Degree distribution and assortativity in line graphs of complex networks

CrossMark

Xiangrong Wang [a,*], Stojan Trajanovski [a], Robert E. Kooij [a,b], Piet Van Mieghem [a]

[a] *Faculty of Electrical Engineering, Mathematics and Computer Science, Delft University of Technology, Delft, The Netherlands*
[b] *TNO (Netherlands Organization for Applied Scientific Research), Delft, The Netherlands*

HIGHLIGHTS

- We investigate the degree distribution and the assortativity of a line graph.
- Degree distributions of line graphs and their original graphs are the same class.
- Line graphs are assortative in most cases.
- Assortativity of line graphs is not linearly related to the one of original graphs.
- We find trees and non-trees with negative assortativity in their line graphs.

ARTICLE INFO

ABSTRACT

Topological characteristics of links of complex networks influence the dynamical processes executed on networks triggered by links, such as cascading failures triggered by links in power grids and epidemic spread due to link infection. The line graph transforms links in the original graph into nodes. In this paper, we investigate how graph metrics in the original graph are mapped into those for its line graph. In particular, we study the degree distribution and the assortativity of a graph and its line graph. Specifically, we show, both analytically and numerically, the degree distribution of the line graph of an Erdős–Rényi graph follows the same distribution as its original graph. We derive a formula for the assortativity of line graphs and indicate that the assortativity of a line graph is not linearly related to its original graph. Additionally, line graphs of various graphs, e.g. Erdős–Rényi graphs, scale-free graphs, show positive assortativity. In contrast, we find certain types of trees and non-trees whose line graphs have negative assortativity.

© 2015 Elsevier B.V. All rights reserved.

1. Introduction

Infrastructures, such as the Internet, electric power grids and transportation networks, are crucial to modern societies. Most researches focus on the robustness of such networks to node failures [1,2]. Specifically, the effect of node failures on the robustness of networks is studied by percolation theory both in single networks [2] and interdependent networks that interact with each other [3]. However, links frequently fail in various real-world networks, such as the failures of transmission lines in electrical power networks, path congestions in transportation networks. The concept of a line graph,

* Corresponding author.
 E-mail address: x.wang-2@tudelft.nl (X. Wang).

that transforms links of the original graph into nodes in the line graph, can be used to understand the influence of link failures on infrastructure networks.

An undirected graph with N nodes and L links can be denoted as $G(N, L)$. The line graph $l(G)$ of a graph G is a graph in which every node in $l(G)$ corresponds to a link in G and two nodes in $l(G)$ are adjacent if and only if the corresponding links in G have a node in common [4]. The graph G is called the original graph of $l(G)$.

Line graphs are applied in various complex networks. Krawczyk et al. [5] propose the line graph as a model of social networks that are constructed on groups such as families, communities and school classes. Line graphs can also represent protein interaction networks where each node represents an interaction between two proteins and each link represents pairs of interaction connected by a common protein [6]. By the line graph transformation, methodologies for nodes can be extended to solve problems related to links in a graph. For instance, the link chromatic number of a graph can be computed from the node chromatic number of its line graph [7]. An Eulerian path (that can be computed rather easily in polynomial time) in a graph transforms to a Hamiltonian path (which is difficult to compute, in fact, the problem is NP-hard) in the line graph. Evan et al. [8] use algorithms that produce a node partition in the line graph to achieve a link partition in order to uncover overlapping communities of a network. Wierman et al. [9] improve the bond (link) percolation threshold of a graph by investigating site (node) percolation in its line graph.

Previous studies focus on various mathematical properties of line graphs. Whitney's Theorem [10] states that, if line graphs of two connected graphs G_1 and G_2 are isomorphic, the graphs G_1 and G_2 are isomorphic unless one is the complete graph K_3 and the other one is the star $K_{1,3}$. Krausz [11], Van Rooij and Wilf [12] have investigated the conditions for a graph to be a line graph. Van Rooij and Wilf [12] have studied the properties of graphs obtained by iterative usage of the line graph transformation, e.g., the line graph $l(G)$ of a graph G, the line graph $l(l(G))$ of the line graph $l(G)$, etc. Furthermore, Harary [13] has shown that for connected graphs that are not path graphs, all sufficiently high numbers of iterations of the line graph transformation produce Hamiltonian graphs.[1] The generation of a random line graph is studied in Ref. [14]. An original graph can be reconstructed [15–17] from its line graph with a computational complexity that is linear in the number of nodes N.

In this paper, we analytically study the degree distribution and the assortativity of line graphs and the relation to the degree distribution and the assortativity of their original networks. We show that the degree distribution in the line graph of the Erdős–Rényi graph follows the same pattern as the degree distribution in Erdős–Rényi. However, the line graph of an Erdős–Rényi graph is not an Erdős–Rényi graph. Additionally, we investigate the assortativity of line graphs and show that the assortativity is not linearly related to the assortativity in the original graphs. The line graphs are assortative in most cases, yet line graphs are not always assortative. We investigate graphs with negative assortativity in their line graphs. The remainder of this paper is organized as follows. The degree distribution of line graphs is presented in Section 2. Section 3 provides the assortativity of line graphs. We conclude in Section 4.

2. Degree distribution

Random graphs are developed as models of real-world networks of several applications, such as peer-to-peer networks, the Internet and the World Wide Web. The degree distribution of Erdős–Rényi random graphs and scale free graphs are recognized by the binomial distribution and the power law distribution, respectively. This section studies the degree distribution of the line graphs of Erdős–Rényi and scale free graphs.

Let $G(N, L)$ be an undirected graph with N nodes and L links. The adjacency matrix A of a graph G is an $N \times N$ symmetric matrix with elements a_{ij} that are either 1 or 0 depending on whether there is a link between nodes i and j or not. The degree d_i of a node i is defined as $d_i = \sum_{k=1}^{N} a_{ik}$. The degree vector $d = (d_1 \ d_2 \ \cdots \ d_N)$ has a vector presentation as $Au = d$, where $u = (1, 1, \ldots, 1)$ is the all-one vector. The adjacency matrix [4] of the line graph $l(G)$ is $A_{l(G)} = R^T R - 2I$, where R is an $N \times L$ unsigned incidence matrix with $R_{il} = R_{jl} = 1$ if there is a link l between nodes i and j, elsewhere 0 and I is the identity matrix. The degree vector $d_{l(G)}$ of the line graph $l(G)$ is $d_{l(G)} = A_{l(G)} u_{L \times 1}$. For an arbitrary node l in the line graph $l(G)$, which corresponds to a link l connecting nodes i and j in graph G (as shown in Fig. 1), the degree d_l of the node l follows

$$d_l = d_i + d_j - 2. \tag{1}$$

The random variable D_i denotes the degree of a randomly chosen node i in Erdős–Rényi graphs $G_p(N)$ and (1) shows that the degree D_l of a link l with end node i in the corresponding line graph is $D_l = D_i + D_j - 2$.

Theorem 1. *The degree distribution of the line graph $l(G_p(N))$ of an Erdős–Rényi graph $G_p(N)$ follows a binomial distribution*

$$\Pr[D_l = k] = \binom{2N - 4}{k} p^k (1 - p)^{(2N-4-k)} \tag{2}$$

with average degree $E[D_{l(G_p(N))}] = (2N - 4)p$.

---

[1] A Hamiltonian graph is a graph possessing a Hamiltonian cycle which is a closed path through a graph that visits each node exactly once.
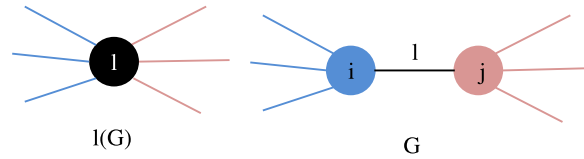
Fig. 1. Node l in line graph $l(G)$ corresponds to the link l in G.

