Emergence of long-range correlations in random networks

Shogo Mizutaka and Takehisa Hasegawa

1 School of Knowledge Science, Japan Advanced Institute of Science and Technology, 1-1 Asahidai, Nomi 924-1292, Japan
2 Department of Mathematics and Informatics, Ibaraki University, 2-1-1 Bunkyo, Mito 310-8512, Japan
* Author to whom any correspondence should be addressed.
E-mail: mizutaka@jaist.ac.jp

Keywords: networks, random graphs, degree correlations, percolation

Abstract

We perform an analysis of the long-range degree correlation of the giant component (GC) in an uncorrelated random network by employing generating functions. By introducing a characteristic length, we find that a pair of nodes in the GC is negatively degree-correlated within the characteristic length and uncorrelated otherwise. At the critical point, where the GC becomes fractal, the characteristic length diverges and the negative long-range degree correlation emerges. We further propose a correlation function for degrees of two nodes separated by the shortest path length $l$, which behaves as an exponentially decreasing function of distance in the off-critical region. The correlation function obeys a power-law with an exponential cutoff near the critical point. The Erdős–Rényi random graph is employed to confirm this critical behavior.

1. Introduction

Most complex systems are described as networks comprising nodes and edges. Real network examples include cells, food webs, the internet, the World Wide Web (WWW), social relationships, and companies' transactions [1]. Such real networks exhibit common structural properties, namely degree correlation, clustering, clique, motif, community structure, core-periphery structure, scale-free property, small-world property, and fractality [1–3]. Network science poses the fundamental question of how these properties relate to each other [4–16].

In some networks, for example, small-world and fractal concept which are seemingly contradicting concepts [1–3]. Network science poses the fundamental question of how these properties relate to each other [4–16].

In general, the fractal objects have no characteristic lengths—their structures are invariant under a length-scale transformation [21]. A network is expected to have something invariant over a wide range of length scales when it is fractal. Applying a renormalization technique to scale-free fractal networks demonstrates that the profiles of the degree distributions are invariant under renormalization [20]. Negative degree correlation in scale-free fractal networks has been observed in various scales [22]. A scaling for degree correlations has been proposed under the assumption that the nearest-neighbor degree correlations of the fractal networks are invariant under renormalization [24]. Previous studies [20, 22–24] focusing on the structures and functions of the renormalized fractal networks have indicated that there is some correlation between its small- and large-scale network metrics, despite the difficulty in handling network renormalization.

With regard to local metrics of fractal networks, several reports addressed the correlation between the degrees of directly connected nodes by edges, i.e., the nearest-neighbor degree correlation. Nearest-neighbor
degree correlations are negative in various fractal networks, including some empirical networks [22], synthetic networks [25, 26], some trees [27, 28], uncorrelated network models in a critical state [29, 30], and percolating clusters of random networks [31] and clustered networks [32]. (The converse is not true: the nearest-neighbor degree correlations do not make networks fractal [25].) Large-scale correlation structures of the fractal networks should be reflected in the degree correlation between nodes beyond their nearest-neighbors, i.e., the long-range degree correlation (LRDC). Fujiki et al have introduced joint and associated conditional probabilities to analyze the LRDCs of networks [33]. They have shown that in the large size limit, an uncorrelated random network satisfies the relation \( P(k, k'|l) = q_k q_{k'} \), where \( P(k, k'|l) \) is the probability that two randomly selected nodes separated by distance \( l \) (two ends of a randomly selected \( l \)-chain) have degrees \( k \) and \( k' \), \( q_k = k p_k / \langle k \rangle \) is the probability that an end of a randomly selected edge has degree \( k \), \( p_k \) is a degree distribution, and \( \langle k \rangle = \sum_k k p_k \). A subsequent study pointed out that various networks, including fractal ones, exhibit LRDCs [34]. In [35], Rybski et al have numerically analyzed the LRDCs of the fractal networks described by the degree fluctuations in \( l \)-chains and indicated that the fractal networks have negative long-range correlations. However, previous studies were performed numerically, and there are no analytical arguments for LRDCs in the fractal networks. In this study, we focus on the giant component (GC) of an uncorrelated random network. By characterizing its LRDC as a function of degrees of a pair of nodes and their distance, we analytically derive the emergence of the negative LRDC in the GC at a critical state.

