Physica A 540 (2020) 123220

Contents lists available at ScienceDirect

Physica A

journal homepage: www.elsevier.com/locate/physa

Prevalence expansion in NIMFA

Zhidong He *, Piet Van Mieghem

Faculty of Electrical Engineering, Mathematics and Computer Science, P.O Box 5031, 2600 GA Delft, The Netherlands

ARTICLE INFO

Article history:
Received 19 February 2019
Received in revised form 2 October 2019
Available online 21 October 2019

Keywords:
SIS prevalence
NIMFA
Taylor expansion
Radius of convergence

ABSTRACT

The N-Intertwined Mean Field Approximation (NIMFA) is a reasonably accurate approximation of the exact SIS epidemic process on a network. The average fraction of infected nodes in the NIMFA steady state, also called the steady-state prevalence, in terms of the effective infection rate can be expanded into a power series around the NIMFA epidemic threshold. In this paper, we investigate the convergence of the steady-state prevalence Taylor expansion. We determine the radius of convergence in some special types of graphs. We also show that the radius of convergence of the steady-state prevalence expansion depends upon the network topology, in particular, the average degree of the network and the spectral gap of the adjacency matrix play a role.

1. Introduction

Epidemic models can describe virus spreading and information propagation in human activities [1]. The Susceptible–Infected–Susceptible (SIS) model is an epidemic model where each infected item can be cured, and becomes susceptible again after recovering from the infection state [2]. In the SIS epidemics on a network, the ratio between the infection rate β and the curing rate δ is called the effective infection rate $\tau = \beta/\delta$. The SIS model features a phase transition [3] around the epidemic threshold τ_c. The spreading process can reach the metastable state [4] if the effective infection rate τ is above the epidemic threshold τ_c. In the metastable state, the viruses can infect a sizeable portion of the population on average and stay in the network for a long time [5]. A first-order mean-field approximation [6] of the epidemic threshold $\tau_c^{(1)} = 1/\lambda_1$, where λ_1 is the spectral radius of the adjacency matrix A, was shown [7] to be a lower bound for the exact epidemic threshold τ_c.

The exact Markovian SIS model [6] in the network G with N nodes consists of 2^N states, which is intractable to solve for large networks. The N-Intertwined Mean-Field Approximation [8] was proposed by introducing the network topology into the deterministic model [9,10], which can approximate the exact SIS epidemics well in some networks. NIMFA approximates the exact Markovian 2^N linear equations into N non-linear differential equations under the assumption that the states of the nodes are uncorrelated. The steady-state infection probabilities $v_{i\infty}(\tau)$ of each node i in NIMFA can be expanded [11] in a power series in terms of the effective infection rate τ at NIMFA epidemic threshold $\tau_c^{(1)}$ and explicitly repeated in Appendix A. Mathematically, the radius of convergence of a power series corresponds to the radius of the largest disk in which the series converges [12], which is of practical significance to validate a Taylor expansion. Practically, we can faster compute the nodal infection probability of the NIMFA steady-state by the truncated expansion with enough terms and an effective infection rate τ within the radius of convergence, instead of numerically solving the governing equation (1). Thus, the radius of convergence of the series, that purely depends upon the underlying topology,

* Corresponding author.
E-mail addresses: Z.He@tudelft.nl (Z. He), P.F.A.VanMieghem@tudelft.nl (P. Van Mieghem).

https://doi.org/10.1016/j.physa.2019.123220
0378-4371/

determines the largest effective infection rate for the Taylor expansion. However, the convergence of the Taylor expansion (A.1) in Appendix A of the nodal infection probability in NIMFA is still an open question and has not been studied yet.

In this paper, we focus on the NIMFA steady-state prevalence, which is defined as the average fraction of infected nodes in the NIMFA steady state. We numerically investigate the radius of convergence of the steady-state prevalence expansion and illustrate that the convergence of the prevalence expansion highly depends on the underlying topology. Specifically, we investigate the behaviour of the radius of convergence in some special types of graphs, e.g., regular graph, star graph, path graph, Erdős–Rényi (ER) random graph and scale-free graph, which helps to estimate the valid range of the effective infection rate τ for the steady-state NIMFA prevalence expansion (4). We also identify the topological properties that influence the radius of convergence in sparse networks and clustered networks.

The outline of the paper is as follows. In Section 2, we briefly review the NIMFA SIS epidemics, provide related notations and describe the method for practical estimation of the radius of convergence. Section 3 investigates the behaviour of the radius of convergence in some special types of graphs. In Section 4, we investigate the topological properties that dominate the radius of convergence. Finally, we summarize the paper in Section 5.

2. Expansion of the NIMFA steady-state prevalence

2.1. NIMFA prevalence

The probability $v_i(t)$ that node i is infected is given in NIMFA [8] by the following first-order nonlinear differential equation

$$\frac{dv_i(t)}{dt} = \beta \sum_{j=1}^{N} a_{ij} v_j(t) - v_i(t)\Big(\beta \sum_{j=1}^{N} a_{ij} v_j(t) + \delta\Big).$$

where $a_{ij} \in \{0, 1\}$ is the entry of the adjacency matrix A of the underlying graph. If $a_{ij} = 1$, then there is a link between node i and node j, and otherwise $a_{ij} = 0$. With the infection probability vector $\boldsymbol{v}(t) = (v_1(t), v_2(t), \ldots, v_N(t))^T$, the governing equation of NIMFA in matrix form [2] is

$$\frac{d\boldsymbol{v}(t)}{dt} = \beta A\boldsymbol{v}(t) - \mathrm{diag}\big(\boldsymbol{v}(t)\big)\Big(\beta A\boldsymbol{v}(t) + \delta u\Big) \tag{1}$$

where $\mathrm{diag}\big(\boldsymbol{v}(t)\big)$ is a diagonal matrix with the infection probability vector $\boldsymbol{v}(t)$ and u is the all-one vector. By solving the governing equation (1) with the initial state $\boldsymbol{v}(0) = (v_1(0), v_2(0), \ldots, v_N(0))^T$, we can obtain the NIMFA prevalence, i.e. the average fraction of infected nodes, as $y(t) = \frac{1}{N}\sum_{i=1}^{N} v_i(t)$.

We denote the eigenvalues of the adjacency matrix A by $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_N$, and denote by x_k the eigenvector, normalized by $x_k^T x_k = 1$, corresponding to eigenvalue λ_k. The infection probabilities $v_{i\infty}$ of node i in the steady state are non-zero if the effective infection rate $\tau = \frac{\beta}{\delta}$ is above the NIMFA epidemic threshold $\tau_c^{(1)} = 1/\lambda_1$. For simplicity without the loss of generality, we set the curing rate equal to $\delta = 1$, and then the infection rate equals $\beta = \tau$. In this paper, we focus on the steady-state prevalence $y_\infty(\tau) = \frac{1}{N}\sum_{i=1}^{N} v_{i\infty}(\tau)$ in terms of the effective infection rate τ.

2.2. Expansion of the NIMFA steady-state prevalence

Before discussing the Taylor expansion of the NIMFA steady-state prevalence $y_\infty(\tau)$, we note that the Taylor series of the steady-state infection probability $v_{i\infty}(\tau)$ in terms of the effective infection rate τ about the epidemic threshold $\tau_c^{(1)}$ does not exist in physical space. The physical space here defines the space of all steady-state infection probability $v_{i\infty}(\tau)$ as solution of (1) that are possible (i.e. a probability must be non-negative). Van Mieghem [6] obtains that the steady-state infection probability $v_{i\infty}(\tau) = 0$ for any effective infection rate $\tau \leq \tau_c^{(1)}$. The derivative of the steady-state infection probability $v_{i\infty}(\tau)$ with respect to the effective infection rate τ on the left-side of $\tau_c^{(1)}$ follows $\frac{dv_{i\infty}(\tau)}{d\tau}\big|_{\tau=(\tau_c^{(1)})^-} = 0$. Invoking the fact [11] that $\frac{dv_i(t,\tau)}{d\tau}\big|_{\tau=(\tau_c^{(1)})^+} = \frac{\lambda_1}{N}\frac{(x_1)_i}{\sum_{i=1}^{N}(x_1)_i^3} > 0$ for the effective infection rate $\tau > \tau_c^{(1)}$, left and right derivatives are not equal, implying that the infection probability $v_{i\infty}(\tau)$ is not an analytic function of the effective infection rate τ at epidemic threshold $\tau_c^{(1)}$. Consequently, the Taylor series does not exist. In other words, the radius of convergence R of the expansion of the NIMFA steady-state prevalence equals $R = 0$.

