Contact process on a graph with communities

David Sivakoff

Departments of Statistics and Mathematics
The Ohio State University
Columbus, OH 43210, USA
E-mail address: dsivakoff@stat.osu.edu
URL: http://www.stat.osu.edu/~dsivakoff/

Abstract. We study the spread of an epidemic between two communities. We model the two communities as independent random graphs, then add a small set of bridge edges between the communities, and the epidemic on this network is the contact process. When the epidemic is supercritical, we prove that it survives for exponentially long in the network size. Further, when the epidemic survives, we prove that the time for the infection to spread to the second community when started from a single individual in the first converges in distribution when suitably normalized. These results generalize to a graph with N communities.

1. Introduction

Let $G_1 = (V_1, E_1)$ and $G_2 = (V_2, E_2)$ be independent instances of $\mathcal{G}(n, p)$, the Erdös-Rényi random graph ensemble with n vertices and edge probability p. Construct the graph $G = (V, E)$ such that $V = V_1 \cup V_2$ and $E = E_1 \cup E_2 \cup B$, where $B \subset V_1 \times V_2$ is a set of 'bridge' edges chosen independently of G_1 and G_2. When B is a small set of edges relative to E_i ($i = 1, 2$), then the graph will have two distinct communities with a higher concentration of edges within each community than between the two communities. We study the contact process (defined below) on this graph as a model for a Susceptible-Infected-Susceptible epidemic in a network with two communities. Figure 1.1 depicts simulations of the contact process on a network with 1000 total vertices ($n = 500$ in each community), mean degree $np = 50$ and $|B| = 1$ bridge edge. Each line represents an independent simulation in which initially there are 2 infected vertices in V_1 and all other vertices are healthy; infected vertices become healthy at rate 1 and transmit the infection to their neighbors at rate $\lambda = 0.06$. In each simulation the infection very quickly reaches a quasi-equilibrium state in the first component V_1, then makes a jump to the second component V_2 at a random time. Our goal is to rigorously prove this

behavior, and to determine the limiting distribution of the random jump time. We
also prove that the contact process survives for exponentially long (in n) on the
random graph before eventually hitting the absorbing state in which all vertices are
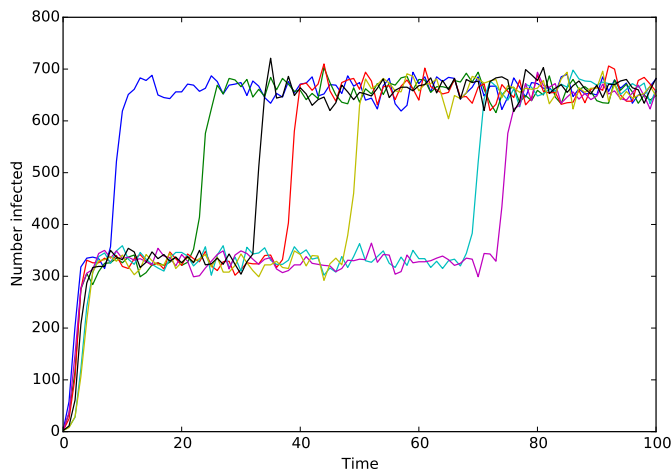healthy.

FIGURE 1.1. Simulation results for the contact process on G with
$|B| = 1$, $n = 500$, $p = 0.1$ and $\lambda = 0.06$ so that $b = 3$. Each line
represents an independent trial started from two initially infected
vertices in V_1.

The contact process on a graph G is a continuous time Markov process $(\xi_t)_{t\geq0}$,
where $\xi_t \subset V$ denotes the set of infected vertices at time t. If $\lambda = \lambda(n, p) > 0$
is the infection rate, then an infected vertex sends the infection to each of its
neighbors in G according to independent Poisson processes with rate λ, and it
becomes uninfected according to an independent Poisson process with rate 1. In
effect, a healthy vertex $v \notin \xi_t$ becomes infected at rate $\lambda |\mathcal{N}(v) \cap \xi_t|$, where $\mathcal{N}(v)$
denotes the set of neighbors of v in G, while an infected vertex becomes healthy at
rate 1. This is made rigorous by Harris' graphical construction, which is described
in Section 4.

The contact process has been studied on two different models of power-law ran-
dom graphs by Chatterjee and Durrett (2009) and Berger et al. (2005). The term
power-law random graph refers to the degree distribution having tails that decay
like $ck^{-\alpha}$. For the random graphs considered by Chatterjee and Durrett (2009)
and Berger et al. (2005) it was shown that the contact process survives on these
graphs for a long time (superpolynomial in $|V|$) for any $\lambda > 0$. These results were
extended by Can and Schapira (2015) to include graphs with unbounded mean
degree, and by Can (2015) who computed the metastable density for the contact
process in the preferential attachment graph. The results of Chatterjee and Durrett
(2009) were further strengthened by Mountford et al. (2016) and Mountford et al.
(2013), who showed that the infection survives for exponentially long in $|V|$ and
computed the metastable density for the contact process on the configuration model
of power-law random graphs. This was in contrast to the mean-field calculations of

Pastor-Satorras and Vespignani ([2001a](#),[b](#)), which predicted that for $\alpha > 3$ there is a $\lambda_c > 0$ so the contact process will die out quickly for $\lambda < \lambda_c$. Subsequently, Peterson ([2011](#)) showed that $\lambda_c > 0$ for the contact process on the complete graph with random vertex weights following a power-law, and in fact explicitly gives the value of λ_c in terms of the second moment of the vertex weights.

1.1. *Main Results.* We study the high-degree regime where the mean degree of the random graph scales as $np = n^a$ where $a \in (0,1]$ is a constant. Since we want the graph to be connected, we need the average degree to be at least $np \geq c \log n$ for $c > 1$. However, part of the proof requires that the random walk on $\mathcal{G}(n,p)$ be very close to uniformly distributed on the vertices after a short amount of time (see Lemmas 2.6 and 4.1). This part of the proof is simplified when $np = n^a$, though a similar method may work for smaller average degrees. Additionally, in the early stages of the infection process we must have large neighborhoods so that transmission events between neighboring infected vertices are rare. For smaller degrees, additional care will be needed to guarantee that the infection spreads out quickly.

Before we state our main results, we will need some notation. Let $\mathbf{P}[\cdot]$ denote the law of $G = G_1 \cup G_2 \cup B$, where $G_1, G_2 \sim \mathcal{G}(n,p)$ are independent random graphs, and the edges in B are chosen independently of G_1 and G_2 such that each edge in B has exactly one endpoint in G_1 and one in G_2 (the specific distribution of the edges in B does not matter for our proofs). The dependence on n and p has been suppressed from our notation. Once G is chosen according $\mathbf{P}[\cdot]$, it is fixed for all time, so we let $\mathbb{P}_A(\cdot)$ be the law of the contact process conditional on G with $\xi_0 = A \subset V$, and abbreviate $\mathbb{P}_v(\cdot) := \mathbb{P}_{\{v\}}(\cdot)$. We will also often want to observe the contact process on only one of the two subcomponents. That is, we will ignore the edges in B, so $\xi_t = \xi_t \cap V_1$ for all t whenever $\xi_0 = A \subset V_1$, and we denote the law of this restricted process conditional on G by $\mathbb{P}_A^{G_1}(\cdot)$. We will use $a_n = O(b_n)$ to mean $a_n \leq C b_n$ for $C > 0$, $a_n = \Omega(b_n)$ to mean $a_n > c b_n$ for $c > 0$, and $a_n = o(b_n)$ to mean $a_n/b_n \to 0$. For sequences of non-negative random variables, we use the notation $X_n = O_\mathbf{P}(a_n)$ to mean there exists a constant $C > 0$ such that $\mathbf{P}[X_n \leq C a_n] \to 1$ as $n \to \infty$, and $X_n = o_\mathbf{P}(a_n)$ to mean that for any $\delta > 0$, $\mathbf{P}[X_n \leq \delta a_n] \to 1$

First we must guarantee that the contact process on the random graph can survive long enough to spread from the first community to the second. If $np\lambda = b < 1$ is constant, then the contact process is dominated by a subcritical branching process, and dies out quickly, so we assume that $np\lambda = b > 1$. In this case we have the following theorem, which says that the supercritical contact process survives exponentially long on the random graph when it survives for at least $\Omega(\log \log n)$ time. We assume survival for at least $\Omega(\log \log n)$ time because at this point the infection has either died out (with probability $\approx 1/b$) or reached order $\log n$ vertices.

Theorem 1.1. *Fix constants $a \in (0,1]$ and $b > 1$. Consider the contact process, ξ_t, on the Erdős-Rényi random graph $G_1 \sim \mathcal{G}(n,p)$ with $np = n^a$, $np\lambda = b > 1$, and $\xi_0 = \{v\} \subset V_1$, and let $r = \frac{2}{b-1} \log \log n$. Then there exist constants $\eta_3, \epsilon, c > 0$ depending on b so that for any $\delta > 0$*

$$\mathbf{P}\left[\mathbb{P}_v^{G_1}\left(\min_{t \in [\eta_3 \log n, e^{cn}]} |\xi_t| \leq \epsilon n \;\middle|\; |\xi_r| > 0\right) > \delta\right] \to 0.$$

The constants η_3 and ϵ in Theorem 1.1 are defined in Lemma 3.2, and the proof appears thereafter. We use an approach similar to the one employed by Peterson (2011) to prove exponential survival of the supercritical contact process on the complete graph with random vertex weights. That is, we show that if the size of the contact process initially exceeds γn for some $\gamma > \epsilon$, then in a small constant amount of time, the size of the contact process is likely (with probability exponentially close to 1) to have increased by the end of the time interval without ever having dropped below size ϵn. The main difference between our proof and Peterson's is that we must rely on an isoperimetric inequality (Lemma 2.3) to guarantee that the contact process has room to expand.

Our main question is, if the infection is started at a single vertex $v_0 \in V_1$, and the infection is able to spread, how long will it be before a positive fraction of the vertices in V_2 are infected? In other words, how long does the infection take to cross a bridge between the two populations and spread throughout the second population? Theorem 1.2 answers this question.

