 < \gamma < 5$, in the denominator, where $\langle k^4 \rangle$ is induced by asymptotically expanding the right-hand side of Eq. (10) near $p = p_c$. For $\gamma < 3$, the $k_{cut}$ dependence of $r_c$ in Fig. 4 seems to suggest that $r_c$ converges to a finite negative value as $k_{cut} \to \infty$. However, $p_c = 0$ in this region and $r_c$ will become 0 for $p > 0$, as mentioned above.

Finally, we consider $r$ for $p > p_c (> 0)$ for the case of $4 < \gamma < 5$. In the inset of Fig. 4, we display the $p$ dependence of $r$ for the scale-free network with $\gamma = 4.5$. We find that $r$ always takes a finite negative value at $p > p_c$, although the assortativity $r_c$ at $p = p_c$ becomes 0. The assortativity $r$ includes the
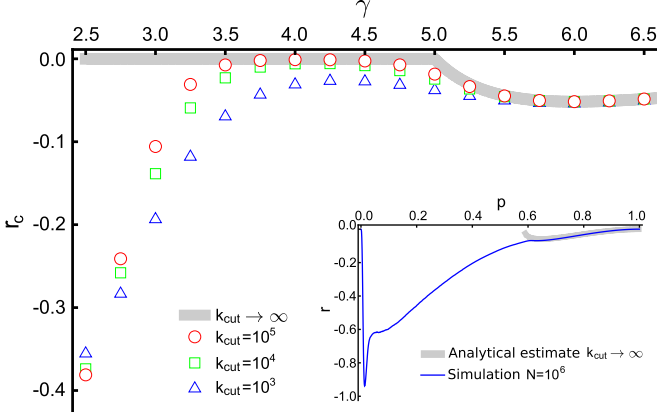
FIG. 4. Main panel: Analytical estimates of assortativity $r_c$ at $p = p_c$ as a function of $\gamma$. The grayscale tube line is the result for scale-free networks with $k_{cut} \to \infty$. Symbols represent the results for scale-free networks with finite cutoff degree $k_{cut} = 10^5$ (red circles), $10^4$ (green squares), and $10^3$ (blue triangles). All data are obtained by Eq. (10) with the aid of Eqs. (2) and (3). Inset: $p$ dependence of the assortativity $r$. The grayscale tube line represents the analytical result for scale-free networks with $\gamma = 4.5$ and $k_{cut} \to \infty$ of the degree distribution. The line is the simulation result for $N = 10^6$. Ten realizations of networks were generated, and the site percolation process was performed $10^3$ times for each network to obtain simulation results.

moments of the degree distribution: $\langle k^3 \rangle$ in $r$ of the whole network or its GC for $p > p_c$, and $\langle k^4 \rangle$ of the GC at $p = p_c$. Therefore, $r$ sometimes becomes useless for scale-free networks because these moments diverge according to the value of the exponent $\gamma$. However, such zero assortativity never means that the GC does not have the degree-degree correlation. We consider the disassortativity of the GC in scale-free networks with an exponent in $3 < \gamma < 5$ in the next section.

## IV. BEHAVIOR OF $\bar{k}_{nn}(k)$

We further discuss the disassortativity of the GC with a different quantity. The average degree $\bar{k}_{nn}(k)$ of nodes adjacent to degree $k$ nodes is more informative than the assortativity $r$. The quantity $\bar{k}_{nn}(k)$ of the GC is calculated from the probability $P_{GC}(k'|k)$ of degree $k'$ nodes adjacent to the degree $k$ nodes in the GC. The probability $P_{GC}(k'|k)$ is given by

$$P_{GC}(k'|k) = \frac{1 - \tilde{g}_0^{k-1} \tilde{g}_0^{k'-1}}{1 - \tilde{g}_0^k} \frac{k' \tilde{P}(k')}{\langle \tilde{k} \rangle}, \qquad (21)$$