Proof. Applying (1), the degree distribution D_l of a node l in a line graph is

$$\Pr[D_l = k] = \Pr[D_i + D_j - 2 = k].$$

Using the law of total probability [18] yields

$$\Pr[D_l = k] = \sum_{m=1}^{k} \Pr[D_j = k - m + 2 | D_i = m]\Pr[D_i = m].$$

Since the random variables D_i and D_j in $G_p(N)$ are independent, we have

$$\Pr[D_l = k] = \sum_{m=1}^{k} \Pr[D_j = k - m + 2]\Pr[D_i = m]. \tag{3}$$

An arbitrarily chosen (i.e. uniformly at random) node l in the line graph $l(G)$ corresponds to an arbitrarily chosen link in G. The degree distribution [18] of the end node i of an arbitrarily chosen link in G is

$$\Pr[D_i = m] = \frac{m\Pr[D = m]}{E[D]} \tag{4}$$

where $\Pr[D = m]$ is the degree distribution of an arbitrarily chosen node in graph G and $E[D]$ is the average degree of an arbitrarily chosen node. In an Erdős–Rényi graph, we have $\Pr[D = m] = \binom{N-1}{m}p^m(1-p)^{N-1-m}$ and $E[D] = (N-1)p$. By substituting (4) into (3) and applying the binomial distribution of random variables D_i and D_j, we have

$$\begin{aligned}
\Pr[D_l = k] &= \sum_{m=1}^{k} \frac{(k-m+2)\Pr[D = k-m+2]}{E[D]} \frac{m\Pr[D = m]}{E[D]} \\
&= \sum_{m=1}^{k} \frac{(k-m+2)\binom{N-1}{k-m-2}p^{k-m+2}(1-p)^{N-1-(k-m+2)}}{(N-1)p} \frac{m\binom{N-1}{m}p^m(1-p)^{N-1-m}}{(N-1)p} \\
&= p^k(1-p)^{2N-4-k} \sum_{m=0}^{k} \binom{N-2}{k-m}\binom{N-2}{m}.
\end{aligned}$$

Using Vandermonde's identity $\binom{m+n}{r} = \sum_{k=0}^{r} \binom{m}{k}\binom{n}{r-k}$, we arrive at (2). □

Theorem 1 illustrates that the degree distribution of the line graph $l(G)$ of an Erdős–Rényi graph G follows a binomial distribution with average degree $E[D_{l(G)}] = (2N - 4)p$. Compared to the average degree $E[D] = (N - 1)p$, the average degree of the line graph of the Erdős–Rényi graph is two times the average degree $E[D]$ of the Erdős–Rényi graph minus $2p$.

Fig. 2 shows the degree distribution of the line graphs of Erdős–Rényi graphs $G_N(p)$ for $N = 100, 200$ and $p = 2p_c$ ($p_c \approx \frac{\ln N}{N}$), where 10^5 Erdős–Rényi graphs are generated. In Fig. 2(a) and (b), the degree distributions of Erdős–Rényi graphs (red circle) follow a binomial distribution. The degree distribution of the corresponding line graph (black square) is fitted by a binomial distribution $B(2N - 4, p)$. The simulation results agree with Theorem 1. Moreover, the average degree $E[D_{l(G)}]$ of the line graph is approximately two times the average degree $E[D]$ of the graph G.

Since the degree distribution of the line graphs of Erdős–Rényi graphs follows a binomial distribution, we pose the question: Is the line graph of an Erdős–Rényi graph also an Erdős–Rényi graph? In order to answer this question, we investigate the eigenvalue distribution of the line graph. Fig. 3 shows the eigenvalue distribution of Erdős–Rényi graphs and their line graphs. As shown in Ref. [4], the eigenvalue distribution of Erdős–Rényi graphs follows a semicircle distribution. The eigenvalue distribution of the line graphs of Erdős–Rényi graphs follows a different distribution than a semicircle distribution. Since the spectrum of a graph can be regarded as the unique fingerprint of that graph to a good approximation [19], we conclude that the line graphs of Erdős–Rényi graphs are not Erdős–Rényi graphs.

Generating functions are powerful to study the degree distribution of networks [18]. Assuming the degree independence of nodes in graph G, Theorem 2 shows the generating function for the line graph $l(G)$ of an arbitrary graph.
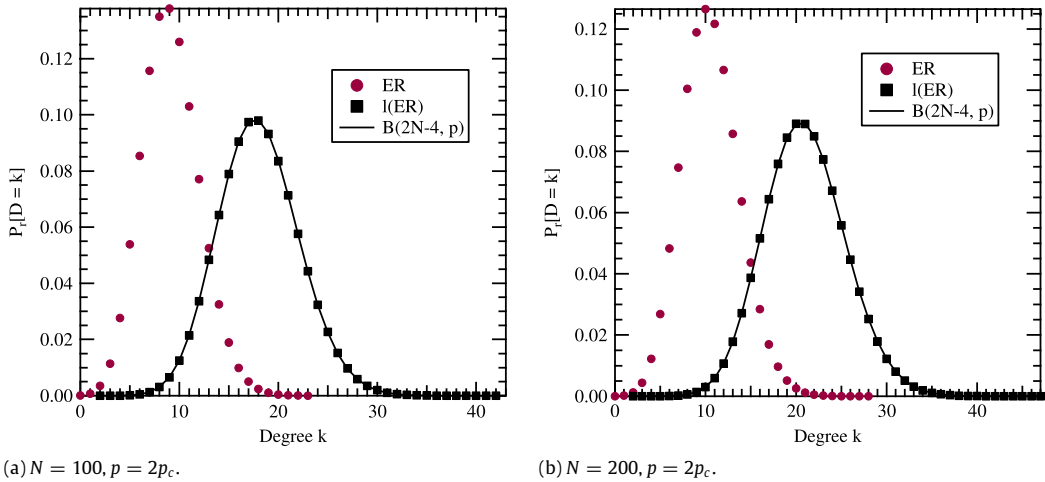
(a) $N = 100, p = 2p_c$. (b) $N = 200, p = 2p_c$.

Fig. 2. The degree distribution of Erdős–Rényi graphs and their corresponding line graphs. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)
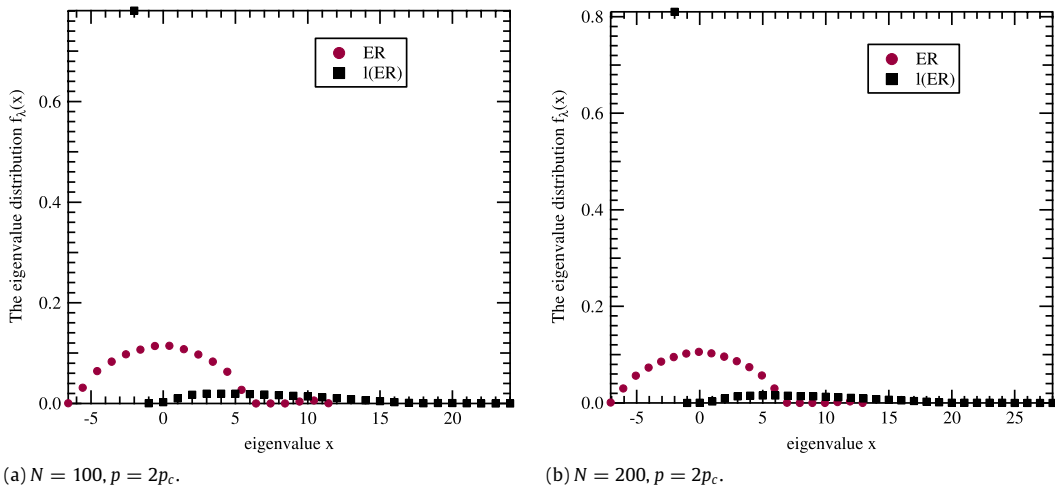

(a) $N = 100, p = 2p_c$. (b) $N = 200, p = 2p_c$.