The expansion of the steady-state infection probability with the effective infection rate $\tau \downarrow \tau_c^{(1)}$ can be derived from

$$\beta A\boldsymbol{v}_\infty(t) - \mathrm{diag}\big(\boldsymbol{v}_\infty(t)\big)\Big(\beta A\boldsymbol{v}_\infty(t) + \delta\boldsymbol{u}\Big) = 0 \tag{2}$$

for the effective infection rate $\tau \to \tau_c^{(1)}$. Mathematically, we extend the solution of (2) to complex τ. Hence, the steady-state infection probability $v_{i\infty}(\tau)$ is allowed to be negative for $\tau < \tau_c^{(1)}$, as illustrated in Fig. 1. Thus, the infection probability $v_{i\infty}(\tau)$ in the extended mathematical space is analytic at $s = \tau^{-1} = (\tau_c^{(1)})^{-1} = \lambda_1$, where the mathematical space includes all solutions of Eq. (2) for $v_{i\infty}(\tau)$, regarded as a complex function of τ, ignoring its probabilistic and physical meaning. Although the negative steady-state infection probability has no physical significance, the Taylor series of the

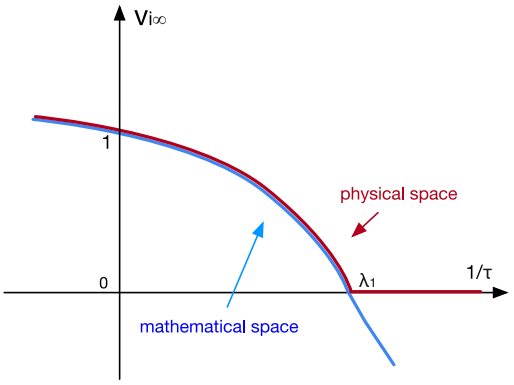


Fig. 1. Illustration of the steady-state infection probability $v_{i\infty}$ of node i as a function of $s = \tau^{-1}$, where node i is an arbitrary node in the network. The red line represents the steady-state infection probability in the physical space, and the blue line represents the extended solution of (2) in the mathematical space.

steady-state infection probability in $z = (\tau_c^{(1)})^{-1} - \tau^{-1}$ around $z = 0$ still coincides with the NIMFA positive infection probability for the effective infection rate $\tau > \tau_c^{(1)}$, as shown in Fig. 1, by analytic continuation [12, Chap. IV].

The NIMFA steady-state infection probability vector $\boldsymbol{v}_\infty(\tau)$ can be expanded in a Taylor series around $z = 0$ as

$$\boldsymbol{v}_\infty(\tau) = \sum_{j=1}^{\infty} \boldsymbol{\alpha}_j z^j = \sum_{j=1}^{\infty} \boldsymbol{\alpha}_j ((\tau_c^{(1)})^{-1} - \tau^{-1})^j. \tag{3}$$

where $z := (\tau_c^{(1)})^{-1} - \tau^{-1}$. The coefficient vector $\boldsymbol{\alpha}_j$ in the expansion equals $\boldsymbol{\alpha}_j = \sum_{k=1}^{N} c_j(k) x_k = X\boldsymbol{c}_j$, where the matrix X contain as columns the eigenvector of the adjacency matrix A, and $c_j(k)$ satisfy [11] the recursions (A.2)–(A.4) in Appendix A. We ignore the trivial solution $v_{i\infty}(\tau) = 0$ with all-zero coefficients $c_j(k)$ and focus on the non-zero infection probability. The expansion of the steady-state infection probability $v_{i\infty}(\tau)$ with the recurrence relation of the coefficients $c_j(k)$ are derived in [11] and revisited (A.2)–(A.4) in Appendix A. Thus, the expansion of the NIMFA steady-state prevalence can be represented by

$$y_\infty(\tau) = \frac{1}{N} u^T \boldsymbol{v}_\infty(\tau) = \sum_{j=1}^{\infty} b_j z^j \tag{4}$$

where the coefficient $b_j = \frac{1}{N} u^T \boldsymbol{\alpha}_j$, and u is the all-one vector.

We define the truncated Taylor series with J terms as

$$y_\infty^{(J)}(\tau) = \sum_{j=1}^{J} b_j z^j \tag{5}$$

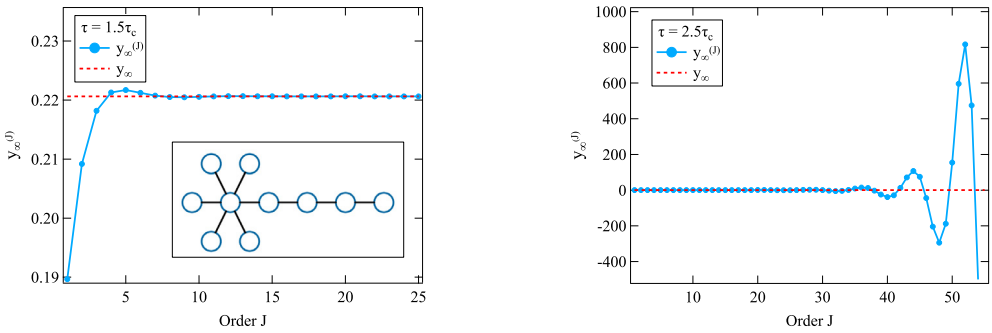
and investigate the accuracy and convergence of the expansion $y_\infty^{(J)}(\tau)$. The steady-state prevalence expansion $y_\infty(\tau)$ converges, if for every arbitrarily small number $\epsilon > 0$, there exists an order J_c such that $|y_\infty^{(J)} - y_\infty| < \epsilon$ for all $J > J_c$. The radius of convergence R is a nonnegative real number or ∞ such that the series (4) converges if $|z| = |(\tau_c^{(1)})^{-1} - \tau^{-1}| < R$ and diverges if $|z| = |(\tau_c^{(1)})^{-1} - \tau^{-1}| > R$ in the complex z-plane. Fig. 2 shows the NIMFA steady-state prevalence expansion $y_\infty^{(J)}(\tau)$ as a function of J for two different effective infection rate τ in an example graph. Fig. 2 illustrates that the prevalence expansion $y_\infty^{(J)}(\tau)$ converges to the NIMFA steady-state prevalence $y_\infty(\tau)$ if the effective infection rate τ is small (e.g., $\tau < 1.5\tau_c^{(1)}$), but diverges if τ is large enough (e.g., $\tau > 2.5\tau_c^{(1)}$) along with a large term z.

2.3. Estimation of the radius of convergence

For the power series $y_\infty(\tau) = \sum_{j=1}^{\infty} b_j z^j$ where $z = (\tau_c^{(1)})^{-1} - \tau^{-1}$ expanded at $z = 0$, the radius of convergence R in the complex z-plane is given by [12]

$$R^{-1} = \lim_{j\to\infty} \left| \frac{b_{j+1}}{b_j} \right| = \limsup_{j\to\infty} \sqrt[j]{|b_j|} \tag{6}$$

Unfortunately, the radius of convergence R of the prevalence expansion in a general graph is intractable to derive analytically. The method of Domb–Sykes plot and Mercer–Roberts plots [13], which are usually applied to estimate the



(a) $y_\infty^{(J)}(\tau)$ with $\tau = 1.5\tau_c^{(1)}$ (b) $y_\infty^{(J)}(\tau)$ with $\tau = 2.5\tau_c^{(1)}$

Fig. 2. Steady-state prevalence expansion $y_\infty^{(J)}(\tau)$ as a function of order J. The example graph topology is illustrated in the subplot (a). The red dash lines represent the NIMFA steady-state prevalence $y_\infty(\tau)$ following the NIMFA equation (2).

radius of convergence of a series if the sign of the coefficients b_j follows a pattern, is also impractical since the coefficients b_j could present an unpredictable sign pattern in a general graph.

