Viral Processes by Random Walks on Random Regular Graphs

Mohammed Abdullah¹, Colin Cooper¹, and Moez Draief^{2,*}

¹ Department of Informatics, King's College London
{mohammed.abdullah,colin.cooper}@kcl.ac.uk

² Department of Electrical and Electronic Engineering, Imperial College London
m.draief@imperial.ac.uk

Abstract. We study the SIR epidemic model with infections carried by k particles making independent random walks on a random regular graph. We give a edge-weighted graph reduction of the dynamics of the process that allows us to apply standard results of Erdős–Renyi random graphs on the particle set. In particular, we show how the parameters of the model produce two phase transitions: In the subcritical regime, $O(\ln k)$ particles are infected. In the supercritical regime, for a constant C determined by the parameters of the model, Ck get infected with probability C , and $O(\ln k)$ get infected with probability $(1 - C)$. Finally, there is a regime in which all k particles are infected. Furthermore, the edge weights give information about when a particle becomes infected. We demonstrate how this can be exploited to determine the completion time of the process by applying a result of Janson on randomly edge weighted graphs.

Keywords: Random walks, random graphs, viral processes, SIR model.

1 Introduction

The spread of an infection throughout a population, often referred to loosely as an *epidemic*, has come to be modelled in various ways in the literature, spurred by the richness of the domains in which the abstract notion of a virus has gone beyond the traditional biological phenomenon. Electronic viruses over computer networks are not the only extension, others include *rumour spreading* [20] or *broadcasting* [3] and *viral marketing* [13, ch. 9]. Models may vary over domains, but the underlying principle is one of spread of some unit of information or change of state through interaction between individuals.

In much of the literature on the spread of epidemics as well as the dissemination of information, individuals reside at fixed locations connected by a graph and the evolution of the state of an individual depends on the state of its neighbours in the graph. In particular if the graph is complete, mean-field (homogeneous mixing) models have been exploited to study the outcome of the diffusion

* Moez Draief gratefully acknowledges the Leverhulme trust funding through the Research Fellowship RF/9/RFG/2010/02/08.

process [9]. More recently, there has been an increasing interest in understanding the impact of the network topology on the spread of epidemics in networks with fixed nodes, see [13] for a review of such results. There has however been little analytical work to date on models where the possible interactions between the nodes are dynamic, i.e., evolves in time.

We explore a particular instance of dynamicity in interaction by assuming that individuals are mobile particles and can only infect each other if they are in sufficiently close proximity. The model is motivated both by certain kinds of biological epidemics, whose transmission may be dominated by sites at which individuals gather in close proximity (e.g., workplaces or public transport for a disease like SARS, cattle markets for foot-and-mouth disease, etc.) and by malware. Furthermore, it is relevant to studying the dissemination of information in opportunistic networks [5] where the information is transmitted between users who happen to be in each others range. As in the case of static networks [20] one may be interested in the time it takes for the rumour to be known to all users.

In our model (elaborated upon below) the virus can be carried by k particles making independent random walks on an n -vertex r -regular random graph G which represents the environment. Each particle is in one of three states: susceptible (S), infected (I), recovered (R). An infected particle can infect a susceptible particle, which remains infected for a fixed *infectious period* ξ before recovering permanently. This is known as the *SIR epidemic model* and is extensively studied in the literature. When $\xi = \infty$ (the *SI model*) particles never go from I to R. Two questions can be asked: (1) When $\xi < \infty$, how many particles ever get infected? (2) How long does the process take to complete? We address both of these questions by reducing the dynamics of the process to what we term an *interaction graph*. This is an Erdős–Renyi (E–R) random graph $\mathcal{G}_{k,\hat{q}}$ on the set of particles, where the edge probability $\hat{q}$ is a function of the parameters of the model. Infected particles are connected components in $\mathcal{G}_{k,\hat{q}}$, and so well known results from the literature on E–R random graphs can be directly applied using our reduction to answer question (1). In particular, we show how the parameters of the model produce two phase transitions: In the subcritical regime, $O(\ln k)$ particles are infected. In the supercritical regime, for a constant C determined by the parameters of the model, Ck get infected with probability C , and $O(\ln k)$ get infected with probability $(1 - C)$. Finally, there is a regime in which all k particles are infected. Statements are with high probability (**whp**), that is, with probability tending to 1 as $n \rightarrow \infty$. Furthermore, the graph reduction assigns weights on the edges that give information about when a particle becomes infected. This information can be used for addressing question (2). As an example, in the case of $\xi = \infty$, we apply a result of Janson [16] on randomly-weighted graphs to deduce that completion time converges in probability to $\frac{2\theta_r n}{k} \ln k$, where $\theta_r = \frac{r-1}{r-2}$. This matches the expectation determined in [8] for the same example.

Whilst the metaphor of an epidemic is a motivating and illustrative one, this work is part of the more general scope of *interacting particle systems* (see, e.g., [2, ch. 14]).

Note, due to space restrictions, we refer to [1] for a more detailed exposition and proofs of lemmas and theorems not presented in this paper.

2 Model

Let $r \geq 3$ be fixed. Let $\mathcal{G}_r$ denote the set of r -regular graphs with vertex set $V = \{1, 2, \dots, n\}$ and the uniform measure. Let $G = (V_G, E_G)$ be chosen uniformly at random (**uar**) from $\mathcal{G}_r$. The results in this paper are always asymptotic in $n = |V_G|$. The notation $A_n \sim B_n$ means that $\lim_{n \rightarrow \infty} A_n/B_n = 1$.

