

## Structure of percolating clusters in random clustered networks

Takehisa Hasegawa\* and Shogo Mizutaka†

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

(Received 25 July 2019; accepted 2 June 2020; published 22 June 2020)

We examine the structure of the percolating cluster (PC) formed by site percolation on a random clustered network (RCN) model. Using the generating functions, we formulate the clustering coefficient and assortative coefficient of the PC. We analytically and numerically show that the PC in the highly clustered networks is clustered even at the percolation threshold. The assortativity of the PC depends on the details of the RCN. The PC at the percolation threshold is disassortative when the numbers of edges and triangles of each node are assigned by Poisson distributions, but assortative when each node in an RCN has the same small number of edges, most of which form triangles. This result seemingly contradicts the disassortativity of fractal networks, although the renormalization scheme unveils the disassortative nature of a fractal PC.

DOI: [10.1103/PhysRevE.101.062310](https://doi.org/10.1103/PhysRevE.101.062310)

### I. INTRODUCTION

A number of studies on complex networks have reported the structural characteristics of a real network ranging from the World Wide Web to food webs [1–4]. Numerous real networks are scale free; i.e., the distribution  $p_k$  of degree  $k$  obeys a power law. Most of the real networks are small world, indicating that the mean shortest path length scales with the logarithm of the number of nodes, and the clustering coefficient, which is the mean probability that two randomly chosen neighbors of a randomly chosen node are adjacent, is high. Real networks would be classified by the degree-degree correlation, i.e., the correlation between the degrees of directly connected nodes. Social networks have a positive degree-degree correlation in which similar degree nodes are more likely to connect to each other while biological and technological networks have a negative degree-degree correlation indicating that dissimilar degree nodes are more likely to connect to each other. Furthermore, Song *et al.* reported on the fractality of real networks [5–7]: some real networks, such as the World Wide Web and protein-protein interaction networks, are fractal in the sense that the number of boxes for tiling a network decreases with the radius of boxes in a power law manner.

It is crucial to understand how structural characteristics are related to each other. Yook *et al.* [8] discovered from real network data that fractal networks have a negative degree-degree correlation, namely, disassortativity. This empirical rule is observed in the synthetic models of fractal networks [6,9], critical branching trees [10], and connected components at a critical state of an uncorrelated network model [11,12]. Furthermore, there are related works concerning the degree-degree correlation of spanning trees in fractal and small-world

networks [13–15] and the converse condition that disassortativity makes a network fractal [16]. However, it still remains unclear why fractal networks possess disassortativity and how robust the empirical rule is.

Site percolation on networks is known to exhibit a phase transition concerning clusters, which are connected components of occupied nodes. When the number of nodes is sufficiently large, the largest cluster is small and finite for  $f < f_c$ ; it occupies a finite fraction of the whole network and is called the percolating cluster (PC) for  $f > f_c$ , and it is a fractal at  $f = f_c$  [17]. Here  $f$  is a fraction of the occupied nodes and  $f_c$  is called the percolation threshold. The analysis of the largest cluster at  $f = f_c$ , which is called the fractal PC, leads us to further examine the relation between the fractality and the disassortativity in complex networks.

In a percolation process, a network splits into multiple connected components. It should be noted that the structural properties of a connected component are different from those of the whole network if the network is not singly connected [11,12,18–20]. Recent studies have focused on the methods to extract the infinitely large connected component from uncorrelated networks and compute its properties (e.g., degree distribution  $p_k$ , average degree  $\bar{k}_{nn}(k)$  of nodes adjacent to degree  $k$  nodes [12], and assortative coefficient  $r$  defined by Pearson's correlation coefficient for degrees of directly connected nodes [11]). Previous work [20] considered a PC formed by site percolation on uncorrelated networks and investigated the properties of the PC. For uncorrelated random networks obeying an arbitrary degree distribution with a finite third moment, the PC possesses a disassortativity above the percolation threshold [20]: the assortative coefficient  $r$  is always less than zero. Moreover, the average degree  $\bar{k}_{nn}(k)$  of the nodes adjacent to the degree  $k$  nodes is proportional to  $k^{-1}$  at  $f \rightarrow f_c$ . These indicate that the fractal PC is disassortative when it is formed by site percolation on uncorrelated networks.

The present study is a continuation of our previous work [20] and discusses whether the disassortativity of PCs is established in correlated networks. It is also interesting

\*takehisa.hasegawa.sci@vc.ibaraki.ac.jp

†Present address: School of Knowledge Science, Japan Advanced Institute of Science and Technology, 1-1 Asahidai, Nomi 924-1292, Japan; mizutaka@jaist.ac.jp

how other structural properties of the PC differ from those of original networks: e.g., how is the PC in a clustered network clustered? Newman [21] and Miller [22] independently introduced a random graph model with clustering, many of whose network properties can be well described via a generating function analysis. This model is highly clustered and even assortative (as shown below)—it is suitable for discussing the above-mentioned question. In this study, we consider site percolation on the random clustered network (RCN) model to investigate the structural properties, the clustering coefficient and the assortative coefficient, of the PC. Our generating function analysis describes the structure of the PC well: it perfectly agrees with the corresponding Monte Carlo simulations. We show that the PC formed by site percolation in highly clustered networks is clustered even at the percolation threshold. With respect to the assortativity, both analytical estimates and simulation results seemingly contradict the disassortativity of fractal networks: the fractal PC is assortative when the nodes in an RCN have the same small number of edges, most of which form triangles. Our discussion focuses on why a positive assortativity is observed in a fractal PC.

The paper is organized as follows. In Sec. II, we introduce the RCN model and recall the generating function approach for deriving its clustering coefficient and assortative coefficient. In Sec. III, we analyze the structure of the PC formed by site percolation on RCNs. We derive the clustering coefficient (Sec. III A) and assortative coefficient (Sec. III B) of the PC, applying our analysis to two types of RCNs (Sec. III C) in order to investigate the structures of the PC. In Sec. IV, we discuss the robustness of the fractal PC's disassortativity.

## II. RANDOM CLUSTERED NETWORK MODEL

The RCN model introduced by Newman [21] generalizes the configuration model to incorporate clustering. We assume that the joint probability,  $p_{s,t}$ , for the mean fraction of nodes with  $s$  single edges and  $t$  triangles is given and assign  $s_i$  edge stubs and  $t_i$  triangle stubs to each node  $i$  according to  $p_{s,t}$  under the constraint that  $\sum_i s_i$  and  $\sum_i t_i$  are multiples of 2 and 3, respectively. Given these stubs, we create a network by randomly selecting pairs of edge stubs and joining them to make single edges and by randomly selecting triples of triangle stubs and joining them to form triangles whose edges are referred to as triangle edges. This results in a random network in which the number of single edges incident to each node and the number of triangles it participates in are distributed according to  $p_{s,t}$ . Note that the total degree  $k$  of a node with  $s$  single edges and  $t$  triangles is  $k = s + 2t$  (Fig. 1).

The clustering coefficient  $C_0$  of the RCN is given by the generating functions [21]. First, we introduce the generating function  $G_p(x, y)$  for the joint probability  $p_{s,t}$ ,

$$G_p(x, y) = \sum_{s=0}^{\infty} \sum_{t=0}^{\infty} p_{s,t} x^s y^t. \quad (1)$$

Because the full degree distribution  $p_k$  is written as  $p_k = \sum_{s,t} p_{s,t} \delta_{k,s+2t}$  using the Kronecker delta  $\delta_{ij}$ , the generating function  $G_{\text{tot}}(z)$  for the full degree distribution  $p_k$  is presented

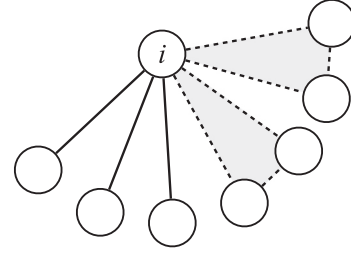


FIG. 1. Node  $i$  in this figure has three single edges and two triangles. The solid lines and dashed lines represent the single edges and triangle edges, respectively. Triangles are filled in with gray. The degree of node  $i$  is  $k_i = 3 + 2 \times 2 = 7$ .

as follows:

$$G_{\text{tot}}(z) = \sum_{k=0}^{\infty} p_k z^k = G_p(z, z^2). \quad (2)$$

The average degree  $\langle k \rangle$  is obtained from  $G_{\text{tot}}(z)$  as follows:

$$\langle k \rangle = \left. \frac{\partial G_{\text{tot}}(z)}{\partial z} \right|_{z=1} = \langle s \rangle + 2\langle t \rangle, \quad (3)$$

where  $\langle s \rangle = \sum_{s,t} s p_{s,t}$  and  $\langle t \rangle = \sum_{s,t} t p_{s,t}$ . For the RCN with  $N$  nodes, the number  $N_3$  of the connected triplets and the number  $N_{\Delta}$  of the triangles are given by the generating functions  $G_p(x, y)$  and  $G_{\text{tot}}(z)$  [21]:

$$N_3 = N \sum_k \frac{k(k-1)}{2} p_k = \frac{1}{2} N \left. \frac{\partial^2 G_{\text{tot}}(z)}{\partial z^2} \right|_{z=1} \quad (4)$$

and

$$3N_{\Delta} = N \sum_{s,t} t p_{s,t} = N \left. \frac{\partial G_p(x, y)}{\partial y} \right|_{x=y=1}. \quad (5)$$

We can then write the clustering coefficient  $C_0$  of the RCN as follows:

$$C_0 = \frac{3N_{\Delta}}{N_3} = 2 \left. \frac{\partial G_p(x, y)}{\partial y} \right|_{x=y=1} \bigg/ \left. \frac{\partial^2 G_{\text{tot}}(z)}{\partial z^2} \right|_{z=1}. \quad (6)$$