In the following sections, we assess the convergence of the expansion numerically. We regard that the prevalence expansion (4) is convergent if the prevalence $y_\infty^{(J)}$ with a large enough order J approximates the NIMFA steady-state prevalence y_∞^* solved by Eqs. (1), i.e., $|y_\infty^{(J)} - y_\infty| < \epsilon$ for all $J > J_c$. In the following simulations, we set $J_c = 150$ and $\epsilon = 0.001 y_\infty$. We have numerically verified that the setting of $J_c = 150$ is large enough to assess the convergence of the expansions in the following simulations.

3. Radius of convergence R of the expansion in some specific graphs

We first investigate the behaviour of the coefficients b_j and the radius of convergence R of the expansion of the steady-state prevalence in some special types of graphs, e.g., complete graphs, regular graphs, star graphs, path graphs, Erdős–Rényi random graphs and scale-free graphs.

3.1. Complete graphs and regular graphs

In the complete graph K_N, the NIMFA steady-state prevalence is

$$y_\infty(s) = 1 - \frac{1}{\tau(N-1)} = \frac{1}{\lambda_1}((\tau_c^{(1)})^{-1} - s) \tag{7}$$

where $s = \tau^{-1}$. The prevalence (7) is analytic in the whole complex plane. Hence, the radius of convergence R is infinity. The coefficient $b_1 = \frac{\sum_{i=1}^N (x_1)_i}{N} c_1(1) = (\lambda_1 \sum_{i=1}^N (x_1)_i^3)^{-1} = \frac{1}{N-1}$, and $b_j = 0$ for $j > 1$, which is consistent with (7). Fig. 3a illustrates that the exact solutions of the steady-state governing equation (2) in a complete graph has two branches, and the branch corresponding to the non-trivial real solution (red line) is analytic for the reciprocal of any effective infection rate $s = \tau^{-1}$.

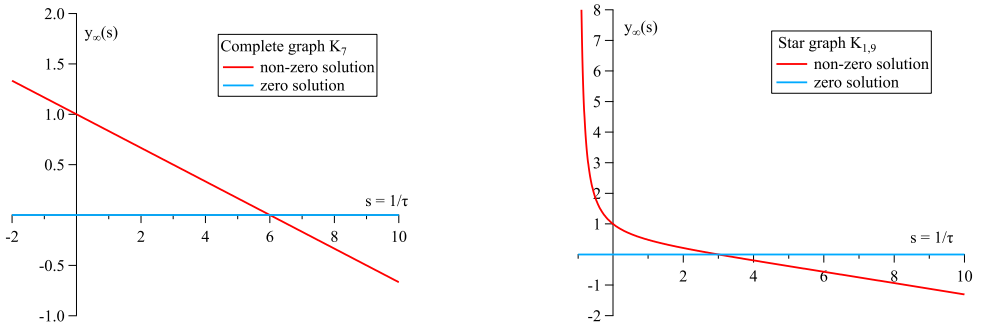
In the r-regular graph with the average degree $E[D] = r$, the steady-state prevalence is $y_\infty(s) = \frac{1}{r}((\tau_c^{(1)})^{-1} - s)$, which implies the steady-state prevalence expansion (4) also has an infinite radius of convergence.

3.2. Star graphs $K_{1,N-1}$

In the star graph $K_{1,N-1}$ with $N-1$ leaf, the NIMFA steady-state infection probability [14] for the centre node equals $v_{c\infty} = \frac{(N-1)-s^2}{(N-1)+s}$ and $v_{c\infty} = 0$, while the NIMFA infection probability for the leaf node equal $v_{l\infty} = \frac{(N-1)-s^2}{(N-1)(s+1)}$ and $v_{l\infty} = 0$, where $s = \tau^{-1}$. Thus, the steady-state prevalence follows $y_\infty = \frac{1}{N}\left(\frac{(N-1)-s^2}{(N-1)+s} + \frac{(N-1)-s^2}{(s+1)}\right)$, which has the nearest singularity $s = -1$. The radius of convergence R of the expansion (4) equals $R = \sqrt{N-1} + 1$ in a star graph, which implies the expansion (4) converges for any effective infection rate $\tau > \tau_c^{(1)}$. Fig. 3b illustrates the NIMFA prevalence $y_\infty(\tau)$ in a star graph, where the non-zero solution corresponds to the expansion (2) has the singularity at $s = -1$.

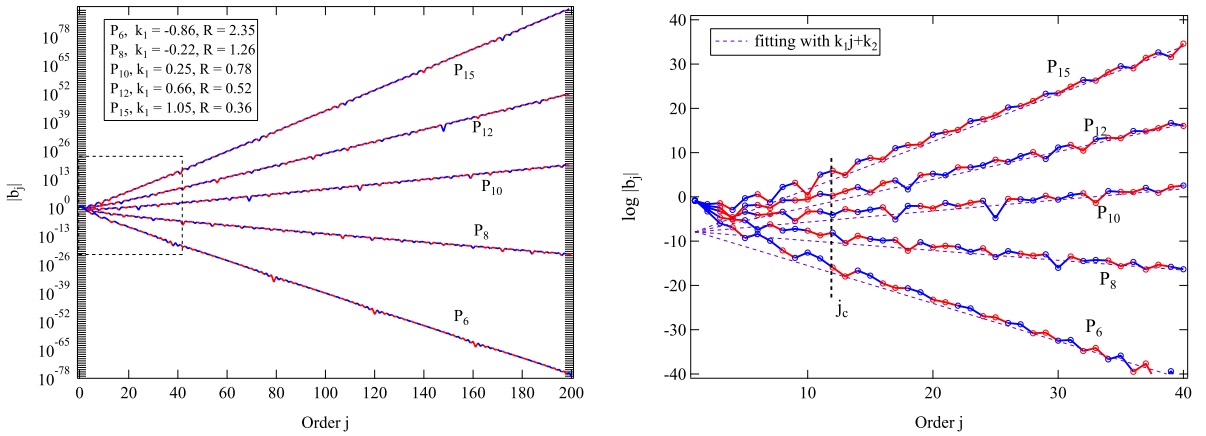
3.3. Path graphs P_N

Since the computation of the NIMFA prevalence in a path graph is intractable, we numerically investigate the radius of convergence of the prevalence expansion. Fig. 4 shows the function of the absolute coefficients $|b_j|$ in the prevalence



(a) y^*_∞ in K_7 (b) y^*_∞ in $K_{1,9}$

Fig. 3. Exact steady-state prevalence y^*_∞ by the solution of Eq. (2) in the complete graph K_7 and the star graph $K_{1,9}$.



(a) (b)

Fig. 4. Coefficient $|b_j|$ as a function of order j in path graphs P_N with N nodes. The red markers represent positive coefficients, and the blue markers represent negative. The function $|b_j|$ is fitted by the function $e^{k_1 j+k_2}$. Subplot (b) shows the function $\log|b_j|$ for small order j.

expansion in path graphs with different network size N, while the period and the sign pattern of the coefficients b_j is too complicated to predict. Fig. 4b shows that the absolute coefficients $|b_j|$ resembles an exponential function $|b_j| \approx e^{k_1 j+k_2}$ if the order j is larger than a critical order j_c. Hence, the radius of convergence R can be estimated by $R \approx e^{-k_1}$, which is almost the same as the radius of convergence estimated in Section 2.3. Further, the NIMFA prevalence with $z = (\tau_c^{(1)})^{-1} - \tau^{-1} < e^{-k_1}$ can be estimated by

$$y^*_\infty(\tau) \approx \sum_{j=1}^{j_c-1} b_j z^j + \sum_{j=j_c}^{\infty} e^{k_1 j+k_2} z^j = \sum_{j=1}^{j_c-1} b_j z^j + e^{k_2} \frac{(e^{k_1} z)^{j_c}}{1 - e^{k_1} z} \tag{8}$$

Fig. 5b shows that the radius of convergence R decreases with increasing network size N in path graphs P_N but not monotonically, i.e., R reaches the maxima at $N = 5$. The radius of convergence R decays as an exponential function of the network size if $N > 8$, and tends to be very small in path networks with a large size, which implies that the effective infection rate τ must be extremely close to the epidemic threshold $\tau_c^{(1)}$ to converge the prevalence expansion (4) in large path networks.