Fig. 3. The eigenvalue distribution of Erdős–Rényi graphs and their corresponding line graphs. The simulations are performed on 10^5 instances.

Theorem 2. *Assuming that the degrees of nodes in a graph G are independent, the generating function for the degree D_l in the line graph l(G) follows*

$$\varphi_{D_l}(z) = \left(\frac{E[z^{D_{l+}}]}{z} \right)^2 \tag{5}$$

where D_{l+} is the degree of the end node of an arbitrarily chosen link l in G.

Proof. The probability generating function for the degree D_l of a node l in the line graph is

$$\varphi_{D_l}(z) = E[z^{D_l}].$$

Using (1), we have

$$\varphi_{D_l}(z) = E[z^{D_i+D_j-2}].$$

Since the condition in the theorem assumes that the random variables D_i and D_j are independent and identically distributed as D_{l+}, we establish Theorem 2. □

We apply the generating function (5) in the line graph of scale-free graphs whose degree distribution follows a power law distribution with the exponent α. In Appendix A, we deduce that, for large N,

$$\Pr[D_l = k] \propto \left(\frac{1}{k+2} \right)^{\alpha^{l(G)}} \tag{6}$$
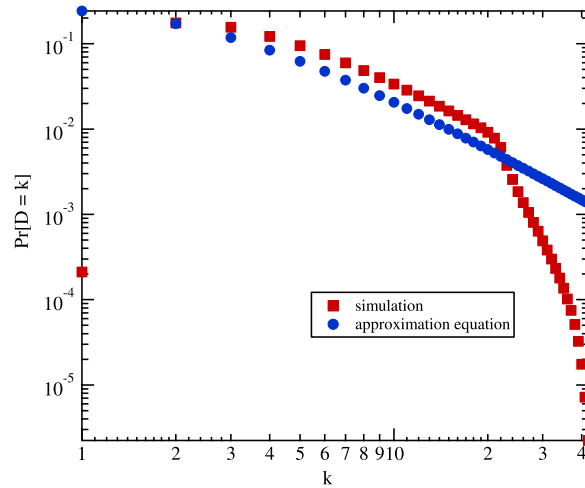
Fig. 4. The degree distribution in the line graph of the Barabási–Albert graph both from simulations and the approximation equation (6). Both the *x*-axis and the *y*-axis are in log scale. The simulations are performed on 10^5 Barabási–Albert graphs with $N = 500$ and average degree 4. The cut-off in the simulation is due to the finite size of the Barabási–Albert graph. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)
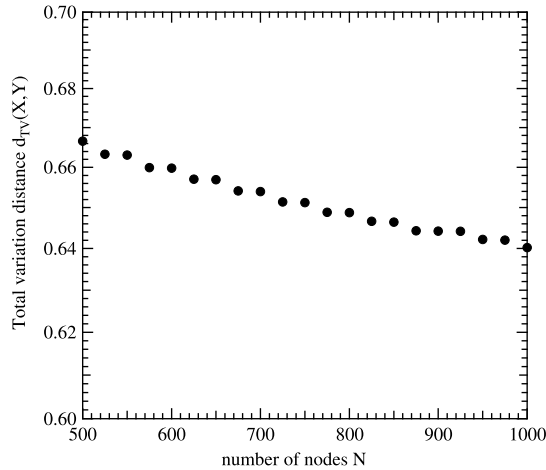

Fig. 5. The total variation distance $d_{TV}(X, Y)$ when the original graph has different number of nodes from 500 to 1000.

where $\alpha^{l(G)} = 2$, whereas in the original graph $\alpha^G = 3$. Eq. (6) illustrates that, when we assume that the degrees in the original graph are independent, the degree distribution in the line graph of Barabási–Albert graph follows a power law degree distribution. However, due to the preferential attachment in scale-free graphs and $2L = \sum_{i=1}^{N} d_i$, the node degrees are dependent rather than independent. Correspondingly, a gap is observed in Fig. 4 between the approximation equation (6) (blue circle) and the simulation result (red square).

The dependency assumption in (5) can be assessed by the total variation distance $d_{TV}(X, Y)$, defined as [18]:

$$d_{TV}(X, Y) = \sum_{k=-\infty}^{\infty} \left| \Pr[X = k] - \Pr[Y = k] \right|$$

where $\Pr[X = k]$ denotes the probability density function for (6) and $\Pr[Y = k]$ for simulations.

Fig. 5 shows the total variation distance when the number of nodes N in Barabási–Albert graphs increases from 500 to 1000 with average degree 4. For each size of the original graph, 10^5 graphs are generated. Fig. 5 demonstrates that $d_{TV}(X, Y)$ decreases with the number of nodes N, starting from 0.667 when $N = 500$ to 0.640 when $N = 1000$. Accordingly, the accuracy of the approximation equation (6) increases with the size of the original graph.

3. Assortativity

Networks with a same degree distribution may have significantly different topological properties [20]. Networks, where nodes preferentially connect to nodes with (dis)similar property, are called (dis)assortative [21]. An overview of the

assortativity in complex networks is given in Ref. [22]. Assortativity is quantified by the linear degree correlation coefficient defined as

$$\rho_{D_{l(G)}} = \frac{E[D_{l^+} D_{l^-}] - E[D_{l^+}]E[D_{l^-}]}{\sigma_{D_{l^+}} \sigma_{D_{l^-}}} \tag{7}$$

where $E[X]$ and σ_X are the mean and standard deviation of the random variable X. The definition (7) has been transformed into a graph formulation in Ref. [20]. In this section, we investigate the assortativity $\rho_{D_{l(G)}}$ of the line graph $l(G)$ and its relation to the assortativity of the graph G.

3.1. Assortativity in the line graph

In this subsection, we derive a formula for the assortativity in a general line graph, represented in Theorem 3. The relation between the assortativity in the line graph and the assortativity in the original graph is shown in Corollary 1.

Theorem 3. *The assortativity in the line graph $l(G)$ of a general graph G is*

$$\rho_{D_{l(G)}} = 1 - \frac{d^T A \Delta d - N_4}{3 d^T A \Delta d + \sum_{k=1}^{N} d_k^4 - 2 \sum_{k=1}^{N} d_k^3 - 2N_3 - \frac{\left(N_3 + \sum_{k=1}^{N} d_k^3 - 2N_2\right)^2}{N_2 - N_1}}$$

where d is the degree vector, $\Delta = \text{diag}(d_i)$ is the diagonal matrix with the nodal degrees in G and $N_k = u^T A^k u$ is the total number of walks of length k.

The proof for Theorem 3 is given in Appendix B. In order to investigate the relation between the assortativity of the line graph $l(G)$ and the assortativity of the graph G, Corollary 1 rephrases the assortativity $\rho_{D_{l(G)}}$ of the line graph $l(G)$ in terms of the assortativity ρ_D of the graph G.