Moreover, the assortative coefficient  $r_0$  of the RCN is formalized using the generating functions. We consider two types of excess degree distributions [21]:  $q_{s,t}$ , which is the probability that a node reached by traversing a single edge has  $s+1$  single edges and  $t$  triangles, and  $r_{s,t}$ , which is the probability that a node reached by traversing a triangle has  $s$  single edges and  $t+1$  triangles. These probabilities are naturally derived as follows:

$$q_{s,t} = \frac{s+1}{\langle s \rangle} p_{s+1,t} \quad (7)$$

and

$$r_{s,t} = \frac{t+1}{\langle t \rangle} p_{s,t+1}. \quad (8)$$

We then introduce the generating functions for  $q_{s,t}$  and  $r_{s,t}$  as

$$G_q(x, y) = \sum_{s=0}^{\infty} \sum_{t=0}^{\infty} q_{s,t} x^s y^t = \frac{1}{\langle s \rangle} \frac{\partial G_p(x, y)}{\partial x} \quad (9)$$

and

$$G_r(x, y) = \sum_{s=0}^{\infty} \sum_{t=0}^{\infty} r_{s,t} x^s y^t = \frac{1}{\langle t \rangle} \frac{\partial G_p(x, y)}{\partial y}, \quad (10)$$

respectively. Here we denote the probability of choosing a single edge by

$$P_s = \frac{\langle s \rangle}{\langle s \rangle + 2\langle t \rangle}, \quad (11)$$

and the probability of choosing a triangle edge by

$$P_t = 1 - P_s = \frac{2\langle t \rangle}{\langle s \rangle + 2\langle t \rangle}. \quad (12)$$

Introducing the probability  $Q_0(k, k')$  that two ends of a randomly chosen edge have degrees  $k+1$  and  $k'+1$  and the probability  $Q_0(k) [= \sum_{k'} Q_0(k, k')]$  that an edge reaches a node with degree  $k+1$ , we can calculate the assortative coefficient  $r_0$  of the RCN from the following equation:

$$r_0 = \frac{\partial_x \partial_y B_0(x, y) - (\partial_x S_0(x))^2}{(x \partial_x)^2 S_0(x) - (\partial_x S_0(x))^2} \Big|_{x=y=1}, \quad (13)$$

where  $B_0(x, y)$  is the generating function for  $Q_0(k_1, k_2)$ ,

$$\begin{aligned} B_0(x, y) &= \sum_{k_1=0}^{\infty} \sum_{k_2=0}^{\infty} Q_0(k_1, k_2) x^{k_1} y^{k_2} \\ &= P_s G_q(x, x^2) G_q(y, y^2) + P_t x y G_r(x, x^2) G_r(y, y^2), \end{aligned} \quad (14)$$

and  $S_0(x) [= B_0(x, 1) = B_0(1, x)]$  is the generating function for  $Q_0(k)$ ,

$$S_0(x) = \sum_{k=0}^{\infty} Q_0(k) x^k = P_s G_q(x, x^2) + P_t x G_r(x, x^2). \quad (15)$$

We note that the RCN model reduces to the configuration model when no triangle stubs exist, i.e.,  $p_{s,t} = p_s$ ,  $q_{s,t} = (s+1)p_{s+1}/\langle s \rangle$ , and  $r_{s,t} = 0$ . The generating functions are then written as  $G_p(x, y) = G_{\text{tot}}(x) = G_0(x)$ ,  $G_q(x, y) = G_1(x)$ ,  $G_r(x, y) = 0$ ,  $B_0(x, y) = G_1(x)G_1(y)$ , and  $S_0(x) = G_1(x)$ , where  $G_0(x) = \sum_k p_k x^k$  and  $G_1(x) = G'_0(x)/G'_0(1)$ . It is easily confirmed from Eqs. (6) and (13) that the configuration model is unclustered and uncorrelated; i.e.,  $C_0 = 0$  and  $r_0 = 0$  [23].

Let us apply the above-mentioned formulations to two types of RCNs. The first example is the Poisson RCN, which has a double Poisson distribution

$$p_{s,t} = e^{-(s)} \frac{\langle s \rangle^s}{s!} e^{-(t)} \frac{\langle t \rangle^t}{t!}. \quad (16)$$

In this case,  $p_{s,t} = q_{s,t} = r_{s,t}$  and the generating functions are simplified to  $G_p(x, y) = G_q(x, y) = G_r(x, y) = e^{(s)(x-1)} e^{(t)(y-1)}$ . The clustering coefficient and the assortative coefficient are obtained from these generating functions as

$$C_0 = \frac{2\langle t \rangle}{2\langle t \rangle + (\langle s \rangle + 2\langle t \rangle)^2} = \frac{2\langle t \rangle}{2\langle t \rangle + \langle k \rangle^2} \quad (17)$$

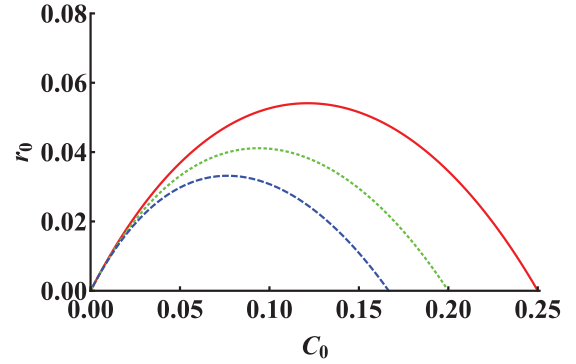


FIG. 2. Assortativity  $r_0$  of the Poisson RCN as a function of the clustering coefficient  $C_0$ . The solid red, dotted green, and dashed blue lines represent  $r_0$  for the cases of  $\langle k \rangle = 3, 4$ , and  $5$ , respectively.

and

$$\begin{aligned} r_0 &= \frac{2\langle s \rangle \langle t \rangle}{(\langle s \rangle + 2\langle t \rangle)^3 + 2\langle t \rangle (\langle s \rangle + 2\langle t \rangle)^2 + 2\langle s \rangle \langle t \rangle} \\ &= \frac{C_0(1 - C_0(1 + \langle k \rangle))}{1 - C_0(1 + (2C_0 - 1)\langle k \rangle)} \geq 0, \end{aligned} \quad (18)$$

respectively. Figure 2 plots  $r_0$  as a function of  $C_0$  for several values of  $\langle k \rangle$ , showing that the Poisson RCN has a weak assortativity in the sense that  $r_0$  takes a very small positive value when  $0 < C_0 < 1/(\langle k \rangle + 1)$ . It has been pointed out in [22,24] that nodes assigned many triangles in the RCN possibly have high degrees compared to those assigned few triangles, although the edge and triangle stubs are randomly connected to stubs of the same type; this bias causes a positive correlation of the nearest degrees.

Another example is the delta RCN, which has a double  $\delta$  function,

$$p_{s,t} = \delta_{s,s_0} \delta_{t,t_0}, \quad (19)$$

indicating that all nodes have  $s_0$  single edges and  $t_0$  triangles. One immediately finds that  $G_p(x, y) = x^{s_0} y^{t_0}$ ,  $G_q(x, y) = x^{s_0-1} y^{t_0}$ , and  $G_r(x, y) = x^{s_0} y^{t_0-1}$ ; thus,  $C_0 = 2t_0/(s_0 + 2t_0)(s_0 + 2t_0 - 1)$  and  $r_0 = 0$ . The delta RCN is clustered for  $t_0 > 0$ ; however, it has no degree correlation because all nodes have the same degree  $s_0 + 2t_0$ .

### III. STRUCTURE OF THE PERCOLATING CLUSTER IN SITE PERCOLATION

We consider site percolation on the RCNs: each node is occupied with probability  $f$  and is unoccupied (removed from the original network) otherwise. For site percolation on a network, the PC emerges at  $f = f_c$ , which is called the percolation threshold. The fraction  $S$  of nodes belonging to the PC becomes  $S > 0$  ( $S = 0$ ) when  $f > f_c$  ( $f \leq f_c$ ).

The analytical treatment for site percolation on the RCN is presented below. We denote by  $u$  the probability that a node reached by traversing a single edge chosen randomly from the original network is not a member of the PC and by  $v$  the probability that a node reached by traversing a triangle edge is not a member of the PC. Probabilities  $u$  and  $v$  are given as

the solution of the following self-consistent equations,

$$u = \tilde{f} + fG_q(u, v^2) \quad (20)$$

and

$$v^2 = (\tilde{f} + fG_r(u, v^2))^2, \quad (21)$$

where  $\tilde{f} = 1 - f$ . The normalized PC size,  $S$ , is given as the probability that a randomly chosen node is a member of the PC; hence, we have the following equation:

$$S = f \left( 1 - \sum_{s,t} p_{s,t} u^s v^{2t} \right) = f(1 - G_p(u, v^2)). \quad (22)$$

To find the percolation threshold  $f_c$ , we assume  $u = 1 - \epsilon_u$  and  $v^2 = 1 - \epsilon_v$  and consider the stability of the trivial solution  $(u, v) = (1, 1)$ . Expanding Eqs. (20) and (21) to leading order in  $\epsilon_u$  and  $\epsilon_v$  gives  $\epsilon = A\epsilon$ , where  $\epsilon = [\epsilon_u, \epsilon_v]^T$  and

$$A = f \begin{bmatrix} \sum_{s,t} sq_{s,t} & \sum_{s,t} tq_{s,t} \\ 2 \sum_{s,t} sr_{s,t} & 2 \sum_{s,t} tr_{s,t} \end{bmatrix}. \quad (23)$$

The percolation threshold  $f_c$  is then given from the condition  $\det |A - I| = 0$  ( $I$  is the identity matrix). The following two sections focus on the PC for  $f > f_c$ , where Eqs. (20) and (21) have a nontrivial solution of  $u$  and  $v$ , i.e.,  $S > 0$ , deriving its clustering coefficient and assortative coefficient.