At step t , let $\mathcal{S}(t), \mathcal{I}(t), \mathcal{R}(t)$ be the set of susceptible, infected and recovered particles respectively. These form a partition of the particle set $\mathcal{P}$. Thus, $|\mathcal{P}| = k$. Let $I_0 = |\mathcal{I}(0)|$.

When two particles x and y meet at some vertex v at time step t , an *interaction* takes place between them at that time step with probability $\rho \in (0, 1]$, which is a fixed constant parameter of the model. We term such an event an *xy interaction* and call ρ the *interaction probability*. If one of the particles is infected and the other is susceptible, the infection takes place upon interaction. We define the *completion time* of the process to be the time step at which the last infection takes place. Consider that the counter has just changed from time step $t - 1$ to t . The sequence of actions is as follows: (i) Every particle makes an independent move in its random walk. (ii) A particle $x \in \mathcal{I}(t - 1)$ that had reached the end of its infectious period by time t changes from state I to state R, so $x \notin \mathcal{I}(t)$ and $x \in \mathcal{R}(t)$. Otherwise, it remains in state I, i.e., $x \in \mathcal{I}(t)$. (iii) Each pair of particles x, y at the same vertex v interact with probability ρ ; the particle-pair interactions are independent of each other. If one of x or y is in $\mathcal{I}(t)$, and the other was in $\mathcal{S}(t - 1)$, the latter becomes infected, and it becomes a member of $\mathcal{I}(t)$. New infections are considered to have started at time t ; we say a particle is infected at step t , and step t counts as one unit in the infectious period. (iv) The counter changes from t to $t + 1$, and the sequence is repeated.

A note on notation and terminology: When we refer to a ‘period’ $[t_1, t_2]$, we mean the set of time steps $\{t_1, t_1 + 1, t_1 + 2, \dots, t_2\}$. When we refer to “time t ” we are referring to step t on the counter - this is a discrete time process. The first time step is $t = 1$, and $t = 0$ is not a time step, just a useful convention to express the state of the system before the first time step.

Observe the pair-wise independence of the interactions: If x, y, z meet at a vertex v at time step t , there is an interaction between each pair ($\{xy, yz, xz\}$) with probability ρ , independently of the other interactions. From the sequence of actions, it can be seen that it is also the case that an infection is not transitive in a single time step. For example, say they meet at the same vertex at time step t , x is infected but y and z are not. If x interacts with y but not z , then y does not pass on the infection to z at that time step, even if they interact at that time step. Let M_k be the total number of particles that ever get infected in the course of the process, and let T_k be the completion time.

3 Results

Theorem 1. *Suppose $\rho = 1$ and $k \rightarrow \infty$ as $n \rightarrow \infty$. Define*

$$\phi \equiv k \left(1 - \left(1 - \frac{1}{\theta_r n} \right)^\xi \right) \quad (1)$$

where $\theta_r = \frac{r-1}{r-2}$. Then

- (i) *If $\phi < 1$ then **whp**, $M_k = O(\ln k) + I_0$*
- (ii) *If $\phi \rightarrow c$ where $c > 1$ is some constant, then if $I_0 = o(k)$,*
 - (a) *with probability α , $M_k = C(k + o(k)) + I_0$,*
 - (b) *with probability $1 - \alpha$, $M_k = O(\ln k) + I_0$,**where C is the unique solution in $(0, 1)$ of the equation $1 - x = e^{-\phi x}$ and $\alpha = (1 + o(1))(1 - (1 - C)^{I_0})$.*
- (iii) *If $\phi > (1 + \varepsilon) \ln k$ where $\varepsilon > 1$ is some constant, then **whp** $M_k = k$ (i.e., all the particles get infected).*

Theorem 1 is for finite ξ , but observe that taking the convention that $x^\infty = 0$ for $|x| < 1$ means that part (iii) is consistent with the SI model, for which all particles get infected almost surely. The theorem effectively gives conditions for transitions between different possible “regimes” of behaviour. The most interesting is the regime of part (ii), which is entered as ϕ transitions from $\phi < 1$ to $\phi > 1$. When $I_0 = 1$, roughly speaking, in this transition the number of infected particles goes from very few ($O(\ln k)$) **whp**, to a constant fraction Ck with probability C , or $O(\ln k)$ with probability $1 - C$. Hence, it’s “all or nothing”, with a constant probability C of being “all”.

Observe further if $\xi = O(n/k)$, then we have $\phi \approx \frac{k\xi}{\theta_r n}$. So $\phi = c > 1$ if $\xi \approx \frac{c\theta_r n}{k}$. This is unsurprising since it can be seen from the tools we use in our analysis that the hitting time of one particle to any of the k others is distributed approximately $\text{Geom}(\frac{k}{\theta_r n})$, and hence has expected value $\frac{\theta_r n}{k}$.

Concerning the completion times, we will demonstrate in Sect. 7.2 that for $\xi = \infty$, (that is, the SI model) and $\rho = 1$, how the edge weightings can be exploited by a fairly straightforward application of a theorem of [16] to get

$$\frac{T_k}{\ln k/k} \xrightarrow{p} 2\theta_r n, \quad (2)$$

where T_k is the completion time for k particles. This is not a new result since it is implied by the matching expectation result from [8], however, our weighting scheme generalises to $\rho < 1$ and/or $\xi < \infty$ and in principle can be exploited in the same manner. In fact, for $\xi = \infty$, $\rho < 1$ we claim that (2) holds as $\frac{T_k}{\ln k/k} \xrightarrow{p} 2\frac{\theta_r n}{\psi}$.

