A Note on the Markovian SIR Epidemic on a Random Graph with Given Degrees

Malwina Luczak[1]

Abstract

We consider a Markovian model of an SIR epidemic spreading on a contact graph that is drawn uniformly at random from the set of all graphs with n vertices and given vertex degrees. Janson, Luczak and Windridge (Random Struct Alg 45(4):724–761, 2014) prove that the evolution of such an epidemic is well approximated by the solution to a simple set of differential equations, thus providing probabilistic underpinnings to the works of Miller (J Math Biol 62(3):349–358, 2011) and Volz (J Math Biol 56(3):293–310, 2008). The present paper provides an additional probabilistic interpretation of the limiting deterministic functions in Janson, Luczak and Windridge (Random Struct Alg 45(4):724–761, 2014), thus clarifying further the connection between their results and the results of Miller and Volz.

1 Introduction

The Markovian SIR epidemic process is a simple model for a disease spreading around a finite population in which each individual is either susceptible, infective or recovered. Individuals are represented by vertices (nodes) in a graph (network) G, with edges corresponding to potentially infectious contacts. Infective vertices become recovered at rate $\rho \geq 0$ and infect each susceptible neighbour at rate $\beta > 0$; those are the only possible transitions, i.e. recovered vertices never become infective.

✉ Malwina Luczak
   malwina.luczak@manchester.ac.uk

1   Department of Mathematics, The University of Manchester, Oxford Road, Manchester M13 9PL, UK

◸ Springer

There have been a number of studies of SIR epidemics on random graphs with a given degree sequence: Volz [8], Miller [7], Decreusefond, Dhersin, Moyal and Tran [3], Barbour and Reinert [1], Janson, Luczak and Windridge [4], Janson, Luczak, Windridge and House [6]. The present note builds on the work of Janson, Luczak and Windridge [4], who prove that the evolution of such an epidemic is well approximated by the solution to a simple set of differential equations; we give a probabilistic interpretation of the results in [4], as well as alternative formulae for the limiting functions derived in that paper.

The analysis in [4] is based on the configuration model for random graphs with a given degree sequence, which we explain in detail later. The half-edges in the model are paired only as needed to determine how the epidemic spreads. In the present paper, we provide a new probabilistic interpretation of the asymptotic probability θ_t that a given half-edge does not transmit infection by time t, used in [4]. We also derive new formulae, easier to interpret, for the numbers of free recovered and infectious half-edges. The equations we obtain should lead to a better understanding of the results in [4], and possibly enable more general results in the future. It is also reasonable to hope that our new equations might enable an alternative, more intuitive, proof of the results of [4] (although it is unlikely that considerable technical details could be avoided).

2 Model, Notation, Assumptions and Summary of Results from Janson, Luczak and Windridge [4]

Let us recall some notation and assumptions from Janson, Luczak and Windridge [4].

For $n \in \mathbb{N}$ and a sequence $(d_i)_1^n$ of non-negative integers, let $G = G(n, (d_i)_1^n)$ be a simple graph (i.e. with no loops or double edges) on n vertices, chosen uniformly at random from among all graphs with degree sequence $(d_i)_1^n$. (It is tacitly assumed that there is some such graph, so $\sum_{i=1}^n d_i$ must be even, at least.)

Given the graph G, the SIR epidemic evolves as a continuous-time Markov chain. At any time, each vertex is either susceptible, infected or recovered. Each infective vertex recovers at rate $\rho \geq 0$ and also infects each susceptible neighbour at rate $\beta > 0$.

There are initially n_S, n_I, and n_R susceptible, infective and recovered vertices, respectively. Further, it is assumed that, for each $k \geq 0$, there are respectively $n_{S,k}, n_{I,k}$ and $n_{R,k}$ of these vertices with degree k. Thus, $n_S + n_I + n_R = n$ and $n_S = \sum_{k=0}^\infty n_{S,k}$, $n_I = \sum_{k=0}^\infty n_{I,k}, n_R = \sum_{k=0}^\infty n_{R,k}$. We write n_k to denote the total number of vertices with degree k; thus, for each k, $n_k = n_{S,k} + n_{I,k} + n_{R,k}$.

Note that all these parameters, as well as the sequence $(d_i)_1^n$, depend on the number n of vertices, although we omit explicit mention of this in the notation. The parameters do not have to be defined for all integers n; a subsequence is enough. We consider asymptotics as $n \to \infty$, possibly through the subsequence where the parameters are defined.