### A. Clustering coefficient of the percolating cluster

We derive the clustering coefficient of the PC by starting with the conditional probability  $P(\text{PC}, s, t_1, t_2 | m, n)$  that a randomly chosen node belongs to the PC and has  $s$  single edges,  $t_1$  triangles with one removed, and  $t_2$  triangles (i.e., the node has  $t_1 + 2t_2$  triangle edges) in the PC, given that it has  $m$  single edges and  $n$  triangles in the original network. This probability is given as

$$P(\text{PC}, s, t_1, t_2 | m, n) = f \binom{m}{s} f^s \tilde{f}^{m-s} \binom{n}{t_2} f^{2t_2} \binom{n-t_2}{t_1} \\ \times (2f\tilde{f})^{t_1} \tilde{f}^{2(n-t_2-t_1)} (1 - \tilde{u}_f^s \tilde{v}_f^{2t_2+t_1}), \quad (24)$$

where

$$\tilde{u}_f = G_q(u, v^2) \quad \text{and} \quad \tilde{v}_f = G_r(u, v^2) \quad (25)$$

are the probability that the occupied node reached by traversing a single edge is not a member of the PC and the probability that the occupied node reached by traversing a triangle edge is not a member of the PC, respectively. Because the probability  $P(\text{PC}, s, t_1, t_2)$  that a randomly chosen node belongs to the PC and has  $s$  single edges,  $t_1$  triangles with one removed, and  $t_2$  triangles in the PC is

$$P(\text{PC}, s, t_1, t_2) = \sum_{m,n} P(\text{PC}, s, t_1, t_2 | m, n) p_{m,n} \\ = f \sum_{m,n} p_{m,n} \binom{m}{s} f^s \tilde{f}^{m-s} \binom{n}{t_2} f^{2t_2} \binom{n-t_2}{t_1} \\ \times (2f\tilde{f})^{t_1} \tilde{f}^{2(n-t_2-t_1)} (1 - \tilde{u}_f^s \tilde{v}_f^{2t_2+t_1}), \quad (26)$$

and the probability  $P(\text{PC})$  that a randomly chosen node belongs to the PC is

$$P(\text{PC}) = \sum_{s,t_1,t_2} P(\text{PC}, s, t_1, t_2) = f(1 - G_p(u, v^2)) = S, \quad (27)$$

we easily obtain the probability  $P_{\text{PC}}(s, t_1, t_2) \equiv P(s, t_1, t_2 | \text{PC})$  that a randomly chosen node has  $s$  single edges,  $t_1$  triangles with one removed, and  $t_2$  triangles conditioned on the node belonging to the PC, from  $P_{\text{PC}}(s, t_1, t_2) = P(\text{PC}, s, t_1, t_2) / P(\text{PC})$ .

Introducing the generating function  $F_{\text{PC}}(x, y, z)$  for  $P_{\text{PC}}(s, t_1, t_2)$  as

$$F_{\text{PC}}(x, y, z) = \sum_{s,t_1,t_2} P_{\text{PC}}(s, t_1, t_2) x^s y^{t_1} z^{t_2} \\ = \frac{1}{1 - G_p(u, v^2)} (G_p(fx + \tilde{f}, f^2z + 2f\tilde{f}y + \tilde{f}^2) \\ - G_p(f\tilde{u}_f x + \tilde{f}, f^2\tilde{v}_f^2 z + 2f\tilde{f}\tilde{v}_f y + \tilde{f}^2)), \quad (28)$$

we obtain the degree distribution  $P_{\text{PC}}(k)$  of the PC and the clustering coefficient  $C_{\text{PC}}$  of the PC as follows:

$$P_{\text{PC}}(k) = \frac{1}{k!} \frac{\partial^k}{\partial x^k} F_{\text{PC}}(x, x, x^2) \Big|_{x=0}, \quad (29)$$

and

$$C_{\text{PC}} = \frac{\partial}{\partial z} F_{\text{PC}}(x, y, z) \Big|_{x=y=z=1} \Big/ \frac{1}{2} \frac{\partial^2}{\partial x^2} F_{\text{PC}}(x, x, x^2) \Big|_{x=1}. \quad (30)$$

We note that for the case of  $p_{s,t} = p_s$ ,  $F_{\text{PC}}(x, y, z)$  is independent of  $y$  and  $z$ , so  $C_{\text{PC}} = 0$ . It means that the PC formed by site percolation on the configuration model is unclustered.

### B. Assortative coefficient of the percolating cluster

Next, we formalize the assortative coefficient of the PC. Our derivation is an extension of [20] in which the assortative coefficient of the PC formed by site percolation on uncorrelated networks was derived.

First, we consider the conditional probability  $Q_s(\text{PC}, s_1, t_1, s_2, t_2 | m_1, n_1, m_2, n_2)$  that a single edge belongs to the PC and its one end has  $s_1$  other single edges and  $t_1$  triangle edges and the other end has  $s_2$  other single edges and  $t_2$  triangle edges in the PC, given that the two ends of the selected single edge have  $m_1$  other single edges and  $n_1$  triangles and  $m_2$  other single edges and  $n_2$  triangles in the original network, respectively. This probability is written as follows:

$$Q_s(\text{PC}, s_1, t_1, s_2, t_2 | m_1, n_1, m_2, n_2) \\ = f^2 \binom{m_1}{s_1} f^{s_1} \tilde{f}^{m_1-s_1} \binom{2n_1}{t_1} f^{t_1} \tilde{f}^{2n_1-t_1} \binom{m_2}{s_2} \\ \times f^{s_2} \tilde{f}^{m_2-s_2} \binom{2n_2}{t_2} f^{t_2} \tilde{f}^{2n_2-t_2} (1 - \tilde{u}_f^{s_1+s_2} \tilde{v}_f^{t_1+t_2}). \quad (31)$$

Here,  $f^2$  in the right-hand side represents the probability that two ends of the focal edge are not removed and thus

the edge remains. The probability  $Q_s(\text{PC}, s_1, t_1, s_2, t_2)$  that a single edge belongs to the PC and its one end has  $s_1$  other

single edges and  $t_1$  triangle edges and the other end has  $s_2$  other single edges and  $t_2$  triangle edges in the PC is

$$Q_s(\text{PC}, s_1, t_1, s_2, t_2) = \sum_{m_1 \geq s_1} \sum_{m_2 \geq s_2} \sum_{n_1 \geq t_1} \sum_{n_2 \geq t_2} Q_s(m_1, n_1, m_2, n_2) Q_s(\text{PC}, s_1, t_1, s_2, t_2 | m_1, n_1, m_2, n_2). \tag{32}$$

Here,  $Q_s(m_1, n_1, m_2, n_2)$  is the probability that a single edge has two ends: one has  $m_1$  other single edges and  $n_1$  triangles and the other has  $m_2$  other single edges and  $n_2$  triangles in the original network. This probability is written as  $Q_s(m_1, n_1, m_2, n_2) = q_{m_1, n_1} q_{m_2, n_2}$  in that the RCN is a random network. The corresponding generating function for  $Q_s(\text{PC}, s_1, t_1, s_2, t_2)$  is as follows:

$$\begin{aligned} H_s(x_1, x_2, y_1, y_2) &= \sum_{s_1=0}^{\infty} \sum_{t_1=0}^{\infty} \sum_{s_2=0}^{\infty} \sum_{t_2=0}^{\infty} Q_s(\text{PC}, s_1, t_1, s_2, t_2) x_1^{s_1} x_2^{t_1} y_1^{s_2} y_2^{t_2} \\ &= f^2(\mathcal{G}_q(x_1, x_2) \mathcal{G}_q(y_1, y_2) - \mathcal{G}_q(\tilde{u}_f x_1, \tilde{v}_f x_2) \mathcal{G}_q(\tilde{u}_f y_1, \tilde{v}_f y_2)), \end{aligned} \tag{33}$$

where

$$\mathcal{G}_q(x, y) = G_q(\tilde{f} + fx, (\tilde{f} + fy)^2). \tag{34}$$

We further introduce the conditional probability  $Q_t(\text{PC}, s_1, t_1, s_2, t_2 | m_1, n_1, m_2, n_2)$  that a triangle edge belongs to the PC and one end of the edge has  $s_1$  single edges and  $t_1$  other triangle edges and the other end has  $s_2$  single edges and  $t_2$  other triangle edges in the PC, respectively, given that the two ends of the selected triangle edge have  $m_1$  single edges and  $n_1$  other triangles (triangles except the one including the selected edge) and  $m_2$  single edges and  $n_2$  other triangles in the original network, respectively, as

$$\begin{aligned} Q_t(\text{PC}, s_1, t_1, s_2, t_2 | m_1, n_1, m_2, n_2) &= f^2 \tilde{f} \binom{m_1}{s_1} f^{s_1} \tilde{f}^{m_1-s_1} \binom{2n_1}{t_1} f^{t_1} \tilde{f}^{2n_1-t_1} \binom{m_2}{s_2} f^{s_2} \tilde{f}^{m_2-s_2} \binom{2n_2}{t_2} f^{t_2} \tilde{f}^{2n_2-t_2} (1 - \tilde{u}_f^{s_1+s_2} \tilde{v}_f^{t_1+t_2}) \\ &+ f^3 \binom{m_1}{s_1} f^{s_1} \tilde{f}^{m_1-s_1} \binom{2n_1}{t_1-1} f^{t_1-1} \tilde{f}^{2n_1-t_1+1} \binom{m_2}{s_2} f^{s_2} \tilde{f}^{m_2-s_2} \binom{2n_2}{t_2-1} f^{t_2-1} \tilde{f}^{2n_2-t_2+1} (1 - \tilde{u}_f^{s_1+s_2} \tilde{v}_f^{t_1+t_2-1}). \end{aligned} \tag{35}$$