where $\tilde{P}(k) = \sum_{m \geqslant k}^{\infty} \binom{m}{k} p^k q^{m-k} P(m)$ and $\langle \tilde{k}^n \rangle = \sum k^n \tilde{P}(k)$ (see the Appendix for details). Note that $\tilde{P}(k)$ corresponds to the degree distribution for the network whose nodes are randomly occupied with probability $p$ on the substrate network. Equation (21) leads to the average degree $\bar{k}_{nn}(k)[= \sum_{k'} k' P_{GC}(k'|k)]$ of nodes adjacent to degree $k$ nodes as

$$\bar{k}_{nn}(k) = \frac{\langle \tilde{k}^2 \rangle}{\langle \tilde{k} \rangle} \left[ \frac{1 - \tilde{g}_0^{k-1} h(\tilde{g}_0)}{1 - \tilde{g}_0^k} \right], \qquad (22)$$
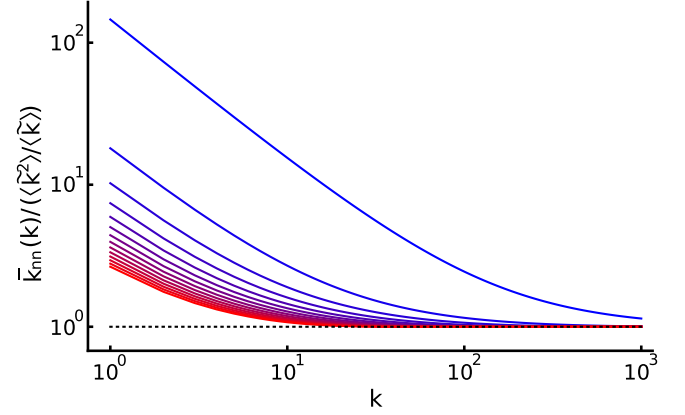


FIG. 5. Rescaled $\bar{k}_{nn}(k)$ by $\frac{\langle \tilde{k}^2 \rangle}{\langle \tilde{k} \rangle}$ as a function of degree $k$ for scale-free networks with $\gamma = 3.5$ and $k_{cut} \to \infty$. Analytical estimates (22) for $p = 0.27, 0.28, \ldots, 0.40$ are shown as the lines from top to bottom. Here $p_c = \langle k \rangle / \langle k(k-1) \rangle = 0.2687$.

where

$$h(\tilde{g}_0) = \sum_k \frac{k^2}{\langle \tilde{k}^2 \rangle} \tilde{P}(k) \tilde{g}_0^{k-1}. \qquad (23)$$

Figure 5 shows analytical estimates of rescaled $\bar{k}_{nn}(k)$ by $\langle \tilde{k}^2 \rangle / \langle \tilde{k} \rangle$ as the function of degree $k$ for a scale-free network with $\gamma = 3.5$ and $k_{cut} \to \infty$ to which the GC shows zero assortativity for $p \geqslant p_c$. The lines for several values of $p(\gtrsim p_c)$ indicate that each rescaled $\bar{k}_{nn}(k)$ decreases monotonically with increasing degree $k$. This means that disassortativity is observed in the GC formed by percolation processes on the scale-free network with $\gamma = 3.5$ and $k_{cut} \to \infty$.

Finally, we study the behavior of $\bar{k}_{nn}(k)$ near $p = p_c$. Using Eq. (15), we expand Eq. (22) as follows:

$$\bar{k}_{nn}(k) \sim \frac{\langle \tilde{k}^2 \rangle}{\langle \tilde{k} \rangle} \left\{ 1 + \frac{1}{k} \frac{\langle \tilde{k}^3 \rangle - 2\langle \tilde{k}^2 \rangle}{\langle \tilde{k}^2 \rangle} \right.$$
$$\left. - \left( 1 - \frac{1}{k} \right) \frac{\langle \tilde{k}^3 \rangle - \langle \tilde{k}^2 \rangle}{\langle \tilde{k}^2 \rangle} \frac{\langle k^2 \rangle - \langle k \rangle}{\langle k \rangle} \epsilon \right\}. \qquad (24)$$