For $\rho < 1$, we claim that Theorem 1 holds when the term $\frac{1}{\theta_r n}$ in (1) is replaced with $\frac{\psi}{\theta_r n}$ where

$$\psi = \frac{\rho(r-1)}{r-2+\rho}. \quad (3)$$

See [1].

4 Related Work

In this section, we briefly describe some of the relevant related work on diffusion processes like epidemic spreading and the dissemination of information in mobile environments.

There has been a growing body of work in the interacting particle systems community analysing epidemic models with mobile particles. In [10] the authors provide a review of the results, techniques and conjectures when the graph is an infinite lattice. In [21], the authors explore by means of mean-field approximations the evolution of the number of infected individuals when individuals perform random motions on the plane. Recent papers by Peres et al [19] and Pettarin et al [22] both analyse mobile networks modeled as multiple random walks; the former as Brownian motion on $\mathbb{R}^d$ and the latter as walks on the grid. In each case, there is a parameter r within which distance a pair of walks can communicate, producing a communication graph (which is disconnected below a certain threshold r_c). [19] studies various aspects of the communication graph, such as how long it takes a particular point to become part of the giant component. [22] studies the broadcast time T_B of a piece of information and originating in one agent in relation to r . Setting $r = 0$ means information is passed only when particles are coincident. In this case, T_B is equivalent to our completion time and [22] give $T_B = \tilde{O}(n/\sqrt{k})$.

Of closer relevance to this work are [12] and [11]. Both of these papers study infections carried by random walks on graphs. In particular [11] analyses an SI model similar to ours; multiple random walks on a graph G that carry a virus, and the quantity of interest is the completion time. They give a general upper bound $\mathbf{E}[T_k] = O(m^* \ln k)$ for any graph G , where m^* is the expected meeting time of a pair of random walks maximised over all pairs of starting vertices. Special cases are analysed too, in particular, they give an upperbound of $\mathbf{E}[T_k] = O(\frac{nr}{k} \ln k \ln n)$ for random r -regular graphs. This is a factor $\ln n$ larger than the precise value of the process considered in this paper.

Finally, we mention [3], which studies flooding on dynamic random networks. A fixed set of n vertices is part of a dynamic random graph process where each edge is an independent two-state Markov chain, either existing or not existing. A single initial vertex initially holds a message and any vertex which receives this message broadcasts it on existing edges for the next k steps. Although flooding is a different process to multiple random walks, the authors develop a reduction to a weighted random graph with some similarity to the interaction graphs we present. It allows them to derive relations between the edge-Markov probabilities and state asymptotic bounds on the number of vertices that receive the message, as well as the broadcast (equivalently, completion) time.

5 Assumptions and Approach

If each particle is distance at least $\omega(k, n) \equiv \Omega(\ln \ln n + \ln k)$ from every other then we say the particles are in *general position* (g.p.). We assume the following:

(i) The number of particles $k \leq n^\epsilon$ where ϵ is some sufficiently small constant. (ii) G is typical. (iii) Particles start in general position. (iv) $I_0 = |\mathcal{I}(0)| \geq 1$ and $|\mathcal{R}(0)| = 0$.

A *typical* graph (formally defined in the appendix) is one that has a number of specific properties. It is connected and non-bipartite, which implies that a random walk will converge to a stationary distribution π on the vertex set. Because the graph is regular, π will be the uniform distribution. A typical graph also has good expansion properties, meaning that a random walk will converge quickly to the stationary distribution; it will be *rapidly mixing*. In fact, for all the graph we consider, $O(k \ln n)$ steps is sufficient for our results. A typical graph also has most of its vertices treelike. A vertex v is *treelike* if there is no cycle in the subgraph $G[v, L_1]$ induced by the set of vertices within distance $L_1 = \lfloor \epsilon_1 \log_r n \rfloor$ of v , where $\epsilon_1 > 0$ is a sufficiently small constant. The near-uniformity and treelike structure throughout the graph allows us to make precise calculations for the behaviour of a random walk from a treelike vertex within the period of the mixing time. This helps us separate the analysis of the period within the mixing time from the period after it. The rapid mixing means that the dynamics of the process is dominated by the long-term behaviour very quickly, and that essentially, short-term behaviour is restricted to groups of local interactions that don't interfere with each other. This provides a degree of homogeneity to the process.

Assumptions (ii) and (iii) are not unreasonable: G is typical **whp** (Lemma 10 in appendix), and it is straightforward to verify that if the positions of each of the k particles are chosen **uar** from the vertices of G , then **whp** they will be in g.p. with respect to each other if ϵ is small enough. We will require additional assumptions beyond those stated in this section and previous ones, and they will be shown to hold **whp** too. Assumption (iv) implies $|\mathcal{S}(0)| = k - I_0$.

The analysis of multiple random walks can be reduced to a single random walk, in a way we informally describe here based on techniques developed in [8] and extended in [1]. We refer to the latter for details and rigorous justification for this section.