Theorem 1.2. *Fix $a \in (0,1]$ and $b > 1$. Suppose $np = n^a$, $np\lambda = b > 1$, and there are $0 < |B| = o(n^a / \log n \log \log n)$ bridges between V_1 and V_2. Choose any $v_0 \in V_1$. There exists $\epsilon > 0$ so that if $\tau := \inf\{t > 0 : |\xi_t \cap V_2| > \epsilon n\}$, then for any $x \in [0, \infty)$ and any $\delta > 0$*

$$\mathbf{P}\left[\left|\mathbb{P}_{v_0}\left(\frac{\tau}{n^a/|B|} \leq x\right) - \left(1 - \frac{1}{b}\right)\left[1 - \exp\left(-\frac{(b-1)^2}{b}x\right)\right]\right| > \delta\right] \to 0.$$

We remark that in Theorems 1.1 and 1.2, the infimum can be taken over all initially infected vertices $v_0 \in V$ inside of the **P**-probability, and the proofs hold as written. Theorem 1.2 says that when the infection survives long enough (which happens with probability $\approx 1 - 1/b$), the distribution of τ, the time at which the infection has spread to a positive fraction of vertices in V_2, is approximately exponential with rate $|B| (b-1)^2/(bn^a)$. The upper bound on the number of bridges that we can accommodate is because our proof requires that the amount of time between the first $\log n$ successive potential transmissions of the infection between the two communities is at least $\Omega(\log \log n)$, and the rate at which such potential transmissions occur is $|B|\lambda = b|B|/n^a$. The maximum number of bridges allowable to guarantee the separation of timescales seen in Figure 1.1 should be $O(n^a)$ as the following mean-field argument demonstrates. For small times t, the number of infected vertices in V_1 at time t is approximately $e^{(b-1)t}$, since each infected vertex has expected degree np and spreads the infection to each neighbor at rate λ, while becoming healthy at rate 1. Therefore, the number of infected vertices in V_1 will reach ϵn at time $s = \log(\epsilon n)/(b - 1)$. The expected rate at which the infection is transmitted from V_1 to V_2 at time t is approximately $e^{(b-1)t}\lambda |B|/n = be^{(b-1)t} |B| /n^{1+a}$, so the expected number of times the infection is transmitted from V_1 to V_2 before time s is

$$\int_0^s \frac{be^{(b-1)t} |B|}{n^{1+a}} dt = \frac{b(e^{(b-1)s} - 1) |B|}{n^{1+a}(b - 1)} \approx \frac{|B| \epsilon b}{n^a(b - 1)}.$$

So, if $|B| \gg n^a$ then the infection is likely to spread to V_2 before reaching size ϵn in V_1. It is worthwhile to note that if V_1 and V_2 are two halves of a homogeneous Erdős-Rényi random graph with mean degree n^a then $|B| \approx n^{1+a}$, so no separation of timescales should be observable.

Theorem 1.2 generalizes easily to N communities, $G_1, \ldots, G_N$, each independent and distributed as $\mathcal{G}(n,p)$ with $np = n^a$ (N is fixed). If B_{ij} is the (possibly empty) set of bridge edges between communities i and j, and $B_i = \cup_j B_{ij}$, then we assume $\max_i |B_i| = o(n^a / \log n \log \log n)$ and for all $1 \le i, j, k, \ell \le N$ there are constants $c, C > 0$ such that whenever $|B_{ij}| > 0$ and $|B_{k\ell}| > 0$ then

$$c \le \frac{|B_{ij}|}{|B_{k\ell}|} \le C.$$

Under this assumption, all pairs of communities have either a comparable number of bridges, or no bridges (are not directly connected), so we let $\beta_{ij} = |B_{ij}| / \max_{k,\ell} |B_{k\ell}|$. We can then define a process $\chi_t \in \{0,1\}^N$ such that $\chi_t(i) = 1$ if and only if community i has at least ϵn infected vertices at time $tn^a / \max_{k,\ell} |B_{k\ell}|$. If $np\lambda = b > 1$ and $\chi_0 = (1,0,0,\ldots,0)$, then χ_t converges in distribution (as $n \to \infty$) on a finite time interval to the monotone stochastic process in which $\chi_t(i)$ flips from 0 to 1 at rate $\sum_j \chi_t(j)\beta_{ij}(b-1)^2/b$. That is, at the community level, the infection process resembles an SI epidemic with inhomogeneous infection rates.

The assumption that the number of bridges between connected communities is of a single order of magnitude is not necessary, except that things become a bit more complicated when the number of bridges between communities can span multiple orders of magnitude. In particular, we need to consider multiple time scales. The time for the infection to pass from community i to community j is proportional to the 'length' of the shortest path between the two, where a path from i to j is a sequence of communities $i = x_0, x_1, \ldots, x_k = j$ such that $\left|B_{x_\ell x_{\ell+1}}\right| > 0$, the 'length' of this path is $1/\min_\ell \left|B_{x_\ell x_{\ell+1}}\right|$ and the shortest path is the one with the minimal 'length'. This is because most time will be spent while the infection attempts to cross the narrowest bottleneck.

1.2. *Proof Outline.* We assume throughout most of the proof of Theorem 1.2 that there is a single bridge edge, $B = \{(u_b, v_b)\}$ with $u_b \in V_1$ and $v_b \in V_2$. The extension to multiple bridges is straightforward, and we comment on how this is done in Section 5.

Step 1. When the contact process starts from a single vertex, $\xi_0 = \{v_0\} \subset G_1$, for a short amount of time, $t \le r = O(\log \log n)$, $|\xi_t|$ is well approximated by a continuous time branching process with survival probability $1 - 1/b$, where $b = np\lambda$ (Lemmas 3.1 and 3.2). By time r, the contact process has either died out, or is destined to survive for exponentially long with positive density by Theorem 1.1. The main difficulty in this part of the proof is managing 'collision' events, where an infected vertex attempts to spread the infection to another already infected vertex. Once the process grows beyond size n^a, these collision events potentially stifle the growth of ξ_t. We overcome this obstacle by the use of the isoperimetric bound given in Lemma 2.3. This bound says that as long as the contact process contains at most a small fraction of vertices in V_1, then the number of edges between ξ_t and $V \setminus \xi_t$ will be large. We use this bound to show that while $|\xi_t|$ is not too large, it stochastically dominates a random walk with positive drift.

Step 2. We use the self-duality of the contact process, which essentially means that the time reversal of the contact process has the same distribution as the contact process – we carefully define the dual process in Section 4. The dual process started at time $t > 0$, $\{\zeta_s^t\}_{s \in [0,t]}$, is constructed so that if it is started from a single vertex $\zeta_0^t = \{v\} \subset V$, then $\zeta_s^t \cap \xi_{t-s} \ne \emptyset$ for some $s \in (0,t)$ if and only if $v \in \xi_t$. In

the graphical construction of the contact process, edges attempt to transmit the infection at rate λ independent of ξ_t, so we observe the dual process started from u_b when the bridge edge (u_b, v_b) attempts to transmit the infection. The dual process will reach size $\Omega(\log n)$ by time r with probability close to $1 - 1/b$. The primary difficulty at this point, and really the main obstacle in the proof, is getting the dual process, which has grown backwards in time, to intersect with the contact process, which has grown forwards in time to size at least ϵn (since G is not the complete graph, it is possible for all of the vertices in the contact process to be far from the vertices in the dual process). We do this by coupling the particles in the dual process with a random walk process in which each particle jumps according to an independent simple random walk on the vertices of G and dies at rate 1. Those particles which survive for time $t_{\mathrm{mix}} = O(\log \log n)$ and do not collide with other particles will all be well mixed (Lemmas 2.6 and 4.1), and each intersects with the contact process with probability larger than ϵ, which implies that the two processes intersect with high probability.

Step 3. At the times when the bridge edge (u_b, v_b) attempts to transmit the infection (a Poisson process $\{T_k^{(u_b, v_b)}\}_k$ with rate λ) we start an independent dual process. Each dual process to survive to time r will intersect with ξ_t and result in the infection spreading to v_b. In turn, v_b will spread the infection to V_2 with probability approaching $1 - 1/b$ by repeating the first step of the proof. Therefore, the number of times that the bridge edge must attempt to transmit the infection approaches a Geometric$([1 - 1/b]^2)$ distribution, and the interarrival times between successive attempts are distributed as independent Exp(λ). So on the event that $\xi_r \neq \emptyset$, the time required to spread the infection to V_2 approaches an Exp$(\lambda[1 - 1/b]^2)$ distribution, and the probability that $\xi_r \neq \emptyset$ approaches $1 - 1/b$; this is the statement in Theorem 1.2.

Section 2 is devoted to proving two key lemmas about Erdős-Rényi random graphs with moderate degree – an isoperimetric bound and a strong mixing time estimate. In Section 3 we compare the early stages of the contact process with a branching process and prove Theorem 1.1. In Section 4 we prove a key lemma about the dual of the contact process, and in Section 5 we put everything together for the proof of Theorem 1.2. The Appendix contains some basic properties of branching processes that are used in Section 3.

2. Isoperimetric Inequality and Mixing Time Bounds

In this section we prove Lemma 2.3, which gives a bound on the ϵ-isoperimetric number (defined below) of an Erdős-Rényi random graph with large mean degree, and Lemma 2.6, which says that the random walk on the high-degree random graph is almost uniformly distributed after a constant number of steps. These two properties of the random graph, which is fixed for all time, hold with **P**-probability tending to 1 as $n \to \infty$, so when we later consider the contact process on a random graph, we can assume it has these properties.

We will make frequent use of the following Bernstein inequality for binomial probabilities (as it appears in Prokhorov (2002)).

Lemma 2.1. *Suppose* $X \sim Bin(m, q)$. *Then for all* $t > 0$,

$$\mathbb{P}\left(|X - \mathbb{E}X| \geq t\right) \leq 2\exp\left\{\frac{-t^2}{2(VarX + t/3)}\right\}.$$

Definition 2.2. *For $U \subset V$, let $\partial U \subset E$ denote the set of edges with exactly one end vertex in U. Define the ϵ-**isoperimetric number**, $i_\epsilon(G)$, of a graph, $G = (V, E)$ as*

$$i_\epsilon(G) = \min \left\{ \frac{|\partial U|}{|U|} \; \middle| \; U \subset V, \; |U| \leq \epsilon |V| \right\}.$$

Lemma 2.3. *If $G \sim \mathcal{G}(n, p)$ is an Erdős-Rényi random graph on n vertices with edge probability p such that $np \geq 28(\log n)^3$, then for any fixed $\epsilon > 0$*

$$\mathbf{P}\left[i_\epsilon(G) \geq (1 - \epsilon)np - (np)^{2/3} \right] \to 1$$

as $n \to \infty$.

Note that this bound is essentially tight since if U is any deterministic set of ϵn vertices, then $|\partial U| \sim \text{Binomial}(|U|\,(1 - \epsilon)n, p)$ so $|\partial U| = |U|\,(1 - \epsilon)np(1 + o_{\mathbf{P}}(1))$. The proof is adapted from the proof of a similar result for the $\frac{1}{2}$-isoperimetric number of random regular graphs due to Bollobás (1988).

Proof: Denote by $P(u, m)$ the probability of the event that G contains a set of vertices $U \subset V$ with $|U| = u$ and $|\partial U| \leq m$. We will have proved the lemma once we show that

$$\sum_{u=1}^{\epsilon n} P(u, m(u)) = o(1)$$

where $m(u) := u\left[(1 - \epsilon)np - (np)^{2/3} \right]$. By Markov's inequality

$$P(u, m(u)) \leq \binom{n}{u} \sum_{s=0}^{m(u)} \binom{u(n - u)}{s} p^s (1 - p)^{u(n-u)-s}$$

$$= \binom{n}{u} \mathbb{P}\left(X_u \leq m(u) \right)$$

where $X_u \sim \text{Binomial}(u(n - u), p)$. Applying Lemma 2.1 with $t = \mathbb{E}X_u - m(u) = u(\epsilon np - up + (np)^{2/3})$, noting that $t \geq u(np)^{2/3}$ (since $u \leq \epsilon n$) gives

$$\mathbb{P}\left(X_u \leq m(u) \right) \leq \mathbb{P}\left(|X_u - \mathbb{E}X_u| \geq u(np)^{2/3} \right)$$

$$\leq 2\exp\left\{ \frac{-u^2(np)^{4/3}}{2\left[u(n-u)p(1-p) + u(np)^{2/3}/3 \right]} \right\}$$

$$\leq 2\exp\left\{ \frac{-u(np)^{2/3}}{2(np)^{1/3} + 2/3} \right\}$$

$$\leq 2\exp\left\{ -\frac{1}{3}u(np)^{1/3} \right\},$$

provided $(np)^{1/3} > 2/3$. Applying the assumption that $np \geq 28(\log n)^3$ completes the proof:

$$\sum_{u=1}^{\epsilon n} P(u, m(u)) \leq 2\sum_{u=1}^{\infty} \left[ne^{-(np)^{1/3}/3} \right]^u$$

$$= 2ne^{-(np)^{1/3}/3}\left(1 - ne^{-(np)^{1/3}/3} \right)^{-1} = o(1).$$

$\square$

We will make use of a random walk on a random graph in the proof of Theorem 1.2, and we will need a bound on the mixing time of this random walk. Loosely speaking, this is the number of steps that it takes for the random walk to be 'close' to its stationary state. Let

$$\|\mu - \nu\|_{TV} := \frac{1}{2} \sum_{v \in V} |\mu(v) - \nu(v)| = \sum_{v \in V} (\mu(v) - \nu(v))_+ = \sup_{A \subset V} |\mu(A) - \nu(A)|$$

denote the total variation distance between two probability measures, μ and ν, on the vertices of a graph $G = (V, E)$.