2. Results

Let us consider an infinitely-large uncorrelated random network with a degree distribution \( p_k \) such that it is locally tree-like and sparse (its average degree is \( O(1) \)). The probability, \( u \), that an edge does not lead to the GC is given as the solution of \( u = G_1(u) \), where \( G_1(u) = \sum_k q_k u^{k-1} \), and it is less than 1 in the presence of the GC. Assuming that a given network contains the GC, i.e., \( u < 1 \), we extract it from this network. The structure of the GC is different from that of the whole uncorrelated network satisfying \( P(k, k'|l) = q_k q_{k'} \), e.g., the GC extracted from an uncorrelated network always takes a negative nearest-neighbor degree correlation unless the network is singly connected (\( u = 0 \)). We start our analysis by introducing probability \( P_{GC}(k, k'|l) \), stating that two ends of a randomly selected \( l \)-chain have degrees \( k \) and \( k' \), given that the chain belongs to the GC. Using the expectation number of nodes with degree \( k' \) at distance \( l \) from a degree-\( k \) node and the expectation number of nodes at distance \( l \) from a random node [see equations (A.7) and (A.10)], we obtain

\[
P_{GC}(k, k'|l) = \frac{1 - v l^{-1} u^{k+k-2} - q_k q_{k'}}{1 - v l^{-1} u^2 - q_k q_{k'}},
\]

(1)

where \( v \) is the probability that a node in between the \( l \)-chain does not reach the GC by traversing any of the edges other than the two that constitute the chain and can be expressed as a function of \( u \), as \( v = G_1(u)/G_1'(1) \) with \( G_1'(u) = dG_1(u)/du = \sum_k (k-1) q_k u^{k-2} / \langle k \rangle \). For \( l = 1 \), it has been reported that \( P_{GC}(k, k'|l = 1) = (1 - u^{k+k-2}) q_k q_{k'} / (1 - u^2) \), which depicts the joint probability that an edge selected randomly from the GC is connected to degree-\( k \) and -\( k' \) nodes [29, 30].

We introduce the characteristic lengths associated with the distance and the degrees of a node pair as

\[
\xi_l = -1 / \log v
\]

(2)

and

\[
\xi_k = -1 / \log u,
\]

(3)

respectively. Consequently, equation (1) is rewritten as

\[
P_{GC}(k, k'|l) = \frac{1 - e^{-(l-1)/\xi_l} e^{-(k+k'-2)/\xi_k} q_k q_{k'}}{1 - e^{-(l-1)/\xi_l} e^{-k/\xi_k} q_k q_{k'}},
\]

(4)

For a finite \( \xi_l, P_{GC}(k, k'|l) \) decays exponentially with increasing \( l \) and converges to \( q_k q_{k'} \) when \( l \gg 1 \). It means that two ends of \( l \)-chains in the GC are degree-uncorrelated for \( l \gg \xi_l \) and degree-correlated for \( l \lesssim \xi_l \).

We further introduce the probability, \( P_{GC}(k'|k, l) \), that one end of a chain has degree \( k' \), given that the chain has a degree-\( k \) node at the other end, has length \( l \), and belongs to the GC. From Bayes’ rule, this probability is given as

\[
P_{GC}(k'|k, l) = \frac{P_{GC}(k, k'|l)}{\sum_{k'} P_{GC}(k, k'|l)}
\]

\[
= \frac{1 - e^{-(l-1)/\xi_l} e^{-(k+k'-2)/\xi_k} q_k q_{k'}}{1 - e^{-(l-1)/\xi_l} e^{-k/\xi_k} q_k q_{k'}},
\]

(5)
The degree distribution is valid of the theoretical analysis (1) by comparing it with the simulation results. In figure 1, we plot the theoretical predictions (lines) of probability and the average degree, corresponding Monte-Carlo simulations (symbols) perfectly.

The critical exponents for \( k \) term of \( k^2 \) where \( \lambda \) and \( \epsilon \) we have \( k^2 \) for a fixed value of \( \lambda \) and \( \epsilon \) is the same as (a). Top panels represent the results for the average degree, \( \lambda = 1 \), and distances (a) \( l = 1 \), (b) \( l = 5 \), and (c) \( l = 10 \); bottom panels depict the average degree, \( \lambda = 2.5 \), and distance (d) \( l = 1 \), (e) \( l = 5 \), and (f) \( l = 10 \). The results for each average degree are obtained from one sampled network of \( 10^7 \) nodes.

For the Erdös-Rényi random graphs as a function of \( k \) and several distances. Lines depict the analytical calculation (1), and symbols represent the corresponding simulation results. In panel (a), solid, dashed, long-dashed, dash-dotted, and dash-double-dotted lines (squares, circles, triangles, inverted triangles, and diamonds) represent analytical results (simulation results) for \( k' = 1, 2, 3, 4, \) and 5, respectively. The legends in (b), (c), . . . , and (f) are the same as (a). Top panels represent the results for the average degree, \( \lambda = 1 \), and distances (a) \( l = 1 \), (b) \( l = 5 \), and (c) \( l = 10 \); bottom panels depict the average degree, \( \lambda = 2.5 \), and distance (d) \( l = 1 \), (e) \( l = 5 \), and (f) \( l = 10 \). The results for each average degree are obtained from one sampled network of \( 10^7 \) nodes.