We define a *product graph* $H = (V_H, E_H)$ with vertex set $V_H = V^k$ and edge set $E_H = E^k$. The vertices $\mathbf{v}$ of H consist of k -tuples $\mathbf{v} = (v_1, v_2, \dots, v_k)$ of vertices $v_i \in V(G)$, $i = 1, \dots, k$, with repeats allowed. Two vertices $\mathbf{v}, \mathbf{w}$ are adjacent if $(v_1, w_1), \dots, (v_k, w_k)$ are edges of G . The graph H replaces k random walks $\mathcal{W}_{u_i}(t)$ on G with current positions v_i and starting positions u_i by a single walk $\mathcal{W}_{\mathbf{u}}^H(t)$.

If $S \subseteq V_H$, then $\Gamma = \Gamma(S)$ is obtained from H by contracting S to a single vertex $\gamma(S)$. All edges, including loops and parallel edges are retained, producing a multigraph. Thus $d_\Gamma(\gamma) = d_H(S)$, where d_F denotes vertex degree in graph F . Moreover Γ and H have the same total degree $(nr)^k$, and the degree of any vertex of Γ , except γ , is r^k .

We can calculate the times for a pair of particles (or one particle and a set of others, or any pair from a set of pairs, or whatever) to meet by calculating the time for the single random walk $\mathcal{W}_{\mathbf{u}}^H(t)$ to visit the relevant set of vertices

$S \subseteq V_H$. This is turn is done by the contraction described above, and calculating the time for the walk $\mathcal{W}_{\mathbf{u}}^\Gamma(t)$ on Γ to visit $\gamma(S)$. These two times asymptotically equal.

The dynamics of the system, is as follows. The random walks have a *mixing time* $T = O(k \ln n)$, and the distribution of a particle at some $\text{timet} \geq T$ is almost the stationary distribution. In this case, because the graph is regular, the stationary distribution is uniform across the vertices of G .

Consider a set A of particle pairs. For a particle pair $(x, y) \in A$, let $\mathcal{B}_{(x,y)}(t)$ be the event $\{\text{no pair in } A \text{ meet in the period } [T, t-1] \text{ and } x \text{ and } y \text{ meet at time } t\}$. Note, (x, y) is unordered. Subject to some restrictions (which are satisfied by the type of interactions relevant to this paper), it is established in [1] that

$$\Pr(\mathcal{B}_{(x,y)}(t)) \sim p(1 - |A|p)^{t-1}, \quad (4)$$

where

$$p = \frac{1}{\theta_r n} \left(1 + O\left(\frac{k \ln n}{n}\right) \right), \quad (5)$$

During the mixing time, we cannot be precise in general. However, if the particles begin in general position with respect to each other then **whp**, the particles don't meet in the mixing time. This is a consequence of Lemma 5. We say a visit to vertex v , or a particle-pair meeting is *T-distinct* if it occurs at least T steps after a previous T -distinct visit/meeting, or the start of the walk. As suggested by the above, we can view the dynamics within the mixing time separately to the dynamics after the mixing time, up until the visit to a vertex or a particle-pair meeting.

6 Two-Particle System

In this section, we discuss how the system behaves when $k = 2$. Let s and x be the two particles, with s being the initial infective. Suppose that the assumptions stated above hold. We allow $\xi < \infty$ and/or $\rho < 1$. The former conditions means that x may never get infected, the latter condition means that it may take more than one meeting between s and x before an interaction takes place. Note, that if s and x were at the same vertex at time t , and happen to move to the same neighbouring vertex in the next step, then this counts as another meeting.

By (4) with $A = \{(s, x)\}$,

$$\Pr(\text{first meeting between } s \text{ and } x \text{ after } T \text{ is at step } t) \sim p(1 - p)^{t-1}, \quad (6)$$

The RHS of (6) is the probability of a $\text{Geom}(p)$ random variable taking value t .

Now, suppose s and x have just stepped to the same vertex v . With probability ρ there will be an interaction. After this, they will move again, either to the same neighbour of v with probability $1/r$ or to different neighbours with probability $(r - 1)/r$. Let

$$\phi = \mathbf{Pr}(\text{No } sx \text{ interaction before they move apart}) \quad (7)$$

$$= \sum_{i \geq 1} (1 - \rho)^i \left(\frac{1}{r}\right)^{i-1} \left(1 - \frac{1}{r}\right) = \frac{(1 - \rho)(r - 1)}{r - 1 + \rho}. \quad (8)$$

The following lemma is from [8]:

Lemma 2. *Let G be a typical r -regular graph, and let v be a vertex of G , treelike to depth $L_1 = \lfloor \epsilon_1 \log_r n \rfloor$. Suppose that at time zero, two independent random walks (W_1, W_2) start from v . Let (a, b) denote the position of the particles at any step. Let $S = \{(u, u) : u \in V\}$. Let f be the probability of a first return to S within $T = O(k \ln n)$ steps given that the walks leave v by different edges at time zero. Then*

$$f = \frac{1}{(r - 1)^2} + O\left(n^{-\Omega(1)}\right).$$

Using this lemma, let

$$\phi_T = \mathbf{Pr}(\text{No } sx \text{ interaction before being apart more than } T \text{ time steps}) \quad (9)$$

$$= \sum_{i \geq 1} \phi^i f^{i-1} (1 - f) = \frac{\phi(1 - f)}{1 - \phi f} \quad (10)$$

Now, assuming s and x start at the same vertex,

$$\mathbf{Pr}(sx \text{ interaction occurs within } T \text{ time steps}) \quad (11)$$

$$\sim \mathbf{Pr}(sx \text{ interaction occurs before } s \text{ and } x \text{ have been apart more than } T \text{ steps}) \quad (12)$$

$$= 1 - \phi_T = 1 - \frac{\phi(1 - f)}{1 - \phi f} \sim \frac{\rho(r - 1)}{r - 2 + \rho} = \psi. \quad (13)$$

Recall ψ was defined in (3), and observe that $\rho \leq \psi \leq 1$ with $\psi = 1$ iff $\rho = 1$.