Corollary 1. *The assortativity $\rho_{D_{l(G)}}$ of the line graph can be written in terms of the assortativity ρ_D of the graph G as*

$$\rho_{D_{l(G)}} = 1 - \frac{(d^T A \Delta d - N_4)\mu^2}{(N_2 - N_1)\left(-4(1 + \rho_D)^2 \left(\frac{1}{N_1} \sum_{i=1}^{N} d_i^3 - \left(\frac{N_2}{N_1}\right)^2\right)^2 + 2\mu^2(1 + \rho_D)\left(\frac{1}{N_1} \sum_{i=1}^{N} d_i^3 - \left(\frac{N_2}{N_1}\right)^2\right) + \mu u_3\right)}$$

where $\mu = E[D_{l(G)}]$ and $u_3 = E[(D_{l(G)} - E[D_{l(G)}])^3]$.

The proof for Corollary 1 is given in Appendix C. Corollary 1 indicates that the assortativity of the line graph is not linearly related to the assortativity of the original graph. For the Erdős–Rényi graphs, a relatively precise relation between the assortativity of the line graph and the one of the original graph is given in Theorem 4.

Theorem 4. *The difference between the assortativity $\rho_{D_{l(G)}}$ of the line graph of an Erdős–Rényi graph $G_N(p)$ and the assortativity ρ_{D_G} of $G_N(p)$ converges to 0.5 in the limit of large graph size N.*

Proof. Based on the definition in Eq. (7) and denoting $l^+ = i \sim c$ and $l^- = c \sim j$, we have

$$\rho_{D_{l(G)}} = \frac{E[(D_i + D_c)(D_j + D_c)] - E[D_i + D_c]E[D_j + D_c]}{\sigma_{D_i + D_c} \sigma_{D_j + D_c}}$$

$$= \frac{E[D_i D_j] - E[D_i]E[D_j] + E[D_i D_c] - E[D_i]E[D_c] + E[D_j D_c] - E[D_j]E[D_c] + E[D_c^2] - E^2[D_c]}{\text{Var}[D_i] + \text{Var}[D_c] + 2E[(D_i - E[D_i])(D_c - E[D_c])]}.$$

In the connected Erdős–Rényi random graph in the limit of large graph size N, the assortativity ρ_{D_G} converges to zero [4] and we have

$$E[D_i D_j] - E[D_i]E[D_j] \approx 0.$$

Similarly, $E[D_i D_c] - E[D_i]E[D_c] \approx 0$ and $E[D_j D_c] - E[D_j]E[D_c] \approx 0$. Combining with $E[(D_i - E[D_i])(D_c - E[D_c])] = E[D_i D_c] - E[D_i]E[D_c] \approx 0$, we arrive at

$$\rho_{D_{l(G)}} \approx \frac{E[D_c^2] - E^2[D_c]}{2\text{Var}[D_c]} = 0.5. \quad \square$$
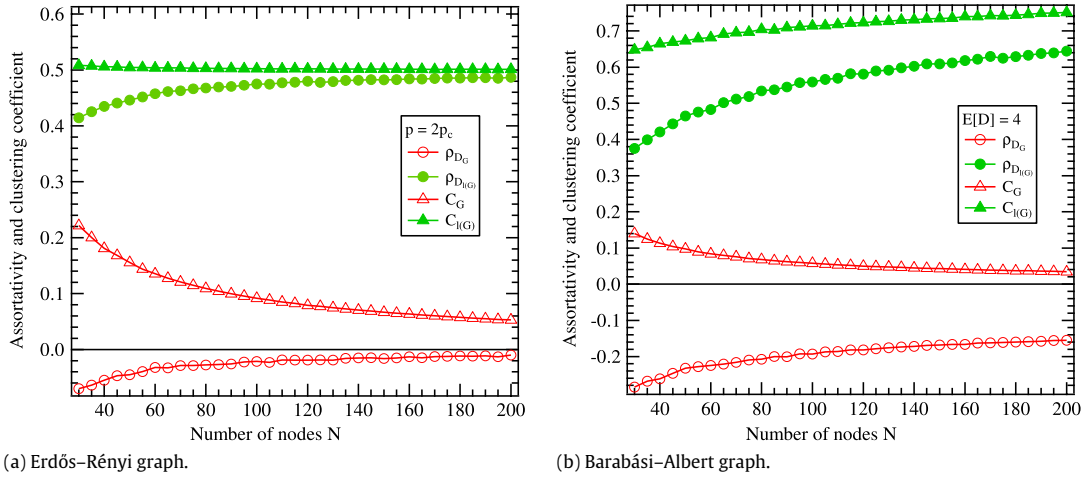
Fig. 6. Assortativity ρ_D and clustering coefficient C_G of the (a) Erdős–Rényi graph $G_p(N)$ with $p = 2p_c$, (b) Barabási–Albert graph with the average degree $E[D] = 4$ and the corresponding line graph $l(G)$.

Table 1
Assortativity of real-world networks and their corresponding line graphs.

Networks	Nodes	Links	ρ_D	$\rho_{D_{l(G)}}$
Co-authorship Network [24]	379	914	−0.0819	0.6899
US airports [25]	500	2 980	−0.2679	0.3438
Dutch Soccer [26]	685	10 310	−0.0634	0.5170
Citation[a]	2678	10 368	−0.0352	0.8127
Power Grid	4941	6 594	−0.0035	0.7007

[a] http://vlado.fmf.uni-lj.si/pub/networks/data/.

In order to verify Theorem 4, Fig. 6 shows the assortativity of (a) Erdős–Rényi graphs, (b) Barabási–Albert graphs, and the assortativity of their corresponding line graphs. In Fig. 6(a), the assortativity of $G_p(N)$ converges to 0 with the increase of the graph size N. Correspondingly, the assortativity in the line graph of $G_p(N)$ converges to 0.5 which confirms Theorem 4. Based on the assortativity ρ_D of a connected Erdős–Rényi graph $G_p(N)$, which is zero [4,21] in the limit of large graph size, we again verify that the line graph of an Erdős–Rényi graph is not an Erdős–Rényi graph. Fig. 6(b) illustrates the assortativity $\rho_{D_{l(G)}}$ of the line graph of the Barabási–Albert graph is also positive and increases with the graph size.

Youssef et al. [23] show that the assortativity is related to the clustering coefficient[2] C_G. Specifically, assortative graphs tend to have a higher number $\blacktriangle_G$ of triangles and thus a higher clustering coefficient compared to disassortative graphs. Fig. 6 shows that the assortative line graphs of both Erdős–Rényi and Barabási–Albert graph have a higher clustering coefficient (above 0.5). The results agree with the findings in Ref. [23].

Table 1 shows the assortativity of real-world networks and their corresponding line graphs. As shown in the table, the line graphs of all the listed networks show assortative mixing even though the original networks show disassortative mixing.

3.2. Negative assortativity in line graphs

Although the assortativity of a line graph is predominantly positive, we cannot conclude that the assortativity in any line graph is positive. This subsection presents graphs, whose corresponding line graphs possesses a negative assortativity.

3.2.1. The line graph of a path graph

A path graph P_N is a tree with two nodes of degree 1, and the other $N - 2$ nodes of degree 2. The line graph $l(P)$ of a path graph P_N is still a path graph but with $N - 1$ nodes. Observation 1 demonstrates that the assortativity in the line graph of a path graph is always negative.

Observation 1. *The assortativity of the line graph $l(P)$ of a path P_N is*

$$\rho_{D_{l(P)}} = -\frac{1}{N - 3}$$

where N is the number of nodes in the original path graph.