and the average, \( k^{GC}_l(k) \), of \( l \)-distant nodes from degree-\( k \) nodes on the GC is given as

\[
k^{GC}_l(k) = \sum_{k'} k' P_{GC}(k' | k, l) = \langle k^2 \rangle \frac{h(u)}{k} + \frac{h(u) e^{-(u-1)/\xi_l} \epsilon^{-k/\xi_k}}{1 - e^{-(u-1)/\xi_l} e^{-k/\xi_k}},
\]

where \( h(u) = \sum_{k} k P_{GC}(k | k, l) \geq 0 \). Note that \( k^{GC}_l(k) \) of \( l = 1 \) corresponds to the average degree of the nearest-neighbor degree-\( k \) nodes on the GC \([29, 30]\). Equation (6) shows that \( k^{GC}_l(k) \) is a decreasing function of \( k' \) for a fixed value of \( l \), indicating that the GC is negatively degree-correlated. Moreover, \( k^{GC}_l(k) \) is a decreasing function of \( l \) for any \( k \), indicating that any degree correlation gradually disappears with increasing \( l \) (as in equation (4)). In summary, the GC in a random network has a negative degree correlation for \( l < \xi_l \), whereas it has no correlations for \( l \gg \xi_l \).

To further discuss the emergence of the LRDC of the GC in detail, we employ the Erdös–Rényi random graph, whose degree distribution is \( p_k = \lambda^k e^{-\lambda} / k! \), where \( \lambda = \langle k \rangle \). Prior to a detailed analysis, we test the validity of the theoretical analysis (1) by comparing it with the simulation results. In figure 1, we plot the theoretical predictions (lines) of probability \( P_{GC}(k, k' | l) \) for the Erdös–Rényi random graphs with \( \lambda = 1 (\lambda \epsilon) \) and \( \lambda = 2.5 \), where \( \epsilon \) is the critical average degree above which the GC exists. The lines in all cases match the corresponding Monte-Carlo simulations (symbols) perfectly.

We assume that \( \lambda = \lambda_c + \delta \) and \( u = v = 1 - \epsilon \), where both \( \delta \) and \( \epsilon \) are infinitesimally small. For \( \lambda \geq \lambda_c \), we have \( \epsilon \sim \lambda - \lambda_c \) and two characteristic lengths, \( \xi_l \) and \( \xi_k \), as

\[
\xi_l = \xi_k \sim \epsilon^{-1} \sim (\lambda - \lambda_c)^{-1}.
\]

The critical exponents for \( \xi_l \) and \( \xi_k \) are unity, which corresponds to the critical exponent of the correlation (chemical) length for the mean size of the finite cluster in the percolation problem \([36]\). For \( \lambda \geq \lambda_c \), the second term of \( k^{GC}_l(k) \) becomes a power-law with an exponential cutoff of both \( l \) and \( k \) within \( \xi_l \) and \( \xi_k \) as

\[
k^{GC}_l(k) = 2 + \frac{1}{(l - 1) + k} e^{-(l-1)/\xi_l} e^{-(k-1)/\xi_k}.
\]
The two characteristic lengths in equation (8) diverge asymptotically in a critical state \( (\lambda \to \lambda_c) \). At \( \lambda = \lambda_c \), \( k_{\text{GC}}^l(k) \) decreases with increasing degree \( k \) in a power-law for any \( l < \infty \):

\[
k_{\text{GC}}^l(k) = 2 + (l + k - 1)^{-1}.
\]

Hence, the negative long-range correlation in the GC stretches entirely at criticality.

Furthermore, we propose a degree-degree correlation function \( C(l) \), which characterizes the critical behavior of the networks. Probability \( P_{\text{GC}}(k, k'|l) \) has full information on the structure of the GC. We define the correlation function for degrees of pairs of nodes separated by the shortest path length \( l \) (\( l \)-distant node pairs) on the GC, as:

\[
C(l) = \langle kk' \rangle_l - \langle k \rangle_l \langle k' \rangle_l,
\]

where \( \langle f(k, k') \rangle_l = \sum_{k,l} f(k,k') P_{\text{GC}}(k,k'|l) \). Combined with equation (4), \( C(l) \) is expressed as

\[
C(l) = -e^{-(l-1)/\xi u} \left( \sum_k k^e (1-u^{k-2}) \right)^2 / (1-e^{-(l-1)/\xi u})^2.
\]

We observe that the degree correlation of \( l \)-distant node pairs on the GC disappears for \( l > \xi \), as \( |C(l)| \) is an exponentially decreasing function of \( l \). The correlation function exhibits critical behavior when the GC exists but very small, i.e., \( \xi_l \gg 1 \): \( C(l) \) in the critical region, and drops according to a power-law with an exponential cutoff,

\[
C(l) \sim -\frac{a}{(b(1-1)+2)^2} e^{-(l-1)/\xi_l},
\]

where \( a = (k^2(k-2))/(k) \) and \( b = (k(k-1)(k-2))/(k) \). Just at the critical point, \( \xi_l \) diverges, and \( C(l) \sim -l^{-2} \) for \( l \gg 1 \). Figure 2 shows the absolute value \( |C(l)| \) of the correlation function for Erdős-Rényi random graphs for several values of \( \lambda \). We observe that the correlation function decays exponentially in the off-critical region (\( \lambda \gg 1.1 \)), and a power-law with exponent \(-2\) exists near criticality (\( \lambda = 1.001 \approx \lambda_c \)).