Clearly, the number of T -distinct meetings in $[0, t]$ is at most t/T . Subject to slightly more stringent restrictions, it can be shown (see [1]) that we have

$$\mathbf{Pr}(\text{there are } i \text{ } T\text{-distinct meetings in } [0, t]) \sim \binom{t}{i} p^i (1 - p)^{t-i}, \quad (14)$$

i.e., it is approximately distributed $\text{Bin}(t, p)$. The probability that there are no interactions in any of the i intervals $[t_j, t_j + T]$ where t_j is the time of the j 'th T -distinct meeting is $(1 - \psi)^i$. Thus

$$\begin{aligned} \mathbf{Pr}(\text{there are no interactions in the period } [0, t]) &\sim \sum_{i=0}^{t/T} \binom{t}{i} (p(1 - \psi))^i (1 - p)^{t-i} \\ &= \sim (1 - \psi p)^t \end{aligned}$$

Hence,

$$\mathbf{Pr}(x \text{ gets infected within time } \xi) \sim 1 - (1 - \psi p)^\xi. \quad (15)$$

When $\rho = \psi = 1$, (15) looks similar to the bracketed terms in (1). This is, of course, not a coincidence since the bracketed term in (1) is essentially the probability that an infection is passed between a pair of particles if one of them had been infected, and therefore, ϕ is effectively the expected number of other particles that are infected by a particular particle.

7 Interaction Graph Framework: SI Model

We remind the reader that proofs of lemmas and theorems not present in what follows are given in [1].

We can study the SI model (the case $\xi = \infty$) by analysing the edge weights of the *interaction graph* $\Upsilon = (V_\Upsilon, E_\Upsilon)$. This is a complete graph on the particle set $\mathcal{P}$, thus $V_\Upsilon = \mathcal{P}$ and $E_\Upsilon = \{(x, y) : x, y \in \mathcal{P}, x \neq y\}$. For a particle x , let $t(x)$ be the time at which x got infected, or ∞ if it never gets infected. For an *interaction edge* $e = (x, y) \in E_\Upsilon$, let $t(e) = \min\{t(x), t(y)\}$. Then the weight $w_\Upsilon(e)$ of the edge is a random variable defined as

$$w_\Upsilon(e) = \min\{t - t(e) : t > t(e) \text{ and there was an } xy \text{ interaction at step } t\}. \quad (16)$$

If the other particle was susceptible at the time of interaction, it becomes infected at the point of interaction. $\mathcal{T}$, therefore, is a space of randomly-weighted complete graphs, and the distribution on $\mathcal{T}$ is the joint distribution of the edge weight random variables $w_\Upsilon(e)$. A member Z of this space - an instance of Υ - is a specific set of values for each random variable $w_\Upsilon(e)$. We may write $Z_\Upsilon \in \mathcal{T}$ to denote a particular instance drawn from the space $\mathcal{T}$ with probability $\mathbf{Pr}(Z_\Upsilon)$.

We claim that the joint distribution of edge weights, is well approximated by the joint distribution of random edge weights of another complete graph on $\mathcal{P}$, $\Lambda = (V_\Lambda, E_\Lambda)$. The weight $w_\Lambda(e)$ of each edge $e \in E_\Lambda$ is an independent random variable with distribution $\text{Geom}(q)$, where $q = \frac{\psi}{\theta_r n}$.

Labelling the edges in Λ as e_i with $1 \leq i \leq \binom{k}{2}$, observe

$$\mathbf{Pr} \left(\bigwedge_{i=1}^{\binom{k}{2}} w_\Lambda(e_i) = z_i \right) = \prod_{i=1}^{\binom{k}{2}} q(1-q)^{z_i-1} \quad (17)$$

Observe that with probability one, neither graph has an edge of infinite weight.

7.1 Interaction Graph Approximation

In what follows we give a proof of the following theorem for the special case in which $\rho = 1$ (implying $\psi = 1$). We claim that it holds when $\rho < 1$ as well.

Theorem 3. *Let*

$$Z_F = \left\{ \bigwedge_{i=1}^{\binom{k}{2}} w_F(e_i) = z_i \right\}. \quad (18)$$

where $F \in \{\Upsilon, \Lambda\}$. Then

$$\Pr(Z_{\mathcal{R}}) = (1 + o(1))\Pr(Z_A). \quad (19)$$

Edge weightings in the interaction graph are pair-wise independent by virtue of the independence of the random walks. However there is a correlation when more than two edges are considered. For example, if particles x and y interacted at some time t and x and z interacted at some time close to t , then y and z are more likely to have interacted at some time close to t .