3.4. ER random graphs

We first present the behaviour of the coefficient b_j in the prevalence expansion in ER random graphs $G_p(N)$ with link density p. Fig. 6 shows that the coefficients b_j can be negative for a small link density p. The coefficients b_j are always positive if the link density p is large enough, which implies that the NIMFA prevalence function in ER graphs has a dominating positive singularity [15]. The absolute coefficients $|b_j|$ also follow an exponential function $|b_j| \sim e^{kj}$, and the radius of convergence R can be estimated by $R \approx e^{-k}$. Fig. 7a shows the distribution of the radius of convergence R in ER random graph $G_p(N)$. The distribution type does not differ much for different link density p, while the tail of the

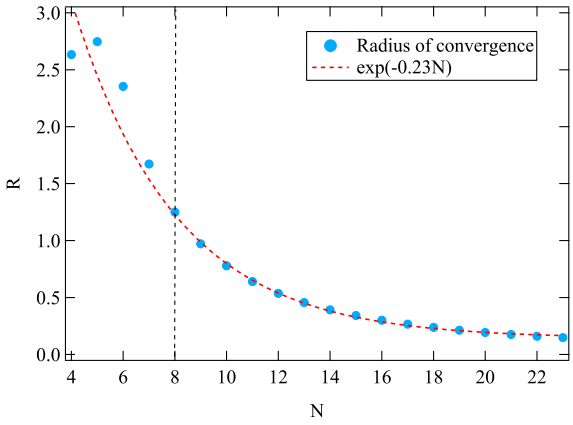


Fig. 5. Radius of convergence R of the steady-state prevalence expansion as a function of network size N in path graphs P_N.

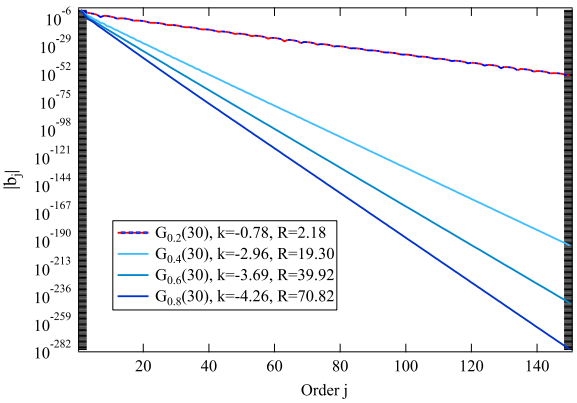


Fig. 6. Coefficient $|b_j|$ as a function of order j in ER random graphs $G_p(N)$ with $N = 30$ nodes and link density p. The red markers represent positive coefficients and the blue markers represent negative. The slope of the function $\log|b_j|$ with order j is denoted by k.

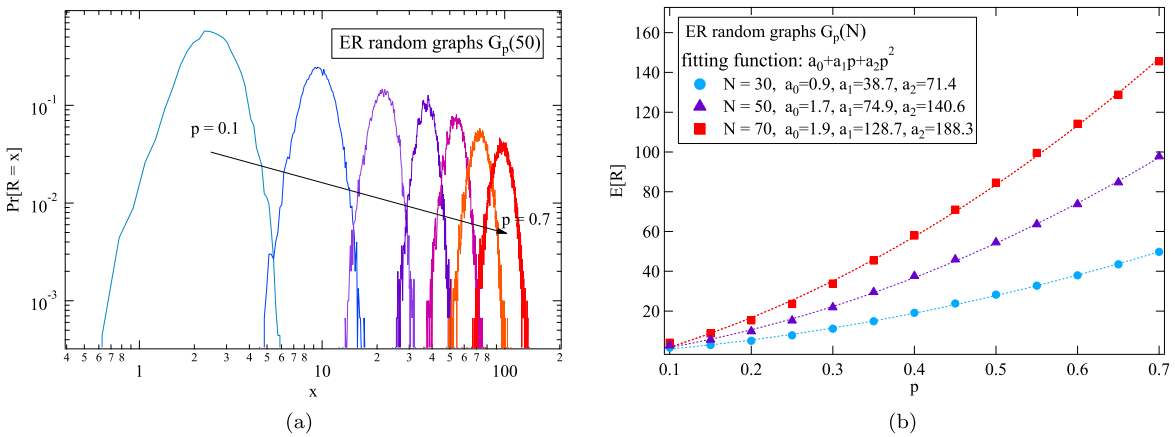


Fig. 7. Subplot (a): Histogram of the radius of convergence R in ER random graphs $G_p(30)$ based on 2×10^4 realizations for each link density p. Subplot (b): Average radius of convergence $E[D]$ in ER random graphs $G_p(N)$ as function of link density p for different network size N.

distribution becomes a little heavier with the increasing link density p. A larger link density p implies more regular ER random graphs, which leads to a higher probability for a large radius of convergence R.

Fig. 7b shows the average radius of convergence $E[R]$ in ER random graphs increases with the link density p, and resembles a quadratic function $E[R] = a_0 + a_1 p + a_2 p^2$ for a network size N. If the link density is large enough, e.g., $p > 0.7$

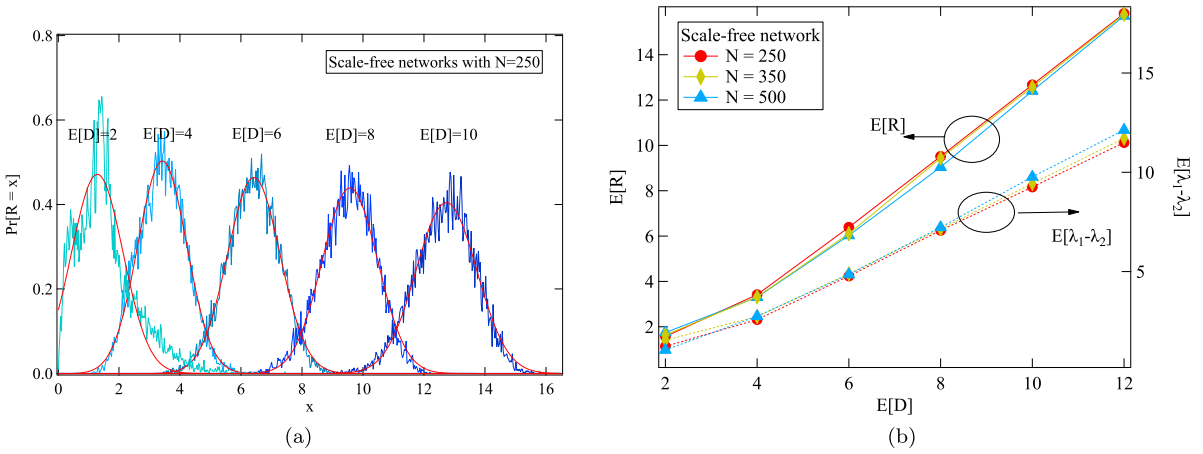


Fig. 8. Subplot (a): Histogram of the radius of convergence R in scale-free graphs with network size $N = 250$ based on 2×10^4 realizations for each average degree $E[D]$. Subplot (b): Average radius of convergence $E[D]$ in scale-free graphs as a function of average degree $E[D]$ for different network size N based on 10^4 realizations. Average spectral gap $E[\lambda_1 - \lambda_2]$ in scale-free graphs as a function of average degree $E[D]$ for different network size N based on 2×10^3 realizations.

in $G_p(70)$, the radius of convergence of the prevalence expansion in ER random graphs can be extremely large with a high probability because the ER graph tends to a regular graph.

3.5. Scale-free graphs

Fig. 8a shows the histogram of the radius of convergence in scale-free networks with network size $N = 250$ and different average degree $E[D]$. The radius of convergence in scale-free networks resembles a Gaussian distribution for the average degree $E[D] > 2$. Fig. 8b shows the average radius of convergence $E[R]$ increases with the average degree $E[D]$, which is in line with the analysis in ER random graphs that the denser and more regular graph usually leads to a higher radius of convergence.