3. Conclusion & discussion

All our analyses conclude that the LRDC in the GC of an uncorrelated random network emerges at the critical point. The correlated structure is described by \( P_{\text{GC}}(k,k'|l) \) and \( k_{\text{GC}}^l(k) \). The GC is negatively correlated for \( l < \xi_l \) whereas it becomes neutral for \( l \gg \xi_l \). At criticality, where \( \xi_l \) diverges and the GC is fractal, the negative degree correlation is observed at any distance. Moreover, the correlation function \( C(l) \) for degrees of \( l \)-distant node pairs decays exponentially in the off-critical region. In contrast, it obeys a power-law with a cutoff, \( C(l) \sim -l^{-2} e^{l/\xi_l} \) for \( l \gg 1 \) near criticality and becomes a power-law, \( C(l) \sim -l^{-2} \) at criticality. In summary, the negative LRDC spontaneously emerges in the fractal networks.

3 The present analysis is restricted to random networks with a finite third moment \( \langle k^3 \rangle \) of the degree distribution, although we have not explicitly mentioned this assumption. The long-range degree correlations for the case of infinitely large \( \langle k^3 \rangle \) must be discussed in the configuration network with \( P(k) \sim k^{-\gamma} \), as in [37], or the percolation problem on a random scale-free network, where a node (or an edge) is retained with a probability or removed.
It should be mentioned that the LRDC of the GC in the present study is an intrinsic one. Fujiiki and Yakubo have recently introduced intrinsic and extrinsic LRDCs [34]. Imposing a (strong) nearest-neighbor degree correlation on a network induces a kind of LRDCs, which is called extrinsic LRDC\(^4\). On one hand, empirical networks are organized by various complex factors and their LRDCs would not be often simple as extrinsic ones. Such correlations are called intrinsic LRDC. In [34], they introduced a method to distinguish these two LRDCs, discussing the finite size effects. It can be presumed that the LRDC of a network is extrinsic if its triplet probability, \(P(k', k''|k)\), that a degree-\(k\) node is connected to a degree-\(k'\) node and a degree-\(k''\) node matches the product of \(P(k'|k)\) and \(P(k''|k)\). With regard to the GC in an uncorrelated random network, the LRDC is considered intrinsic in that the triplet probability \(P_{GC}(k', k''|k)\) indicates \(P_{GC}(k', k''|k) = (1 - u^{k+k'+k''-4})q_k^2q_k^2/(1 - u^k) \neq P_{GC}(k'|k)P_{GC}(k''|k)\) within \(\xi_k\).

Bialas and Oleš have investigated the correlation function for generic trees [28]. Their correlation function (equation (12) in [28]) behaves as a power-law, which is similar to \(C(l)\) in the present study. Our LRDC of the GC is attributed to the divergence of the correlation length in percolation transition while the LRDC in the generic trees is not associated with the transition. The tree structures found in both generic trees and the GC at criticality may result in a power-law behavior of \(C(l)\), although further studies are required to gain an understanding of this common mechanism.

Previous works have captured nearest-neighbor degree correlations of fractal networks which are renormalized at several length scales, implying a correlation between small- and large-scale degree correlations [22, 24]. We did not attain the problem how the LRDC treated in this study associates with the nearest-neighbor degree correlations of renormalized networks. This study deepens our understanding of the relation between the emergence of the LRDC and criticality/fractality.

Data availability statement

All data that support the findings of this study are included within the article (and any supplementary information files).

Acknowledgments

The authors would like to thank K Yakubo for fruitful discussions. SM. and TH acknowledge the financial support from JSPS (Japan) KAKENHI Grant Number JP18K03648. SM was supported by a Grant-in-aid for JSPS Research Fellow (No. 18J00527) from the Japan Society for the Promotion of Science (JSPS). TH acknowledges the financial support from JSPS (Japan) KAKENHI Grant Number JP18K13473.