The result means that $\bar{k}_{nn}(k)$ near $p = p_c$ is proportional to $k^{-1}$ independently of the original degree distribution $P(k)$. In addition, in the limit of $\epsilon \to 0$, i.e., $p \to p_c$, $\bar{k}_{nn}(k)$ is rewritten as

$$\bar{k}_{nn}(k) = \frac{\langle \tilde{k}^2 \rangle}{\langle \tilde{k} \rangle} \left\{ 1 + \frac{1}{k} \frac{\langle \tilde{k}^3 \rangle - 2\langle \tilde{k}^2 \rangle}{\langle \tilde{k}^2 \rangle} \right\}. \qquad (25)$$

The exact expression (25) of $\bar{k}_{nn}(k)$ at $p = p_c$ holds if $\langle k^3 \rangle < \infty$ because $\langle \tilde{k}^3 \rangle$ contains $\langle k^3 \rangle$. To summarize, $\bar{k}_{nn}(k)$ at and above $p_c$ shows the disassortativity of the GC for scale-free networks with $4 < \gamma < 5$, although $r_c$ failed to capture it. For $3 < \gamma < 4$, $\bar{k}_{nn}(k)$ is useless just at $p_c$ but again shows the disassortativity of the GC above $p_c$.

## V. SUMMARY AND DISCUSSION

In this work, the degree-degree correlations of GCs formed by the site percolation process on uncorrelated random networks have been analyzed. By formulating the joint probability of degrees on a GC by means of the generating

function, we have shown the following general properties of GCs formed by the percolation process in random networks: (1) The assortativity $r$ defined by Pearson's correlation coefficient for degrees satisfies an inequality $r \leqslant 0$ in the percolating phase if the third moment $\langle k^3 \rangle$ of the degree distribution is finite. (2) The average degree $\bar{k}_{nn}(k)$ of nodes adjacent to degree $k$ nodes at the percolation threshold is proportional to $k^{-1}$ if $\langle k^3 \rangle < \infty$.

As has been shown through this work, the negative degree-degree correlation (disassortativity) naturally emerges when we focus on a component of an uncorrelated network. It should be noted that one cannot understand the degree-degree correlations of whole networks even if we analyze their components, and one may not be able to understand correctly the behavior of dynamics on networks even if we investigate the dynamics on the components. This probably holds true for real networks constructed by data: the difference between the degree correlations of the whole network and of a component would emerge in real-world networks. It is necessary to pay attention to the lack of data when we analyze real-world networks because a lack of data, expressed as the removals of nodes or edges in percolation processes, would enhance the degree-degree correlations.

The results in this study are consistent with the previous result concerning the relation between fractality and disassortativity of real-world networks [18]. The disassortativity of GCs might be established even if an original network has a certain strength of positive degree-degree correlation. However, the behavior of degree-degree correlations of the GCs in assortative networks is not so simple, as will be argued elsewhere [19].

We did not discuss the degree-degree correlations of GCs in scale-free networks with $\langle k^3 \rangle = \infty$. To evaluate the correlations of such networks, Spearman's rank correlation coefficient of degrees has been utilized [16,20,21]. It is interesting to evaluate degree-degree correlations of GC using Spearman's rank correlation coefficient, although we expect the generality of disassortativity of percolating clusters.

### APPENDIX: DERIVATION OF SEVERAL QUANTITIES

The generating function $B(x, y)$ for $P_{GC}(k, k')$ in Eq. (7) is calculated as follows:

$$
\begin{aligned}
B(x, y) &= \sum_{k \geqslant 1, k' \geqslant 1} P_{GC}(k, k') x^{k-1} y^{k'-1} \\
&= \sum_{k \geqslant 1, k' \geqslant 1} \frac{1 - G_1^{k-1}(u) G_1^{k'-1}(u)}{1 - G_1^2(u)} \sum_{l \geqslant k, m \geqslant k'} \binom{l-1}{k-1} p^{k-1} q^{l-k} \binom{m-1}{k'-1} p^{k'-1} q^{m-k'} \frac{l P(l)}{\langle k \rangle} \frac{m P(m)}{\langle k \rangle} x^{k-1} y^{k'-1} \\
&= \frac{1}{1 - G_1^2(u)} \sum_{l,m}^{\infty,\infty} \frac{l P(l)}{\langle k \rangle} \frac{m P(m)}{\langle k \rangle} \sum_{k,k'}^{l,m} \left[ 1 - G_1^{k-1}(u) G_1^{k'-1}(u) \right] \binom{l-1}{k-1} p^{k-1} q^{l-k} \binom{m-1}{k'-1} p^{k'-1} q^{m-k'} x^{k-1} y^{k'-1} \\
&= \frac{1}{1 - G_1^2(u)} \sum_{l,m}^{\infty,\infty} \frac{l P(l)}{\langle k \rangle} \frac{m P(m)}{\langle k \rangle} \{ (q + px)^{l-1} (q + py)^{m-1} - [q + pG_1(u)x]^{l-1} [q + pG_1(u)y]^{m-1} \} \\
&= \frac{G_1(q + px) G_1(q + py) - G_1[q + pG_1(u)x] G_1[q + pG_1(u)y]}{1 - G_1^2(u)}.
\end{aligned}
\tag{A1}
$$

Using Eqs. (A1) and (8), we obtain components constructing the assortativity $r$ as

$$
\begin{aligned}
\partial_x \partial_y B(x, y)|_{x=y=1} &= \frac{p^2 G_1'(q + px) G_1'(q + py) - p^2 G_1^2(u) G_1'[q + pG_1(u)x] G_1'[q + pG_1(u)y]}{1 - G_1^2(u)} \bigg|_{x=y=1} \\
&= \frac{p^2 G_1'^2(1) - p^2 G_1^2(u) G_1'^2(u)}{1 - G_1^2(u)},
\end{aligned}
\tag{A2}
$$

$$
\partial_x S(x)|_{x=y=1} = \frac{p G_1'(q + px) - p G_1^2(u) G_1'[q + pG_1(u)x]}{1 - G_1^2(u)} \bigg|_{x=y=1} = \frac{p G_1'(1) - p G_1^2(u) G_1'(u)}{1 - G_1^2(u)},
\tag{A3}
$$

$$
\partial_x^2 S(x)|_{x=y=1} = \frac{p^2 G_1''(q + px) - p^2 G_1^3(u) G_1''[q + pG_1(u)x]}{1 - G_1^2(u)} \bigg|_{x=y=1} = \frac{p^2 G_1''(1) - p^2 G_1^3(u) G_1''(u)}{1 - G_1^2(u)},
\tag{A4}
$$

where Eq. (3) holds. By substituting Eqs. (A2), (A3), and (A4) into Eq. (9), we have Eq. (10).

The probability $P_{\text{GC}}(k'|k)$ is obtained as follows. As $Q_{\text{GC}}(k) = \sum_{k'} P_{\text{GC}}(k, k')$ is given by

$$Q_{\text{GC}}(k) = \frac{1 - G_1^k(u)}{1 - G_1^2(u)} \sum_{l \geq k} \binom{l-1}{k-1} p^{k-1} q^{l-k} \frac{l P(l)}{\langle k \rangle},$$ (A5)

we have $P_{\text{GC}}(k'|k) = \frac{P_{\text{GC}}(k,k')}{Q_{\text{GC}}(k)}$ using Eqs. (6) and (A5) as

$$P_{\text{GC}}(k'|k) = \frac{1 - G_1^{k-1}(u) G_1^{k'-1}(u)}{1 - G_1^k(u)} \sum_{m \geq k'} \binom{m-1}{k'-1} p^{k'-1} q^{m-k'} \frac{m P(m)}{\langle k \rangle}$$