We observe that the network size N almost does not influence the average radius of convergence $E[R]$, which reminds invariant topological properties with the network size N in scale-free networks. Fig. 8b shows that the average spectral gap $E[\lambda_1 - \lambda_2]$ also differs little with increasing the network size N in scale-free networks, which hints that the radius of convergence R in scale free graphs may be dominated by the spectral gap $\lambda_1 - \lambda_2$.

4. Effect of the topological properties on R

In the last section, we illustrate that the prevalence expansion in a dense or regular network usually has a large radius of convergence R. The prevalence expansion is valid for any effective infection rate τ if the radius of convergence $R > \lambda_1$. We investigate the relatively small radius convergence $R < \lambda_1$, which often occurs in sparse networks and clustered networks.

4.1. Effect of the network topology on R

Since the radius of convergence R of the steady-state prevalence expansion (4) can be estimated by the inverse of the convergence order of the absolute coefficient $|b_j|$, i.e., $R = \frac{1}{\lim_{j\to\infty} \sqrt[j]{|b_j|}}$. We first investigate the behaviour of the absolute coefficient $|b_j|$.

Lemma 1. *The convergence order of the coefficients $|b_j| = |u^T \boldsymbol{\alpha}_j|$ in the prevalence expansion (4) is upper bounded by the convergence order of the norm of coefficients $\|\boldsymbol{\alpha}_j\|$ and the norm of coefficients $\|\boldsymbol{c}_j\|$, i.e.*

$$\lim_{j\to\infty} |b_j|^{\frac{1}{j}} \le \lim_{j\to\infty} \|\boldsymbol{\alpha}_j\|^{\frac{1}{j}} = \lim_{j\to\infty} \|\boldsymbol{c}_j\|^{\frac{1}{j}} \tag{9}$$

where the coefficient vector $\boldsymbol{\alpha}_j = X\boldsymbol{c}_j$.

Proof. See Appendix B. □

Then we investigate the convergence order of the coefficient $|b_j|$ by making use of the coefficient vector $\boldsymbol{c}_j$. Defining $z = (\tau_c^{(1)})^{-1} - \tau^{-1}$ and substituting the infection probability vector $\boldsymbol{v}_\infty(\tau) = \sum_{j=1}^{\infty} \boldsymbol{\alpha}_j z^j$ into (2), we can obtain that

$$A\sum_{j=1}^{\infty}\boldsymbol{\alpha}_j z^j - (\lambda_1 - z)\sum_{j=1}^{\infty}\boldsymbol{\alpha}_j z^j = diag\left(\sum_{j=1}^{\infty}\boldsymbol{\alpha}_j z^j\right) A \left(\sum_{j=1}^{\infty}\boldsymbol{\alpha}_j z^j\right) \tag{10}$$

Equating the coefficients of the term z^j for $j \geq 2$ yields

$$\lambda_1\boldsymbol{\alpha}_j - A\boldsymbol{\alpha}_j = \boldsymbol{\alpha}_{j-1} - \sum_{k=1}^{j-1} diag(\boldsymbol{\alpha}_k)A\boldsymbol{\alpha}_{j-k} \tag{11}$$

Invoking the coefficient $\boldsymbol{\alpha}_j = X\boldsymbol{c}_j$, we multiplying X^T on both sides of (11) and arrive

$$(\lambda_1 I - \Lambda)\boldsymbol{c}_j = \boldsymbol{c}_{j-1} - X^T\sum_{k=1}^{j-1} diag(X\boldsymbol{c}_k)AX\boldsymbol{c}_{j-k} \tag{12}$$

where the adjacency matrix $A = X\Lambda X^T$, and Λ is the diagonal matrix with $\Lambda_{ii} = \lambda_i$.

For $m = 1$, we introduce the basis vector e_m and obtain that $(\boldsymbol{c}_j)_1 = e_1^T\left(X^T\sum_{k=1}^{j} diag(\boldsymbol{\alpha}_k)A\boldsymbol{\alpha}_{j+1-k}\right)$, which yields that

$$\begin{aligned}|(\boldsymbol{c}_j)_1| &= \left|e_1^T\left(X^T\sum_{k=1}^{j} diag(X\boldsymbol{c}_k)AX\boldsymbol{c}_{j+1-k}\right)\right| \leq \left\|\sum_{k=1}^{j} diag(\boldsymbol{c}_k)A\boldsymbol{c}_{j+1-k}\right\| \\ &\leq \lambda_1\sum_{k=1}^{j}\|(\boldsymbol{c}_k)\|\cdot\|\boldsymbol{c}_{j+1-k}\| = 2\lambda_1\|\boldsymbol{c}_1\|\cdot\|\boldsymbol{c}_j\| + \lambda_1\sum_{k=2}^{j-1}\|\boldsymbol{c}_k\|\cdot\|\boldsymbol{c}_{j+1-k}\|\end{aligned} \tag{13}$$

Supposing that $\lim_{j\to\infty}|(\boldsymbol{c}_j)_1|^{\frac{1}{j}} = \max_{i\in N}\{\lim_{j\to\infty}|(\boldsymbol{c}_j)_i|^{\frac{1}{j}}\}$, i.e., $\|\boldsymbol{c}_j\| \leq \sqrt{N}|(\boldsymbol{c}_j)_1|$, we arrive at

$$|(\boldsymbol{c}_j)_1| \leq \frac{\lambda_1}{1 - \lambda_1\|\boldsymbol{c}_1\|\sqrt{N}}\sum_{k=2}^{j-1}\|(\boldsymbol{c}_k)\|\cdot\|\boldsymbol{c}_{j+1-k}\| \tag{14}$$

For $m > 1$, we similarly obtain that

$$\begin{aligned}|(\boldsymbol{c}_j)_m| &= \frac{1}{\lambda_1 - \lambda_m}\left|e_m^T\left(\boldsymbol{c}_{j-1} - X^T\sum_{k=1}^{j-1} diag(X\boldsymbol{c}_k)AX\boldsymbol{c}_{j-k}\right)\right| \\ &\leq \frac{1}{\lambda_1 - \lambda_2}\|\boldsymbol{c}_{j-1}\| + \frac{1}{\lambda_1 - \lambda_2}\left\|\sum_{k=1}^{j-1} diag(\boldsymbol{c}_k)A\boldsymbol{c}_{j-k}\right\| \\ &\leq \frac{1}{\lambda_1 - \lambda_2}\|\boldsymbol{c}_{j-1}\| + \frac{\lambda_1}{\lambda_1 - \lambda_2}\sum_{k=1}^{j-1}\|\boldsymbol{c}_k\|\cdot\|\boldsymbol{c}_{j-k}\|\end{aligned} \tag{15}$$

The recurrence relations (14) and (15) show that the upper bound of the maximal convergence order of the coefficient $|(\boldsymbol{c}_j)_m|$, i.e. $\max_{i\in N}\{\lim_{j\to\infty}|(\boldsymbol{c}_j)_i|^{\frac{1}{j}}\}$, as well as the upper bound of $\lim_{j\to\infty}|b_j|^{\frac{1}{j}}$ and the lower bound of the radius of convergence R, are coupled to the largest eigenvalue λ_1 and the spectral gap $\lambda_1 - \lambda_2$. We further propose a heuristic lower bound of the radius of convergence of the steady-state prevalence expansion (4) in Appendix C.

From the equation of coefficients (11), we also obtain the inequality

$$\|\lambda_1 I - A\|\cdot\|\boldsymbol{\alpha}_j\| \geq \|(\lambda_1 I - A)\boldsymbol{\alpha}_j\| \geq \|\boldsymbol{\alpha}_{j-1} - \sum_{k=1}^{j-1} diag(\boldsymbol{\alpha}_k)A\boldsymbol{\alpha}_{j-1-k}\| \tag{16}$$

which leads to the recurrence relation

$$\|\boldsymbol{\alpha}_j\| \geq \frac{1}{\lambda_1 - \lambda_N}\left|\|\boldsymbol{\alpha}_{j-1}\| - \left|\left|\sum_{k=1}^{j-1} diag(\boldsymbol{\alpha}_k)A\boldsymbol{\alpha}_{j-1-k}\right|\right|\right| \tag{17}$$

The recurrence relation (17) implies that the lower bound of the norm coefficient $\|\boldsymbol{\alpha}_j\|$, as well as the upper bound of the radius of convergence R, is related to the maximal difference of eigenvalues $\lambda_1 - \lambda_N$.