Let X_k be a discrete-time, simple random walk on the vertices of $G \sim \mathcal{G}(n, p)$ where $np = n^a$ and $a \in (0, 1]$. We denote the k-step transition probabilities of X_k by $P^k(u, v)$ for $u, v \in V$. Also, we will denote the probability measure on V that corresponds to the kth step of the random walk started at u by $P^k(u, \cdot)$. Let π be the stationary distribution for this random walk whenever it is uniquely defined. Note that the random walk is ergodic (so π is unique) with **P**-probability tending to 1 because X_k is aperiodic (G is not bipartite) and irreducible (G is connected) with probability superpolynomially close to 1.

Typically, the mixing time for a random walk, X_k, is defined to be the smallest k such that $\sup_{u \in V} \left\| P^k(u, \cdot) - \pi \right\| \leq \alpha$ for some fixed $\alpha \in (0, 1/2)$. In the proof of Lemma 4.1 we need tight control on $P^k(u, \cdot)$ for many vertices, u, simultaneously, so our goal is to show that for some constant k_{mix} depending only on a,

$$\sup_{u \in V} \left| P^{k_{\text{mix}}}(u, A) - \frac{|A|}{n} \right| \leq |A| \, o_{\mathbf{P}}(n^{-1})$$

for every $A \subset V$, where the $o_{\mathbf{P}}(n^{-1})$ term does not depend on A. For this we need the following lemma, which says that the stationary distribution of X_k is almost uniform.

Lemma 2.4. *Let $G \sim \mathcal{G}(n, p)$ with $np = n^a$, $a \in (0, 1]$, and let π be the stationary distribution for the simple random walk on G (conditional on its existence and uniqueness). If μ is the uniform probability distribution on G ($\mu(\{v\}) = n^{-1}$ for all $v \in V$), then*

$$\mathbf{P}\left[|\pi(A) - \mu(A)| \leq 3n^{-(1+a/3)} |A| \quad \forall A \subset V \right] = 1 - \exp[-\Omega(n^{a/3})].$$

Proof: It is easily verified that $\pi(v) = \deg(v)/(2 |E|)$ for all $v \in V$. By Lemma 2.1, $\deg(v) \in [n^a - n^{2a/3}, n^a + n^{2a/3}]$ for all $v \in V$ with **P**-probability $1 - \exp[-\Omega(n^{a/3})]$. This implies that, for all $A \subset V$ and all sufficiently large n ($\geq 2^{3/a}$),

$$\pi(A) \leq \frac{|A|}{n} \left[\frac{1 + n^{-a/3}}{1 - n^{-a/3}} \right]$$

$$\leq \frac{|A|}{n} \left[1 + 3n^{-a/3} \right]$$

with **P**-probability $1 - \exp[-\Omega(n^{a/3})]$. Likewise, we have the corresponding lower bound $\pi(A) \geq (|A|/n)[1 - 3n^{-a/3}]$ for all A with **P**-probability $1 - \exp[-\Omega(n^{a/3})]$, which proves the lemma, since $\mu(A) = |A|/n$. $\qquad\square$

Note that Lemma 2.4 is slightly stronger than the statement that

$$\|\pi - \mu\|_{TV} \leq 3n^{-a/3}$$

with probability $1 - \exp[-\Omega(n^{a/3})]$.

Dou (1992) proved that when $G \sim \mathcal{G}(n, d/(n-1))$, $d \in \{1, \ldots, n-1\}$, and $k \geq 0$ and $v \in V$ are fixed, then

$$\mathbb{E} \left\| P^k(v, \cdot) - \mu \right\|_{TV} \leq c \left(\frac{n}{d^k} + \frac{1}{d} \right)^{1/2}$$

for some absolute constant c. (Note that μ, the uniform probability distribution, appears in the statement, and not π.) The problem is that this result does not provide a bound on the mixing time, which entails taking the supremum over all initial locations for the random walk. This presents a problem for us, since we want to say that many independent random walks, started from different locations, will all be well mixed at the same time. To remedy this for high degree random graphs, we have the next lemma.

Lemma 2.5. *Let $G \sim \mathcal{G}(n, p)$ with $np = n^a$ for a constant $a \in (0, 1]$. If $\kappa = \lfloor 1 + 1/a \rfloor$, then*

$$\mathbf{P} \left[\sup_{v \in V} \| P^\kappa(v, \cdot) - \pi \|_{TV} \leq n^{-(\kappa a - 1)/3} \right] \to 1.$$

The definition of κ in Lemma 2.5 is such that κ is the smallest integer strictly larger than $1/a$.

Proof: We will employ a simple path counting argument similar to an argument used by Lubetzky and Sly (2010) to prove cutoff for the random walk on random regular graphs. Let $\mathcal{P}_k(u, v)$ denote the number of paths in G of length k that start at u and end at v. As in the proof of Lemma 2.4, we use Lemma 2.1 to bound the maximum and minimum degrees of G as $\deg(v) \in [n^a - n^{2a/3}, n^a + n^{2a/3}]$ for all $v \in V$ with $\mathbf{P}$-probability $1 - \exp[-\Omega(n^{a/3})]$. Therefore, the probability that the random walk traverses any path of length k, $(v_0, v_1), (v_1, v_2), \ldots, (v_{k-1}, v_k) \in E$, is

$$P^1(v_0, v_1) \cdot P^1(v_1, v_2) \cdots P^1(v_{k-1}, v_k) \geq \frac{1}{n^{ka}}(1 - kn^{-a/3})$$

with $\mathbf{P}$-probability $1 - \exp[-\Omega(n^{a/3})]$. Thus, with $\mathbf{P}$-probability $1 - \exp[-\Omega(n^{a/3})]$, for each $u, v \in V$

$$P^k(u, v) \geq \mathcal{P}_k(u, v) \frac{1}{n^{ka}}(1 - kn^{-a/3}).$$

To obtain a lower bound on $\mathcal{P}_k(u, v)$, we introduce the following notation for balls of radius k in G. For any $u \in V$ and $k \geq 0$, let

$$B_k(u) := \{v \in V \mid \text{dist}(u, v) \leq k\}$$
$$\partial_V B_k(u) := B_k(u) \setminus B_{k-1}(u)$$

where $\text{dist}(u, v)$ denotes the length of the shortest path in G from u to v, and $\partial_V B_0(u) := \{u\}$.

For a pair of vertices $u, v \in V$, let $k_u = \left\lceil \frac{\kappa - 1}{2} \right\rceil$ and $k_v = \left\lfloor \frac{\kappa - 1}{2} \right\rfloor$, so that $k_u + k_v = \kappa - 1$. We will construct a ball of radius k_u around u, then remove this ball from V and construct a second ball around v of radius k_v whenever $v \notin B_{k_u}(u)$. Since most vertices in V are not within k_u of u, we will have two disjoint sets of vertices for most pairs u, v. The number of paths between u and v in G is at least the number of edges between the vertex boundaries of these two balls, so we seek a uniform lower bound on this quantity.

To obtain bounds on $|\partial_V B_{k_u}(u)|$, we start at u and reveal edges layer by layer. To start, by Lemma 2.1,

$$\mathbf{P}\left[n^a - n^{2a/3} \leq |\partial_V B_1(u)| \leq n^a + n^{2a/3}\right] = 1 - \exp[-\Omega(n^{a/3})].$$

By induction on k, we wish to show that there is $C > 0$ such that $\left||\partial_V B_k(u)| - n^{ak}\right|$ $\leq Cn^{ak-a/3}$ for all $k \leq k_u$ with $\mathbf{P}$-probability $1 - \exp[-\Omega(n^{a/3})]$. Assume for some $k \leq k_u$ that there is a constant $C_{k-1} > 0$ such that

$$\mathbf{P}\left[|\partial_V B_{k-1}(u)| \geq n^{(k-1)a}(1 - C_{k-1}n^{-a/3})\right] = 1 - \exp[-\Omega(n^{a/3})], \qquad \text{and}$$

$$\mathbf{P}\left[|B_{k-1}(u)| \leq n^{(k-1)a}(1 + C_{k-1}n^{-a/3})\right] = 1 - \exp[-\Omega(n^{a/3})].$$

Let E_{k-1} be the intersection of the two events above; we will show there exists a constant $C_k > 0$ such that $\mathbf{P}[E_k^c \cap E_{k-1}] = \exp[-\Omega(n^{a/3})]$. Since $k \leq k_u$, we have that $(k-1)a < 1/2$, so on E_{k-1}, $|V \setminus B_{k-1}(u)| \geq n - n^{1/2}$ for large n. This means that for large n, on E_{k-1}, there are at least $n^{(k-1)a+1}(1 - 2C_{k-1}n^{-a/3})$ and at most $n^{(k-1)a+1}(1 + C_{k-1}n^{-a/3})$ potential edges between $\partial_V B_{k-1}(u)$ and $V \setminus B_{k-1}(u)$. Therefore, by Lemma 2.1,

$$\mathbf{P}\left[E_{k-1} \cap \left\{|\partial_V B_k(u)| < n^{(k-1)a+1}p(1 - 2C_{k-1}n^{-a/3}) - \left[n^{(k-1)a+1}p\right]^{2/3}\right\}\right]$$
$$= \exp[-\Omega(n^{a/3})],$$

which implies

$$\mathbf{P}\left[E_{k-1} \cap \left\{|\partial_V B_k(u)| < n^{ka}(1 - 3C_{k-1}n^{-a/3})\right\}\right] = \exp[-\Omega(n^{a/3})].$$

Likewise,

$$\mathbf{P}\left[E_{k-1} \cap \left\{|\partial_V B_k(u)| > n^{ka}(1 + 2C_{k-1}n^{-a/3})\right\}\right] = \exp[-\Omega(n^{a/3})].$$

It follows that if $C_k = 3C_{k-1}$, then $\mathbf{P}[E_k^c \cap E_{k-1}] = \exp[-\Omega(n^{a/3})]$, which concludes the induction argument, and implies we can take $C = 3^{k_u}$.