In order to obtain results about the behaviour of the process in the limit as $n \to \infty$, we need some regularity conditions on the asymptotics of the degree sequence and of the initial conditions (for instance that the proportion of initially susceptible vertices of each fixed degree k tends to a limit). The following conditions are imposed in [4]: all limits are as $n \to \infty$.

D1 The fractions of initially susceptible, infective and recovered vertices converge to some $\alpha_S, \alpha_I, \alpha_R \in [0, 1]$, i.e.

$$n_S/n \to \alpha_S, \qquad n_I/n \to \alpha_I, \qquad n_R/n \to \alpha_R. \tag{2.1}$$

Further, $\alpha_S > 0$.

D2 The degree of a randomly chosen initially susceptible vertex converges to a probability distribution $(p_k)_0^\infty$, i.e.

$$n_{S,k}/n_S \to p_k, \qquad k \geq 0. \tag{2.2}$$

Further, this limiting distribution has a finite and positive mean

$$\lambda := \sum_{k=0}^\infty k p_k \in (0, \infty). \tag{2.3}$$

D3 The average degree of a randomly chosen susceptible vertex converges to λ, i.e.

$$\sum_{k=0}^\infty k n_{S,k}/n_S \to \lambda. \tag{2.4}$$

D4 The average degree over all vertices converges to $\mu > 0$, i.e.

$$\sum_{k=0}^\infty k n_k/n = \sum_{i=1}^n d_i/n \to \mu, \tag{2.5}$$

and, in more detail, for some μ_S, μ_I, μ_R,

$$\sum_{k=0}^\infty k n_{S,k}/n \to \mu_S, \tag{2.6}$$

$$\sum_{k=0}^\infty k n_{I,k}/n \to \mu_I, \qquad \sum_{k=0}^\infty k n_{R,k}/n \to \mu_R. \tag{2.7}$$

D5 The maximum degree of the initially infective vertices is not too large:

$$\max\{k : n_{I,k} > 0\} = o(n). \tag{2.8}$$

D6 Either $p_1 > 0$ or $\rho > 0$ or $\mu_R > 0$.

Clearly, $\alpha_S + \alpha_I + \alpha_R = 1$ and $\mu_S + \mu_I + \mu_R = \mu$. Further, assumptions D1–D3 imply $\sum_{k=0}^\infty k n_{S,k}/n \to \alpha_S \lambda$, and so $\mu_S = \alpha_S \lambda$.

Let $G^*(n, (d_i)_1^n)$ be the random multigraph with given degree sequence $(d_i)_1^n$ defined by the *configuration model*: we take a set of d_i half-edges for each vertex i and combine half-edges into edges by a uniformly random matching (see e.g. Bollobás [2]). Conditioned on the multigraph being simple, we obtain $G = G(n, (d_i)_1^n)$, the uniformly distributed random graph with degree sequence $(d_i)_1^n$.

Janson, Luczak and Windridge [4] first prove their results for the SIR epidemic on G^*, and, by conditioning on G^* being simple, they deduce that these results also hold for the SIR epidemic on G. Their argument relies on the probability that G^* is simple being bounded away from zero as $n \to \infty$. By the main theorem of Janson [5], this occurs provided the following condition holds.

G1 The degree of a randomly chosen vertex has a bounded second moment, i.e.

$$\sum_{k=0}^{\infty} k^2 n_k = O(n). \tag{2.9}$$

The authors of [4] study the SIR epidemic on the multigraph G^*, revealing its edges dynamically while the epidemic spreads. The process analysed in [4] works as follows. A half-edge is said to be *free* if it is not yet paired to another half-edge. A half-edge is called susceptible, infective or recovered according to the type of vertex it belongs to.

At time 0, there are d_i half-edges attached to vertex i, for each i, and all half-edges are free. Subsequently, each free infective half-edge chooses a free half-edge at rate β, uniformly at random from among all the other free half-edges. Together the pair form an edge, and are removed from the pool of free half-edges. If the chosen half-edge belongs to a susceptible vertex, then that vertex becomes infective, and thus all of its half-edges become infective also. Infective vertices also recover at rate ρ.

The process stops when there are no free infective half-edges, at which point the epidemic stops spreading. Some infective vertices may remain but they will recover at i.i.d. exponential times without affecting any other vertex, and are irrelevant from of the point of view of the epidemic. Some susceptible and recovered half-edges may also remain, and these are paired off uniformly at time ∞ to reveal the remaining edges in G^*. This step is unimportant for the spread of the epidemic, but is performed for the purpose of transferring the results from the multigraph G^* to the simple graph G.