We say an edge e is *active* in the period $[t(e), t(e) + w_{\mathcal{R}}(e) - 1]$. The event $Z_{\mathcal{R}}$ can be restated as follows: For each edge $e_i = (x, y)$, the first xy interaction after time $t(e_i)$ is at time $t(e_i) + z_i$.

For a pair of particles x, y and some weighted graph F on the particle set, let $D_F(x, y)$ represent the weighted distance between x and y in F . Let $d_F(x) = \min\{D_F(s, x) : s \in \mathcal{I}(0)\}$. Furthermore, for an edge $e = (x, y)$, let $d_F(e) = \min\{d_F(x), d_F(y)\}$. $\square$

Lemma 4. *For a particle x , $t(x) = d_{\mathcal{R}}(x)$, and so for an edge $e \in E_{\mathcal{R}}$, $t(e) = d_{\mathcal{R}}(e)$.*

Therefore,

$$Z_{\mathcal{R}} = \{\text{For each edge } e_i = (x, y), \text{ there is no } xy \text{ interaction in the period } [d_{\mathcal{R}}(e_i) + 1, d_{\mathcal{R}}(e_i) + z_i - 1] \text{ and there is an } xy \text{ interaction at time } d_{\mathcal{R}}(e_i) + z_i\}. \quad (20)$$

Let A_0 be the set of active edges at time $\tau_0 = 0$. For example, in the case $\mathcal{I}(0) = \{s\}$, $A_0 = \{(s, x) : x \in \mathcal{P}, x \neq s\}$ whence $|A_0| = k - 1$. Let τ_i , $1 \leq i \leq R$ for some R , denote the times at which the set of active edges changes. Let A_i be the set of edges active in the period $[\tau_i, \tau_{i+1} - 1]$. We call the τ_i 's *epochs*. A particular edge $e_i = (x, y)$ will become active at some epoch $\tau = d_{\mathcal{R}}(e_i)$ and remain so until $\tau' - 1$ (inclusive), where $\tau' = d_{\mathcal{R}}(e_i) + z_i$ is some epoch after τ . Its period of activation may, therefore, cross a number of epochs, in which case it will be a member of a number of edge sets A_j .

We analyse each period $[\tau_j + 1, \tau_{j+1}]$ individually through application of (4). In the product graph H , we identify the set of vertices $S(A_j) \subseteq V_H$ such that $\mathbf{v} \in S(A_j)$ iff for some $(x, y) \in A_j$ $\mathbf{v}$ represents x and y being at the same vertex in G . Thus, if $\mathbf{v} = (v_1, v_2, \dots, v_k)$, then $v_r = v_s$ where r and s are the vector indices for x and y respectively. Since each pair is not to interact in the period $[\tau_j + 1, \tau_{j+1} - 1]$, and since we assume $\rho = 1$, the walk $\mathcal{W}$ on H must avoid the set $S(A_j)$ until $\tau_{j+1} - 1$ (inclusive). Then, at τ_{j+1} it must visit $S_j \subseteq S(A_j)$. $\mathbf{v} \in S_j$ iff it represents any of the $(x, y) \in A_j$ that interact at time τ_{j+1} being at the same vertex in G . We deal with (non-)visits to $S(A_j)$ and S_j by applying (4) to their contractions $\gamma(S(A_j))$ and $\gamma(S_j)$.

We do not prove (19) for all possible $Z_{\mathcal{R}}$, only those that are termed *good*. Let $T = O(k \ln n)$ be a mixing time and let $L = T^3$, and let $\ell = 2(T + L)$.

Call $Z_{\mathcal{T}}$ *good* if it satisfies these criteria: (a) $k^3 \ln n \sum_{i=1}^{\binom{k}{2}} z_i = o(n^2)$ (b) None of the $\binom{k}{2}$ interactions that form the edges of $\mathcal{T}$ take place within ℓ steps of each other.

Lemma 5. *With high probability, $Z_{\mathcal{T}} \in \mathcal{T}$ is good.*

We shall assume the $Z_{\mathcal{T}}$ that is drawn is good. Note part (b) implies $R = \binom{k}{2}$ and so we only need to consider cases of $Z_{\mathcal{T}}$ where, for each j , S_j is a single particle pair. It also implies that during the period $[\tau_j + 1, \tau_j + \ell]$, we do not need to account for interactions. The following lemma applies (4) sequentially for each period $[\tau_j + 1, \tau_{j+1}]$. By the Markov property, we can multiply the probabilities for each period. Theorem 3 follows from this.

Lemma 6. *Let $p_j = |A_j|p(1 - O(k^3 \ln n/n))$ where p is from (5),*

$$\Pr(Z_{\mathcal{T}}) = \prod_{j=0}^R (1 + O(L\pi_j) + O(\epsilon_j))p(1 - p_j)^{\tau_{j+1} - \tau_j - 1} \quad (21)$$

$$= (1 + o(1))\Pr(Z_A). \quad (22)$$

In (21) we have included correction factors which we left out in (4) for clarity.

7.2 Completion Time

In [8] an expectation of $\frac{2\theta_{\mathcal{T}}n}{k} \ln k$ was determined for the completion time of a broadcasting model on k particles that is equivalent to the SI model with $I_0 = 1$ and $\rho = 1$. We apply a theorem from [16] to get a convergence in probability to the same value. Assign each edge (i, j) of a complete graph on n vertices a random weight Y_{ij} . The weights are assumed to be independent and identically distributed, non-negative and satisfying $\Pr(Y_{ij} \leq t) = t + o(t)$ as $t \rightarrow 0$. Let X_{ij} be the minimal total weight of a path between a given pair of vertices i, j . The theorem states $\frac{\max_{j < n} X_{ij}}{\ln n/n} \xrightarrow{P} 2$. By scaling, we can apply the result to an exponential analogue of the edge probabilities of Λ , and show that the distribution is well approximated, thereby giving (2). See [1] for details.