After exposing $B_{k_u}(u)$ we can employ the same argument starting from $v \in V \setminus B_{k_u}(u)$, but at each step avoiding the vertices in $B_{k_u}(u)$. Let

$$B_k^u(v) := \{w \in V \setminus B_{k_u}(u) \mid \text{dist}_u(v, w) \leq k\}$$
$$\partial_V B_k^u(v) := B_k^u(v) \setminus B_{k-1}^u(v),$$

where $\text{dist}_u(v, w)$ denotes the graph distance on the maximal subgraph of G with vertex set $V \setminus B_{k_u}(u)$. Note that when $a > 1/2$, $k_v = 0$, so $B_{k_v}^u(v) = \{v\}$. When $a \leq 1/2$, removing the vertices in $B_{k_u}(u)$ from V does not affect any of the estimates made above, because $|B_{k_u}(u)| = O_{\mathbf{P}}(n^{k_u a})$ and $k_u a \leq (1+a)/2 < 1 - a/3$, so at the induction step $\left|V \setminus (B_{k_u}(u) \cup B_{k-1}^u(v))\right| \geq n(1 - n^{-a/3})$ for large n. Therefore, $\left||\partial_V B_{k_v}^u(v)| - n^{k_v a}\right| \leq Cn^{k_v a - a/3}$ with probability $1 - \exp[-\Omega(n^{a/3})]$.

Now we observe that every edge between the sets $\partial_V B_{k_u}(u)$ and $\partial_V B_{k_v}^u(v)$ contributes at least one path of length κ to $\mathcal{P}_\kappa(u, v)$. On the event $v \notin B_{k_u}(u)$ and the above bounds on $\partial_V B_{k_u}(u)$ and $\partial_V B_{k_v}^u(v)$ hold, there are at least $|\partial_V B_{k_u}(u)| \cdot |\partial_V B_{k_v}^u(v)| \geq n^{(\kappa-1)a}(1 - 2Cn^{-a/3})$ potential edges between the two sets. So by again applying Lemma 2.1 we have that

$$\mathbf{P}\left[v \notin B_{k_u}(u), \; \mathcal{P}_\kappa(u, v) < n^{\kappa a - 1}(1 - 2Cn^{-a/3} - n^{-(\kappa a - 1)/3})\right]$$

$$= \exp[-\Omega(n^{(\kappa a-1)/3})],$$

where the constant in the exponent $\Omega(n^{(\kappa a-1)/3})$ does not depend on u or v. Since $a \geq \kappa a - 1 > 0$, by applying a union bound over all pairs u, v, we conclude that

$$\mathbf{P}\left[P^{\kappa}(u,v) \geq \frac{1}{n}(1-n^{-(\kappa a-1)/3}) \; \forall u \in V \text{ and } v \in V \setminus B_{k_u}(u)\right]$$

$$= 1 - \exp[-\Omega(n^{(\kappa a-1)/3})].$$

Therefore, with $\mathbf{P}$-probability $1 - \exp[-\Omega(n^{(\kappa a-1)/3})]$, for all large n

$$\sup_{u \in V} \|\mu - P^{\kappa}(u, \cdot)\|_{TV} = \sup_{u \in V} \sum_{v \in V} \left(\frac{1}{n} - P^{\kappa}(u,v)\right)_+$$

$$\leq \sup_{u \in V} \sum_{v \in V \setminus B_{k_u}(u)} \left(\frac{1}{n} - \frac{1}{n}(1-n^{-(\kappa a-1)/3})\right)_+ + \frac{1}{n}|B_{k_u}(u)|$$

$$\leq n^{-(\kappa a-1)/3}.$$

Applying Lemma 2.4 and the triangle inequality completes the proof. □

An immediate consequence of Lemmas 2.4 and 2.5 is the following lemma, which we will use in the proof of Theorem 1.2.

Lemma 2.6. *Let* $G \sim \mathcal{G}(n,p)$ *with* $np = n^a$ *for* $a \in (0,1]$, *and* $P^k(u, \cdot)$ *be probability measure on* V *corresponding to the* k^{th} *step of the simple random walk on* G *started at* u. *Let* π *be the stationary distribution of the random walk, and* μ *be the uniform probability measure on* V. *If* $k_{mix} = 7 \lfloor 1 + 1/a \rfloor \cdot \left\lceil \frac{1}{\lfloor 1+1/a \rfloor a - 1} \right\rceil$ *then*

$$\mathbf{P}\left[\sup_{u \in V} \left\|P^{k_{mix}}(u, \cdot) - \pi\right\|_{TV} \leq n^{-2}\right] \to 1.$$

Furthermore,

$$\mathbf{P}\left[\sup_{u \in V} \left|P^{k_{mix}}(u, A) - \mu(A)\right| \leq 4n^{-(1+a/3)}|A| \quad \forall A \subset V\right] \to 1.$$

Proof: The first equation is due to the following fact, which is a consequence of Lemma 4.11 and equation (4.31) of Levin et al. (2009). If $\sup_{u \in V} \left\|P^k(u, \cdot) - \pi\right\| \leq \alpha$ for some $k \in \mathbb{N}$ and $\alpha \in (0, 1/2)$, then for any integer $\ell > 0$,

$$\sup_{u \in V} \left\|P^{\ell k}(u, \cdot) - \pi\right\| \leq (2\alpha)^{\ell}.$$

In our case, by Lemma 2.5, $\alpha = n^{-(\kappa a-1)}$, $k = \kappa = \lfloor 1 + 1/a \rfloor$ and $\ell = 7 \left\lceil \frac{1}{\lfloor 1+1/a \rfloor a - 1} \right\rceil$; note that $\|P^k(u, \cdot) - \pi\|_{TV}$ is monotone decreasing in k on the event that the random walk is ergodic, which occurs with $\mathbf{P}$-probability tending to 1. The second equation follows from the first and Lemma 2.4. □

3. Survival of the Infection

In this section we will show that the contact process on an Erdős-Rényi random graph survives for exponentially long whenever it does not die out very quickly. We will do this by making use of three comparisons with branching processes. We will use lower and upper bounding branching processes to carefully control the behavior of the contact process started from a single vertex in its early stages. Then we will

use a second lower bounding branching process to show that the contact process will survive to linear size whenever it survives beyond its initial growth stage. The relevant facts about these branching processes can be found in the Appendix. Finally, once the contact process occupies a positive fraction of the vertices, we will compare it to a random walk with positive drift to show that it will continue to occupy a positive fraction of vertices for exponentially long.

Lemma 3.1. *Let* $r = \frac{2}{b-1-2(np)^{-1/3}} \log \log n = \frac{2}{b-1} \log \log n + o(1)$ *and* $\eta_2 = 3/\log b$. *If the contact process starts with a single vertex,* $\xi_0 = \{v\}$, *then*

$$\mathbb{P}_v \left(\sup_{t \leq 2r} |\xi_t| > (\log n)^6 \right) = O_{\mathbf{P}}(n^{-2}), \tag{3.1}$$

$$\mathbb{P}_v \left(0 < |\xi_r| \leq \eta_2 \log n \right) = O_{\mathbf{P}}((\log n)^{-1}), \tag{3.2}$$

$$\left| \mathbb{P}_v \left(|\xi_r| = 0 \right) - \frac{1}{b} \right| = O_{\mathbf{P}}((\log n)^{-1}). \tag{3.3}$$

Proof: By Lemma 2.1, the degree of any vertex in G_1 or G_2 is in the interval $[np - (np)^{2/3}, np + (np)^{2/3}]$ with $\mathbf{P}$-probability $1 - \exp[-\Omega((np)^{1/3})]$. For each n, let Y_t^n be a branching process in which each individual gives birth to a single offspring at rate $\beta_Y^n = b - 2(np)^{-1/3}$ and each individual dies at rate 1. Likewise, let Z_t^n be a branching process in which each individual gives birth to a single offspring at rate $\beta_Z^n = b + (np)^{-1/3}$ and each individual dies at rate 1. It is clear that Z_t^n stochastically dominates $|\xi_t|$ restricted to the graph G_1 (ignoring edges between G_1 and G_2) for all time provided $Z_0^n = |\xi_0|$. We now apply Lemma A.4 with $r_1 = \frac{2}{\beta_Z^n - 1} \log \log n$, $\eta_1 = \frac{6b}{(b-1)\log b} > \frac{6\beta_Z^n}{(\beta_Z^n - 1)\log \beta_Z^n}$ and $\gamma = 7/3$, and note that for large n, $\gamma r_1 > 2r$ and $\eta_1 (\log n)^{17/3} < (\log n)^6$, giving

$$\mathbf{P} \left[\mathbb{P}_v \left(\sup_{t \leq \gamma r_1} |\xi_t| > \eta_1 (\log n)^{2\gamma+1} \right) \leq \mathbb{P}_1 \left(\sup_{t \leq \gamma r_1} Z_t^n > \eta_1 (\log n)^{2\gamma+1} \right) \right] \to 1, \qquad \text{so}$$

$$\mathbb{P}_v \left(\sup_{t \leq 2r} |\xi_t| > (\log n)^6 \right) \leq \mathbb{P}_v \left(\sup_{t \leq \gamma r_1} |\xi_t| > \eta_1 (\log n)^{2\gamma+1} \right) = O_{\mathbf{P}}(n^{-2}).$$

This gives equation (3.1).

This means that with $\mathbf{P}$-probability $1 - O_{\mathbf{P}}(n^{-2})$, we have that

$$\sup_{t \leq r} \max_{u \in V} |\mathcal{N}(u) \cap \xi_t| \leq (\log n)^6 < (np)^{2/3}$$

for large n. Therefore, for all $t \in [0, r]$, $|\mathcal{N}(v) \setminus \xi_t| > np - 2(np)^{2/3}$, and $|\xi_t|$ stochastically dominates Y_t^n. If $\eta_2 = 3/\log b < 3/\log \beta_Y^n$, then by equation (3.1), Lemma A.2 and the fact that the extinction probability for $(Y_t^n)_{t \geq 0}$ is $1/\beta_Y^n$,

$$\mathbb{P}_v \left(|\xi_r| \leq \eta_2 \log n \right) \leq \mathbb{P}_1 \left(Y_r^n \leq \eta_2 \log n \right) + O_{\mathbf{P}}(n^{-2})$$
$$= \mathbb{P}_1 \left(Y_r^n = 0 \right) + O_{\mathbf{P}}((\log n)^{-1})$$
$$\leq \frac{1}{b - 2(np)^{-1/3}} + O_{\mathbf{P}}((\log n)^{-1})$$
$$= \frac{1}{b} + O_{\mathbf{P}}((\log n)^{-1}).$$

Likewise, by an exact calculation of the transition probability for $Z_{r_1}^n$ on the third line below, which appears in equation (A.5) of the Appendix, we have (with **P**-probability tending to 1 for the first line)

$$
\begin{aligned}
\mathbb{P}_v \left(|\xi_r| \leq \eta_2 \log n \right) &\geq \mathbb{P}_1 \left(Z_r^n = 0 \right) + \mathbb{P}_v \left(0 < |\xi_r| \leq \eta_2 \log n \right) \\
&\geq \mathbb{P}_1 \left(Z_{r_1}^n = 0 \right) + \mathbb{P}_v \left(0 < |\xi_r| \leq \eta_2 \log n \right) \\
&= \frac{1 - e^{-(\beta_Z^n - 1)r_1}}{\beta_Z^n - e^{-(\beta_Z^n - 1)r_1}} + \mathbb{P}_v \left(0 < |\xi_r| \leq \eta_2 \log n \right) \\
&= \frac{1}{b} - O_{\mathbf{P}}((\log n)^{-2}) + \mathbb{P}_v \left(0 < |\xi_r| \leq \eta_2 \log n \right).
\end{aligned}
$$