Clearly, if all the pairings are completed then the resulting graph is the multigraph G^*. Moreover, the quantities of interest (numbers of susceptible, infective and recovered vertices at each time t) have the same distribution as if we were to reveal the multigraph G^* first and run the SIR epidemic on G^* afterwards.

For $t \geq 0$, let S_t, I_t and R_t denote the numbers of susceptible, infective and recovered vertices, respectively, at time t. Thus S_t is decreasing and R_t is increasing. Also $S_0 = n_S$, $I_0 = n_I$ and $R_0 = n_R$.

For the dynamics described above (with half-edges paired off dynamically, as described), for $t \geq 0$, let $X_{S,t}$, $X_{I,t}$ and $X_{R,t}$ be the number of free susceptible, infective and recovered half-edges at time t, respectively. Thus $X_{S,t}$ is decreasing, $X_{S,0} = \sum_{k=0}^{\infty} k n_{S,k}$, $X_{I,0} = \sum_{k=0}^{\infty} k n_{I,k}$ and $X_{R,0} = \sum_{k=0}^{\infty} k n_{R,k}$.

For the uniformly random graph G with degree sequence $(d_i)_1^n$, the variables $X_{S,t}$, $X_{I,t}$ and $X_{R,t}$, for $t \geq 0$, are defined as above conditioned on the final multigraph G^* being a simple graph.

It is shown in [4] that, upon suitable scaling, the processes S_t, I_t, R_t, $X_{S,t}$, $X_{I,t}$, $X_{R,t}$ converge to deterministic functions. The limiting functions are written in terms of a parameterisation $\theta_t \in [0, 1]$ of time solving an ordinary differential equation given below. In [4], the function θ_t is interpreted as the limiting probability that a given initially susceptible half-edge has not been selected for pairing with a (necessarily infective) half-edge by time t. Let

$$v_S(\theta) := \alpha_S \sum_{k=0}^{\infty} p_k \theta^k, \qquad \theta \in [0, 1], \tag{2.10}$$

so the limiting fraction of susceptible vertices is $v_S(\theta_t)$ at time t (since the events of being selected for pairing will be approximately independent for different half-edges, when n is large). Similarly, for susceptible half-edges, the limiting function is

$$h_S(\theta) := \alpha_S \sum_{k=0}^{\infty} k\theta^k p_k = \theta v_S'(\theta), \qquad \theta \in [0, 1]. \tag{2.11}$$

For the total number of free half-edges, let

$$h_X(\theta) := \mu \theta^2, \qquad \theta \in [0, 1]. \tag{2.12}$$

For the numbers of half-edges of the remaining types, for $\theta \in [0, 1]$, let

$$h_R(\theta) := \mu_R \theta + \frac{\mu\rho}{\beta} \theta(1 - \theta), \tag{2.13}$$

$$h_I(\theta) := h_X(\theta) - h_S(\theta) - h_R(\theta). \tag{2.14}$$

Thus $h_X(\theta) = h_S(\theta) + h_I(\theta) + h_R(\theta)$. Note that

$$v_S(1) = \alpha_S, \tag{2.15}$$

$$h_S(1) = \alpha_S \lambda = \mu_S, \qquad h_R(1) = \mu_R, \qquad h_I(1) = \mu - \mu_S - \mu_R = \mu_I. \tag{2.16}$$

The analysis in [4] covers two separate cases.

The first is where $\mu_I > 0$, meaning that the limiting proportion of initially infective individuals is positive. It is shown in [4] that there is a unique $\theta_\infty \in (0, 1)$ with $h_I(\theta_\infty) = 0$. Further, h_I is strictly positive on $(\theta_\infty, 1]$ and strictly negative on $(0, \theta_\infty)$. Defining the 'infective pressure'

$$p_I(\theta) := \frac{h_I(\theta)}{h_X(\theta)}, \tag{2.17}$$

there is a unique solution $\theta_t : [0, \infty) \to (\theta_\infty, 1]$ to the differential equation

$$\frac{d}{dt}\theta_t = -\beta\theta_t \mathrm{p_I}(\theta_t), \tag{2.18}$$

subject to the initial condition $\theta_0 = 1$.