---

[2] The clustering coefficient $C_G = \frac{3\blacktriangle_G}{N_2}$ is defined as three times the number $\blacktriangle_G$ of triangles divided by the number N_2 of connected triples.
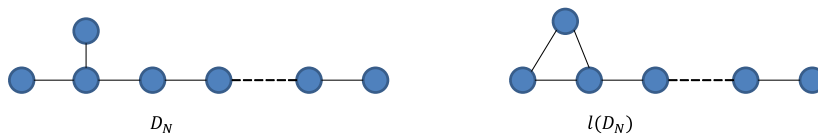
Fig. 7. The graph D_N whose line graph has the negative assortativity.

Proof. The reformulation [4] of the assortativity can be written as

$$\rho_D = 1 - \frac{\sum_{i \sim j}(d_i - d_j)^2}{\sum_{i=1}^{N-1}(d_i)^3 - \frac{1}{2L}\left(\sum_{i=1}^{N-1}d_i^2\right)^2}. \tag{8}$$

Since the line graph of a path with N nodes is a path graph with $N - 1$ nodes, where 2 nodes have node degree 1 and the other $(N - 1) - 2$ nodes have degree 2, we have that

$$\sum_{i=1}^{N-1} d_i^k = 2 \times 1^k + ((N - 1) - 2) \times 2^k \tag{9}$$

and

$$\sum_{i \sim j}(d_i - d_j)^2 = 2 \times 1^2. \tag{10}$$

Applying Eqs. (9) and (10) into (8), we establish the Observation 1. □

The negative assortativity $\rho_{D_{l(P)}}$ of the line graph $l(P)$ of a path graph is an exception to the positive assortativity of the line graphs of the Erdős–Rényi graph, Barabási–Albert graph and real-world networks given in Table 1. Moreover, the assortativity of the line graph $l(P)$ is a fingerprint for the line graph $l(P)$ to be a path graph.

3.2.2. The line graph of a path-like graph

Let $P_{n_1, n_2, \ldots, n_t, p}^{m_1, m_2, \ldots, m_t}$ be a path of p nodes $(1 \sim 2 \sim \cdots \sim p)$ with pendant paths of n_i links at nodes m_i, following the definition in Ref. [27]. We define the graph D_N through $D_N = P_{1, N-1}^2$ as drawn in Fig. 7. Observation 2 shows that the assortativity in the corresponding line graph $l(D_N)$ is always negative.

Observation 2. *The assortativity of the line graph $l(D_N)$ of the graph D_N in Fig. 7 is*

$$\rho_{D_{l(D_N)}} = -\frac{1}{2N - 3}$$

where N is the number of nodes in the graph D_N.

Proof. Since 1 node has node degree 1, 1 node has node degree 3 and the other $(N - 1) - 2$ nodes have degree 2, we have that

$$\sum_{i=1}^{N-1} d_i^k = 1 \times 1^k + 1 \times 3^k + ((N - 1) - 2) \times 2^k \tag{11}$$

and

$$\sum_{i \sim j}(d_i - d_j)^2 = 1 \times 1^2 + 3 \times 1^2. \tag{12}$$

Applying Eqs. (11) and (12) into (8), we establish the Observation 2. □

We define the graph E_N through $E_N = P_{1, N-1}^3$ as drawn in Fig. 8. The graph E_N is obtained from D_N by moving the pendant path from node 2 to node 3. The assortativity of the line graph $l(E_N)$ of the graph E_N is

$$\rho_{D_{l(E_N)}} = -\frac{1}{N - 2}.$$

For the graphs $P_{1, N-1}^{m_i}$ with one pendant path of 1 link at node m_i ($i = 2, 3, \ldots, N - 2$), there are $N - 3$ positions to attach the pendant path. Since the position for adding the pendant path is symmetric at $\lceil \frac{N-1}{2} \rceil$. We only consider i from 2 to $\lceil \frac{N-1}{2} \rceil$. Among all the graphs $P_{1, N-1}^{m_i}$ where $i = 2, 3, \ldots, \lceil \frac{N-1}{2} \rceil$, the line graphs of the graph D_N and E_N always have

Fig. 8. The graph E_N whose line graph has the negative assortativity.
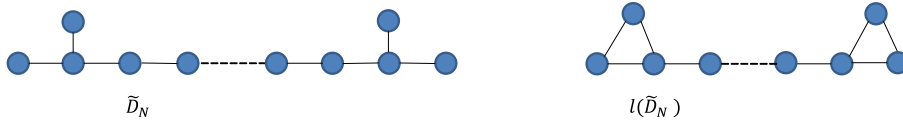

Fig. 9. The graph $\widetilde{D}_N$ whose line graph has the negative assortativity.

negative assortativity. The line graph of the graph $P_{1,\,N-1}^{m_i}$, where $i = 4,\ 5,\ \ldots,\ \lceil \frac{N-1}{2} \rceil$, has negative assortativity when the size N is small and has positive assortativity as N increases.

The graph $\widetilde{D}_N$ is defined through $\widetilde{D}_N = P_{1,\,1,\,N-2}^{2,\,N-3}$ as drawn in Fig. 9. Observation 3 shows that the assortativity in the corresponding line graph $l(\widetilde{D}_N)$ is always negative.

Observation 3. *The assortativity of the line graph $l(\widetilde{D}_N)$ of the graph $\widetilde{D}_N$ in Fig. 9 is*

$$\rho_{D_{l(\widetilde{D}_N)}} = -\frac{3}{N-3}$$

where N is the number of nodes in $\widetilde{D}_N$.

Proof. Since 2 nodes have node degree 3 and the other $(N-1) - 2$ nodes have degree 2, we have that

$$\sum_{i=1}^{N-1} d_i^k = 2 \times 3^k + ((N-1) - 2) \times 2^k \tag{13}$$

and

$$\sum_{i \sim j} (d_i - d_j)^2 = 6 \times 1^2. \tag{14}$$

Applying Eqs. (13) and (14) into (8), we establish the Observation 3.   □

The graphs $\widetilde{E}_N$ and $\widetilde{F}_N$ are defined through $\widetilde{E}_N = P_{1,\,1,\,N-2}^{2,\,N-4}$ and $\widetilde{F}_N = P_{1,\,1,\,N-2}^{3,\,N-4}$ as drawn in Fig. 10. The assortativity for the line graph of $\widetilde{E}_N$ is

$$\rho_{D_{l(\widetilde{E}_N)}} = -\frac{16}{5N-16}.$$

The assortativity for the line graph of $\widetilde{F}_N$ is

$$\rho_{D_{l(\widetilde{F}_N)}} = -\frac{25}{7N-25}.$$

Graphs $\widetilde{D}_N, \widetilde{E}_N, \widetilde{F}_N$ are the graphs whose line graphs always have the negative assortativity. For the remaining graphs $P_{1,\,1,\,N-2}^{m_i,\,m_j}, i \neq j$, their line graphs have negative assortativity when N is small. As N increases, the assortativity of the line graphs is positive.

3.2.3. Line graph of non-trees

Both the path graphs and path-like graphs are trees. In this subsection, we study whether there exist non-trees whose line graphs have negative assortativity.

We start by studying the non-trees $l(D_N)$, $l(E_N)$ and $l(\widetilde{D}_N)$, $l(\widetilde{E}_N)$, $l(\widetilde{F}_N)$ in Figs. 7–10. The non-tree graphs consist of cycles of 3 nodes connected by disjoint paths. The line graph of the non-tree $l(D_N)$ is denoted as $l(l(D_N))$, which is also the line graph of D_N. By simulations we determine the non-tree graphs whose line graphs have negative assortativity. The results are given in Figs. 11 and 12.