8 Interaction Graph Framework: SIR Model

To analyse the SIR model, we build an interaction graph but we modify the process by considering two phases. Phase 1 assigns edge weights as before: an edge $e = (x, y)$ is weighted by the time it takes for an xy interaction after one of them has become infected. It is possible, of course, that neither ever get infected when ξ is finite, in which case the edge (x, y) does not become active (and is not given a weight) in this phase. Phase 1 ends at τ_{end} when there are no more active edges. At this point, there remains a set of susceptible particles $\mathcal{S}(\tau_{end})$ that were never infected in phase 1. Phase 2 begins with an arbitrary $z \in \mathcal{S}(\tau_{end})$

being given a pseudo-infection with an infinite infectious period, and we proceed with an SI process on the particle set $\mathcal{S}(\tau_{end})$, giving them edge weights, as per the normal SI process. A pseudo-infection is the same process as an infection but we do not count pseudo-infected particles as infected; only particles infected in phase 1 are counted as infect. Phase 2 exists to maintain the same probability space as in the SI model.

Call the resulting (complete) graph Ψ . We remove an edge e in Ψ if $w_\Psi(e) > \xi$ and call the resulting graph Ψ' .

Let $\mathcal{C}(x)$ be the component of Ψ' containing a particle x . Let $D_{\Psi'}(x, y)$ be the weighted distance between particles x and y if $y \in \mathcal{C}(x)$, otherwise let $D_{\Psi'}(x, y) = \infty$. Let $d_{\Psi'}(y) = \min\{D_{\Psi'}(x, y) : x \in \mathcal{I}(0)\}$.

Theorem 7. *In Ψ' , (i) A particle y is infected if and only if $y \in \mathcal{C}(x)$ for some initial infective $x \in \mathcal{I}(0)$, (ii) If a particle y gets infected, the infection time is $t(y) = d_{\Psi'}(y)$.*

We show that Theorem 3 holds for Ψ . Let Λ and Z_F for some graph space F be as before.

Lemma 8.

$$\Pr(Z_\Psi) = (1 + o(1))\Pr(Z_\Lambda). \quad (23)$$

As with Ψ , Ψ' is a space of weighted graphs. Furthermore, we can define an equivalence relation $\sim$ on Ψ' such that for $Y, Z \in \Psi'$, we have $Y \sim Z$ iff $E(Y) = E(Z)$, that is, they have the same edge set. Let $\Psi'_{/\sim}$ be the graph space induced by Ψ' and the equivalence relation $\sim$. Thus, the probability a graph G is drawn from $\Psi'_{/\sim}$, that is, $\Pr(G \in \Psi'_{/\sim}) = \Pr(\text{some } Z \text{ is drawn from } \Psi' \text{ such that } E(Z) = E(G))$. We show $\Psi'_{/\sim}$ is approximated by $\mathcal{G}_{k, \hat{q}}$, the space of Erdős–Renyi random graphs with edge probability $\hat{q} = 1 - (1 - q)^\xi$, where $q = \frac{\psi}{\theta_{r,n}}$ for the special case $\rho = \psi = 1$.

Lemma 9. *Let G be a graph on the particle set $\mathcal{P}$.*

$$\Pr(G \in \Psi'_{/\sim}) = (1 + o(1))\Pr(G \in \mathcal{G}_{k, \hat{q}}). \quad (24)$$

Proof of Theorem 1. For an Erdős–Renyi random graph space $\mathcal{G}_{n,p}$ on n vertices and edge probability p , the connectivity results are well known (see, e.g., [13]). By Lemma 9, the **whp** statements carry through to $\Psi'_{/\sim}$. Since Theorem 1 deals with the case of one initial infective s , the infected particles will be those in $\mathcal{C}(s)$. Cases (i) and (iii) of Theorem 1 are straightforward to see from the above. For case (ii), there is a unique giant component g of size C_1 in $G \in \Psi'_{/\sim}$. By the symmetry of the model, the probability any particular particle x being in g is $|g|/k = (1 + o(1))C$. Thus, the giant component is infected with this probability. Otherwise, s will be placed in a component of size at most $O(\ln k)$. $\square$

References

1. Abdullah, M., Cooper, C., Draief, M.: Viral Processes by Random Walks on Random Regular Graphs (2011), <http://arxiv.org/abs/1104.3789>