Appendix A. Derivation of \(P_{GC}(k_s, k_l)\)

Let us consider an infinitely-large uncorrelated random network with a locally tree-like structure and focus on \(l\)-distant nodes, which are nodes at distance \(l\) from a randomly chosen node (seed) (see Fig. A1). We suppose that for a randomly chosen seed the numbers of \(l\)-distant nodes with degrees \(1, 2, \ldots, k, \ldots\), and \(K\) are \(n_1, n_2, \ldots, n_k, \ldots, \) and \(n_K\), respectively, where \(K\) is the maximum degree. We denote the sequence of the number of \(l\)-distant nodes with each degree by \(n = (n_1, n_2, \ldots, n_K)\). Let \(P_l(n, GC|N_l)\) be the probability that the sequence for the \(l\)-distant nodes is \(n\) and the component is the GC under the condition that the total number of \(l\)-distant nodes \(N_l(= \sum_{k=1}^{K} n_k)\) is given. Using the multinomial distribution, \(P_l(n, GC|N_l)\) is given as

\[
P_l(n, GC|N_l) = \frac{N_l!}{n_1!n_2! \cdots n_K!} \left( 1 - \prod_{k=1}^{K} (u^{k-1})^{n_k} \right)^{K} \prod_{k=1}^{K} q_{n_k}^{n_k}, \tag{A.1}
\]

where \(u\) is the probability that an edge does not lead to the GC, \(p_k\) is the degree distribution, and \(q_{nk} = \sum_k kp_k/k\) is the neighbor’s degree distribution. The term \(1 - \prod_{k=1}^{K} (u^{k-1})^{n_k}\) corresponds with the probability that at least one of outgoing edges from \(l\)-distant nodes connects to the GC. By combining \(P_l(n, GC|N_l)\) and \(P_l(N_l|k_l)\), which is the probability that the number of \(l\)-distant nodes is \(N_l\) given that the seed has \(k_l\) edges, we construct the probability \(P_l(n, GC|k_l)\) that the sequence for the \(l\)-distant nodes is \(n\) and the component is

\[^4\] For example, when a given network has only a strong negative nearest-neighbor degree correlation on \(P(k, k')\), it will exhibit an extrinsic LRDC, such that the degrees of node pairs are positively correlated at \(l = 2, 4, \ldots\)
Here, the GC given that the seed has \( k_s \) edges as

\[
P_l(n, GC|k_s) = \sum_{N_l} P_l(N_l|k_s)P_l(n, GC|N_l)
\]

\[
= \sum_{N_l} P_l(N_l|k_s) \frac{N_l!}{n_1!n_2! \cdots n_k!} \left( 1 - \prod_{k=1}^{K} (u^{-1})^{n_k} \right) \prod_{k=1}^{K} q_k^{n_k}. \tag{A.2}
\]

We introduce the generating function \( \tilde{F}_{GC}(x|k_s) \) for \( P_l(n, GC|k_s) \) as

\[
\tilde{F}_{GC}(x|k_s) = \sum_n P_l(n, GC|k_s) \prod_{k=1}^{K} x_k^{n_k}, \quad \text{where} \quad x = (x_1, x_2, \ldots, x_K).
\]

This function is calculated as follows:

\[
\tilde{F}_{GC}(x|k_s) = \sum_n P_l(n, GC|k_s) \frac{N_l!}{n_1!n_2! \cdots n_k!} \prod_{k=1}^{K} q_k^{n_k} \prod_{k=1}^{K} (1 - (u^{-1})^{n_k}) \prod_{k=1}^{K} q_k^{n_k}.
\]

\[
= \sum_{N_l} \sum_n P_l(n, GC|k_s) \frac{N_l!}{n_1!n_2! \cdots n_k!} \left( \prod_{k=1}^{K} (q_k x_k)^{n_k} - \prod_{k=1}^{K} (q_k x_k u^{-1})^{n_k} \right) \prod_{k=1}^{K} q_k^{n_k}.
\]

\[
= \tilde{G}_l \left( \sum_k q_k x_k |k_s \right) - \tilde{G}_l \left( \sum_k q_k x_k u^{-1} |k_s \right). \tag{A.4}
\]

Here, \( \tilde{G}_l(x|k_s) \) is the generating function for the probability distribution \( P_l(n|k_s) \) and is given as

\[
\tilde{G}_l(x|k_s) = (G_l(G_l(\cdots (G_l(x)) \cdots)))^{k_s}, \tag{A.5}
\]

which is known as the generating function for the distribution of the number of children in \( l \) generation under a given offspring distribution \( q_k \) and initial population \( k_s \) \[38\]. The probability distribution \( P_l(n, GC|k_s) \) that the sequence for the \( l \)-distant nodes is \( n \), the seed has \( k_s \) edges, and they belong to the GC is \( P_l(n, GC|k_s) = P_l(n, GC|k_s)p_s \), and the corresponding generating function \( F_{GC}(x, k_s) = \sum_n P_l(n, GC|k_s) \prod_{k=1}^{K} x_k^{n_k} \) is thus given by the product of \( \tilde{F}_{GC}(x|k_s) \) and \( p_s \),

\[
F_{GC}(x, k_s) = p_s \tilde{F}_{GC}(x|k_s)
\]

\[
= p_s \tilde{G}_l \left( \sum_k q_k x_k |k_s \right) - \tilde{G}_l \left( \sum_k q_k x_k u^{-1} |k_s \right). \tag{A.6}
\]
Differentiating equation (A.6) with respect to $\times_k$ and substituting $x = 1$ into it, we have the expectation number $\langle N^{k_k}(l, GC) \rangle$ of $l$-distant nodes with degree $k_i$ where a randomly chosen seed has $k_i$ edges and its component belongs to the GC:

$$\langle N^{k_k}(l, GC) \rangle = \frac{\partial F_{GC}(x, k_i)}{\partial x_k} \bigg|_{x=1} = k_p q_k G_1^2(1) \left(1 - v^{l-1} u^{k_i + k_i - 2}\right),$$

(A.7)

where $v = G_1(u)/G_1(1)$.