Furthermore, there is a unique solution $\hat{I}_t$ to

$$\frac{d}{dt}\hat{I}_t = \frac{\beta h_I(\theta_t)h_S(\theta_t)}{h_X(\theta_t)} - \rho\hat{I}_t, \ t \geq 0, \qquad \hat{I}_0 = \alpha_I. \tag{2.19}$$

Defining also $\hat{R}_t := 1 - v_S(\theta_t) - \hat{I}_t$, Theorem 2.6 in [4] states that, for the epidemic on the multigraph G^*, under conditions D1–D6, uniformly on $[0, \infty)$,

$$S_t/n \xrightarrow{\text{P}} v_S(\theta_t), \qquad I_t/n \xrightarrow{\text{P}} \hat{I}_t, \qquad R_t/n \xrightarrow{\text{P}} \hat{R}_t, \tag{2.20}$$

$$X_{S,t}/n \xrightarrow{\text{P}} h_S(\theta_t), \qquad X_{I,t}/n \xrightarrow{\text{P}} h_I(\theta_t), \qquad X_{R,t}/n \xrightarrow{\text{P}} h_R(\theta_t), \tag{2.21}$$

$$X_t/n \xrightarrow{\text{P}} h_X(\theta_t). \tag{2.22}$$

Moreover, the number $S_\infty := \lim_{t\to\infty} S_t$ of susceptibles that escape infection satisfies

$$S_\infty/n \xrightarrow{\text{P}} v_S(\theta_\infty).$$

The same holds on the graph G under the additional assumption G1 (Theorem 2.7 in [4]).

This means that the entire course of the epidemic, including the final number of individuals affected by the outbreak is approximately deterministic, following the solution to equations (2.18) and (2.19) (with suitable formulae derived from these equations for the remaining variables).

The other case covered is where there are initially a small number of infectives – of order less than n, so that $\mu_I = 0$ (i.e. the initial proportion of infectives tends to 0 as $n \to \infty$). We recall from [4] that

$$\mathfrak{R}_0 := \left(\frac{\beta}{\rho + \beta}\right)\left(\frac{\alpha_S}{\mu}\right)\sum_{k=0}^{\infty}(k-1)kp_k; \tag{2.23}$$

the *basic reproductive ratio* of the epidemic. If $\mathfrak{R}_0 > 1$, then the epidemic takes off with positive probability, even from a small initial number of infectives.

It is shown in [4] that, when $\mathfrak{R}_0 > 1$, even if $\mu_I = 0$, then there is a unique $\theta_\infty \in (0, 1)$ with $h_I(\theta_\infty) = 0$, and that h_I is strictly positive on $(\theta_\infty, 1)$ and strictly negative on $(0, \theta_\infty)$.

The initial condition of the limiting differential equation, now defined on $(-\infty, \infty)$, is shifted so that $t = 0$ corresponds to the time T_0 in the random process, which is the infimum of times t such that the fraction of susceptible individuals has fallen from

about $\alpha_S = v_S(1)$ to some fixed smaller s_0 by time t. Note that T_0 is finite if and only if a large outbreak occurs, and such an outbreak is already established by time T_0.

It is shown in [4] that there is a unique continuously differentiable $\theta_t : \mathbb{R} \to (\theta_\infty, 1)$ such that

$$\frac{d}{dt}\theta_t = -\beta\theta_t p_I(\theta_t), \qquad \theta_0 = v_S^{-1}(s_0). \tag{2.24}$$

Furthermore, $\theta_t \searrow \theta_\infty$ as $t \to \infty$ and $\theta_t \nearrow 1$ as $t \to -\infty$.

The processes are extended to be defined on $(-\infty, \infty)$ by taking $S_t = S_0$ for $t < 0$, and similarly for the other processes.

The following is proven in [4] (Theorems 2.9 and 2.10), for both the simple graph G and the multigraph G^*. Suppose that conditions D1–D6 and G1 hold. Assume that $\mathfrak{R}_0 > 1$. Suppose also that $\alpha_I = \mu_I = 0$ but there is initially at least one infective vertex with non-zero degree.