The two ends of a triangle edge have a common neighbor to form a triangle. The first and the second terms of the right-hand side are the contributions when this neighbor is unoccupied and occupied, respectively. The probability  $Q_t(\text{PC}, s_1, t_1, s_2, t_2)$  that a triangle edge belongs to the PC and its ends have  $s_1$  single edges and  $t_1$  other triangle edges and  $s_2$  single edges and  $t_2$  other triangle edges in the PC, respectively, is

$$Q_t(\text{PC}, s_1, t_1, s_2, t_2) = \sum_{m_1 \geq s_1} \sum_{m_2 \geq s_2} \sum_{n_1 \geq t_1} \sum_{n_2 \geq t_2} Q_t(m_1, n_1, m_2, n_2) Q_t(\text{PC}, s_1, t_1, s_2, t_2 | m_1, n_1, m_2, n_2), \tag{36}$$

where the probability  $Q_t(m_1, n_1, m_2, n_2)$  that the ends of a triangle edge have  $m_1$  single edges and  $n_1$  other triangles and  $m_2$  single edges and  $n_2$  other triangles in the original network, respectively, is  $Q_t(m_1, n_1, m_2, n_2) = r_{m_1, n_1} r_{m_2, n_2}$  for the RCN. The corresponding generating function for  $Q_t(\text{PC}, s_1, t_1, s_2, t_2)$  is as follows:

$$\begin{aligned} H_t(x_1, x_2, y_1, y_2) &= \sum_{s_1=0}^{\infty} \sum_{t_1=0}^{\infty} \sum_{s_2=0}^{\infty} \sum_{t_2=0}^{\infty} Q_t(\text{PC}, s_1, t_1, s_2, t_2) x_1^{s_1} x_2^{t_1} y_1^{s_2} y_2^{t_2} \\ &= f^2((\tilde{f} + fx_2 y_2) \mathcal{G}_r(x_1, x_2) \mathcal{G}_r(y_1, y_2) - (\tilde{f} + f\tilde{v}_f x_2 y_2) \mathcal{G}_r(\tilde{u}_f x_1, \tilde{v}_f x_2) \mathcal{G}_r(\tilde{u}_f y_1, \tilde{v}_f y_2)), \end{aligned} \tag{37}$$

where

$$\mathcal{G}_r(x, y) = G_r(\tilde{f} + fx, (\tilde{f} + fy)^2). \tag{38}$$

The probability  $Q(\text{PC}, s_1, t_1, s_2, t_2)$  that an edge belongs to the PC and the ends of the selected edge have  $s_1$  single edges and  $t_1$  triangle edges and  $s_2$  single edges and  $t_2$  triangle edges except the selected edge in the PC, respectively, is

$$Q(\text{PC}, s_1, t_1, s_2, t_2) = P_s Q_s(\text{PC}, s_1, t_1, s_2, t_2) + P_t Q_t(\text{PC}, s_1, t_1, s_2, t_2), \tag{39}$$

in that an edge chosen randomly from the original network is either of a single edge (with probability  $P_s$ ) or a triangle edge (with probability  $P_t$ ). The corresponding generating function is given as follows:

$$H(x_1, x_2, y_1, y_2) = \sum_{s_1=0}^{\infty} \sum_{t_1=0}^{\infty} \sum_{s_2=0}^{\infty} \sum_{t_2=0}^{\infty} Q(\text{PC}, s_1, t_1, s_2, t_2) x_1^{s_1} x_2^{t_1} y_1^{s_2} y_2^{t_2} = P_s H_s(x_1, x_2, y_1, y_2) + P_t H_t(x_1, x_2, y_1, y_2). \tag{40}$$

The corresponding generating function for the probability  $Q(\text{PC}, k_1, k_2)$  that an edge belongs to the PC and has two ends with degrees  $k_1 + 1$  and  $k_2 + 1$  in the PC is given as  $H(x, x, y, y)$ . The probability  $Q(\text{PC})$  that a randomly chosen edge belongs to a



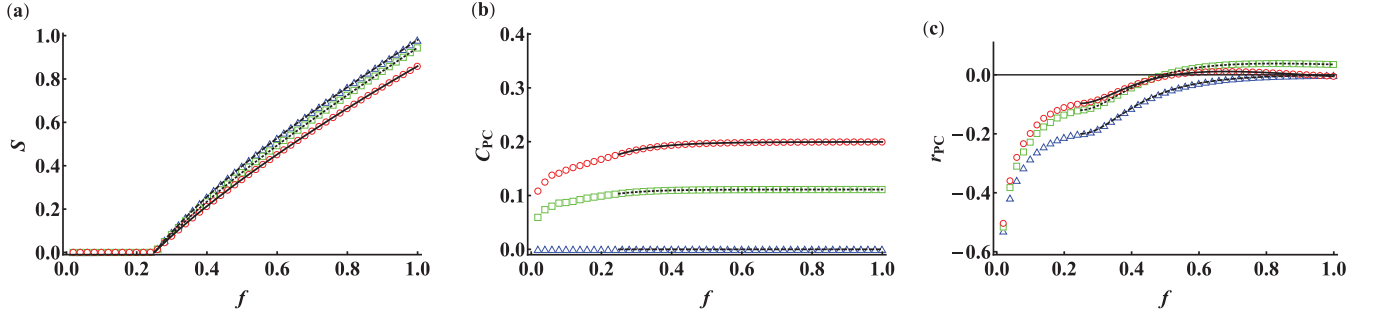


FIG. 3. Comparison with the analytical treatments and simulation results for the structures of the PC formed by site percolation: (a) normalized size  $S$ , (b) clustering coefficient  $C_{PC}$ , and (c) assortative coefficient  $r_{PC}$ . The Poisson RCN with  $\langle k \rangle = 4$  is utilized as the original network. The simulation results are for the cases of  $\langle s \rangle = 0$  and  $\langle t \rangle = 2$  (red circles),  $\langle s \rangle = 2$  and  $\langle t \rangle = 1$  (green squares), and  $\langle s \rangle = 4$  and  $\langle t \rangle = 0$  (blue triangles). The corresponding analytical estimates are represented by the solid lines, dotted lines, and dashed lines.

PC is given as follows:

$$Q(PC) = H(1, 1, 1, 1) = f^2(P_s(1 - \tilde{u}_f^2) + P_t(1 - (\tilde{f} + f\tilde{v}_f)\tilde{v}_f^2)). \quad (41)$$

Hence, we obtain the probability  $Q_{PC}(k_1, k_2) \equiv Q(k_1, k_2|PC)$  that an edge chosen from the PC has two ends with degree  $k_1 + 1$  and  $k_2 + 1$  in the PC from  $Q_{PC}(k_1, k_2) = Q(PC, k_1, k_2)/Q(PC)$ .

Using the generating function  $B_{PC}(x, y)$  for  $Q_{PC}(k_1, k_2)$  given as

$$B_{PC}(x, y) = \sum_{k_1 \geq 0} \sum_{k_2 \geq 0} Q_{PC}(k_1, k_2) x^{k_1} y^{k_2} = \frac{H(x, x, y, y)}{Q(PC)} = \frac{f^2 P_s}{Q(PC)} [\mathcal{G}_q(x, x) \mathcal{G}_q(y, y) - \mathcal{G}_q(\tilde{u}_f x, \tilde{v}_f x) \mathcal{G}_q(\tilde{u}_f y, \tilde{v}_f y)] \\ + \frac{f^2 P_t}{Q(PC)} [(\tilde{f} + fxy) \mathcal{G}_r(x, x) \mathcal{G}_r(y, y) - (\tilde{f} + f\tilde{v}_f xy) \mathcal{G}_r(\tilde{u}_f x, \tilde{v}_f x) \mathcal{G}_r(\tilde{u}_f y, \tilde{v}_f y)], \quad (42)$$

and the generating function  $S_{PC}(x) [= B_{PC}(x, 1) = B_{PC}(1, x)]$  for the probability  $Q_{PC}(k) = \sum_{k'} Q_{PC}(k, k')$  of an edge in the PC reaching a node with degree  $k + 1$ , given as

$$S_{PC}(x) = \sum_{k \geq 0} Q_{PC}(k) x^k = \frac{H(x, x, 1, 1)}{Q(PC)} \\ = \frac{f^2 P_s}{Q(PC)} [\mathcal{G}_q(x, x) - \tilde{u}_f \mathcal{G}_q(\tilde{u}_f x, \tilde{v}_f x)] + \frac{f^2 P_t}{Q(PC)} [(\tilde{f} + fx) \mathcal{G}_r(x, x) - (\tilde{f} + f\tilde{v}_f x) \tilde{v}_f \mathcal{G}_r(\tilde{u}_f x, \tilde{v}_f x)], \quad (43)$$

we obtain the assortative coefficient  $r_{PC}$  of the PC as

$$r_{PC} = \frac{\partial_x \partial_y B_{PC}(x, y) - (\partial_x S_{PC}(x))^2}{(x \partial_x)^2 S_{PC}(x) - (\partial_x S_{PC}(x))^2} \Big|_{x=y=1}. \quad (44)$$

This is an extension of the formulation for uncorrelated networks [20]. For the case of  $p_{s,t} = p_s$ , one finds, after tedious but simple algebra, that the above formulation yields the assortative coefficient of the PC on the uncorrelated networks:  $r_{PC}$  is given by Eq. (37) with  $B_{PC}(x, y) = [G_1(\tilde{f} + fx)G_1(\tilde{f} + fy) - G_1(\tilde{f} + f\tilde{u}_f x)G_1(\tilde{f} + f\tilde{u}_f y)]/(1 - \tilde{u}_f^2)$  and  $S_{PC}(x) = [G_1(\tilde{f} + fx) - \tilde{u}_f G_1(\tilde{f} + f\tilde{u}_f x)]/(1 - \tilde{u}_f^2)$ , where  $\tilde{u}_f = G_1(u)$  and  $u$  is the solution of  $u = \tilde{f} + fG_1(u)$ .