As shown in Figs. 11 and 12, for the line graphs of the non-trees to have negative assortativity, there can be either 1 or 2 cycles in the non-trees. In Fig. 11, the line graph $l(l(E_N))$ of $l(E_N)$ has 1 cycle connected by two paths and the maximal path length is 2. In Fig. 12, two cycles are connected by maximal 3 paths and the maximal path length is 4 in the line graph $l(l(\widetilde{F}_N))$. Moreover, for a line graph to have negative assortativity, the size of the original graph is in general small, less than 14 nodes in our simulations.
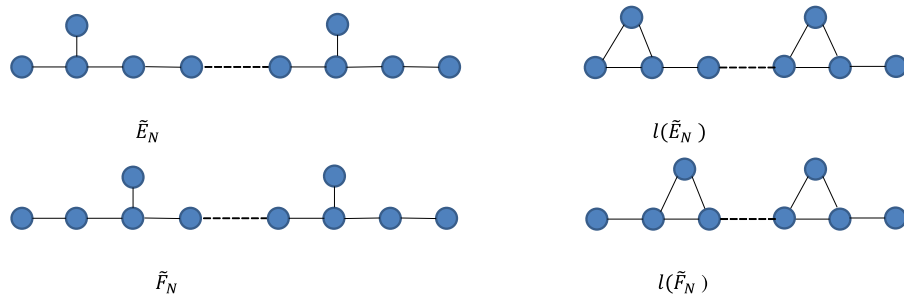
Fig. 10. The graphs $\widetilde{E}_N$ and $\widetilde{F}_N$ whose line graphs have the negative assortativity.

Fig. 11. Non-tree graphs $l(D_N)$, $l(E_N)$ whose line graphs $l(l(D_N))$, $l(l(E_N))$ have negative assortativity.
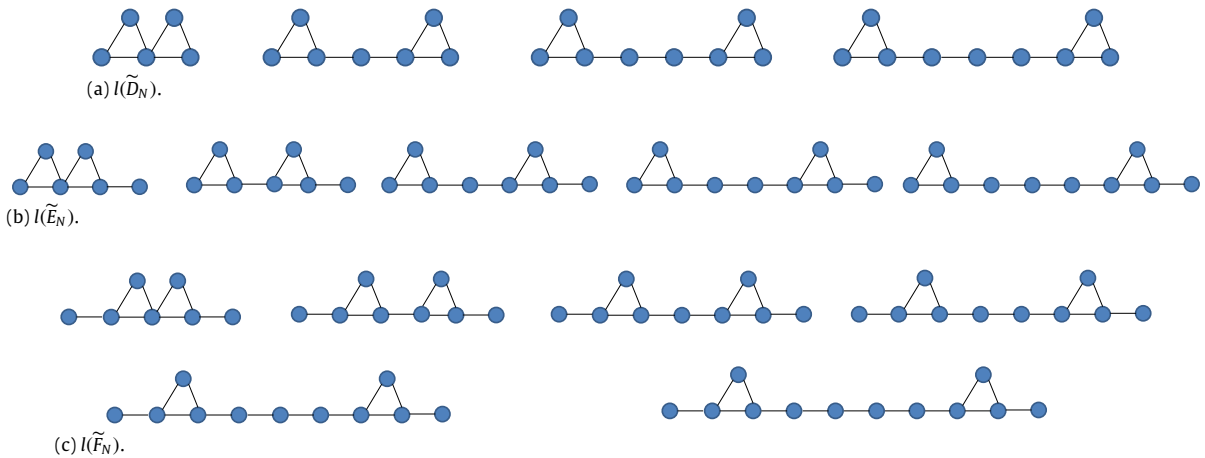

(a) $l(\widetilde{D}_N)$.

(b) $l(\widetilde{E}_N)$.

(c) $l(\widetilde{F}_N)$.

Fig. 12. Non-tree graphs $l(\widetilde{D}_N)$, $l(\widetilde{E}_N)$, $l(\widetilde{F}_N)$ whose line graphs $l(l(\widetilde{D}_N))$, $l(l(\widetilde{E}_N))$, $l(l(\widetilde{F}_N))$ have negative assortativity.

4. Conclusion

Topological characteristics of links influence the dynamical processes executed on complex networks triggered by links. The line graph, which transforms links from a graph to nodes in its line graph, generalizes the topological properties from nodes to links. This paper investigates the degree distribution and the assortativity of line graphs. The degree distribution of the line graph of an Erdős–Rényi random graph follows the same pattern of the degree distribution as the original graph. We derive a formula for the assortativity of the line graph. We indicate that the assortativity of the line graph is not linearly related to the assortativity of the original graph. Moreover, the assortativity is positive for the line graphs of Erdős–Rényi graphs, Barabási–Albert graphs and most real-world networks. In contrast, certain types of trees, path and path-like graphs, have negative assortativity in their line graphs. Furthermore, non-trees consisting of cycles and paths can also have negative assortativity in their line graphs.

Acknowledgment

Appendix A. Proof of Eq. (6)

The degree distribution in scale free graphs G is

$$\Pr[D = k] = \frac{k^{-\alpha}}{c_1}, \quad k = s, \ldots, K \tag{A.1}$$

where $c_1 = \sum_{k=s}^{K} k^{-\alpha}$ is the normalization constant and s is the minimum degree and K is the maximum degree in G. Assuming the node degrees in the scale free graph are independent, the generating function for the line graph of scale free graphs can be written as Eq. (5). Substituting the derivative of the generating function $\varphi'_D(z) = \frac{1}{E[D]} \sum_{k=0}^{N-1} k z^{k-1} \Pr[D = k]$ and the average degree $E[D] = \sum_{k=0}^{N-1} k \Pr[D = k] = \frac{c_2}{c_1}$, where $c_2 = \sum_{k=s}^{K} k^{1-\alpha}$, into Eq. (5) yields

$$\varphi_{D_l}(z) = \left(\frac{c_1}{c_2}\right)^2 \left(\varphi'_D(z)\right)^2 \tag{A.2}$$

and the Taylor coefficients obey

$$\Pr[D_l = k] = \left(\frac{c_1}{c_2}\right)^2 \frac{1}{k!} \frac{d^k \left(\varphi'_D(z)\right)^2}{dz^k}\bigg|_{z=0}.$$

Using the Leibniz's rule $(fg)^{(k)} = \sum_{m=0}^{k} \binom{k}{m} f^{(m)} g^{(k-m)}$, where $f = g = \varphi'_D(z)$, yields

$$\Pr[D_l = k] = \left(\frac{c_1}{c_2}\right)^2 \frac{1}{k!} \sum_{m=0}^{k} \binom{k}{m} \frac{d^{m+1}(\varphi_D(z))}{dz^{m+1}} \frac{d^{k-m+1}(\varphi_D(z))}{dz^{k-m+1}}\bigg|_{z=0}.$$

Substituting $k! \Pr[D = k] = \frac{d^k(\varphi_D(z))}{dz^k}\bigg|_{z=0}$, we arrive at

$$\Pr[D_l = k] = \frac{1}{k!} \left(\frac{c_1}{c_2}\right)^2 \sum_{m=0}^{k} \frac{k!}{m!(k-m)!} (m+1)! \Pr[D = m+1](k-m+1)! \Pr[D = k-m+1].$$

Applying the power law degree distribution in Eq. (A.1), we have

$$\Pr[D_l = k] = \frac{1}{c_2^2} \sum_{m=1}^{k+1} \left(m(k+2-m)\right)^{1-\alpha}. \tag{A.3}$$