Combining the last two inequalities (and equation (3.1) to guarantee that ξ_t dominates Y_t^n) completes the proof. $\qquad\square$

By Lemma 2.3, the ϵ-isoperimetric number of G_j for $j = 1, 2$ is bounded below as

$$
\mathbf{P} \left[i_{2\epsilon}(G_j) \geq (1 - 2\epsilon)np - (np)^{2/3} \right] \to 1.
$$

We choose $\epsilon = \frac{1}{4}(1 - 1/b^{1/3}) > 0$ so that

$$
(1 - 2\epsilon - (np)^{-1/3} - O((\log n)^{8b^2/(b-1)}(np)^{-1}))b > (1 - 3\epsilon)b > 1 \qquad (3.4)
$$

for sufficiently large n. Then Lemma 2.3 applied to $i_{2\epsilon}(G_j)$ implies that, with **P**-probability tending to 1, for all t such that $|\xi_t| \leq 2\epsilon n$,

$$
|\partial(\xi_t)| \geq (1 - 3\epsilon)np |\xi_t|. \qquad (3.5)
$$

The term $O((\log n)^{8b^2/(b-1)}(np)^{-1})$ appears in (3.4) because in the proof of Theorem 1.2 it will be necessary to avoid as many as $(\log n)^{8b^2/(b-1)}$ vertices that have been observed by the dual process defined in the next section. Since (3.5) is a property of the graph, which is fixed for all time, we will assume that this inequality holds for the remainder of the proof. This means that during the time interval $[r, T_{2\epsilon}]$, where

$$
T_{2\epsilon} = \inf\{t > r : |\xi_t| > 2\epsilon n\},
$$

$|\xi_t|$ stochastically dominates a branching process with per-capita birth rate $(1 - 3\epsilon)b > 1$ and death rate 1. Denote this branching process by W_t. Then

$$
\begin{aligned}
\mathbb{P} \left(T_{2\epsilon} \geq s \mid \eta_2 \log n < |\xi_r| < (\log n)^{6b} \right) &\leq \mathbb{P} \left(W_s \leq 2\epsilon n \mid W_r = \eta_2 \log n \right) \\
&\leq \mathbb{P} \left(W_s \leq 2\epsilon n \mid W_r = 1 \right)^{\eta_2 \log n},
\end{aligned}
$$

because the event that $W_s \leq 2\epsilon n$ given that $W_r = \eta_2 \log n$ implies that all of the $\eta_2 \log n$ families at time r must not have exceeded size $2\epsilon n$ by time s. If we let $s = \frac{3}{(1-3\epsilon)b - 1} \log n + r$, then Lemma A.2 and the observation that $-3\log_b(1 - 3\epsilon) < 1$ imply

$$
\begin{aligned}
\mathbb{P} \left(W_s \leq 2\epsilon n \mid W_r = 1 \right)^{\eta_2 \log n} &\leq \left(\frac{1}{(1 - 3\epsilon)b} + O(n^{-2}) \right)^{\eta_2 \log n} \\
&\leq n^{-3 - 3\log_b(1 - 3\epsilon)} \exp[O(n^{-2} \log n)] \\
&= O(n^{-2}).
\end{aligned}
$$

Lemma 3.1, (3.5) and the last two inequalities imply the next lemma.

Lemma 3.2. *With r defined as in Lemma 3.1, $\epsilon = \frac{1}{4}(1 - 1/b^{1/3})$ and $\eta_3 = \frac{4}{(1-4\epsilon)b-1}$ then*

$$\mathbb{P}_v \left(T_{2\epsilon} \geq \eta_3 \log n \mid |\xi_r| > 0 \right) = O_{\mathbf{P}}((\log n)^{-1}). \tag{3.6}$$

This means that if the contact process is able to survive to time r, which happens with $\mathbf{P}$-probability $1 - 1/b + O_{\mathbf{P}}((\log n)^{-1})$, then it will reach size $2\epsilon n$ before time $\eta_3 \log n$. Next we prove Theorem 1.1, which says that if the contact process, (ξ_t), survives to time r, then it will continue to survive for exponentially long. In particular, it will survive long enough to spread to the second component.

Before the proof, we remark that in the proof of Theorem 1.2, we will apply Theorem 1.1, but with as many as $O((\log n)^{6b})$ vertices removed from G_1 at any time during the process. This added restriction has no effect on the proof of Theorem 1.1. Our approach is similar to that of Peterson (2011) for the contact process on the complete graph with random edge weights.

Proof of Theorem 1.1: We intend to show that there exist constants $\rho \in (\epsilon, 2\epsilon]$ and $\tau, C > 0$ so that

$$\mathbf{P}\left[\inf_{A \subset V : |A| > \rho n} \mathbb{P}\left(|\xi_\tau| > \rho n, \ \min_{t \in [0,\tau]} |\xi_t| > \epsilon n \ \Big| \ \xi_0 = A \right) \geq 1 - e^{-Cn} \right] \to 1, \tag{3.7}$$

This means that with $\mathbb{P}$-probability exponentially close to 1, if the size of the contact process initially exceeds ρn, then at time τ the size of the contact process will again exceed ρn and will not have dropped below ϵn along the way. By subdividing the time interval $[0, e^{Cn/2}]$ into $e^{Cn/2}/\tau$ intervals of length τ, this implies that

$$\mathbf{P}\left[\sup_{A \subset V : |A| > \rho n} \mathbb{P}\left(\min_{t \in [0, e^{Cn/2}]} |\xi_t| \leq \epsilon n \ \Big| \ \xi_0 = A \right) \leq \frac{1}{\tau} e^{-Cn/2} \right] \to 1. \tag{3.8}$$

By Lemma 3.2, $T_{2\epsilon} < \eta_3 \log n$ with probability $1 - o(1)$ conditional on $|\xi_r| > 0$. The Strong Markov Property and equation (3.8) imply the result with $c = C/2$.

To prove equation (3.7), first we observe that by monotonicity of the contact process, it suffices to prove the statement with initial size $|A| = \rho n$. We will actually prove the stronger statement that for $\gamma = \min\{\epsilon(1 - 4\epsilon)b, \ 2\epsilon\}$ (recall that $(1 - 4\epsilon)b = b^{2/3} > 1$), there exist $\tau, C > 0$ and $\rho \in (\epsilon, \gamma)$, not depending on A, such that

$$\mathbf{P}\left[\mathbb{P}\left(|\xi_\tau| > \rho n, \ |\xi_t| \in [\epsilon n, \gamma n] \ \forall t \in [0,\tau] \mid \xi_0 = A \right) \geq 1 - e^{-Cn} \right] \to 1. \tag{3.9}$$

The difference between the events in equations (3.7) and (3.9) is that in the latter the size of the contact process is also not allowed to exceed size γn.

The total jump rate of $|\xi_t|$ is at most $O_{\mathbf{P}}(n)$ for all t, since all vertices have degrees close to $n^a = b/\lambda$ with $\mathbf{P}$-probability tending to 1. So, for sufficiently small $\tau > 0$ (depending on b), the $\mathbb{P}$-probability that the number of jumps in the time interval $[0, \tau]$ exceeds $\frac{1}{4}(\gamma - \epsilon)n$ is at most $e^{-C'n}$ for some $C' > 0$ depending on τ and b. If we choose $\rho = \frac{1}{2}(\epsilon + \gamma)$, then $\xi_0 = A$ with $|A| = \rho n$ implies that $|\xi_t| \in [\epsilon n, \gamma n]$ for all $t \leq \tau$ with probability exceeding $1 - e^{-C'n}$. While $|\xi_t| \in [\epsilon n, \gamma n]$, the maximum rate at which $|\xi_t| \to |\xi_t| - 1$ is $\gamma n \leq \epsilon(1 - 4\epsilon)bn < \epsilon(1 - 3\epsilon)bn$, which is the minimum rate at which $|\xi_t| \to |\xi_t| + 1$ in this interval by (3.5). Therefore, on the event that $|\xi_t| \in [\epsilon n, \gamma n]$ for all $t \leq \tau$, $|\xi_t|$ stochastically dominates a random walk with positive drift up to time τ. Specifically, $|\xi_t|$ stochastically dominates X_t, where X_t is the continuous time random walk that jumps to $X_t + 1$ at rate $\epsilon(1 - 3\epsilon)bn$ and

to $X_t - 1$ at rate $\epsilon(1 - 4\epsilon)bn$. By a standard large deviations argument for random walks, $\mathbb{P}_{\rho n}(X_\tau \le \rho n) \le e^{-C''n}$, where $C'' > 0$ depends on τ and b. Choosing $C < \min\{C', C''\}$, we have demonstrated equation (3.9), and thus completed the proof of Theorem 1.1. $\qquad\square$

4. Duality

It is well known that the contact process is self-dual in the following sense. For any two sets of vertices $A, B \subset V$ and $t > 0$

$$\mathbb{P}\left(\xi_t^A \cap B \ne \emptyset\right) = \mathbb{P}\left(\xi_t^B \cap A \ne \emptyset\right). \tag{4.1}$$

For our purposes, the way to understand the contact process duality is through the graphical representation of the process (see, for example, Part I of Liggett (1999), which we paraphrase here). To each vertex $v \in V$ we assign a rate 1 Poisson process with jump times $\{T_k^v\}_{k=1}^\infty$, and to each ordered pair of vertices (u, v) joined by an edge in G ($\{u, v\} \in E$) we assign a rate λ Poisson process with jump times $\{T_k^{(u,v)}\}_{k=1}^\infty$. All of these Poisson processes are independent of one another. To construct the contact process graphically, we begin by drawing the space-time axes $G \times [0, \infty)$. For each $k \in \mathbb{N}$ and $v \in V$, we draw a recovery dot, $\bullet$, at the point $(v, T_k^v) \in G \times [0, \infty)$. For each $k \in \mathbb{N}$ and ordered pair (u, v) such that $\{u, v\} \in E$, we draw an infection arrow, $\rightarrow$, from $(u, T_k^{(u,v)})$ to $(v, T_k^{(u,v)})$.

We say that there is an active path from (v_0, t_1) to (v_ℓ, t_2), with $t_1 < t_2$, if there is a sequence of arrows between v_0 and v_ℓ,

$$t_1 \le T_{k_1}^{(v_0, v_1)} < T_{k_2}^{(v_1, v_2)} < \cdots < T_{k_\ell}^{(v_{\ell-1}, v_\ell)} \le t_2,$$

such that there are no recovery dots encountered along the way,

$$T_k^{v_i} \notin [T_{k_i}^{(v_{i-1}, v_i)}, T_{k_{i+1}}^{(v_i, v_{i+1})}) \qquad \forall k \in \mathbb{N} \text{ and } i = 1, \ldots, \ell - 1,$$

$$T_k^{v_0} \notin [t_1, T_{k_1}^{(v_0, v_1)}) \text{ and } T_k^{v_\ell} \notin [T_{k_\ell}^{(v_{\ell-1}, v_\ell)}, t_2] \qquad \forall k \in \mathbb{N}.$$

We now have that $v \in \xi_t$ if and only if there is a vertex $u \in \xi_0$ such that there is an active path from $(u, 0)$ to (v, t). Therefore, by tracing the arrows in reverse, we can determine for each $v \in V$ whether $v \in \xi_t$. Because the Poisson processes determining arrows in each direction between pairs of adjacent vertices are iid, we have that the sets

$$\{u \in V \mid \exists v \in A \text{ s.t. there is an active path from } (u, 0) \text{ to } (v, t)\} \text{ and}$$

$$\{v \in V \mid \exists u \in A \text{ s.t. there is an active path from } (u, 0) \text{ to } (v, t)\}$$

are equal in distribution, which is equivalent to the self duality of the contact process as stated in equation (4.1). Define the *dual process*, $(\zeta_s)_{s \in [0,t]}$, of the contact process started at $(v, t) \in V \times [0, \infty)$ by

$$\zeta_s = \{u \in V \mid \text{there is an active path from } (u, t - s) \text{ to } (v, t)\}.$$

We can view (ζ_s) as being the contact process constructed by time reversal of the arrows in the graphical representation for (ξ_t). Note that (ζ_s) and (ξ_t) live on the same probability space, and are such that for any $s \in [0, t]$, $v \in \xi_t$ if and only if $\zeta_s \cap \xi_{t-s} \ne \emptyset$. Moreover, $\mathbb{P}(\zeta_s \cap B \ne \emptyset \mid \zeta_0 = \{v\}) = \mathbb{P}(\xi_s \cap B \ne \emptyset \mid \xi_0 = \{v\})$ for all $B \subset V$ and $s \in [0, t]$ by the duality equation (4.1).