In a similar way, we easily find the generating function $F_{GC}(x)$ for the probability distribution $P_{l}(n, GC)$ that for a randomly chosen node, the sequence of the number of $l$-distant nodes with each degree is $n$ and these nodes are member of the GC as

$$F_{GC}(x) = \sum_{k_i=1}^{K} F_{GC}(x, k_i) = G_1 \left(\sum_k q_k x_k\right) - G_1 \left(\sum_k q_k x_k u^{k_i - 1}\right),$$

(A.8)

where

$$G_1(x) = \sum_k p_k G_1(x|k) = G_0(G_1(\cdots(G_1(x)\cdots))).$$

(A.9)

From equation (A.8), the expectation number $\langle N(l, GC) \rangle$ of $l$-distant nodes from a random seed which belong to the GC is calculated as

$$\langle N(l, GC) \rangle = \sum_{k_i=1}^{K} \frac{\partial F_{GC}(x)}{\partial x_k} \bigg|_{x=1} = \langle k \rangle G_1^2(1) \left(1 - v^{l-1} u^2\right)$$

(A.10)

By dividing equation (A.7) by equation (A.10), we obtain the probability $P_{GC}(k_s, k_i|l)$ that two ends of a randomly chosen $l$-chain from the GC have degree $k_s$ and $k_i$ as

$$P_{GC}(k_s, k_i|l) = \frac{\langle N^{k_s,k_i}(l, GC) \rangle}{\langle N(l, GC) \rangle},$$

$$= \frac{G_1^2(1) \left(1 - v^{l-1} u^{k_i + k_i - 2}\right)}{G_1^2(1) \left(1 - v^{l-1} u^2\right)} q_k q_k,$$

$$= \frac{1 - v^{l-1} u^{k_i + k_i - 2}}{1 - v^{l-1} u^2} q_k q_k,$$

(A.11)

where we call a connected path with length $l$ as an $l$-chain. Here, the denominator is proportional to the number of $l$-chains in the GC in that $\langle k \rangle G_1(1)^{l-1}$ $\langle k \rangle G_1(1)^{l-1} v^{l-1} u^2$ represents the average number of nodes at distance $l$ from a randomly chosen node in the whole network (finite components). When the networks are singly connected, i.e., $u = 0$, equation (A.11) reduces to $P(k_s, k_i|l) = q_k q_k$ which is a known result for uncorrelated random networks [33]. Taking the limit $u \to 1$ ($v \to 1$), we obtain $P_{GC}(k_s, k_i|l)$ at criticality as

$$\lim_{u \to 1, v \to 1} P_{GC}(k_s, k_i|l, GC) = \frac{G_1^2(1)}{G_1^2(1)} \frac{(l - 1) + (k_i + k_s - 2)}{(l - 1) + 2} q_k q_k.$$

(A.12)

Replacing $u$ and $v$ by $\xi_k = -1/\log u$ and $\xi_i = -1/\log v$, respectively, we can rewrite equation (A.11) as

$$P_{GC}(k_s, k_i|l) = \frac{1 - e^{-(l-1)/\xi_i}}{1 - e^{-(l-1)/\xi_i}} \frac{e^{-(k_i + k_s - 2)/\xi_k} - q_k q_k}{e^{-(k_i + k_s - 2)/\xi_k} - q_k q_k}.$$

(A.13)

Let us introduce the probability $P(k|k_s, l, GC)$ that one end of a chain has degree $k_i$ given that the chain has length $l$ and has a degree-$k_s$ node as a starting node, and belongs to the GC. The probability is given from equation (A.11) as

$$P(k_s|k_s, l, GC) = \frac{1 - v^{l-1} u^{k_s + k_i - 2}}{1 - v^{l-1} u^2} q_k.$$

(A.14)
We get the average degree $k_{l}^{GC}(k_{s})$ of $l$-distant nodes from a degree-$k_{s}$ node on GC as

$$k_{l}^{GC}(k_{s}) = \sum_{k_{l}} k_{l}P(k_{l}|k_{s}, l, GC)$$

$$= \frac{\langle k^{2} \rangle}{\langle k \rangle} + \frac{h(u)u^{l-1}u^{k_{s}}}{1 - u^{l-1}u^{k_{s}}}$$