### C. Examples with numerical check

In this section, we applied our analysis to two RCNs, namely, the Poisson RCN and the delta RCN, discussing the structural properties of the PC formed by site percolation. Moreover, we performed Monte Carlo simulations to verify the validity of our analytical estimates. In our simulations, we generated ten network realizations consisting of  $N = 3 \times 10^6$  nodes and carried out the Newman-Ziff algorithm [25]

for site percolation  $10^3$  times on each realization. On each run, we specified the largest cluster corresponding to the PC for  $f > f_c$  and evaluated its size, clustering coefficient, and assortative coefficient to compare each average value with the corresponding analytical estimate.

Figures 3(a)–3(c) show the  $f$  dependence of the normalized PC size,  $S$ , the clustering coefficient of the PC,  $C_{PC}$ , and the assortative coefficient of the PC,  $r_{PC}$ , respectively, for site percolation on the Poisson RCN with  $\langle k \rangle = 4$  and several combinations of  $\langle s \rangle$  and  $\langle t \rangle$ . The solid, dotted, and dashed lines given only for  $f > f_c$  represent the analytical estimates and the symbols (red circles, green squares, and blue triangles) represent the Monte Carlo results. Our analytical estimates perfectly matched with the simulation results for  $f > f_c$  in all cases. As shown in Fig. 3(a), the PC on the Poisson RCN emerges at  $f_c = 1/4$ , irrespective of the value of  $C_0$  [26]. (Note that for bond percolation  $f_c$  depends on the value of  $C_0$  [27], as shown in Fig. 8(a).) Figures 3(b) and 3(c) show

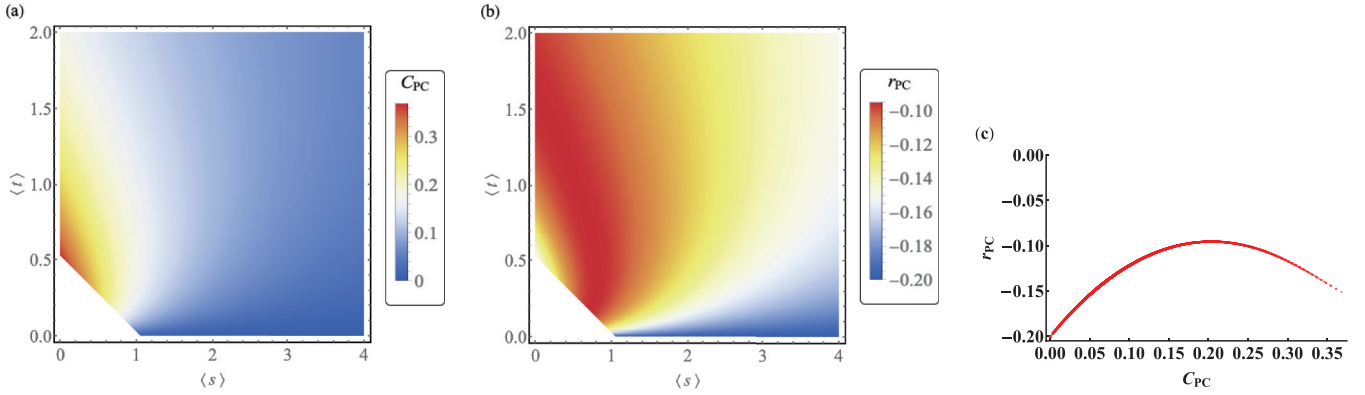


FIG. 4. Dependence of (a) clustering coefficient  $C_{PC}$  and (b) assortative coefficient  $r_{PC}$  of the fractal PC formed by site percolation on  $\langle s \rangle$  and  $\langle t \rangle$  of the Poisson RCN. Here  $C_{PC}$  and  $r_{PC}$  at  $f \simeq f_c$  are obtained from analytical estimates. Blank areas reflect the absence of the PC. (c) Scatter plot of  $C_{PC}$  (color-coded values in (a)) and  $r_{PC}$  (color-coded values in (b)) on the fractal PC.

that both the clustering coefficient  $C_{PC}$  and the assortative coefficient  $r_{PC}$  of the PC exhibit no singular behaviors at and around the percolation threshold  $f_c$ . Moreover, we notice that  $C_{PC} > 0$  at  $f = f_c$  (i.e., the PC is already highly clustered when it emerged), although it is a fractal (i.e., a fractal PC). For the case of  $\langle t \rangle = 0$ , the assortative coefficient,  $r_{PC}$ , of the PC is always negative [see the blue triangles in Fig. 3(c)] and takes  $-1/5$  at  $f = f_c$ , as already derived in [20]. The assortative coefficient of the PC becomes positive for a large  $f$  if the original RCN is assortative [see the red circles and green squares in Fig. 3(c)], although  $r_{PC}$  always becomes negative at and around  $f_c$ , irrespective of the assortativity of the original network. In Figs. 4(a) and 4(b), we plot the clustering coefficient and assortative coefficient of the fractal PC, i.e.,  $C_{PC}$  and  $r_{PC}$  at  $f = f_c$ , in the  $(\langle s \rangle, \langle t \rangle)$  plane. The clustering coefficient of the fractal PC becomes smaller at larger  $\langle s \rangle$  and larger  $\langle t \rangle$  [Fig. 4(a)]. Figures 4(b) and 4(c) clearly show that the assortative coefficient of the fractal PC is negative for any  $\langle s \rangle$  and  $\langle t \rangle$ . This supports the disassortativity of the fractal PC, according with our previous work [20] for uncorrelated networks.

The analysis of the RCN, however, does not necessarily give the disassortativity of fractal PCs. Figure 5 shows the results of  $S$ ,  $C_{PC}$ , and  $r_{PC}$ , for the delta RCN of  $s_0 + t_0 = 4$ . We again observe that  $C_{PC}$  and  $r_{PC}$  show no singular behaviors at  $f_c$  and  $C_{PC} > 0$  already at  $f = f_c$  (if  $C_0 > 0$ ). However,

both analytical estimate and simulations yield a different conclusion as regards the assortativity:  $r_{PC} > 0$  even at  $f_c$  if the delta RCN is clustered [see the red circles and green squares in Fig. 5(c)]. In Fig. 6, we calculate  $C_{PC}$  and  $r_{PC}$  at  $f = f_c$  for the delta RCNs with changing  $s_0$  and  $t_0$  in the range  $k = 3-12$ , and display their scatter plot. Figures 6(a) and 6(b) show that  $C_{PC}$  and  $r_{PC}$  of the fractal PC tend to decrease as degree  $k$  increases. Furthermore, in the delta RCNs with a fixed value of  $k$ ,  $C_{PC}$  and  $r_{PC}$  of the fractal PC decrease as the number of triangles per node  $t_0$  decreases. As the blue circles in Fig. 6(c) indicate, the fractal PC in the delta RCN can be assortative when it is highly clustered or, equivalently, the original delta RCN has small  $k$  and large  $C_0$ .

#### IV. DISCUSSION

In this study, we derived the clustering coefficient  $C_{PC}$  and assortative coefficient  $r_{PC}$  of the percolating cluster formed by site percolation on the random clustered network, thereby validating the disassortativity of fractal networks. Applying our formulation to the RCN whose joint probability of single edges and triangles obeys a double Poisson distribution (Poisson RCN) and the RCN whose nodes have the same numbers of single edges and triangles (delta RCN), we confirmed that our analytical estimates for  $C_{PC}$  and  $r_{PC}$  perfectly agree with the simulation results. Our results signified that both

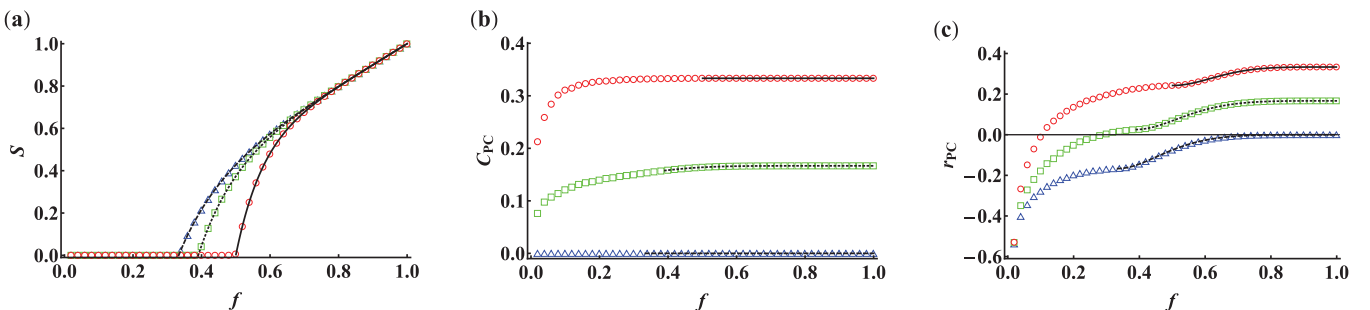


FIG. 5. Comparison with the analytical treatments and simulation results for the structures of the PC formed by site percolation: (a) normalized size  $S$ , (b) clustering coefficient  $C_{PC}$ , and (c) assortative coefficient  $r_{PC}$ . The delta RCN with degree 4 is utilized as the original network. The simulation results are for the cases of  $s_0 = 0$  and  $t_0 = 2$  (red circles),  $s_0 = 2$  and  $t_0 = 1$  (green squares), and  $s_0 = 4$  and  $t_0 = 0$  (blue triangles). The corresponding analytical estimates are represented by the solid lines, dotted lines, and dashed lines.