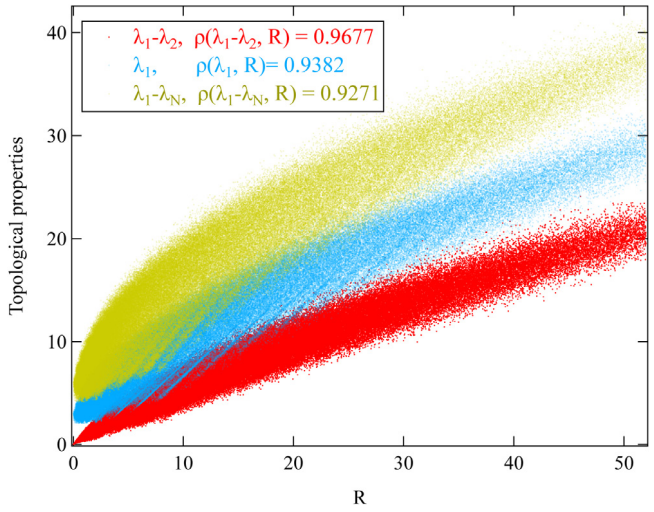


Fig. 9. Correlation between the radius of convergence R and the topological properties, e.g., the largest eigenvalue λ_1, the spectral gap $\lambda_1 - \lambda_2$, the maximal difference of eigenvalues $\lambda_1 - \lambda_N$. The plot is based on 2×10^5 realizations of sparse random graphs. The random sparse graph is generated with the network size uniformly chosen in $N \in [10, 120]$ and the link density uniformly chosen in $p \in [\frac{2}{N}, \frac{2}{N} + 0.3]$, i.e., the additional number of links on the spanning tree is $(p - \frac{2}{N})\frac{N(N-1)}{2}$.

4.2. Numerical tests

Since a large number of real-world networks are sparse networks with a small average degree, we first investigate the radius of convergence R in random sparse networks with N nodes and L links, which are generated in two steps to guarantee the connectivity. First, at step 1, we generate a (uniformly chosen) random spanning tree based on a complete graph with N nodes [16]; and then, at step 2, we add additional $L - N + 1$ links, randomly and uniformly, on the complementary graph of the spanning tree. Fig. 9 shows the correlation between the radius of convergence R and the topological properties, e.g., the largest eigenvalue λ_1, the spectral gap $\lambda_1 - \lambda_2$, the maximal difference of eigenvalues $\lambda_1 - \lambda_N$. The above three topological properties all present a high correlation with the radius of convergence, while the spectral gap $\lambda_1 - \lambda_2$ has the highest correlation coefficient.

We also present the radius of convergence R in modular ER random graphs $G(N, m, p, p_m)$, where N is the number of nodes, m is the number of modules, p is the overall link density and p_m is the proportion of links within modules [17,18]. Fig. 10 shows that the correlation between the spectral gap and the radius of convergence is still high in modular ER random graphs, while the correlation between the radius of convergence R and the other properties, e.g., the largest eigenvalue λ_1 and the maximal difference of eigenvalues $\lambda_1 - \lambda_N$, degrades much.

We further explore the relation between the radius of convergence R and the topological properties in 10 real-world networks. The topologies of these networks are obtained from Newman network collection[1] and Pajek data sets.[2] Fig. 11 shows that the spectral gap $\lambda_1 - \lambda_2$ has a stronger correlation than the largest eigenvalue λ_1 with the radius of convergence R. Moreover, the spectral gap $\lambda_1 - \lambda_2$ can act as an indicator to estimate the radius of convergence R in sparse networks.

We verify that the spectral gap of the adjacency matrix plays a critical role in the radius of convergence. The spectral gap is also an indicator of the spectral expansion [19], which describes the goodness of connectivity and Cheeger constant of a graph [20]. A larger spectral gap could increase the radius of convergence and improve the validity of the NIMFA stead-state prevalence expansion, which coincides with the fact that NIMFA approaches the exact SIS model better in well-connected networks.

5. Summary

In this paper, we investigate the convergence of the NIMFA steady-state prevalence expansion (4) in terms of the effective infection rate at NIMFA epidemic threshold. The network topology alters the radius of convergence R of the steady-state prevalence expansion (4), which is infinite in regular graphs and becomes finite in irregular graphs. The average radius of convergence increases with the density (the average degree) in random graphs, e.g., ER random graph and scale-free graphs. The radius of convergence R is also coupled to the eigenvalues of the adjacency matrix, especially, a smaller spectral gap $\lambda_1 - \lambda_2$ usually decreases the radius of convergence in sparse networks and clustered networks.

1 http://www-personal.umich.edu/mejn/netdata/.
2 http://vlado.fmf.uni-lj.si/pub/networks/data/.

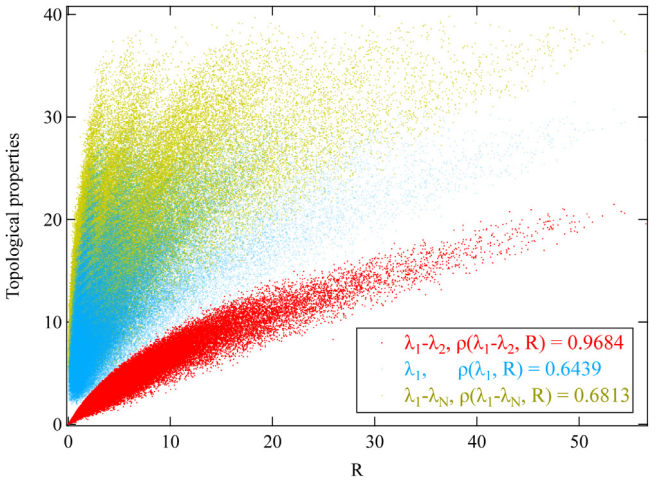
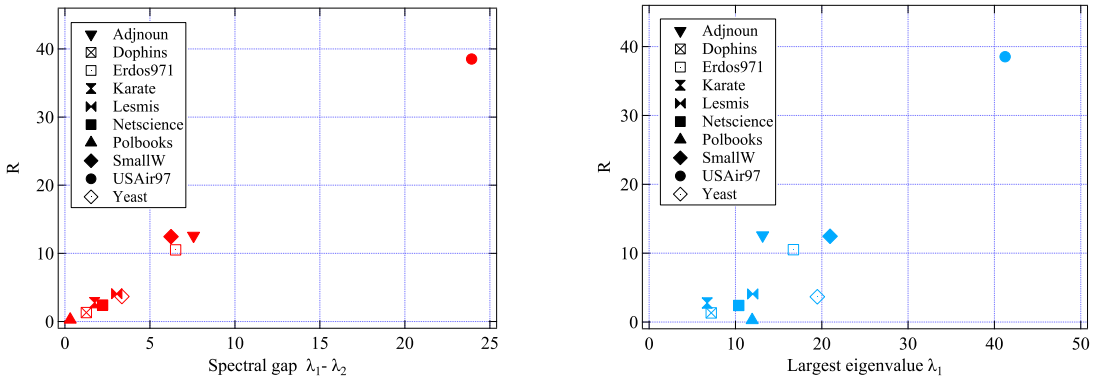


Fig. 10. Correlation between the radius of convergence R and the topological properties, e.g., the largest eigenvalue λ_1, the spectral gap $\lambda_1 - \lambda_2$, the maximal difference of eigenvalues $\lambda_1 - \lambda_N$. The plot is based on 2×10^5 realizations of modular ER random graphs $G(N, m, p, p_m)$ with $N \in [10, 100]$, $m \in [1, 6]$, $p \in [0.05, 0.3]$ and $p_m \in [0.5, 0.9]$.