Then, $\liminf_{n \to \infty} \mathbb{P}(T_0 < \infty) > 0$. Also, conditional on $T_0 < \infty$, uniformly on $(-\infty, \infty)$,

$$S_{T_0+t}/n \xrightarrow{\text{p}} v_S(\theta_t), \qquad I_{T_0+t}/n \xrightarrow{\text{p}} \hat{I}_t, \qquad R_{T_0+t}/n \xrightarrow{\text{p}} \hat{R}_t, \tag{2.25}$$

$$X_{S,T_0+t}/n \xrightarrow{\text{p}} h_S(\theta_t), \quad X_{I,T_0+t}/n \xrightarrow{\text{p}} h_I(\theta_t), \quad X_{R,T_0+t}/n \xrightarrow{\text{p}} h_R(\theta_t), \tag{2.26}$$

$$X_{T_0+t}/n \xrightarrow{\text{p}} h_X(\theta_t). \tag{2.27}$$

Also, conditional on $T_0 < \infty$, the number of susceptibles that escape infection satisfies

$$S_\infty/n \xrightarrow{\text{p}} v_S(\theta_\infty).$$

Here, $\hat{I}_t$ is the unique solution to

$$\frac{d}{dt}\hat{I}_t = \frac{\beta h_I(\theta_t)h_S(\theta_t)}{h_X(\theta_t)} - \rho\hat{I}_t, \qquad \lim_{t \to -\infty} \hat{I}_t = 0, \tag{2.28}$$

and $\hat{R}_t := 1 - v_S(\theta_t) - \hat{I}_t$.

This means that, on the event that a large outbreak occurs, the evolution of the epidemic is approximately deterministic, following the solution to equations (2.24) and (2.28), with suitable formulae for the remaining variables derived from these equations. The point of translating the time variable by T_0 is that, without the translation, at time 0, we have $p_I(\theta_0) = p_I(1) = 0$ (since $\mu_I = 0$), which leads to the trivial solution $\theta_t = 1$ for all t, which is the disease-free equilibrium and corresponds to a stochastic epidemic that does not take off.

3 New Probabilistic Interpretation of θ and Alternative Formulae for Limiting Deterministic Functions

We will now give a more complete probabilistic interpretation of the function θ_t used to define the deterministic limit for the SIR epidemic.

As stated in the previous section, the function θ_t used to define the deterministic limits satisfies

$$\frac{d\theta_t}{dt} = -\beta\theta_t \frac{h_I(\theta_t)}{h_X(\theta_t)}.$$

Substituting $h_X(\theta) = \mu\theta^2$, $h_I(\theta) = \mu\theta^2 - \mu_R\theta - \frac{\mu\rho}{\beta}\theta(1-\theta) - \alpha_S \sum_k kp_k\theta^k$, we can rewrite this as

$$\frac{d\theta_t}{dt} = -\beta \frac{\mu\theta_t^2 - \mu_R\theta_t - \frac{\mu\rho}{\beta}\theta_t(1-\theta_t) - \alpha_S \sum_k kp_k\theta_t^k}{\mu\theta_t}$$

$$= -(\beta + \rho)\theta_t + \rho + \frac{\beta\mu_R}{\mu} + \frac{\beta\alpha_S}{\mu} \sum_k kp_k\theta_t^{k-1}.$$

It follows that

$$\frac{d}{dt}(\theta_t e^{(\beta+\rho)t}) = \left(\rho + \frac{\beta\mu_R}{\mu}\right)e^{(\beta+\rho)t} + \frac{\beta\alpha_S}{\mu}\left(\sum_k kp_k\theta_t^{k-1}\right)e^{(\beta+\rho)t},$$

and so, integrating,

$$\theta_t = \left(1 - \frac{\rho + \frac{\beta\mu_R}{\mu}}{\beta+\rho}\right)e^{-(\beta+\rho)t} + \frac{\rho + \frac{\beta\mu_R}{\mu}}{\beta+\rho} + \frac{\beta\alpha_S}{\mu}\sum_k kp_k \int_0^t \theta_s^{k-1}e^{-(\beta+\rho)(t-s)}ds$$

$$= \frac{\rho}{\beta+\rho} + \frac{\beta}{\beta+\rho}\frac{\mu_R}{\mu} + \frac{\beta}{\beta+\rho}\frac{\mu - \mu_R}{\mu}e^{-(\beta+\rho)t}$$

$$+ \frac{\beta\alpha_S}{\mu}\sum_k kp_k \int_0^t \theta_s^{k-1}e^{-(\beta+\rho)(t-s)}ds,$$

and so

$$\theta_t = \frac{\mu_R}{\mu} + \frac{\mu - \mu_R}{\mu}F(t) + \frac{\beta\alpha_S}{\mu}\sum_k kp_k \int_0^t \theta_s^{k-1}e^{-(\beta+\rho)(t-s)}ds, \qquad (3.1)$$