2. Aldous, D., Fill, J.: Reversible Markov Chains and Random Walks on Graphs, <http://stat-www.berkeley.edu/pub/users/aldous/RWG/book.html> (in preparation)
3. Baumann, H., Crescenzi, P., Fraigniaud, P.: Parsimonious Flooding in Dynamic Graphs. In: Proceedings of the 28th ACM Symposium on Principles of Distributed Computing, pp. 260–269 (2009)
4. Buscarino, A., Fortuna, L., Frasca, M., Latora, V.: Disease Spreading in Populations of Moving Agents. *EPL (Europhysics Letters)* 82(3) (2008)
5. Chaintreau, A., Hui, P., Scott, J., Gass, R., Crowcroft, J., Diot, C.: Impact of Human Mobility on Opportunistic Forwarding Algorithms. *IEEE Transactions Mobile Computing* 6(6), 606–620 (2007)
6. Cooper, C., Frieze, A.: The Cover Time of Random Regular Graphs. *SIAM Journal on Discrete Mathematics* 18, 728–740 (2005)
7. Cooper, C., Frieze, A.M.: The Cover Time of the Giant Component of a Random graph. *Random Structures and Algorithms* 32, 401–439 (2008)
8. Cooper, C., Frieze, A.M., Radzik, T.: Multiple Random Walks in Random Regular Graphs. *SIAM Journal on Discrete Mathematics* 23, 1738–1761 (2009)
9. Daley, D., Gani, J.: Epidemic Modelling: An Introduction. *Studies in Mathematical Biology*. Cambridge University Press, Cambridge (2001)
10. Dickman, R., Rolla, L., Sidoravicius, V.: Activated Random Walkers: Facts, Conjectures and Challenges. *Journal of Statistical Physics* 138, 126–142 (2010)
11. Dimitriou, T., Nikolettseas, S., Spirakis, P.: The Infection Time of Graphs. *Discrete Applied Mathematics* 154, 18 (2006)
12. Draief, M., Ganesh, A.: A random Walk Model for Infection on Graphs: Spread of Epidemics and Rumours with Mobile Agents. *Discrete Event Dynamic Systems* 21, 1 (2011)
13. Draief, M., Massoulié, L.: Epidemics and Rumours in Complex Networks. *London Mathematical Society Series*, vol. 369. Cambridge University Press, Cambridge (2010)
14. Feller, W.: An Introduction to Probability Theory, 2nd edn., vol. I. Wiley, New York (1960)
15. Friedman, J.: A Proof of Alon’s Second Eigenvalue Conjecture and Related Problems, vol. 195. *Memoirs of the American Mathematical Society*, Providence (2008)
16. Janson, S.: One, Two and Three Times $\log n/n$ for Paths in a Complete Graph with Random Weights. *Combinatorics, Probability and Computing* 8, 347–361 (1999)
17. Kurtz, T.G., Lebensztayn, E., Leichsenring, A.R., Machado, F.P.: Limit Theorems for an Epidemic Model on the Complete Graph. *Latin American Journal of Probability and Mathematical Statistics* 4, 45–55 (2008)
18. Lovász, L.: Random Walks on Graphs: A Survey, *Combinatorics*. In: Paul Erdős is Eighty, vol. 2, pp. 1–46. Bolyai Society Mathematical Studies (1993)
19. Peres, Y., Sinclair, A., Sousi, P., Stauffer, A.: Mobile Geometric Graphs: Detection, Coverage and Percolation. In: Proceedings of the Twenty-Second Annual ACM-SIAM Symposium on Discrete Algorithms, pp. 412–428 (2011)
20. Pittel, B.: On spreading a rumor. *SIAM Journal of Applied Mathematics* 47(1), 213–223 (1987)
21. Rhodes, C., Nekovee, M.: The Opportunistic Transmission of Wireless Worms Between Mobile Devices. *Physica A: Statistical Mechanics and Its Applications* 387(27), 6837–6840 (2008)
22. Pettarin, A., Pietracaprina, A., Pucci, G., Upfal, E.: Tight Bounds on Information Dissemination in Sparse Mobile Networks. In: Proceedings of the 30th Annual ACM SIGACT-SIGOPS Symposium on Principles of Distributed Computing (2011)

23. Zhou, J., Liua, Z.: Epidemic Spreading in Communities with Mobile Agents. *Physica A: Statistical Mechanics and its Applications* 388(7), 1, 1228–1236 (2009)

Appendix

Typical r -regular Graphs

We say an r -regular graph G is *typical* if it satisfies the properties P1–P4 listed below. We first give some definitions.

Let

$$L_1 = \lfloor \epsilon_1 \log_r n \rfloor, \quad (25)$$

where $\epsilon_1 > 0$ is a sufficiently small constant.

A vertex v is *treelike* if there is no cycle in the subgraph $G[v, L_1]$ induced by the set of vertices within distance L_1 of v .

A cycle C is *small* if $|C| \leq L_1$.

P1. G is connected and not bipartite.

P2. The second eigenvalue of the adjacency matrix of G is at most $2\sqrt{r-1} + \epsilon$, where $\epsilon > 0$ is an arbitrarily small constant.

P3. There are at most $n^{2\epsilon_1}$ vertices on small cycles.

P4. No pair of cycles C_1, C_2 with $|C_1|, |C_2| \leq 100L_1$ are within distance $100L_1$ of each other.

Note that P3 implies that at most n^{ϵ_C} vertices of a typical r -regular graph are not treelike, where

$$n^{\epsilon_C} = O(r^{L_1} n^{2\epsilon_1}) = O(n^{3\epsilon_1}). \quad (26)$$

Lemma 10. *Let $\mathcal{G}'_r \subseteq \mathcal{G}_r$ be the set of typical r -regular graphs. Then $|\mathcal{G}'_r| \sim |\mathcal{G}_r|$.*

For Lemma 10, that P2 holds **whp** is a very difficult result of Friedman [15]. The other properties are straightforward to establish; see, e.g., [6].