For $\alpha = 3$, we transform Eq. (A.3) in the following form:

$$c_2^2 \Pr[D_l = k] = \frac{1}{(k+2)^3} \sum_{i=1}^{k+1} \frac{1}{\left(\frac{i}{k+2}\right)^2 \left(1 - \frac{i}{k+2}\right)^2} \frac{1}{k+2}. \tag{A.4}$$

We use the following expression between a sum in the limit to infinity and a definite integral [28]

$$\int_a^b f(x) dx = \lim_{n \to \infty} \sum_{k=1}^{n} f(x_k) \Delta x.$$

We set $\Delta x = \frac{1}{k+2}, x_i = i \Delta x = \frac{i}{k+2}, f(x) = \frac{1}{x^2(1-x)^2}$ and (A.4) boils down to

$$c_2^2 \Pr[D_l = k] = \frac{1}{(k+2)^3} \sum_{i=1}^{k+1} f(x_i) \Delta x. \tag{A.5}$$

We consider the case of limit to infinity for k ($k \to \infty$) or k very large and evaluate the sum $\sum_{i=1}^{k+1} f(x_i) \Delta x$, which can be transformed into

$$\sum_{i=1}^{k+1} f(x_i) \Delta x \approx \int_{\frac{1}{k+2}}^{\frac{k+1}{k+2}} f(x) dx. \tag{A.6}$$

Now,

$$\int_{\frac{1}{k+2}}^{\frac{k+1}{k+2}} f(x) dx = \int_{\frac{1}{k+2}}^{\frac{k+1}{k+2}} \frac{1}{x^2(1-x)^2} dx$$

$$= \int_{\frac{1}{k+2}}^{\frac{k+1}{k+2}} \left(\frac{2}{x} + \frac{2}{1-x} + \frac{1}{x^2} + \frac{1}{(1-x)^2}\right) dx$$

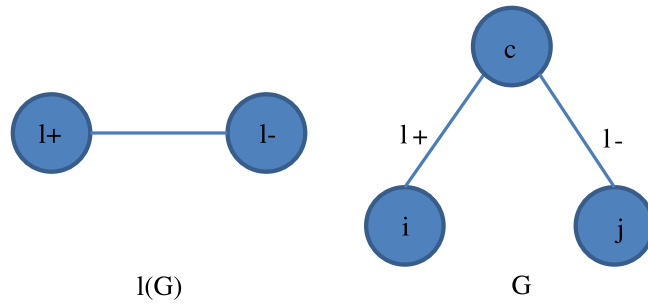$$= 2\left(2 \ln(k+1) + \frac{k(k+2)}{k+1}\right). \tag{A.7}$$

Fig. B.13. Link transformation.

Using (A.7) and (A.6) into (A.5), leads to

$$c_2^2 \Pr[D_l = k] \approx \frac{2}{(k+2)^2}\left(\frac{2\ln(k+1)}{k+2} + \frac{k}{k+1}\right).$$

Since $\lim_{k\to\infty} \frac{\ln(k+1)}{k+2} = 0$ and $\lim_{k\to\infty} \frac{k}{k+1} = 1$, we arrive at

$$\Pr[D_l = k] \approx \frac{1}{c_2^2}(k+2)^{-2}. \tag{A.8}$$

Appendix B. Proof for Theorem 3

Proof. A link l with end nodes l^+ and l^- in the line graph $l(G)$ corresponds to a connected triplet in G. Without loss of generality, we assume that nodes l^+ and l^- in the line graph correspond to links $l^+ = i \sim c$ and $l^- = j \sim c$, where links $i \sim c$ and $j \sim c$ share a common node c, in the original graph as shown in Fig. B.13. The degree in line graph is $d_{l^+} = d_i + d_c - 2$ and $d_{l^-} = d_j + d_c - 2$. Since subtracting 2 everywhere does not change the linear correlation coefficient, we proceed with $d_{l^+} = d_i + d_c$ and $d_{l^-} = d_j + d_c$. First, we compute the joint expectation

$$E[D_{l^+}D_{l^-}] = \frac{\displaystyle\sum_{i=1}^N \sum_{\substack{j=1 \\ j\neq i}}^N \sum_{c=1}^N (d_i + d_c)(d_j + d_c)a_{ic}a_{jc}}{2L_{l(G)}}$$

$$= \frac{\displaystyle\sum_{i=1}^N \sum_{j=1}^N \sum_{c=1}^N (d_i + d_c)(d_j + d_c)a_{ic}a_{jc} - \sum_{i=1}^N \sum_{c=1}^N (d_i + d_c)^2 a_{ic}^2}{2L_{l(G)}}$$

$$= \frac{\displaystyle\sum_{i=1}^N \sum_{j=1}^N \sum_{c=1}^N d_i a_{ic}a_{jc}d_j + 2\sum_{i=1}^N \sum_{j=1}^N \sum_{c=1}^N d_i a_{ic}a_{jc}d_c + \sum_{i=1}^N \sum_{j=1}^N \sum_{c=1}^N d_c^2 a_{ic}a_{jc}}{2L_{l(G)}} - \frac{2\displaystyle\sum_{i=1}^N \sum_{c=1}^N d_i^2 a_{ic}^2 + 2\sum_{i=1}^N \sum_{c=1}^N d_i a_{ic}^2 d_c}{2L_{l(G)}}.$$

With $\sum_{j=1}^N a_{jc} = d_c$, we arrive at

$$E[D_{l^+}D_{l^-}] = \frac{d^T A^2 d + 2d^T A\Delta d + \displaystyle\sum_{i=c}^N d_c^4 - 2\sum_{i=1}^N d_i^3 - 2d^T Ad}{2L_{l(G)}}. \tag{B.1}$$

The average degree $E[D_{l^+}] = E[D_i + D_c]$ is the average degree of two connected nodes i and c from a triplet (see Fig. B.13) in the original graph. Thus,

$$E[D_{l^+}] = \frac{\displaystyle\sum_{i=1}^N \sum_{\substack{j=1 \\ j\neq i}}^N \sum_{c=1}^N (d_i + d_c)a_{ic}a_{jc}}{2L_{l(G)}}$$

$$= \frac{\displaystyle\sum_{i=1}^N \sum_{j=1}^N \sum_{c=1}^N d_i a_{ic}a_{jc} + \sum_{i=1}^N \sum_{j=1}^N \sum_{c=1}^N d_c a_{ic}a_{jc} - \sum_{i=1}^N \sum_{c=1}^N d_i a_{ic}^2 - \sum_{i=1}^N \sum_{c=1}^N d_c a_{ic}^2}{2L_{l(G)}}$$

from which

$$E[D_{l+}] = \frac{d^T A d + \sum\limits_{c=1}^{N} d_c^3 - 2d^T d}{2L_{l(G)}}. \tag{B.2}$$

The variance $\sigma_{D_{l+}}^2 = \mathrm{Var}[D_{l+}] = E[D_{l+}^2] - (E[D_{l+}])^2$ and

$$E[D_{l+}^2] = \frac{\sum\limits_{i=1}^{N} \sum\limits_{\substack{j=1 \\ j \neq i}}^{N} \sum\limits_{c=1}^{N} (d_i + d_c)^2 a_{ic} a_{jc}}{2L_{l(G)}}$$

$$= \frac{\sum\limits_{i=1}^{N} \sum\limits_{j=1}^{N} \sum\limits_{c=1}^{N} d_i^2 a_{ic} a_{jc} + 2\sum\limits_{i=1}^{N} \sum\limits_{j=1}^{N} \sum\limits_{c=1}^{N} d_i a_{ic} a_{jc} d_c + \sum\limits_{i=1}^{N} \sum\limits_{j=1}^{N} \sum\limits_{c=1}^{N} d_c^2 a_{ic} a_{jc}}{2L_{l(G)}} - \frac{2\sum\limits_{i=1}^{N} \sum\limits_{c=1}^{N} d_i^2 a_{ic}^2 + 2\sum\limits_{i=1}^{N} \sum\limits_{c=1}^{N} d_i a_{ic}^2 d_c}{2L_{l(G)}}$$

which we rewrite as

$$E[D_{l+}^2] = \frac{3d^T A \Delta d + \sum\limits_{c=1}^{N} d_c^4 - 2\sum\limits_{i=1}^{N} d_i^3 - 2d^T A d}{2L_{l(G)}}. \tag{B.3}$$