where

$$F(t) = \frac{\rho}{\beta+\rho} + \frac{\beta}{\beta+\rho}e^{-(\beta+\rho)t}.$$

Noting that

$$\int_0^t \theta_s^{k-1} \frac{dF(t-s)}{ds} ds + \int_0^t \frac{d}{ds}(\theta_s)^{k-1} F(t-s) ds = \left[\theta_s^{k-1} F(t-s) \right]_0^t,$$

we see that, for each $k \geq 1$,

$$\beta \int_0^t \theta_s^{k-1} e^{-(\beta+\rho)(t-s)} ds = \theta_t^{k-1} - F(t) - \int_0^t \frac{d}{ds}(\theta_s)^{k-1} F(t-s) ds$$

$$= \theta_t^{k-1} - F(t) + \beta(k-1) \int_0^t \theta_s^{k-1} \frac{h_I(\theta_s)}{h_X(\theta_s)} F(t-s) ds.$$

This then implies, using $\alpha_S \sum_k k p_k = \mu_S$, that

$$\theta_t = \frac{\mu_R}{\mu} + \frac{\mu - \mu_R}{\mu} F(t) + \frac{\alpha_S}{\mu} \sum_k k p_k \theta_t^{k-1} - \frac{\mu_S}{\mu} F(t)$$

$$+ \frac{\beta \alpha_S}{\mu} \sum_k k(k-1) p_k \int_0^t \theta_s^{k-1} \frac{h_I(\theta_s)}{h_X(\theta_s)} F(t-s) ds$$

and hence that

$$\theta_t = \frac{\mu_R}{\mu} + \frac{\mu_I}{\mu} F(t) + \frac{\alpha_S}{\mu} \sum_k k p_k \theta_t^{k-1}$$

$$+ \frac{\beta \alpha_S}{\mu} \sum_k k(k-1) p_k \int_0^t \theta_s^{k-1} \frac{h_I(\theta_s)}{h_X(\theta_s)} F(t-s) ds. \qquad (3.2)$$

Considering formula (3.2), we will now discuss how the function θ_t is the asymptotic probability that a half-edge does not transmit infection (i.e. initiate a pairing) by time t. This should be the same as the limiting probability that a given initially susceptible half-edge has not been paired with a (necessarily infective) half-edge by time t, as interpreted in [4], since that probability is that its eventual partner has not transmitted infection by time t.

Given a random half-edge, conditional on it being initially recovered, which has probability μ_R/μ, it does not transmit infection by time t with probability 1.

Conditional on the half-edge being initially infected, which has probability μ_I/μ, it does not transmit by time t with probability $F(t)$. In the formula for $F(t)$, the term $\frac{\rho}{\beta+\rho}$ is the probability that recovery of the vertex occurs before the half-edge initiates a pairing. The term $\frac{\beta}{\beta+\rho} e^{-(\beta+\rho)t}$ is the probability that the half-edge initiates a pairing before recovery but neither of these events happens by time t.

Conditional on the half-edge being initially susceptible, which happens with probability μ_S/μ, we need to further consider the degree of its vertex. With conditional probability $\frac{\alpha_S k p_k}{\mu_S}$, it has degree k, and then the edge cannot transmit if the vertex does not get infected by time t or only gets infected by transmitting the infection to the half-edge itself, which happens with probability θ_t^{k-1}. The half-edge

also cannot transmit by time t if one of the other $k - 1$ half-edges gets infected at some time $s \leq t$, but then the half-edge in question does not initiate a pairing before vertex recovery over a period of length $t - s$; this happens with probability $-\int_0^t \frac{d}{ds}(\theta_s)^{k-1} F(t - s) ds = \beta(k - 1) \int_0^t \theta_s^{k-1} \frac{h_I(\theta_s)}{h_X(\theta_s)} F(t - s) ds$.

Alternatively, we have

$$n\mu\theta_t^2 = n\mu_R\theta_t + n\mu_I\theta_t F(t) + nh_S(\theta_t)$$
$$+ n\beta\alpha_S\theta_t \sum_k k(k - 1) p_k \int_0^t \theta_s^{k-1} \frac{h_I(\theta_s)}{h_X(\theta_s)} F(t - s) ds.$$

The left hand-side here is approximately the total number of free half-edges at time t. The term $n\mu_R\theta_t$ is approximately the total number of initially recovered half-edges that are still free at time t. The term $n\mu_I\theta_t F(t)$ is approximately the total number of initially infective half-edges that are still free at time t. The term $nh_S(\theta_t)$ is approximately the total number of free susceptible half-edges at time t. The term

$$n\beta\alpha_S\theta_t \int_0^t \sum_k k(k - 1) p_k \theta_s^{k-1} \frac{h_I(\theta_s)}{h_X(\theta_s)} F(t - s) ds$$

is approximately the total number of half-edges belonging to initially susceptible vertices that got infected before time t and are still free at time t.