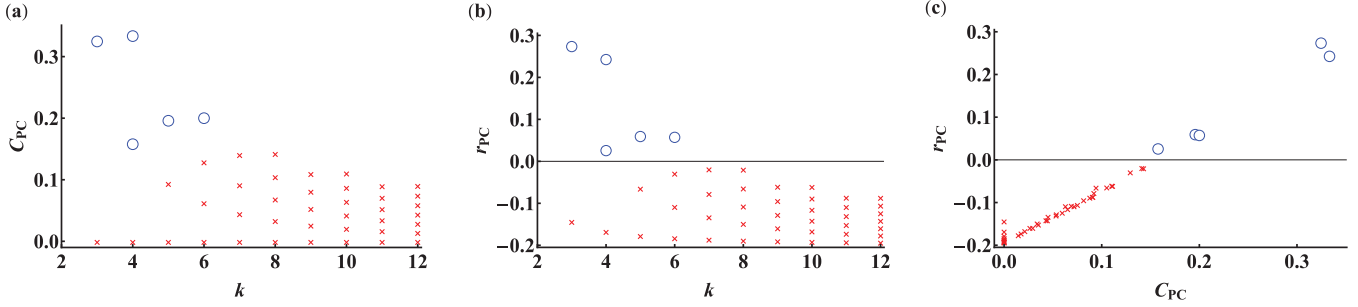


FIG. 6. Result for (a) clustering coefficient  $C_{PC}$  and (b) assortative coefficient  $r_{PC}$  of the fractal PC formed by site percolation on the delta RCN with degree  $k$  and (c) their scatter plot. Here  $C_{PC}$  and  $r_{PC}$  at  $f \simeq f_c$  are obtained from analytical estimates. All possible combinations of  $s_0$  and  $t_0$  for given degree  $k (= 3, 4, \dots, 12)$  are considered. The blue circles (red crosses) represent data indicating  $r_{PC} > 0$  ( $r_{PC} < 0$ ).

the clustering coefficient and the assortative coefficient of the PC do not exhibit any singular behavior near the percolation threshold, and the PC at the percolation threshold, namely, the fractal PC, is clustered as long as an underlying RCN is clustered. As regards the assortativity of the PC, the result seemingly contradicts the disassortativity of the fractal networks: the fractal PCs exhibit  $r_{PC} < 0$  for all Poisson RCNs and most delta RCNs, but  $r_{PC} > 0$  for only a few delta RCNs.

The question remains as regards to whether the last result immediately denies the disassortativity of the fractal networks. We should note that the positive assortativity of the delta RCN is easily lost. For example, we revisit site percolation on the delta RCN with  $p_{s,t} = \delta_{s,0}\delta_{t,2}$ , in which  $r_{PC} > 0$  for a fractal PC [red circles in Fig. 5(c)]. This network consists of only triangles; triangles are a basic unit giving a characteristic scale. Let us consider applying the box covering scheme [5,6] to a fractal PC formed on this network [Fig. 7(a)]. Tiling a fractal PC with the estimated minimum number of boxes of a linear size  $l_B = 4$  and renormalizing it so that each box is replaced as a supernode [Fig. 7(b)], we recalculated the clustering coefficient and the assortative coefficient of the renormalized ones. Renormalization breaks the characteristic scale (triangle) and unveils a disassortative structure: the clustering coefficient and the assortative coefficient are changed from  $C_{PC} \approx 0.333$  and  $r_{PC} \approx 0.245$  to  $C_{PC} \approx 0.112$  and  $r_{PC} \approx -0.129$ , respectively, under renormalization [28]. It indicates that the assortativity of a fractal

PC in the delta RCN is attributed to the characteristic scale of the triangles. The fractal PC formed by site percolation on the delta RCN appears disassortative for larger scales. As was shown in Fig. 6, the assortativity of the fractal PC is observed in only five combinations of  $s_0$  and  $t_0$ , i.e.,  $(s_0, t_0) = (1, 1), (2, 1), (0, 2), (1, 2), (0, 3)$ . In these combinations, every node has equally few edges, most of which form triangles. As with the present example ( $p_{s,t} = \delta_{s,0}\delta_{t,2}$ ), it is likely that the assortativity observed for delta RCNs is attributed to the triangles giving a characteristic scale and is easily broken by rescaling.

Moreover, the disassortativity of a fractal PC arises in the delta RCN when it is formed by bond percolation. Similar to site percolation, we can derive the clustering coefficient and the assortative coefficient of the PC formed by bond percolation on the RCN (see the Appendix). Figures 8 and 9 show the structural properties of the PC formed by bond percolation on the Poisson and delta RCNs, respectively. The triangles are easily broken in bond percolation in contrast to the formation of the PC: the clustering coefficient of the fractal PC is relatively small when compared with site percolation [Figs. 8(b) and 9(b)]. As indicated by both analytical estimates and simulation results,  $r_{PC}$  is negative at  $f = f_c$  for not only the Poisson RCN but also for the delta RCN [Figs. 8(c) and 9(c)]. The fractal PC is thus disassortative on the delta RCN when formed by bond percolation.

Having considered these results, it can be presumed that the fractal networks formed by percolation processes on networks are disassortative in essence. Further studies on the disassortativity of fractal networks should be conducted. This study has concentrated on the degree-degree correlation of networks, but has not asked for higher-order degree correlations, such as the long-range degree correlation [29–32]. Our formulation in this study may be extended to compute the long-range correlation, and it is expected to be provided by future work.

#### ACKNOWLEDGMENTS

S.M. and T.H. acknowledge financial support from JSPS (Japan) KAKENHI Grant No. JP18KT0059. S.M. was supported by a Grant-in-Aid for Early-Career Scientists (Grant No. 18K13473) and a Grant-in-Aid for JSPS Research Fellow (Grant No. 18J00527) from the Japan Society for the Promotion of Science (JSPS) for performing this work. T.H. acknowledges financial support from JSPS (Japan) KAKENHI Grants No. JP16H03939 and No. JP19K03648.

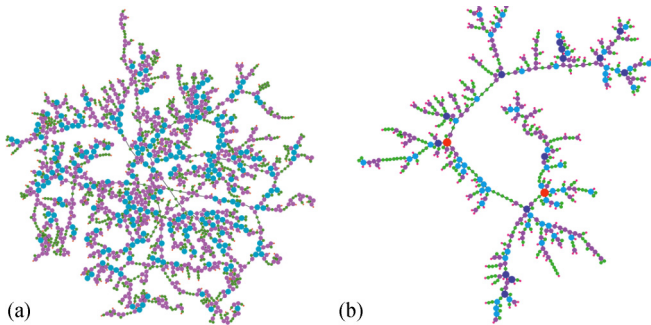


FIG. 7. Snapshot of (a) the largest cluster at the percolation threshold and (b) the renormalized one using the box counting scheme with  $l_B = 4$ . The numbers of nodes and triangles decrease from 1913 and 480 in (a) to 509 and 37 in (b), respectively. The assortative coefficient of the largest cluster is then changed from 0.215 to  $-0.126$ .



### APPENDIX: STRUCTURE OF THE PERCOLATING CLUSTER FORMED BY BOND PERCOLATION IN RANDOM CLUSTERED NETWORK

In this Appendix, we treat bond percolation on the RCN: each edge is open (not removed) with probability  $f$  and closed (removed from the original network) otherwise (probability  $\tilde{f} = 1 - f$ ). The normalized PC size,  $S$ , is given by solving the following equations [21]:

$$S = 1 - G_p(u, v^2), \quad (\text{A1})$$

with

$$u = \tilde{f} + fG_q(u, v^2) \quad (\text{A2})$$

and

$$\begin{aligned} v^2 &= 1 - 2f\tilde{f}^2(1 - G_r(u, v^2)) - f^2(3\tilde{f} + f)(1 - G_r(u, v^2))^2 \\ &= \tilde{f}^2 + 2\tilde{f}^2fG_r(u, v^2) + f^2(3\tilde{f} + f)G_r(u, v^2)^2. \end{aligned} \quad (\text{A3})$$

Here  $v^2$  is the probability that the two adjacent nodes forming a triangle with a node are not members of the PC. The percolation threshold  $f_c$  is given as the point above which  $u < 1$  and  $v < 1$  are the solution of the above-mentioned equations. By linearizing Eqs. (A2) and (A3) around  $(u, v) = (1, 1)$  and examining the condition  $\det|A - I| = 0$ , where

$$A = \begin{bmatrix} f \sum sq_{s,t} & f \sum tq_{s,t} \\ 2f(1 + f\tilde{f}) \sum sr_{s,t} & 2f(1 + f\tilde{f}) \sum tr_{s,t} \end{bmatrix} \quad (\text{A4})$$

and  $I$  is the identity matrix, we obtain the percolation threshold  $f_c$ . We hereafter assume  $f > f_c$ , focusing on the PC.