Let $\{u_b, v_b\} \in E$ be the bridge edge between V_1 and V_2, such that $u_b \in V_1$ and $v_b \in V_2$. We will use the graphical representation of the dual of the contact

process to determine whether the vertex u_b is infected at the times $T_k^{(u_b,v_b)}$. The idea is that by Theorem 1.1 there are at least ϵn vertices infected in V_1 at the times $T_k^{(u_b,v_b)} - r - t_{\mathrm{mix}}$ (where $t_{\mathrm{mix}} = O(\log \log n)$ will be defined later). By Lemma 3.1, the probability that the dual process started at $(u_b, T_k^{(u_b,v_b)})$ survives to time $T_k^{(u_b,v_b)} - r$ is approximately $1 - 1/b$. If the dual process does survive, then by coupling the dual process with a random walk process, we will show that in time t_{mix} (going backwards in time still) at least one of the active vertices in the dual process will coincide with one of the ϵn infected vertices in the contact process at time $T_k^{(u_b,v_b)} - r - t_{\mathrm{mix}}$ with high probability. Therefore, u_b will be infected at time $T_k^{(u_b,v_b)}$, which will immediately result in v_b becoming infected. In turn, v_b starts a widespread infection in V_2 with probability close to $1 - 1/b$ by Lemma 3.1 and Theorem 1.1.

To begin, for each $k \in \mathbb{N}$ we let $(\zeta_t^k)_{t \in [0, T_k^{(u_b,v_b)}]}$ be the dual process on G_1 started at $(u_b, T_k^{(u_b,v_b)})$. Note that $\{T_k^{(u_b,v_b)}\}_k$ are random times, but they are independent of $\{T_k^{(u,v)}\}_k$ and $\{T_k^v\}_k$ for all ordered edges (u,v) and vertices v in G_1. The interpretation of ζ_t^k is that $u_b \in \xi_{T_k^{(u_b,v_b)}}$ if and only if $\zeta_t^k \cap \xi_{T_k^{(u_b,v_b)} - t} \neq \emptyset$. By the self-duality of the contact process, we can apply Lemma 3.1 to (ζ_t^k) to say that $\left| \zeta_r^k \right| > 0$ with $\mathbb{P}$-probability differing from $(1 - 1/b)$ by at most $O_{\mathbf{P}}((\log n)^{-1})$, and when $\left| \zeta_r^k \right| > 0$, $\left| \zeta_r^k \right| > \eta_2 \log n$ with $\mathbb{P}$-probability $1 - O_{\mathbf{P}}((\log n)^{-1})$. The purpose of the next lemma is to show, via a coupling with random walks, that whenever $\left| \zeta_r^k \right| > \eta_2 \log n$, we have $u_b \in \xi_{T_k^{(u_b,v_b)}}$ with high probability.

Lemma 4.1. *If ζ_t^k is the dual of the contact process started at $(u_b, T_k^{(u_b,v_b)})$, $\epsilon > 0$ and $t_{mix} = \frac{2}{b+1} \log \log n$ then for all $\delta > 0$*

$$\mathbf{P}\left[\sup_{A \subset V_1, |A| \geq \epsilon n} \mathbb{P}^{G_1}\left(\zeta_{r+t_{mix}}^k \cap A = \emptyset \mid \left| \zeta_r^k \right| \geq \eta_2 \log n \right) > \delta \right] \to 0.$$

For the proof of Lemma 4.1 we will first need to construct the coupling between ζ_t^k and a random walk process, then prove some facts about this process. Let $\mathbf{X}_t = \{X_t^1, X_t^2, \ldots, X_t^{\eta_2 \log n}\}$ be the locations of $\eta_2 \log n$ independent, continuous time, simple random walks on G_1 that independently die at rate 1. That is, for each i, X_t^i is the continuous time random walk on G_1 that holds at a vertex $u \in V_1$ for time $\mathrm{Exp}(\lambda \deg(u))$, then jumps to $v \sim u$ with probability $1/\deg(u)$, and which has a total life span distributed as $\mathrm{Exp}(1)$. When a walker dies, we remove it from the set $\mathbf{X}_t$.

We can couple $\mathbf{X}_t$ with ζ_{r+t}^k until $t = T_{\mathrm{die}} \wedge T_{\mathrm{collide}}$, where $T_{\mathrm{die}} = \inf\{t : \mathbf{X}_t = \emptyset\}$ is the time at which all of the random walkers die out, and $T_{\mathrm{collide}} = \inf\{t : X_t^i = X_t^j \text{ for some } i \neq j\}$ is the first time at which two random walks collide. First, let X_0^i be the ith element of the lexicographical ordering of the vertices in ζ_r^k. We then update X_t^i according to the infection arrows by always following the first arrow to arrive in the correct orientation (going backwards in time, so this is actually the most recent arrow infecting the current vertex at which the walker resides). That is, if X_t^i is at the vertex u at time t, then X_t^i will jump to v at (random walk) time $T_k^{(u_b,v_b)} - r - T_\ell^{(v,u)}$ if the arrow from v to u is the next encountered:

$$T_\ell^{(v,u)} = \max_{w \sim u} \max_j \{T_j^{(w,u)} \mid T_j^{(w,u)} < T_k^{(u_b,v_b)} - r - t\}.$$

If, while at the vertex u, X_t^i encounters a recovery dot before an arrow into u (there is a j so that $T_k^{(u_b,v_b)} - r - t > T_j^u > T_\ell^{(v,u)}$) then the random walk dies. It is clear that this constructs the random walk X_t^i as described above, since the waiting time until the first infection arrow into the vertex u is distributed as the minimum of $\deg(u)$ random variables with $\mathrm{Exp}(\lambda)$ distribution, which is $\mathrm{Exp}(\lambda \deg(u))$, and the recovery dots appear at rate 1.

Under this coupling $\mathbf{X}_t \subset \zeta_{r+t}^k$. When a collision occurs between two random walkers, however, our coupling would cause those walks to stick together for all time. To avoid this, at the time of the first collision, T_{collide}, we stop the coupling between $\mathbf{X}_t$ and ζ_{r+t}^k, and instead let each of the random walks proceed independently (of the other random walks and of ζ_{r+t}^k). This way, at time t_{mix} the locations of the surviving random walks are independent, and as long as $t_{\mathrm{mix}} < T_{\mathrm{collide}}$ they are still coupled with ζ_{r+t}^k.

The proof of Lemma 4.1 has three basic ingredients. First, many of the random walkers will survive for time t_{mix}. Second, most of the random walkers will make at least k_{mix} jumps by time t_{mix}, so by Lemma 2.6 their locations will be almost uniformly distributed, so intersection with A is imminent. Finally, a collision is unlikely to occur before time t_{mix}, so the random walk process will still be coupled with ζ_{r+t}^k.

Proof: We assume that $\left|\zeta_r^k\right| \geq \eta_2 \log n$, $A \subset V_1$ with $|A| \geq \epsilon n$ is fixed, and we have the coupling described above between $\mathbf{X}_t$ and ζ_{r+t}^k.

First we observe that the random walkers are mutually independent, and so are their death clocks. Since each random walker dies at rate 1, the number of random walks that survive to time $t_{\mathrm{mix}} = \frac{2}{b+1} \log \log n$ is $\mathrm{Binomial}(\eta_2 \log n, (\log n)^{-2/(b+1)})$. So by Chebychev's inequality, the number of random walks that survive until time t_{mix} is at least $\frac{1}{2}\eta_2 (\log n)^{(b-1)/(b+1)}$ with $\mathbb{P}$-probability $1 - O_{\mathbf{P}}((\log n)^{-(b-1)/(b+1)})$.

Next, we want all of the random walkers to make at least k_{mix} steps so we can apply Lemma 2.6. Since the minimum degree of G_1 is at least $n^a - n^{2a/3}$ with $\mathbf{P}$-probability tending to 1, the $\mathbb{P}$-probability that the ith walker, X_t^i, jumps fewer than k_{mix} times by time t_{mix} (ignoring whether the walker survives to time t_{mix}, as these events are independent) is at most

$$\sum_{k < k_{\mathrm{mix}}} \frac{1}{k!} e^{-b(1-n^{-a/3})t_{\mathrm{mix}}} [b(1-n^{-a/3})t_{\mathrm{mix}}]^k$$

$$\leq e^{1-b(1-n^{-a/3})t_{\mathrm{mix}}} [b(1-n^{-a/3})t_{\mathrm{mix}}]^{k_{\mathrm{mix}}}$$

$$= O_{\mathbf{P}} \left((\log n)^{-2b/(b+1)} (\log \log n)^{k_{\mathrm{mix}}} \right).$$

Therefore, all of the surviving random walks will make at least k_{mix} jumps by time t_{mix} with $\mathbb{P}$-probability $1 - O_{\mathbf{P}}((\log n)^{-(b-1)/(b+1)} (\log \log n)^{k_{\mathrm{mix}}})$.