(a) Spectral gap (b) Largest eigenvalue

Fig. 11. The relation between the radius of convergence R and the topological properties (i.e., the spectral gap $\lambda_1 - \lambda_2$ and the largest eigenvalue λ_1) in 10 real-world networks. The networks Adjnoun, Dophins, Karate, Lesmis, Netscience, Polbooks are collected from Newman network collection, and the networks Erdos971, USAir97, and Yeast are collected from Pajek data sets. We extract the giant component of each network and regard all links as undirected and unweighted.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

This research is supported by the China Scholarship Council (CSC). We are grateful to Hale Cetinay and Albert Senén Cerdà for useful comments.

Appendix A. Recurrence of the coefficients $c_j(k)$ in the expansion of $v_{i\infty}(\tau)$

The steady-state infection probability $v_{i\infty}(\tau)$ of node i in SIS epidemics with the effective infection rate $\tau \downarrow \tau_c^{(1)}$ follows

$$v_{i\infty}(\tau) = \sum_{j=1}^{\infty} \sum_{k=1}^{N} c_j(k)(x_k)_i((\tau_c^{(1)})^{-1} - \tau^{-1})^j \tag{A.1}$$

where all coefficients $c_j(k)$ for the non-trivial solution $v_{i\infty}(\tau)$ are determined in a recursive way as follows [11].

Defining $T(m,l,k)=\sum_{q=1}^{N}(x_m)_q(x_l)_q(x_k)_q$, the coefficients $c_j(m)$ obey, for $m>1$ and $j>2$, the recursion

$$c_j(m)=\frac{c_{j-1}(m)}{\lambda_1-\lambda_m}\{1-c_1(1)(\lambda_1+\lambda_m)T(m,m,1)\}-\frac{c_1(1)}{\lambda_1-\lambda_m}\sum_{k=1;k\neq m}^{N}(\lambda_1+\lambda_k)c_{j-1}(k)T(m,k,1)$$
$$-\frac{1}{\lambda_1-\lambda_m}\sum_{n=2}^{j-2}\sum_{l=1}^{N}\sum_{k=1}^{N}c_{j-n}(l)c_n(k)\lambda_k T(m,l,k) \tag{A.2}$$

while, for $j=2$ and $m>1$,

$$c_2(m)=-\frac{1}{\lambda_1-\lambda_m}\frac{T(m,1,1)}{\lambda_1 T^2(1,1,1)} \tag{A.3}$$

and $c_1(m)=0$. For $m=1$, there holds that $c_1(1)=(\lambda_1\sum_{j=1}^{N}(x_1)_j^3)^{-1}$ and for $j=1$, the coefficients $c_j(1)$ satisfy the recursion

$$c_j(1)=-\frac{1}{\lambda_1 T(1,1,1)}\sum_{k=2}^{N}(\lambda_1+\lambda_k)c_j(k)T(1,1,k)-\sum_{n=2}^{j-1}\sum_{l=1}^{N}\sum_{k=1}^{N}c_{j+1-n}(l)c_n(k)\lambda_k T(1,l,k) \tag{A.4}$$

Appendix B. Proof for Lemma 1

Proof. Without loss of generality, we assume that the convergence order of the elements $|(\boldsymbol{\alpha}_j)_i|$ for $i=1,\ldots,N$ are different, and the coefficient $|(\boldsymbol{\alpha}_j)_k|$ has the maximum convergence order, i.e., $\lim_{j\to\infty}|(\boldsymbol{\alpha}_j)_k|^{\frac{1}{j}}=max_{i\in N}\{\lim_{j\to\infty}|(\boldsymbol{\alpha}_j)_i|^{\frac{1}{j}}\}$. Then there exists a critical order j_c such that $|(\boldsymbol{\alpha}_j)_k|=\max_{i\in N}|(\boldsymbol{\alpha}_j)_i|$ for $j>j_c$. The absolute coefficients $|b_j|$ follows

$$\lim_{j\to\infty}|b_j|^{\frac{1}{j}}=\lim_{j\to\infty}\left|\frac{1}{N}u^T\boldsymbol{\alpha}_j\right|^{\frac{1}{j}}\leq\lim_{j\to\infty}\left(\frac{1}{N}\sum_{i=1}^{N}|(\boldsymbol{\alpha}_j)_i|\right)^{\frac{1}{j}}$$
$$=\lim_{j\to\infty}|(\boldsymbol{\alpha}_j)_k|^{\frac{1}{j}}\left(\frac{1}{N}+\frac{1}{N}\sum_{i=1,i\neq k}^{N}\frac{|(\boldsymbol{\alpha}_j)_i|}{|(\boldsymbol{\alpha}_j)_k|}\right)^{\frac{1}{j}}=\lim_{j\to\infty}|(\boldsymbol{\alpha}_j)_k|^{\frac{1}{j}} \tag{B.1}$$

Meanwhile, the norm of the coefficients $\|\boldsymbol{\alpha}_j\|$ follows

$$\lim_{j\to\infty}\|\boldsymbol{\alpha}_j\|^{\frac{1}{j}}=\lim_{j\to\infty}|(\boldsymbol{\alpha}_j)_k|^{\frac{1}{j}}\left(1+\sum_{i=1,i\neq k}^{N}\frac{|(\boldsymbol{\alpha}_j)_i|^2}{|(\boldsymbol{\alpha}_j)_k|^2}\right)^{\frac{1}{2j}}=\lim_{j\to\infty}|(\boldsymbol{\alpha}_j)_k|^{\frac{1}{j}} \tag{B.2}$$

Hence, we obtain the relation that $\lim_{j\to\infty}|b_j|^{\frac{1}{j}}\leq\lim_{j\to\infty}\|\boldsymbol{\alpha}_j\|^{\frac{1}{j}}$.

Invoking that the coefficient $\boldsymbol{\alpha}_j=X\boldsymbol{c}_j$ and supposing $\lim_{j\to\infty}|(\boldsymbol{c}_j)_m|^{\frac{1}{j}}=max_{i\in N}\{\lim_{j\to\infty}|(\boldsymbol{c}_j)_i|^{\frac{1}{j}}\}$, we can obtain that

$$\lim_{j\to\infty}\|\boldsymbol{\alpha}_j\|^{\frac{1}{j}}=\lim_{j\to\infty}\|\boldsymbol{c}_j\|^{\frac{1}{j}}=\lim_{j\to\infty}|(\boldsymbol{c}_j)_m|^{\frac{1}{j}}=\lim_{j\to\infty}|(\boldsymbol{\alpha}_j)_k|^{\frac{1}{j}} \tag{B.3}$$

which implies that $\|\boldsymbol{\alpha}_j\|$ and $\|\boldsymbol{c}_j\|$ have the same convergence order. □

Appendix C. A lower bound of the radius of convergence *R*

We hereby heuristically propose a lower bound of the radius of convergence R. We rewrite (12) as

$$(\lambda_1 I-\Lambda)\boldsymbol{c}_j=X^T\boldsymbol{\alpha}_{j-1}-X^T\sum_{k=1}^{j-1}diag(\boldsymbol{\alpha}_k)A\boldsymbol{\alpha}_{j-1-k} \tag{C.1}$$

and define the vector $\tilde{\boldsymbol{c}}_j$ with $(\tilde{\boldsymbol{c}}_j)_1=0$ and $(\tilde{\boldsymbol{c}}_j)_m=(\boldsymbol{c}_j)_m$ for $m>2$ and $j>2$. We suppose that the norm of the coefficients $\|\tilde{\boldsymbol{c}}_j\|$ approximates $\|\boldsymbol{c}_j\|$ well for a large network size N. Further denoting the matrix $S:=diag(0,\frac{1}{\lambda_1-\lambda_2},\ldots,\frac{1}{\lambda_1-\lambda_N})$ and invoking $\|\boldsymbol{\alpha}_j\|=\|\boldsymbol{c}_j\|$, we can obtain the norm of coefficients $\|\boldsymbol{\alpha}_j\|$ follows

$$\|\boldsymbol{\alpha}_j\|\approx\|\tilde{\boldsymbol{c}}_j\|=\|SX^T\boldsymbol{\alpha}_{j-1}-SX^T\sum_{k=1}^{j-1}diag(\boldsymbol{\alpha}_k)A\boldsymbol{\alpha}_{j-1-k}\|$$