$$= \frac{1 - G_1^{k-1}(u) G_1^{k'-1}(u)}{1 - G_1^k(u)} \frac{k'}{p\langle k \rangle} \sum_{m \geq k'} \binom{m}{k'} p^{k'} q^{m-k'} P(m) = \frac{1 - G_1^{k-1}(u) G_1^{k'-1}(u)}{1 - G_1^k(u)} \frac{k' \tilde{P}(k')}{\langle \tilde{k} \rangle},$$ (A6)

where $\tilde{P}(k)$ is the degree distribution of connected components consisting of occupied nodes in a network with degree distribution $P(k)$:

$$\tilde{P}(k) = \sum_{m \geq k} \binom{m}{k} p^k q^{m-k} P(m).$$ (A7)

Then, the first three moments of $\tilde{P}(k)$ are as follows:

$$\langle \tilde{k} \rangle = p\langle k \rangle,$$ (A8)

$$\langle \tilde{k}^2 \rangle = p^2 \langle k^2 \rangle + pq\langle k \rangle,$$ (A9)

$$\langle \tilde{k}^3 \rangle = p^3 \langle k^3 \rangle + 3p^2 q\langle k^2 \rangle + pq(q-p)\langle k \rangle.$$ (A10)

Substituting $\tilde{g}_0 = G_1(u)$ into Eq. (A6), we have Eq. (21).

[1] R. Albert and A.-L. Barabási, Rev. Mod. Phys. **74**, 47 (2002).

[2] S. N. Dorogovtsev and J. F. F. Mendes, *Evolution of Networks: From Biological Nets to the Internet and WWW* (Oxford University Press, Oxford, 2003).

[3] M. E. J. Newman, Phys. Rev. Lett. **89**, 208701 (2002).

[4] M. E. J. Newman, Phys. Rev. E **67**, 026126 (2003).

[5] J. P. Gleeson, Phys. Rev. E **77**, 046117 (2008).

[6] A. V. Goltsev, S. N. Dorogovtsev, and J. F. F. Mendes, Phys. Rev. E **78**, 051105 (2008).

[7] A. Arenas, A. Díaz-Guilera, J. Kurths, Y. Moreno, and C. Zhou, Phys. Rep. **469**, 93 (2008).

[8] S. Melnik, M. A. Porter, P. J. Mucha, and J. P. Gleeson, Chaos **24**, 023106 (2014).

[9] R. Pastor-Satorras, C. Castellano, P. Van Mieghem, and A. Vespignani, Rev. Mod. Phys. **87**, 925 (2015).

[10] P. Bialas and A. K. Oleś, Phys. Rev. E **77**, 036124 (2008).

[11] I. Tishby, O. Biham, E. Katzav, and R. Kűhn, Phys. Rev. E **97**, 042318 (2018).

[12] D. Stauffer and A. Aharony, *Introduction to Percolation Theory* (Taylor & Francis, London, 1991).

[13] A. Srivastava, B. Mitra, N. Ganguly, and F. Peruani, Phys. Rev. E **86**, 036106 (2012).

[14] M. E. J. Newman, S. H. Strogatz, and D. J. Watts, Phys. Rev. E **64**, 026118 (2001).

[15] R. Cohen, K. Erez, D. Ben-Avraham, and S. Havlin, Phys. Rev. Lett. **85**, 4626 (2000).

[16] N. Litvak and R. van der Hofstad, Phys. Rev. E **87**, 022801 (2013).

[17] M. E. J. Newman and R. M. Ziff, Phys. Rev. E **64**, 016706 (2001).

[18] S.-H. Yook, F. Radicchi, and H. Meyer-Ortmanns, Phys. Rev. E **72**, 045105 (2005).

[19] T. Hasegawa and S. Mizutaka (unpublished).

[20] W.-Y. Zhang, Z.-W. Wei, B.-H. Wang, and X.-P. Han, Physica A **451**, 440 (2016).

[21] Y. Fujiki, S. Mizutaka, and K. Yakubo, Eur. Phys. J. B **90**, 126 (2017).