Conditional on the events that X_t^i survives to time t_{mix} and makes at least k_{mix} jumps, then by Lemma 2.6, with $\mathbf{P}$-probability tending to 1 uniformly in $A \subset V_1$,

$$\mathbb{P}\left(X_{t_{\mathrm{mix}}}^i \in A\right) \geq \frac{|A|}{n}(1 - 4n^{-a/3}) \geq \epsilon(1 - 4n^{-a/3}).$$

Let E be the event that there are at least $\frac{1}{2}\eta_2 (\log n)^{(b-1)/(b+1)}$ random walkers that survive to time t_{mix} and make at least k_{mix} jumps. Then the probability that none of the random walkers hits the set A is (with $\mathbf{P}$-probability tending to 1 for the

second inequality)

$$\mathbb{P}\left(\boldsymbol{X}_{t_{\mathrm{mix}}} \cap A = \emptyset\right) \leq \mathbb{P}\left(\boldsymbol{X}_{t_{\mathrm{mix}}} \cap A = \emptyset | E\right) + \mathbb{P}\left(E^c\right)$$

$$\leq (1 - \epsilon(1 - 4n^{-a/3}))^{\frac{1}{2}\eta_2 (\log n)^{(b-1)/(b+1)}} + \mathbb{P}\left(E^c\right)$$

$$= O_{\mathbf{P}}((\log n)^{-(b-1)/(b+1)}(\log \log n)^{k_{\mathrm{mix}}}). \tag{4.2}$$

Now we only need to check that the coupling between $\boldsymbol{X}_t$ and ζ^k_{r+t} has not been violated before time t_{mix}. That is, we need to check that $T_{\mathrm{collision}} > t_{\mathrm{mix}}$. Observe that $T_{\mathrm{collision}} \succ T'$ where $T' \sim \mathrm{Exp}(\lambda(\eta_2 \log n)^2)$, since at any time there are fewer than $|\boldsymbol{X}_t|^2 \leq (\eta_2 \log n)^2$ directed edges connecting vertices occupied by the random walkers, and each directed edge has a rate λ Poisson clock. Therefore, the probability that a collision occurs by time t_{mix} is

$$\mathbb{P}\left(T_{\mathrm{collision}} \leq t_{\mathrm{mix}}\right) \leq \mathbb{P}\left(T' \leq t_{\mathrm{mix}}\right)$$

$$= 1 - \exp\left[-t_{\mathrm{mix}}\lambda(\eta_2 \log n)^2\right]$$

$$= 1 - \exp\left[-\frac{2b\eta_2^2}{b+1}n^{-a}(\log n)^2 \log \log n\right]$$

$$= O(n^{-a}(\log n)^2 \log \log n). \tag{4.3}$$

Combining equations (4.2) and (4.3) shows that the coupling between ζ^k_{r+t} and $\boldsymbol{X}_t$ will not be violated before time t_{mix}, so

$$\mathbb{P}^{G_1}\left(\zeta^k_{r+t_{\mathrm{mix}}} \cap A \neq \emptyset \mid |\zeta^k_r| \geq \eta_2 \log n\right) = 1 - O_{\mathbf{P}}((\log n)^{-(b-1)/(b+1)}(\log \log n)^{k_{\mathrm{mix}}}),$$

which completes the proof. $\qquad\square$

5. Proof of Theorem 1.2

First, observe that by equation (3.3) of Lemma 3.1

$$\left|\mathbb{P}^{G_1}_{v_0}\left(|\xi_r| > 0\right) - \left(1 - \frac{1}{b}\right)\right| = O_{\mathbf{P}}((\log n)^{-1}).$$

Then by Theorem 1.1, for all $\delta > 0$

$$\mathbf{P}\left[\mathbb{P}^{G_1}_{v_0}\left(\min_{t \in [\eta_3 \log n, e^{cn}]} |\xi_t| \leq \epsilon n \mid |\xi_r| > 0\right) > \delta\right] \to 0,$$

so if $|\xi_r| > 0$, then $|\xi_t| > \epsilon n$ for all times $t \in [\eta_3 \log n, e^{cn}]$ with $\mathbb{P}$-probability tending to 1.

Now we consider the sequence of times at which the bridge edge may transmit the infection from u_b to v_b, $\{T^{(u_b, v_b)}_k\}_k$. The interarrival times, $T^{(u_b, v_b)}_1$ and $T^{(u_b, v_b)}_k - T^{(u_b, v_b)}_{k-1}$ for $k \geq 2$, are independent with distribution $\mathrm{Exp}(\lambda)$. If any of the first $(\log n)^{1/2}$ interarrival times are smaller than $2r + t_{\mathrm{mix}}$ or if $T^{(u_b, v_b)}_1 < \eta_3 \log n + r + t_{\mathrm{mix}}$, then we will not have enough time to allow the dual process, (ζ^k_s) started at $(u, b, T^{(u_b, v_b)}_k)$ to intersect with (ξ_t), and to guarantee that $\{u_b \in \xi_{T^{(u_b, v_b)}_k}\}$ are essentially independent events for $k = 1, \ldots, (\log n)^{1/2}$. The reason for the $2r$ term is that once the infection is transmitted to the second component, we need additional time r to determine whether it survives in that component. Let IA be the event that all of the first $(\log n)^{1/2}$ interarrival times are each longer than $2r + t_{\mathrm{mix}}$,

that $T_1^{(u_b,v_b)} \geq \eta_3 \log n + r + t_{\mathrm{mix}}$, and that $T_{(\log n)^{1/2}}^{(u_b,v_b)} < e^{cn}$. Then by a simple union bound, $\mathbb{P}(\mathtt{IA}^c) = O_{\mathbf{P}}(n^{-a} \log n)$.

Next, successively for each $k = 1, 2, \ldots, (\log n)^{1/2}$ we observe the dual process $(\zeta_s^k)_{s \in [0, T_k^{(u_b,v_b)}]}$ started at $(u_b, T_k^{(u_b,v_b)})$. We observe the kth dual process from time $s = 0$ to time $s = r$, and during this time, by Lemma 3.1, (ζ_s^k) has either died out or exceeded size $\eta_2 \log n$:

$$\mathbb{P}_{v_0}\left(0 < |\zeta_r^k| \leq \eta_2 \log n\right) = O_{\mathbf{P}}((\log n)^{-1}),$$

$$\left|\mathbb{P}_{v_0}\left(|\zeta_r^k| = 0\right) - \frac{1}{b}\right| = O_{\mathbf{P}}((\log n)^{-1}).$$

If $|\zeta_r^k| \geq \eta_2 \log n$, $|\xi_r| > 0$ and $\mathtt{IA}$ occurs, then by Lemma 4.1, $\zeta_{r+t_{\mathrm{mix}}}^k \cap \xi_{T_k^{(u_b,v_b)} - r - t_{\mathrm{mix}}} \neq \emptyset$ with $\mathbb{P}$-probability tending to 1 in $\mathbf{P}$-probability.

We remark here that on the event $\mathtt{IA}$, no two dual processes (ζ_t^k) for different values of $k \leq (\log n)^{1/2}$ are observed simultaneously. By considering the expected number of offspring in a pure-birth branching process with birth rate $2b > b(1 + (np)^{-1/3})$ starting with a single individual and run to time $\frac{b+2}{b-1} \log \log n > r + t_{\mathrm{mix}}$, the number of vertices ever infected by any dual process (ζ_t^k) for $k \leq (\log n)^{1/2}$ is at most $(\log n)^{1/2} \cdot (\log n)^{1+2b(b+2)/(b-1)} \leq (\log n)^{8b^2/(b-1)}$ with $\mathbb{P}$-probability $O_{\mathbf{P}}((\log n)^{-1/2})$. If the kth dual process either fails to survive to time r or fails to intersect with $\xi_{T_k^{(u_b,v_b)} - r - t_{\mathrm{mix}}}$, then the contact process (ξ_t) for $t \in [T_k^{(u_b,v_b)} - r - t_{\mathrm{mix}}, T_k^{(u_b,v_b)}]$, must avoid any vertices ever infected by $(\zeta_s^k)_{s \in [0, r+k_{\mathrm{mix}}]}$, so as to avoid observing the same Poisson processes in the graphical construction more than once. Avoiding these vertices is not a problem, since we have accounted for the fact that (ξ_t) must avoid these vertices in equation (3.4), which determines the lower-bounding branching process used to prove Theorem 1.1.

To this point, by Lemmas 3.1 and 4.1 and Theorem 1.1, the dual process $\zeta_{r+t_{\mathrm{mix}}}^k$, started at space-time $(u_b, T_k^{(u_b,v_b)})$, will intersect with $\xi_{T_k^{(u_b,v_b)} - r - t_{\mathrm{mix}}}$ with $\mathbb{P}$-probability differing from $1 - \frac{1}{b}$ by at most $O_{\mathbf{P}}((\log n)^{-1})$. If successful, the result is that u_b is infected at time $T_k^{(u_b,v_b)}$, and therefore $v_b \in V_2$ becomes infected at this time. By applying Lemma 3.1 and Theorem 1.1 to the infection in G_2 started at v_b, this will lead to either extinction of the infection within G_2 by time $T_k^{(u_b,v_b)}$, or else wide-spread infection of G_2 (with greater than ϵn infected vertices in G_2) by time $T_k^{(u_b,v_b)} + \eta_3 \log n$ with $\mathbb{P}$-probability differing from $1 - \frac{1}{b}$ by at most $O_{\mathbf{P}}((\log n)^{-1})$. Let K denote the smallest value $k \leq (\log n)^{1/2}$ such that $\xi_{T_k^{(u_b,v_b)} + r} \cap V_2 \neq \emptyset$, and let $K = \infty$ if this does not occur. We then have that for all $\ell < (\log n)^{1/2}$

$$\left|\mathbb{P}_{v_0}(K > \ell \mid |\xi_r| > 0) - \left[1 - \left(1 - \frac{1}{b}\right)^2\right]^\ell\right| = O_{\mathbf{P}}((\log n)^{-1/2}),$$

which implies that K (given $|\xi_r| > 0$) converges in distribution to Geometric$\left(\left(1 - \frac{1}{b}\right)^2\right)$ with $\mathbf{P}$-probability tending to 1. Since the distribution of the sum of a Geometric$\left(\left(1 - \frac{1}{b}\right)^2\right)$ number of independent Exp(b) random variables has Exp$\left(\frac{(b-1)^2}{b}\right)$ distribution, and since $n^{-a} T_1^{(u_b,v_b)}$ and $n^{-a}(T_k^{(u_b,v_b)} - T_{k-1}^{(u_b,v_b)})$

for $k \geq 2$ are i.i.d. $\mathrm{Exp}(b)$ random variables, we have for fixed $x \geq 0$ that

$$\left| \mathbb{P}_{v_0} \left(n^{-a} T_K^{(u_b, v_b)} > x \mid |\xi_r| > 0 \right) - \exp\left(-\frac{(b-1)^2}{b} x \right) \right| = o_{\mathbf{P}}(1).$$

The time

$$\tau := \inf\{t > 0 : |\xi_t^{v_0} \cap V_2| > \epsilon n\},$$

differs from $T_K^{(u_b, v_b)}$ by $O_{\mathbf{P}}(\log n)$, which is the time required for the contact process to grow from a single infected vertex to size ϵn in each of G_1, G_2. Therefore, with equation (3.3) of Lemma 3.1, this completes the proof in the case where there is a single bridge edge.

The proof for the case of multiple bridges, $|B| = o(n^a / \log n \log \log n)$, is nearly identical if we replace the sequence of bridge transmission times $T_k^{(u_b, v_b)}$ with $T_k^B := \min\{T_\ell^{(u_b, v_b)} > T_{k-1}^B : (u_b, v_b) \in B, u_b \in V_1, v_b \in V_2, \ell \geq 1\}$ for $k \geq 1$ (and $T_0^B = 0$), the sequence of times at which *some* bridge edge attempts to transmit the infection. All that needs to be checked is that the first $(\log n)^{1/2}$ interarrival times of consecutive transmission attempts are larger than $2r + t_{\mathrm{mix}} = O(\log \log n)$ and the first attempt occurs after time $\eta_3 \log n$. This occurs with $\mathbb{P}$-probability $O_{\mathbf{P}}((\log \log n)^{-1})$, as the latter event has the larger probability. The remainder of the proof proceeds as above. □

Appendix A. Branching Processes

In this section we prove Lemmas A.2 and A.4 regarding the size of a supercritical, continuous-time branching process, Z_t, in which each individual gives birth to a single offspring at rate $\beta > 1$ and dies at rate 1. Let $\mathbf{Pr}(\cdot)$ be the probability measure associated with this process, and for all $i, j \in \mathbb{Z}_{\geq 0}$ and $t \geq 0$, let $P_{ij}(t) := \mathbf{Pr}(Z_t = j \mid Z_0 = i) =: \mathbf{Pr}_i(Z_t = j)$.