The number of links $L_{l(G)}$ in a line graph is [4]

$$L_{l(G)} = \frac{1}{2} d^T d - L = \frac{1}{2}(N_2 - N_1). \tag{B.4}$$

After substituting Eqs. (B.1)–(B.4) into (7), we establish the theorem. □

Appendix C. Proof for Corollary 1

Proof. Using the variance $\sigma_{D_{l+}}^2 = \mathrm{Var}[D_{l+}] = E[D_{l+}^2] - (E[D_{l+}])^2$, we rewrite the definition of assortativity (7) as

$$\rho_{D_{l(G)}} = 1 + \frac{E[D_{l+} D_{l-}] - E[D_{l+}^2]}{\sigma_{D_{l+}}^2}. \tag{C.1}$$

According to Eqs. (B.1) and (B.3), we have that

$$E[D_{l+} D_{l-}] - E[D_{l+}^2] = \frac{d^T A^2 d - d^T A \Delta d}{2L_{l(G)}}$$

$$= \frac{N_4 - d^T A \Delta d}{2L_{l(G)}}. \tag{C.2}$$

The variance $\mathrm{Var}[D_{l+}]$ of the end node of an arbitrarily chosen link can be written in terms of the variance $\mathrm{Var}[D_{l(G)}]$ of an arbitrarily chosen node [29]

$$\sigma_{D_{l+}}^2 = \frac{\mu u_3 - (\mathrm{Var}[D_{l(G)}])^2 + \mu^2 \mathrm{Var}[D_{l(G)}]}{\mu^2} \tag{C.3}$$

where $\mu = E[D_{l(G)}]$ and $u_3 = E[(D_{l(G)} - E[D_{l(G)}])^3]$. The variance $\mathrm{Var}[D_{l(G)}]$ of an arbitrarily chosen node can be written in terms of the assortativity [4]

$$\mathrm{Var}[D_{l(G)}] = 2(1 + \rho_D)\left(\frac{1}{N_1}\sum\limits_{i=1}^{N} d_i^3 - \left(\frac{N_2}{N_1}\right)^2\right). \tag{C.4}$$

Substituting (C.2)–(C.4) into (C.1), we prove the Corollary 1. □

References

[1] R. Albert, I. Albert, G.L. Nakarado, Structural vulnerability of the North American power grid, Phys. Rev. E 69 (2004) 025103.
[2] R. Cohen, K. Erez, D. ben Avraham, S. Havlin, Resilience of the Internet to random breakdowns, Phys. Rev. Lett. 85 (2000) 4626–4628.
[3] S.V. Buldyrev, R. Parshani, G. Paul, H.E. Stanley, S. Havlin, Catastrophic cascade of failures in interdependent networks, Nature 464 (7291) (2010) 1025–1028.
[4] P. Van Mieghem, Graph Spectra for Complex Networks, Cambridge University Press, Cambridge, UK, 2011.
[5] M.J. Krawczyk, L. Muchnik, A. Manka-Krason, K. Kulakowski, Line graphs as social networks, Physica A 390 (2011) 2611–2618.
[6] J.B. Pereira-Leal, A.J. Enright, C.A. Ouzounis, Detection of functional modules from protein interaction networks, Proteins: Struct. Funct. Bioinform. 54 (1) (2004) 49–57.
[7] R. Diestel, Graph Theory, Grad. Texts in Math., 1997.
[8] T.S. Evans, R. Lambiotte, Line graphs, link partitions, and overlapping communities, Phys. Rev. E 80 (1) (2009) 016105.
[9] J.C. Wierman, D.P. Naor, J. Smalletz, Incorporating variability into an approximation formula for bond percolation thresholds of planar periodic lattices, Phys. Rev. E 75 (1) (2007) 011114.
[10] H. Whitney, Congruent graphs and the connectivity of graphs, Amer. J. Math. 54 (1932) 150–168.
[11] J. Krausz, Démonstration nouvelle d'un théorème de Whitney sur les réseaux, Mat. Fiz. Lapok 50 (1943) 75–85.
[12] A.C.M. van Rooij, H.S. Wilf, The interchange graph of a finite graph, Acta Math. Acad. Sci. Hungar. 16 (1965) 263–269.
[13] F. Harary, Graph Theory, Addison-Wesley, 1969.
[14] D. Liu, S. Trajanovski, P. Van Mieghem, Random line graphs and a linear law for assortativity, Phys. Rev. E 87 (1) (2013) 012816.
[15] N.D. Roussopoulos, A $max\{m, n\}$ algorithm for detecting the graph h from its line graph g, Inform. Process. Lett. 2 (1973) 108–112.
[16] P.G.H. Lehot, An optimal algorithm to detect a line graph and output its root graph, J. ACM 21 (1974) 569–575.
[17] D. Liu, S. Trajanovski, P. Van Mieghem, ILIGRA: an efficient inverse line graph algorithm, J. Math. Model. Algorithms Oper. Res. 14 (1) (2014) 13–33.
[18] P. Van Mieghem, Performance Analysis of Complex Networks and Systems, Cambridge University Press, 2014.
[19] E.R. van Dam, W.H. Haemers, Which graphs are determined by their spectrum? Linear Algebra Appl. 373 (2003) 241–272.
[20] P. Van Mieghem, H. Wang, X. Ge, S. Tang, F.A. Kuipers, Influence of assortativity and degree-preserving rewiring on the spectra of networks, Eur. Phys. J. B 76 (4) (2010) 643–652.
[21] M.E.J. Newman, Assortative mixing in networks, Phys. Rev. Lett. 89 (20) (2002) 208701.
[22] R. Noldus, P. Van Mieghem, Assortativity in complex networks, J. Complex Netw. (2015).
[23] M. Youssef, Y. Khorramzadeh, S. Eubank, Network reliability: the effect of local network structure on diffusive processes, Phys. Rev. E 88 (5) (2013) 052810.
[24] M.E.J. Newman, Finding community structure in networks using the eigenvectors of matrices, Phys. Rev. E 74 (3) (2006) 036104.
[25] V. Colizza, R. Pastor-Satorras, A. Vespignani, Reaction–diffusion processes and metapopulation models in heterogeneous networks, Nat. Phys. 3 (4) (2007) 276–282.
[26] R.E. Kooij, A. Jamakovic, F. van Kesteren, T. de Koning, I. Theisler, P. Veldhoven, The Dutch soccer team as a social network, Connections 29 (1) (2009) 4–14.
[27] E.R. van Dam, R.E. Kooij, The minimal spectral radius of graphs with a given diameter, Linear Algebra Appl. 423 (2) (2007) 408–419.
[28] P.J. Davis, P. Rabinowitz, Methods of Numerical Integration, Courier Corporation, 2007.
[29] H. Wang, W. Winterbach, P. Van Mieghem, Assortativity of complementary graphs, Eur. Phys. J. B 83 (2) (2011) 203–214.