First, we derive the clustering coefficient of the PC. A randomly chosen node has  $m$  single edges and  $n$  triangles in an original network with probability  $p_{m,n}$ . We consider the probability  $P(\text{PC}, s, t_1, t_2, t_3, t_4)$  that a randomly chosen node belongs to the PC and has  $s$  single edges and  $t_1, t_2, t_3$ , and  $t_4$  motifs (shown in Fig. 10) in the PC. This probability is presented as follows:

$$P(\text{PC}, s, t_1, t_2, t_3, t_4) = \sum_{m,n} p_{m,n} \binom{m}{s} f^s \tilde{f}^{m-s} \binom{n}{t_1, t_2, t_3, t_4} f^{3t_4} (f^2 \tilde{f})^{t_3} (2f^2 \tilde{f})^{t_2} (2f \tilde{f}^2)^{t_1} \tilde{f}^{2(n-t_1-t_2-t_3-t_4)} (1 - \tilde{u}_f^s \tilde{v}_f^{t_1+2t_2+2t_3+2t_4}), \quad (\text{A5})$$

where  $\tilde{u}_f = G_q(u, v^2)$  and  $\tilde{v}_f = G_r(u, v^2)$ . We denote by  $F_{\text{PC}}(x, y_1, y_2, z_1, z_2)$  the generating function for the probability  $P_{\text{PC}}(s, t_1, t_2, t_3, t_4)$  that a node randomly chosen from the PC has  $s$  single edges and  $t_1, t_2, t_3$ , and  $t_4$  motifs in Fig. 10. Using  $P_{\text{PC}}(s, t_1, t_2, t_3, t_4) = P(\text{PC}, s, t_1, t_2, t_3, t_4)/P(\text{PC})$ , where  $P(\text{PC})$  is the probability that a randomly chosen node is a member of the PC,

$$P(\text{PC}) = 1 - G_p(u, v^2) = S, \quad (\text{A6})$$

and after some transformations we obtain

$$\begin{aligned} F_{\text{PC}}(x, y_1, y_2, z_1, z_2) &= \sum_{s=0}^{\infty} \sum_{t_1=0}^{\infty} \sum_{t_2=0}^{\infty} \sum_{t_3=0}^{\infty} \sum_{t_4=0}^{\infty} P_{\text{PC}}(s, t_1, t_2, t_3, t_4) x^s y_1^{t_1} y_2^{t_2} z_1^{t_3} z_2^{t_4} \\ &= \frac{1}{P(\text{PC})} G_p(\tilde{f} + fx, \tilde{f}^2 + 2\tilde{f}^2fy_1 + 2\tilde{f}f^2y_2 + \tilde{f}f^2z_1 + f^3z_2) \\ &\quad - \frac{1}{P(\text{PC})} G_p(\tilde{f} + f\tilde{u}_fx, \tilde{f}^2 + 2\tilde{f}^2f\tilde{v}_fy_1 + 2\tilde{f}f^2\tilde{v}_f^2y_2 + \tilde{f}f^2\tilde{v}_f^2z_1 + f^3\tilde{v}_f^2z_2). \end{aligned} \quad (\text{A7})$$

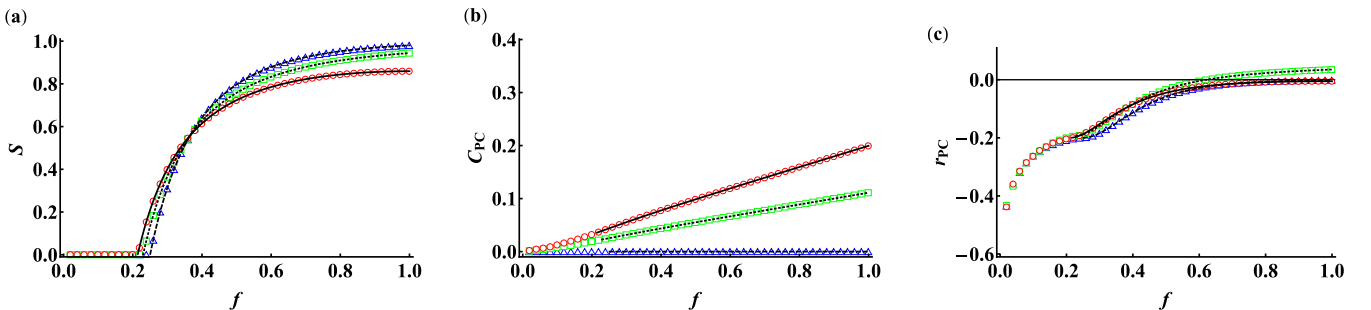


FIG. 8. Comparison with the analytical treatments and simulation results for the structures of the PC formed by bond percolation: (a) normalized size  $S$ , (b) clustering coefficient  $C_{\text{PC}}$ , and (c) assortative coefficient  $r_{\text{PC}}$ . The Poisson RCN with  $\langle k \rangle = 4$  is utilized as the original network. The simulation results are for the cases of  $\langle s \rangle = 0$  and  $\langle t \rangle = 2$  (red circles),  $\langle s \rangle = 2$  and  $\langle t \rangle = 1$  (green squares), and  $\langle s \rangle = 4$  and  $\langle t \rangle = 0$  (blue triangles). The corresponding analytical estimates, whose derivations are given in the Appendix, are represented by the solid lines, dotted lines, and dashed lines.

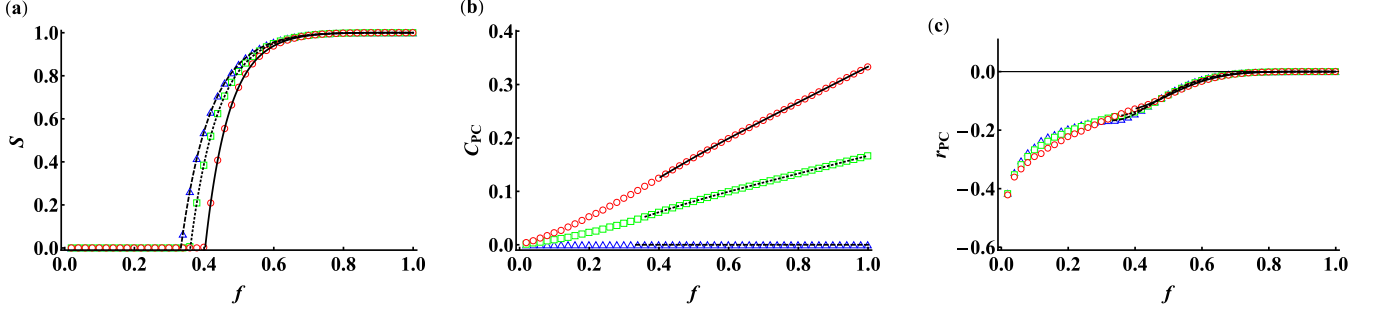


FIG. 9. Comparison with the analytical treatments and simulation results for the structures of the PC formed by bond percolation: (a) normalized size  $S$ , (b) clustering coefficient  $C_{PC}$ , and (c) assortative coefficient  $r_{PC}$ . The delta RCN with degree 4 is utilized as the original network. The simulation results are for the cases of  $s_0 = 0$  and  $t_0 = 2$  (red circles),  $s_0 = 2$  and  $t_0 = 1$  (green squares), and  $s_0 = 4$  and  $t_0 = 0$  (blue triangles). The corresponding analytical estimates, whose derivations are given in the Appendix, are represented by the solid lines, dotted lines, and dashed lines.

The degree distribution of the PC,  $P_{PC}(k)$ , is given as

$$P_{PC}(k) = \frac{1}{k!} \frac{\partial^k}{\partial x^k} F_{PC}(x, x, x, x^2, x^2) \Big|_{x=0}, \quad (\text{A8})$$

and the clustering coefficient of the PC,  $C_{PC}$ , is given as  $C_{PC} = 3N_{\Delta}/N_3$ , where

$$3N_{\Delta} = \frac{\partial}{\partial z_2} F_{PC}(x, y_1, y_2, z_1, z_2) \Big|_{x=y_1=y_2=z_1=z_2=1} \quad (\text{A9})$$

and

$$N_3 = \frac{1}{2} \frac{\partial^2}{\partial x^2} F_{PC}(x, x, x, x^2, x^2) \Big|_{x=1}. \quad (\text{A10})$$

Next, we formalize the assortative coefficient of the PC,  $r_{PC}$ . The derivation of  $r_{PC}$  for bond percolation is the same as that for site percolation in Sec. III B.

The generating function for  $Q(PC, s_1, t_1, s_2, t_2)$ , which is the probability that an edge belongs to the PC and its ends have  $s_1$  single edges and  $t_1$  triangle edges and  $s_2$  single edges and  $t_2$  triangle edges except the selected edge in the PC, respectively, is

$$H(x_1, x_2, y_1, y_2) = P_s H_s(x_1, x_2, y_1, y_2) + P_t H_t(x_1, x_2, y_1, y_2), \quad (\text{A11})$$

where

$$H_s(x_1, x_2, y_1, y_2) = f G_q(g_1(x_1), g_2(x_2)) G_q(g_1(y_1), g_2(y_2)) - f G_q(h_1(x_1), h_2(x_2)) G_q(h_1(y_1), h_2(y_2)) \quad (\text{A12})$$

and

$$H_t(x_1, x_2, y_1, y_2) = f(\tilde{f}^2 + \tilde{f}f(x_2 + y_2) + f^2 x_2 y_2) G_r(g_1(x_1), g_2(x_2)) G_r(g_1(y_1), g_2(y_2)) - f(\tilde{f}^2 + \tilde{f}f \tilde{v}_f(x_2 + y_2) + f^2 \tilde{v}_f x_2 y_2) G_r(h_1(x_1), h_2(x_2)) G_r(h_1(y_1), h_2(y_2)). \quad (\text{A13})$$

We used the following notations in Eqs. (A12) and (A13):

$$g_1(x) = \tilde{f} + fx, \quad g_2(x) = \tilde{f}^2 + 2\tilde{f}fx + f^2 x^2 \quad (\text{A14})$$

and

$$h_1(x) = \tilde{f} + f\tilde{u}_f x, \quad h_2(x) = \tilde{f}^2 + 2\tilde{f}^2 f \tilde{v}_f x + 2\tilde{f} f^2 \tilde{v}_f^2 x + f^2 \tilde{v}_f^2 x^2. \quad (\text{A15})$$

The generating function for the probability  $Q(PC, k, k')$  that an edge belongs to the PC and has two ends with degrees  $k + 1$  and  $k' + 1$  in the PC is then presented as  $\sum_{k, k'} Q(PC, k, k') x^k y^{k'} = H(x, x, y, y)$  and the generating function for the probability

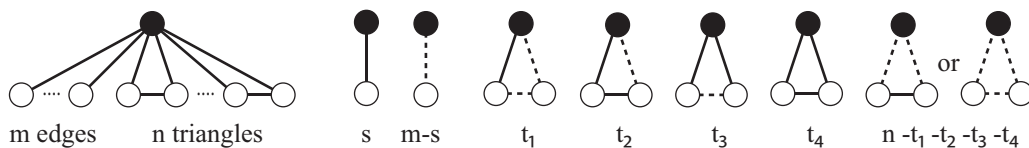


FIG. 10. Motifs of the single edges and triangles in a cluster formed by bond percolation. The solid and dashed lines represent the open and closed edges, respectively.