Lemma A.1. *Let Z_t be a branching process where each individual gives birth to a single offspring at rate $\beta > 1$ and dies at rate 1. If $Z_0 = \lceil 3 \log_\beta n \rceil =: \alpha$ then*

$$\mathbf{Pr}_\alpha (Z_t = 0 \text{ eventually}) \leq n^{-3}.$$

Proof: If $Z_0 = 1$, then $\mathbf{Pr}_1(Z_t = 0 \text{ eventually}) = 1/\beta$ Harris (1963). If $Z_0 = \alpha$ then all of the families are mutually independent, so

$$\mathbf{Pr}_\alpha (Z_t = 0 \text{ eventually}) = \beta^{-\alpha} \leq n^{-3}.$$

□

Lemma A.2. *Let Z_t be a branching process where each individual gives birth to a single offspring at rate $\beta > 1$ and dies at rate 1. When $Z_0 = 1$, if $r = \frac{2}{\beta - 1} \log \log n$ and $s = \frac{3}{\beta - 1} \log n$, then for any constants $C, c > 0$ and $\gamma \geq 1$*

$$\mathbf{Pr}_1 (0 < Z_r \leq C \log n) = O((\log n)^{-1}), \tag{A.1}$$

$$\mathbf{Pr}_1 \left(Z_{\gamma r} > \frac{2\beta}{\beta - 1} (\log n)^{2\gamma + 1} \right) = O(n^{-2}), \quad \text{and} \tag{A.2}$$

$$\mathbf{Pr}_1 (0 < Z_s \leq cn) = O(n^{-2}). \tag{A.3}$$

The proof uses the following standard bounds.

Lemma A.3. *If* $0 < y < \frac{1}{2}$ *then*

$$- y - y^2 \leq \log(1 - y) \leq -y - \frac{y^2}{2}. \tag{A.4}$$

We now prove Lemma A.2.

Proof: The transition probabilities for Z_t can be computed exactly, as in Chapter V of Harris (1963):

$$P_{10}(t) = \frac{1 - e^{-(\beta - 1)t}}{\beta - e^{-(\beta - 1)t}}$$

$$P_{1k}(t) = [1 - P_{10}(t)]\,[1 - \eta(t)]\,\eta(t)^{k-1} \tag{A.5}$$

$$\eta(t) = \frac{1 - e^{-(\beta - 1)t}}{1 - \frac{1}{\beta}e^{-(\beta - 1)t}}.$$

So we have at time $r = \frac{2}{\beta - 1}\log\log n$ that

$$\mathbf{Pr}_1\left(0 < Z_r \leq C \log n\right) = \sum_{1 \leq k \leq C \log n} P_{1k}(r)$$

$$\leq [1 - P_{10}(r)] \cdot \left[1 - \eta(r)^{C \log n}\right]. \tag{A.6}$$

We apply Lemma A.3 to obtain:

$$\log\left[1 - e^{-(\beta - 1)r}\right] = \log\left[1 - (\log n)^{-2}\right] \geq -(\log n)^{-2} - (\log n)^{-4}$$

$$\log\left[1 - \frac{1}{\beta}e^{-(\beta - 1)r}\right] = \log\left[1 - \frac{1}{\beta}(\log n)^{-2}\right] \leq -\frac{1}{\beta}(\log n)^{-2} - \frac{1}{2\beta^2}(\log n)^{-4}$$

whenever $n \geq 5$. Subtracting the second line from the first, then multiplying both sides by $C \log n$ gives

$$\log \eta(r)^{C \log n} \geq -C\left(1 - \frac{1}{\beta}\right)(\log n)^{-1} - O\left((\log n)^{-3}\right).$$

Combining equation (A.6) with this bound and the fact that $1 - e^{-x} \leq x$ proves equation (A.1):

$$\mathbf{Pr}_1\left(0 < Z_r \leq C \log n\right) \leq 1 - \exp\left[\log \eta(r)^{C \log n}\right]$$

$$\leq C\left(1 - \frac{1}{\beta}\right)(\log n)^{-1} + O\left((\log n)^{-3}\right).$$

Equation (A.3) is proved in the same way. To prove equation (A.2) we begin with

$$\mathbf{Pr}_1\left(Z_{\gamma r} > \frac{2\beta}{\beta - 1}(\log n)^{2\gamma + 1}\right) = \sum_{k > \frac{2\beta}{\beta - 1}(\log n)^{2\gamma + 1}} P_{1k}(\gamma r)$$

$$\leq [1 - P_{10}(\gamma r)] \cdot \eta(\gamma r)^{\frac{2\beta}{\beta - 1}(\log n)^{2\gamma + 1} - 1}.$$

Now applying Lemma A.3 as above (but with the upper and lower bounds reversed) yields:

$$\mathbf{Pr}_1\left(Z_{\gamma r} > \frac{2\beta}{\beta - 1}(\log n)^{2\gamma + 1}\right) \leq \exp\left[\log \eta(\gamma r)^{\frac{2\beta}{\beta - 1}(\log n)^{2\gamma + 1} - 1}\right]$$

$$\leq \exp\left[-\frac{2\beta}{\beta - 1}\left(1 - \frac{1}{\beta}\right)\log n + O\left((\log n)^{-1}\right)\right] = O(n^{-2}).$$

$\square$

Lemma A.4. *Using the same setup as in Lemma A.2, let $\gamma \geq 1$ and let $\eta_1 = \frac{6\beta}{(\beta-1)\log\beta}$, then*

$$\mathbf{Pr}_1\left(\sup_{0\leq t\leq \gamma r} Z_t \geq \eta_1(\log n)^{2\gamma+2}\right) = O\left(n^{-2}\right).$$

Proof: For the duration of this proof let $T := \inf\{t \mid Z_t \geq \eta_1(\log n)^{2\gamma+2}\}$. Then the claim in Lemma A.4 is equivalent to $\mathbb{P}_1\left(T \leq \gamma r\right) = O\left(n^{-2}\right)$. Our strategy is to use equation (A.2) in Lemma A.2 to say that $Z_{\gamma r}$ can be at most $O((\log n)^{2\gamma+1})$, then use the Strong Markov Property to say that if Z_t exceeds $\eta_1(\log n)^{2\gamma+2}$ at any time $t \leq \gamma r$, then it is unlikely to drop below size $O((\log n)^{2\gamma+1})$ by time γr. In the calculation below, we apply the Strong Markov Property and translation invariance of the branching process at the third line. The fourth line will follow by observing that for the branching process to transition from size $\eta_1(\log n)^{2\gamma+2}$ to size $\frac{2\beta}{\beta-1}(\log n)^{2\gamma+1}$, then at least one of the $\frac{2\beta}{\beta-1}(\log n)^{2\gamma+1}$ sets of $3\log_\beta n$ individuals must go extinct. The last line will follow from Lemma A.1.

$$\mathbf{Pr}_1\left(T \leq \gamma r\right) \leq \mathbf{Pr}_1\left(T \leq \gamma r,\ Z_{\gamma r} \leq \frac{2\beta}{\beta-1}(\log n)^{2\gamma+1}\right)$$

$$+ \mathbf{Pr}_1\left(Z_{\gamma r} > \frac{2\beta}{\beta-1}(\log n)^{2\gamma+1}\right)$$

$$\leq \mathbf{Pr}_1\left(Z_{\gamma r} \leq \frac{2\beta}{\beta-1}(\log n)^{2\gamma+1} \,\middle|\, T \leq \gamma r\right) + O(n^{-2})$$

$$\leq \mathbf{Pr}_{\eta_1(\log n)^{2\gamma+2}}\left(\inf_{t\leq \gamma r} Z_t \leq \frac{2\beta}{\beta-1}(\log n)^{2\gamma+1}\right) + O(n^{-2})$$

$$\leq 1 - [1 - \mathbf{Pr}_{3\log_\beta n}\left(Z_{\gamma r} = 0\right)]^{\frac{2\beta}{\beta-1}(\log n)^{2\gamma+1}} + O(n^{-2})$$

$$\leq 1 - \left[1 - \frac{1}{n^3}\right]^{\frac{2\beta}{\beta-1}(\log n)^{2\gamma+1}} + O(n^{-2}) = O\left(n^{-2}\right).$$

$\square$

Acknowledgements

References

N. Berger, C. Borgs, J. T. Chayes and A. Saberi. On the spread of viruses on the internet. In *Proceedings of the Sixteenth Annual ACM-SIAM Symposium on Discrete Algorithms*, pages 301–310. ACM, New York (2005). MR2298278.

B. Bollobás. The isoperimetric number of random regular graphs. *European J. Combin.* **9** (3), 241–244 (1988). MR947025.

V. H. Can. Metastability for the contact process on the preferential attachment graph. *ArXiv Mathematics e-prints* (2015). arXiv: 1502.05633.

V. H. Can and B. Schapira. Metastability for the contact process on the configuration model with infinite mean degree. *Electron. J. Probab.* **20**, no. 26, 22 (2015). MR3325096.

S. Chatterjee and R. Durrett. Contact processes on random graphs with power law degree distributions have critical value 0. *Ann. Probab.* **37** (6), 2332–2356 (2009). MR2573560.

C. C. Z. Dou. *Studies of random walks on groups and random graphs.* Ph.D. thesis, Massachusetts Institute of Technology, Dept. of Mathematics (1992).

T. E. Harris. *The theory of branching processes.* Die Grundlehren der Mathematischen Wissenschaften, Bd. 119. Springer-Verlag, Berlin; Prentice-Hall, Inc., Englewood Cliffs, N.J. (1963). MR0163361.

D. A. Levin, Y. Peres and E. L. Wilmer. *Markov chains and mixing times.* American Mathematical Society, Providence, RI (2009). ISBN 978-0-8218-4739-8. MR2466937.

T. M. Liggett. *Stochastic interacting systems: contact, voter and exclusion processes*, volume 324 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer-Verlag, Berlin (1999). ISBN 3-540-65995-1. MR1717346.

E. Lubetzky and A. Sly. Cutoff phenomena for random walks on random regular graphs. *Duke Math. J.* **153** (3), 475–510 (2010). MR2667423.

T. Mountford, J.-C. Mourrat, D. Valesin and Q. Yao. Exponential extinction time of the contact process on finite graphs. *Stochastic Process. Appl.* **126** (7), 1974–2013 (2016). MR3483744.

T. Mountford, D. Valesin and Q. Yao. Metastable densities for the contact process on power law random graphs. *Electron. J. Probab.* **18**, No. 103, 36 (2013). MR3145050.

R. Pastor-Satorras and A. Vespignani. Epidemic dynamics and endemic states in complex networks. *Phys. Rev. E* **63**, 066117 (2001a). DOI: 10.1103/PhysRevE.63.066117.

R. Pastor-Satorras and A. Vespignani. Epidemic spreading in scale-free networks. *Phys. Rev. Lett.* **86**, 3200–3203 (2001b). DOI: 10.1103/PhysRevLett.86.3200.

J. Peterson. The contact process on the complete graph with random vertex-dependent infection rates. *Stochastic Process. Appl.* **121** (3), 609–629 (2011). MR2763098.

A.V. Prokhorov. Bernstein inequality. In M.. Hazewinkel, editor, *Encyclopaedia of Mathematics.* Springer (2002). http://eom.springer.de/B/b015700.htm.