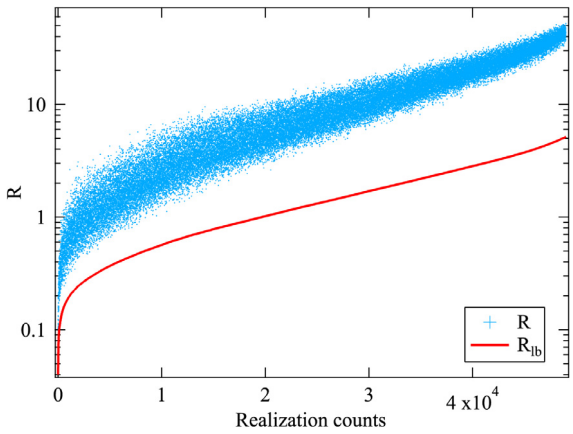


Fig. C.12. Radius of convergence R for 4.5×10^4 realizations of connected ER random graphs sorted by the estimated lower bound R_{lb}. The red line represents the proposed lower bound of the radius of convergence R_{lb}. The ER random graph $G_p(N)$ is generated with the network size uniformly chosen in $N \in [10, 120]$ and the link density uniformly chosen in $p \in [0.05, \frac{2}{N} + 0.3]$.

$$\begin{aligned}
&\leq \|S\boldsymbol{\alpha}_{j-1}\| + \|S\sum_{k=1}^{j-1} diag(\boldsymbol{\alpha}_k)A\boldsymbol{\alpha}_{j-1-k}\| \\
&\leq \|S\| \cdot \|\boldsymbol{\alpha}_{j-1}\| + \|S\| \sum_{k=1}^{j-1} \|diag(\boldsymbol{\alpha}_k)\| \cdot \|A\| \cdot \|\boldsymbol{\alpha}_{j-1-k}\| \\
&\leq \frac{1}{\lambda_1 - \lambda_2}\|\boldsymbol{\alpha}_{j-1}\| + \frac{\lambda_1}{\lambda_1 - \lambda_2} \sum_{k=1}^{j-1} \|\boldsymbol{\alpha}_k\| \cdot \|\boldsymbol{\alpha}_{j-1-k}\|
\end{aligned} \tag{C.2}$$

Defining q_j as the upper bound of $\|\boldsymbol{\alpha}_j\|$, we have a new recurrence formula of q_j for $j > 2$ as

$$q_j = hq_{j-1} + g\sum_{k=0}^{j} q_k q_{j-k} \tag{C.3}$$

with $q_0 = 0$, $q_1 = \|\boldsymbol{\alpha}_1\|$ and $q_2 = \|\boldsymbol{\alpha}_2\|$, where $h := \frac{1}{\lambda_1 - \lambda_2}$ and $g := \frac{\lambda_1}{\lambda_1 - \lambda_2}$. Further defining the generating function $G(z) = \sum_{j=0}^{\infty} q_j z^j$, we can derive that the generating function $G(z)$ follows

$$gG^2(z) = (1 - hz)G(z) + (gq_1^2 + hq_1 - d_2)z^2 - q_1 z \tag{C.4}$$

We define the functional equation $F(z, G) = gG^2(z) - (1 - hz)G(x) - (gd_1^2 + hd_1 - d_2)z^2 + d_1 z$. Bender [21] shows that if there exist real positive numbers $r > 0$ such that the function equation $F(r, G(r)) = 0$ and $\frac{\partial F(z,G(z))}{\partial G}\big|_{z=r} = 0$, the convergence order of positive coefficients q_j follows $\lim_{j\to\infty} \sqrt[j]{q_j} = r^{-1}$. In our case, we can compute that $r = \frac{h+2gd_1-2\sqrt{gd_2}}{h^2+4hgd_1+4g^2d_1^2-4gd_2}$. According to Lemma 1, the radius of convergence for the prevalence expansion follows

$$R = \frac{1}{\lim_{j\to\infty}\sqrt[j]{|b_j|}} \geq \frac{1}{\lim_{j\to\infty}\sqrt[j]{\|\boldsymbol{\alpha}_j\|}} \geq \frac{1}{\lim_{j\to\infty}\sqrt[j]{q_j}} = r \tag{C.5}$$

Thus, we can obtain a lower bound of the radius of convergence for the prevalence expansion

$$R_{lb} = \frac{h + 2gd_1 - 2\sqrt{gd_2}}{h^2 + 4hgd_1 + 4g^2d_1^2 - 4gd_2} \tag{C.6}$$

Fig. C.12 shows the radius of convergence R and the proposed lower bound R_{lb} of (C.6) in more than 4.5×10^4 realizations in random sparse graphs. The radius of convergence R and the lower bound R_{lb} present similar behaviours, but the proposed R_{lb} only provides a loose lower bound for the radius of convergence R, which hints that the precise radius of convergence R cannot be inferred only by the eigenvalues of the adjacency matrix.

References

[1] R. Pastor Satorras, C. Castellano, P. Van Mieghem, A. Vespignani, Epidemic processes in complex networks, Rev. Modern Phys. 87 (3) (2015) 925.

[2] P. Van Mieghem, Performance Analysis of Complex Networks and Systems, Cambridge University Press, 2014.
[3] C. Castellano, R. Pastor-Satorras, Thresholds for epidemic spreading in networks, Phys. Rev. Lett. 105 (21) (2010) 218701.
[4] Z. He, P. Van Mieghem, The spreading time in SIS epidemics on networks, Physica A 494 (2018) 317–330.
[5] P. Van Mieghem, Decay towards the overall-healthy state in SIS epidemics on networks, 2013, arXiv preprint arXiv:1310.3980.
[6] P. Van Mieghem, J. Omic, R. Kooij, Virus spread in networks, IEEE/ACM Trans. Netw. 17 (1) (2009) 1–14.
[7] P. Van Mieghem, R. Van de Bovenkamp, Non-Markovian infection spread dramatically alters the susceptible-infected-susceptible epidemic threshold in networks, Phys. Rev. Lett. 110 (10) (2013) 108701.
[8] P. Van Mieghem, The N-intertwined SIS epidemic network model, Computing 93 (2–4) (2011) 147–169.
[9] K. Devriendt, P. Van Mieghem, Unified mean-field framework for SIS epidemics on networks, based on graph partitioning and the isoperimetric inequality, Phys. Rev. E 96 (5) (2017) 052314.
[10] A. Lajmanovich, J.A. Yorke, A deterministic model for Gonorrhea in a nonhomogeneous population, Math. Biosci. 28 (3–4) (1976) 221–236.
[11] P. Van Mieghem, The viral conductance of a network, Comput. Commun. 35 (12) (2012) 1494–1506.
[12] E.C. Titchmarsh, The theory of Functions, Oxford University Press, 1968.
[13] G. Mercer, A. Roberts, A centre manifold description of contaminant dispersion in channels with varying flow properties, SIAM J. Appl. Math. 50 (6) (1990) 1547–1565.
[14] E. Cator, P. Van Mieghem, Susceptible-infected-susceptible epidemics on the complete graph and the star graph: Exact analysis, Phys. Rev. E 87 (1) (2013) 012811.
[15] E.J. Hinch, Perturbation Methods, Cambridge University Press, 1991.
[16] D.B. Wilson, Generating random spanning trees more quickly than the cover time, in: Proceedings of the Twenty-Eighth Annual ACM Symposium on Theory of Computing, ACM, 1996, pp. 296–303.
[17] R.V. Solé, S. Valverde, Information theory of complex networks: on evolution and architectural constraints, in: Complex Networks, Springer, 2004, pp. 189–207.
[18] M.E. Newman, Modularity and community structure in networks, Proc. Nat. Acad. Sci. 103 (23) (2006) 8577–8582.
[19] S. Hoory, N. Linial, A. Wigderson, Expander graphs and their applications, Bull. Amer. Math. Soc. 43 (4) (2006) 439–561.
[20] F.D. Malliaros, V. Megalooikonomou, Expansion properties of large social graphs, in: International Conference on Database Systems for Advanced Applications, Springer, 2011, pp. 311–322.
[21] E.A. Bender, Asymptotic methods in enumeration, SIAM Rev. 16 (4) (1974) 485–515.