$Q(\text{PC}, k) = \sum_{k'} Q(\text{PC}, k, k')$  that an edge belongs to the PC and reaches a node with degree  $k + 1$  is given as  $\sum_k Q(\text{PC}, k)x^k = H(x, x, 1, 1)$ . The probability that a randomly chosen edge is open and belongs to the PC is

$$\begin{aligned} Q(\text{PC}) &= H(1, 1, 1, 1) \\ &= fP_s(1 - \tilde{u}_f^2) + fP_t(1 - (\tilde{f}^2 + f(2\tilde{f} + f)\tilde{v}_f)\tilde{v}_f^2). \end{aligned} \quad (\text{A16})$$

Then the assortative coefficient of PC,  $r_{\text{PC}}$ , is given as

$$r_{\text{PC}} = \left. \frac{\partial_x \partial_y B_{\text{PC}}(x, y) - (\partial_x S_{\text{PC}}(x))^2}{(x \partial_x)^2 S_{\text{PC}}(x) - (\partial_x S_{\text{PC}}(x))^2} \right|_{x=y=1}, \quad (\text{A17})$$

where  $B_{\text{PC}}(x, y)$  is the generating function for  $Q_{\text{PC}}(k, k') = Q(\text{PC}, k, k')/Q(\text{PC})$ ,

$$\begin{aligned} B_{\text{PC}}(x, y) &= \frac{H(x, x, y, y)}{Q(\text{PC})} \\ &= \frac{fP_s}{Q(\text{PC})} [G_q(g_1(x), g_2(x))G_q(g_1(y), g_2(y)) - G_q(h_1(x), h_2(x))G_q(h_1(y), h_2(y))] \\ &\quad + \frac{fP_t}{Q(\text{PC})} [(\tilde{f}^2 + \tilde{f}f(x+y) + f^2xy)G_r(g_1(x), g_2(x))G_r(g_1(y), g_2(y)) \\ &\quad - (\tilde{f}^2 + \tilde{f}f\tilde{v}_f(x+y) + f^2\tilde{v}_fxy)G_r(h_1(x), h_2(x))G_r(h_1(y), h_2(y))], \end{aligned} \quad (\text{A18})$$

and  $S_{\text{PC}}(x)$  is the generating function for  $Q_{\text{PC}}(k) = \sum_{k'} Q_{\text{PC}}(k, k')$ ,

$$\begin{aligned} S_{\text{PC}}(x) &= \frac{H(x, x, 1, 1)}{Q(\text{PC})} \\ &= \frac{fP_s}{Q(\text{PC})} [G_q(g_1(x), g_2(x)) - G_q(h_1(x), h_2(x))\tilde{u}_f] \\ &\quad + \frac{fP_t}{Q(\text{PC})} [(\tilde{f} + fx)G_r(g_1(x), g_2(x)) - (\tilde{f}^2 + \tilde{f}f\tilde{v}_f + f\tilde{v}_fx)G_r(h_1(x), h_2(x))\tilde{v}_f]. \end{aligned} \quad (\text{A19})$$

- 
- [1] A.-L. Barabási, *Network Science* (Cambridge University Press, Cambridge, UK, 2016).
- [2] M. Newman, *Networks* (Oxford University Press, Oxford, UK, 2018).
- [3] R. Pastor-Satorras, C. Castellano, P. Van Mieghem, and A. Vespignani, *Rev. Mod. Phys.* **87**, 925 (2015).
- [4] S. N. Dorogovtsev, A. V. Goltsev, and J. F. F. Mendes, *Rev. Mod. Phys.* **80**, 1275 (2008).
- [5] C. Song, S. Havlin, and H. A. Makse, *Nature (London)* **433**, 392 (2005).
- [6] C. Song, S. Havlin, and H. A. Makse, *Nat. Phys.* **2**, 275 (2006).
- [7] C. Song, L. K. Gallos, S. Havlin, and H. A. Makse, *J. Stat. Mech.: Theory Exp.* (2007) P03006.
- [8] S.-H. Yook, F. Radicchi, and H. Meyer-Ortmanns, *Phys. Rev. E* **72**, 045105(R) (2005).
- [9] H. D. Rozenfeld, S. Havlin, and D. Ben-Avraham, *New J. Phys.* **9**, 175 (2007).
- [10] J. S. Kim, B. Kahng, and D. Kim, *Phys. Rev. E* **79**, 067103 (2009).
- [11] I. Tishby, O. Biham, E. Katzav, and R. Kühn, *Phys. Rev. E* **97**, 042318 (2018).
- [12] P. Bialas and A. K. Oleś, *Phys. Rev. E* **77**, 036124 (2008).
- [13] K.-I. Goh, G. Salvi, B. Kahng, and D. Kim, *Phys. Rev. Lett.* **96**, 018701 (2006).
- [14] D.-H. Kim, J. D. Noh, and H. Jeong, *Phys. Rev. E* **70**, 046126 (2004).
- [15] Z.-W. Wei and B.-H. Wang, *Phys. Rev. E* **94**, 032309 (2016).
- [16] Y. Fujiki, S. Mizutaka, and K. Yakubo, *Eur. Phys. J. B* **90**, 126 (2017).
- [17] D. Stauffer and A. Aharony, *Introduction to Percolation Theory* (Taylor & Francis, Philadelphia, 1994).
- [18] I. Tishby, O. Biham, E. Katzav, and R. Kühn, *Phys. Rev. E* **99**, 042308 (2019).
- [19] I. Tishby, O. Biham, R. Kühn, and E. Katzav, *Phys. Rev. E* **98**, 062301 (2018).
- [20] S. Mizutaka and T. Hasegawa, *Phys. Rev. E* **98**, 062314 (2018).
- [21] M. E. J. Newman, *Phys. Rev. Lett.* **103**, 058701 (2009).
- [22] J. C. Miller, *Phys. Rev. E* **80**, 020901(R) (2009).
- [23] The configuration model can be highly clustered if the degree distribution is heavy tailed [2]. In the present study, we did not consider such a case.
- [24] J. P. Gleeson, S. Melnik, and A. Hackett, *Phys. Rev. E* **81**, 066114 (2010).
- [25] M. E. J. Newman and R. M. Ziff, *Phys. Rev. E* **64**, 016706 (2001).

- [26] For Poisson RCNs where  $G_p(x, y) = G_q(x, y) = G_r(x, y)$ , from  $\det |A - I| = 0$ ,  $f_c(\langle s \rangle + 2\langle t \rangle) = f_c \langle k \rangle = 1$  is obtained. Thus, the percolation threshold of site percolation is  $f_c = 1/\langle k \rangle$ , which depends only on the average degree  $\langle k \rangle$ .
- [27] Bond percolation on the RCNs has already been examined [21,22,24]. Newman [21] demonstrated that for the Poisson RCN, the clustering decreases the bond percolation threshold  $f_c$ . On the other hand, by comparing  $f_c$  for clustered networks and unclustered networks of the same degree correlation, Miller [22] clarified that this decrease in  $f_c$  is not caused by the influence of the clustering, but by the influence of the assortative correlation specific to the Poisson RCN. Furthermore, Gleeson *et al.* [24] arrived at the same conclusion by using their clustered  $\gamma$ -theory networks. Thus, the network clustering increases  $f_c$  for the delta RCN, which has no degree-degree correlation, although it decreases  $f_c$  for the Poisson RCN.
- [28] To take the average, we simulated site percolation of  $f = f_c$  ten times on each of the ten network realizations, picked up the largest clusters, and applied the box cover scheme of  $l_B = 4$  ten times to each largest cluster. The number of nodes in the delta RCN was 90 000. The average number of nodes in the largest clusters was changed from 2571 to 675 under renormalizations.
- [29] D. Rybski, H. D. Rozenfeld, and J. P. Kropp, *Europhys. Lett.* **90**, 28002 (2010).
- [30] Y. Fujiki, T. Takaguchi, and K. Yakubo, *Phys. Rev. E* **97**, 062308 (2018).
- [31] M. Mayo, A. Abdelzaher, and P. Ghosh, *Comput. Soc. Networks* **2**, 4 (2015).
- [32] C. Orsini, M. M. Dankulov, P. Colomer-de Simón, A. Jamakovic, P. Mahadevan, A. Vahdat, K. E. Bassler, Z. Toroczkai, M. Boguná, G. Caldarelli *et al.*, *Nat. Commun.* **6**, 8627 (2015).