The function θ_t is closely related to the corresponding function in [7, 8], but these papers do not engage in the same way as [4] with the construction of the configuration model multigraph and simple graph by pairing half-edges and revealing them as they are needed while the epidemic spreads. Instead, for instance, Miller [7] defines a function $\theta(t)$ as follows. An edge is chosen uniformly at random, with endpoints v and u, and then a direction for the edge, say from v ('base') to u ('target'). Then the spread of the epidemic is modified so that infectious contacts from u to v are disallowed. Then $\theta(t)$ is defined to be the probability that there has not been an infectious contact from v to u by time t in the modified process.

We saw in (2.13) that $X_{R,t}/n$ is asymptotically close to

$$h_R(\theta_t) := \mu_R\theta_t + \frac{\mu\rho}{\beta}\theta_t(1 - \theta_t).$$

The term $\frac{\mu\rho}{\beta}\theta_t(1 - \theta_t)$ in the above formula is compact but does not appear readily interpretable.

We claim that the limiting function can instead be expressed in the form

$$\tilde{h}_R(t) = \mu_R\theta_t + \mu_I\theta_t \frac{\rho}{\beta + \rho}(1 - e^{-(\beta+\rho)t})$$
$$+ \alpha_S \sum_k kp_k\theta_t \left(-\int_0^t \frac{d}{ds}(\theta_s)^{k-1} \frac{\rho}{\beta + \rho}(1 - e^{-(\beta+\rho)(t-s)}) ds \right).$$

$$(3.3)$$

To understand this formula, note that $n\mu_R\theta_t$ is approximately the number of free recovered half-edges that were initially recovered.

Also,

$$\frac{\rho}{\beta + \rho}(1 - e^{-(\beta+\rho)t})$$

is the probability that a vertex infectious at time 0 recovers by time t and that its recovery happens before an infectious half-edge attached to this vertex initiates a pairing. This implies that

$$n\mu_I\theta_t \frac{\rho}{\beta + \rho}(1 - e^{-(\beta+\rho)t})$$

is approximately the total number of free recovered half-edges that were infectious at time 0.

Finally,

$$n\alpha_S \sum_k kp_k\theta_t\left(-\int_0^t \frac{d}{ds}(\theta_s)^{k-1}\frac{\rho}{\beta + \rho}(1 - e^{-(\beta+\rho)(t-s)})ds\right)$$

is approximately the total number of free recovered half-edges whose vertices were susceptible at time 0, got infected and recovered by time t.

We are now going to verify that $\tilde{h}_R(t) = h_R(\theta_t)$. This means that we need to verify that

$$\frac{\mu\rho}{\beta}(1 - \theta_t) = \mu_I \frac{\rho}{\beta + \rho}(1 - e^{-(\beta+\rho)t})$$
$$+ \alpha_S \sum_k kp_k\left(-\int_0^t \frac{d}{ds}(\theta_s)^{k-1}\frac{\rho}{\beta + \rho}(1 - e^{-(\beta+\rho)(t-s)})ds\right).$$

To do that, first note that, integrating by parts,

$$-\int_0^t \frac{d}{ds}(\theta_s)^{k-1}\frac{\rho}{\beta+\rho}(1 - e^{-(\beta+\rho)(t-s)})ds$$
$$= \left[-\theta_s^{k-1}\frac{\rho}{\beta+\rho}(1 - e^{-(\beta+\rho)(t-s)})\right]_0^t - \rho\int_0^t \theta_s^{k-1}e^{-(\beta+\rho)(t-s)}ds$$
$$= \frac{\rho}{\beta+\rho}(1 - e^{-(\beta+\rho)t}) - \rho\int_0^t \theta_s^{k-1}e^{-(\beta+\rho)(t-s)}ds.$$

This means we actually need to verify that

$$\frac{\mu\rho}{\beta}(1 - \theta_t) = \mu_I \frac{\rho}{\beta + \rho}(1 - e^{-(\beta+\rho)t})$$
$$+ \mu_S\frac{\rho}{\beta + \rho}(1 - e^{-(\beta+\rho)t})$$
$$- \rho\alpha_S \sum_k kp_k \int_0^t \theta_s^{k-1}e^{-(\beta+\rho)(t-s)}ds.$$