(A.15)

where $h(u) = \sum_{k} k_{s}^{l}u^{k_{s} - 2}$. We consider the situation that the GC exists but infinitely small i.e., $u \sim 1 - \epsilon$ and $v \sim 1 - (k(k-1)(k-2))\epsilon/(k(k-1))$. In the situation, equation (A.15) approximates

$$k_{l}^{GC}(k_{s}) \sim \frac{\langle k^{2} \rangle}{\langle k \rangle} \left(1 + \frac{\langle k^{3} \rangle - 2\langle k \rangle^{2}}{\langle k \rangle} \frac{1}{(l-1)(k(k-1)(k-2))/(k(k-1)) + k_{s}}\right).$$

(A.16)

**Appendix B. Correlation function $C(l)$**

We consider the correlation function $C(l)$ defined as

$$C(l) = \langle kk' \rangle_{l} - \langle k \rangle_{l}\langle k' \rangle_{l},$$

(B.1)

where $\langle f(k, k') \rangle_{l} = \sum_{k, k'} f(k, k')P_{GC}(k, k'|l)$. From equations (A.11) and (B.1), we have

$$C(l) = \sum_{k, k'} kk' \frac{1 - v^{l-1}u^{k+k'-2}}{1 - v^{l-1}u^{2}}q_{k}q_{k' - k} - \left(\sum_{k, k'} \frac{1 - v^{l-1}u^{k+k'-2}}{1 - v^{l-1}u^{2}}q_{k}q_{k'}\right)^{2}$$

$$= \frac{\langle \langle k^{2} \rangle_{l} - v^{l-1}k_{s}u^{k_{s}-1} \rangle_{l}^{2}}{1 - v^{l-1}u^{2}} - \left(\frac{\langle \langle k \rangle_{l} \rangle_{l}^{2} - v^{l-1}\sum_{k} k_{s}u^{k_{s}-1}u}{1 - v^{l-1}u^{2}}\right)^{2}.$$  

(B.2)

The first term of the right hand side in equation (B.2) is

$$\frac{\langle \langle k^{2} \rangle_{l} \rangle_{l}^{2} - v^{l-1}(\sum_{k} k_{s}u^{k_{s}-1})^{2}}{1 - v^{l-1}u^{2}}$$

$$= \frac{(1 - v^{l-1}u^{2})\left(\frac{\langle \langle k^{2} \rangle_{l} \rangle_{l}^{2} - v^{l-1}(\sum_{k} k_{s}u^{k_{s}-1})^{2}}{(1 - v^{l-1}u^{2})}\right)^{2}}{(1 - v^{l-1}u^{2})^{2}}$$

$$= \frac{\langle \langle k^{2} \rangle_{l} \rangle_{l}^{2} - \langle \langle k \rangle_{l} \rangle_{l}^{2}v^{l-1}u^{2} + v^{2(l-1)}(\sum_{k} k_{s}u^{k_{s}-1})^{2}u^{2} - v^{l-1}(\sum_{k} k_{s}u^{k_{s}-1})^{2}}{(1 - v^{l-1}u^{2})^{2}},$$

(B.3)

and the second term is

$$\left(\frac{\langle \langle k \rangle_{l} \rangle_{l}^{2} - v^{l-1}(\sum_{k} k_{s}u^{k_{s}-1})u}{1 - v^{l-1}u^{2}}\right)^{2}$$

$$= \frac{\langle \langle k \rangle_{l} \rangle_{l}^{2} - 2\langle \langle k \rangle_{l} \rangle_{l}v^{l-1}(\sum_{k} k_{s}u^{k_{s}-1})u + v^{2(l-1)}(\sum_{k} k_{s}u^{k_{s}-1})^{2}u^{2}}{(1 - v^{l-1}u^{2})^{2}}.$$  

(B.4)

Then, we have
where we used $\nu = e^{-1/\xi_l}$. Expanding $u$ and $e^{-(l-1)/\xi_l}$ in the denominator as $u \sim 1 - \epsilon$ and $e^{-(l-1)/\xi_l} \sim 1 - (l - 1)/\xi_l$, we drive the correlation function in the critical region as

$$C(l) \sim \frac{a^2}{(b(l - 1) + 2)^2} e^{-(l-1)/\xi_l},$$

where $a = \langle k^2(k - 2) \rangle / \langle k \rangle$ and $b = \langle k(k - 1)(k - 2) \rangle / \langle k \rangle$.