But, as seen in (3.1),

$$\theta_t = \frac{\mu_R}{\mu} + \frac{\mu - \mu_R}{\mu} F(t) + \frac{\beta \alpha_S}{\mu} \sum_k k p_k \int_0^t \theta_s^{k-1} e^{-(\beta+\rho)(t-s)} ds,$$

where

$$F(t) = \frac{\rho}{\beta + \rho} + \frac{\beta}{\beta + \rho} e^{-(\beta+\rho)t},$$

and so

$$-\rho \alpha_S \sum_k k p_k \int_0^t \theta_s^{k-1} e^{-(\beta+\rho)(t-s)} ds = -\frac{\mu \rho}{\beta} \theta_t + \frac{\rho(\mu - \mu_R)}{\beta + \rho} e^{-(\beta+\rho)t}$$
$$+ \frac{\rho}{\beta} \frac{\rho \mu}{\beta + \rho} + \frac{\rho \mu_R}{\beta + \rho}.$$

This means that we need to verify that

$$\frac{\mu \rho}{\beta} (1 - \theta_t) = \mu_I \frac{\rho}{\beta + \rho} (1 - e^{-(\beta+\rho)t}) + \mu_S \frac{\rho}{\beta + \rho} (1 - e^{-(\beta+\rho)t})$$
$$- \frac{\mu \rho}{\beta} \theta_t + \frac{\rho(\mu - \mu_R)}{\beta + \rho} e^{-(\beta+\rho)t} + \frac{\rho}{\beta} \frac{\rho \mu}{\beta + \rho} + \frac{\rho \mu_R}{\beta + \rho},$$

which holds, noting that $\mu_S + \mu_I + \mu_R = \mu$.

Similarly, we have an alternative formula for the limit of $X_{I,t}/n$, the asymptotic scaled number of free infectious half-edges at time t:

$$\tilde{h}_I(t) = \mu_I \theta_t e^{-(\beta+\rho)t} + \alpha_S \theta_t \sum_k k p_k \left(- \int_0^t \frac{d}{ds} (\theta_s)^{k-1} e^{-(\beta+\rho)(t-s)} ds \right). \quad (3.4)$$

For infectious vertices, we have

$$\hat{I}_t = \alpha_I e^{-\rho t} + \beta \int_0^t h_I(\theta_s) \frac{h_S(\theta_s)}{h_X(\theta_s)} e^{-\rho(t-s)} ds$$
$$= \alpha_I e^{-\rho t} - \int_0^t \frac{d v_S(\theta_s)}{ds} e^{-\rho(t-s)} ds, \quad (3.5)$$

and, for recovered vertices,

$$\hat{R}_t = \alpha_R + \alpha_I (1 - e^{-\rho t}) + \beta \int_0^t h_I(\theta_s) \frac{h_S(\theta_s)}{h_X(\theta_s)} (1 - e^{-\rho(t-s)}) ds$$
$$= \alpha_R + \alpha_I (1 - e^{-\rho t}) - \int_0^t \frac{d v_S(\theta_s)}{ds} (1 - e^{-\rho(t-s)}) ds. \quad (3.6)$$

Declarations

Conflict of interest There are no conflicting interests.

References

1. Barbour, A.D., Reinert, G.: Approximating the epidemic curve. Electr. J. Probab. **18**(54), 30 (2013)
2. Bollobás, B.: Random Graphs, 2nd edn. Cambridge University Press, Cambridge (2001)
3. Decreusefond, L., Dhersin, J., Moyal, P., Tran, V.C.: Large graph limit for an SIR process in random network with heterogeneous connectivity. Ann. Appl. Probab. **22**(2), 541–575 (2012)
4. Janson, S., Luczak, M., Windridge, P.: Law of large numbers for the SIR epidemic on a random graph with given degrees. Random Struct. Alg. **45**(4), 724–761 (2014)
5. Janson, S.: The probability that a random multigraph is simple. Combin. Probab. Comput. **18**(1–2), 205–225 (2009)
6. Janson, S., Luczak, M., Windridge, P., House, T.: Near-critical SIR epidemic on a random graph with given degrees. J. Math. Biol. **74**(4), 843–886 (2016)
7. Miller, J.C.: A note on a paper by Erik Volz: SIR dynamics in random networks. J. Math. Biol. **62**(3), 349–358 (2011)
8. Volz, E.: SIR dynamics in random networks with heterogeneous connectivity. J. Math. Biol. **56**(3), 293–310 (2008)