**ORCID iDs**

Shogo Mizutaka 🐑 https://orcid.org/0000-0002-0494-0556

**References**

[1] Caldarelli, G. 2007 Scale-free Networks: Complex Webs in Nature and Technology (Oxford: Oxford University Press)
[2] Newman M 2018 Networks (Oxford: Oxford University Press)
[3] Rombach M P, Porter M A, Fowler J H and Mucha P J 2014 Core-periphery structure in networks SIAM J. Appl. Math. 74 167–90
[4] Chung F and Lu L 2002 The average distances in random graphs with given expected degrees Phys. Rev. E 65 062308
[5] Cohen R and Havlin S 2003 Scale-free networks are ultrasmall Phys. Rev. Lett. 90 058701
[6] Ravasz E and Barabási A L 2003 Hierarchical organization in complex networks Phys. Rev. E 67 026112
[7] Vázquez A, Dobrin R, Sergi D, Eckmann J P, Oltvai Z N and Barabási A L 2004 The topological relationship between the large-scale attributes and local interaction patterns of complex networks Proc. Natl Acad. Sci. 101 17940–5
[8] Stegehuis C, Van Der Hofstad R, Janssen A E M and Van Leeuwaarden J S H 2017 Clustering spectrum of scale-free networks Phys. Rev. E 96 042309
[9] Xuvi-Brunet R and Sokolov I M 2004 Reshuffling scale-free networks: from random to assortative Phys. Rev. E 70 66102
[10] Soffer S N and Vázquez A 2005 Network clustering coefficient without degree–correlation bias Phys. Rev. E 71 057101
[11] Angeles Serrano M and Boguñá M 2005 Tuning clustering in random networks with arbitrary degree distributions Phys. Rev. E 72 036133
[12] Radicchi F, Castellano C, Cecconi F, Loretto V and Parisi D 2004 Defining and identifying communities in networks Proc. Natl Acad. Sci. 101 2658–63
[13] Palla G, Derényi I, Farkas I and Vicsek T 2005 Uncovering the overlapping community structure of complex networks in nature and society Nature 435 814–8
[14] Arenas A, Fernández A, Fortunato S and Gómez S 2008 Motif-based communities in complex networks J. Phys. A: Math. Theor. 41 224001
[15] Fortunato S 2010 Community detection in graphs Phys. Rep. 486 75–174
[16] Orsini C et al 2015 Quantifying randomness in real networks Nat. Commun. 6 8627
[17] Kawasaki F and Yakubo K 2010 Reciprocal relation between the fractal and the small-world properties of complex networks Phys. Rev. E 82 036113
[18] Rozenfeld H D, Song C and Makse H A 2010 Small-world to fractal transition in complex networks: a renormalization group approach Phys. Rev. Lett. 104 025701
[19] Watts D J and Strogatz S H 1998 Collective dynamics of small-world networks Nature 393 440–2
[20] Song C, Havlin S and Makse H A 2005 Self-similarity of complex networks Nature 433 392
[21] Brandes U and Havlin S 2012 Fractals and Disordered Systems (Berlin: Springer)
[22] Yook S-H, Radicchi F and Meyer-Ortmanns H 2005 Self-similar scale-free networks and disassortativity Phys. Rev. E 72 045105
[23] Gallos L K, Song C, Havlin S and Makse H A 2007 Scaling theory of transport in complex biological networks Proc. Natl Acad. Sci. 104 7746–51
[24] Gallos L K, Song C and Makse H A 2008 Scaling of degree correlations and its influence on diffusion in scale-free networks Phys. Rev. Lett. 100 248701
[25] Fujiki Y, Mizutaka S and Yakubo K 2017 Fractality and degree correlations in scale-free networks Eur. Phys. J. B 90 126
[26] Song C, Havlin S and Makse H A 2006 Origins of fractality in the growth of complex networks Nat. Phys. 2 275
[27] Goh K-I, Sabi G, Kahng B and Kim D 2006 Skeleton and fractal scaling in complex networks Phys. Rev. Lett. 96 018701
[28] Bialas P and Olej A K 2010 Long-range disassortative correlations in generic random trees Phys. Rev. E 81 041136
[29] Bialas P and Olej A K 2008 Correlations in connected random graphs Phys. Rev. E 77 036124
[30] Tishby I, Bham O, Katzav E and Kühn R 2018 Revealing the microstructure of the giant component in random graph ensembles Phys. Rev. E 97 042318
[31] Mizutaka S and Hasegawa T 2018 Disassortativity of percolating clusters in random networks Phys. Rev. E 98 062314
[32] Hasegawa T and Mizutaka S 2020 Structure of percolating clusters in random clustered networks Phys. Rev. E 101 062310
[33] Fujiki Y, Takaguchi T and Yakubo K 2018 General formulation of long-range degree correlations in complex networks Phys. Rev. E 97 062308
[34] Fujiki Y and Yakubo K 2020 Identification of intrinsic long-range degree correlations in complex networks Phys. Rev. E 101 032308
[35] Rybski D, Rozenfeld H D and Kropp J P 2010 Quantifying long-range correlations in complex networks beyond nearest neighbors *Europhys. Lett.* 90 28002

[36] Cohen R and Havlin S 2010 *Complex Networks: Structure, Robustness and Function* (Cambridge: Cambridge University Press)

[37] Ivan Kryven 2017 General expression for the component size distribution in infinite configuration networks *Phys. Rev. E* 95 052303

[38] Harris T E 1989 *The Theory of Branching Processes* (New York: Dover)