Electron. J. Probab. **21** (2016), no. 56, 1–22.
ISSN: 1083-6489 DOI: 10.1214/16-EJP6

A lower bound for p_c in range-R bond percolation in two and three dimensions*

Spencer Frei† Edwin Perkins‡

Abstract

We use the connection between bond percolation and SIR epidemics to establish lower bounds for the critical percolation probability in 2 and 3 dimensions as the range becomes large. The bound agrees with the conjectured asymptotics for the long range critical probability, refines results of M. Penrose, and complements results of van der Hofstad and Sakai in dimensions greater than 6.

Keywords: long range bond percolation; critical probability; SIR epidemic.
AMS MSC 2010: Primary 60K35, Secondary 60J68; 60J80; 92D30.
Submitted to EJP on March 13, 2016, final version accepted on August 10, 2016.
Supersedes arXiv:1603.04130.

1 Introduction

1.1 Range-R bond percolation

We study the critical probability in range-R bond percolation. For a parameter $R \in \mathbb{N}$ called the *range*, define $\mathbb{Z}^d/R = \{x/R : x \in \mathbb{Z}^d\}$. We construct an undirected graph $\mathbb{Z}^d_R$ with vertex set $\mathbb{Z}^d/R$ and assign edges between two vertices $x, y \in \mathbb{Z}^d/R$ if $0 < \|x - y\|_\infty \leq 1$, where $\|\cdot\|_\infty$ denotes the ℓ^∞ norm on $\mathbb{R}^d$. Write $x \sim y$ if there exists an edge between x and y in $\mathbb{Z}^d/R$, let $\mathcal{N}(x)$ denote the set of neighbours of x, and denote its size by

$$V(R) := |\mathcal{N}(x)| = |\{y \in \mathbb{Z}^d/R, y \sim 0 : 0 < \|y\|_\infty \leq 1\}| = (2R+1)^d - 1 \ ,$$

where $|S|$ denotes the cardinality of a finite set S. Let $E(\mathbb{Z}^d_R)$ denote the set of edges in $\mathbb{Z}^d_R$. The structure of this graph is unchanged if one scales the lattice so that the vertex set is $\mathbb{Z}^d$ and there are edges between points $x, y \in \mathbb{Z}^d$ when $0 < \|x - y\|_\infty \leq R$, but for the remainder of this paper we shall focus solely on the graph $\mathbb{Z}^d_R$ for $d \leq 3$. If $x \sim y$, we let (x, y) or (y, x) denote the edge between x and y.

*Supported by an NSERC Discovery grant.
†Department of Statistics, University of California, Los Angeles, United States of America.
E-mail: spencerfrei@ucla.edu
‡Department of Mathematics, University of British Columbia, Vancouver, Canada.
E-mail: perkins@math.ubc.ca

We construct a random, undirected subgraph $G = G_R$ with vertex set $\mathbb{Z}_R^d$ by considering a collection of i.i.d. Bernoulli random variables $\{B(e) : e \in E(\mathbb{Z}_R^d)\}$, each with parameter $p > 0$. An edge e is *open* iff $B(e) = 1$, while edges with $B(e) = 0$ are *closed*. G is the resulting subgraph with edge set equalling the set of open edges. Two vertices x and y in $\mathbb{Z}_R^d$ are *connected* if $x = y$ or there is a path between x and y consisting of open edges; we denote this event by $x \leftrightarrow y$. Let $\lambda = V(R)p > 0$ denote the mean number of neighbours in G of any vertex. The cluster $\mathcal{C}_x$ in G containing x is

$$\mathcal{C}_x := \{y \in \mathbb{Z}^d/R : x \leftrightarrow y\},$$

and the *percolation probability*, $q(p)$, is denoted by

$$q(p) = \mathbb{P}_p(|\mathcal{C}_0| = \infty).$$

The *critical probability* is defined by

$$p_c = p_c(R) := \inf\{p : q(p) > 0\} \ ,$$

and the associated critical value of λ is

$$\lambda_c = \lambda_c(R) := p_c(R)V(R) \ .$$

The obvious monotonicity in p shows that $q(p) = 0$ for $p \in [0, p_c)$ and $q(p) > 0$ for $p \in (p_c, 1]$. M. Penrose [13] showed that

$$\lim_{R\to\infty} \lambda_c(R) = 1 \ . \tag{1.1}$$

If we view the set of vertices distance n in the G-graph distance from the origin as a set-valued "interactive branching process" η_n, then for R large and λ bounded, η_n should be well-approximated by an ordinary Galton-Watson branching process with offspring mean λ. This is because with so many potential percolation steps from generation n to $n + 1$, it is unlikely that the process will take a step to a previously visited site. As a result one expects the critical mean $\lambda_c(R)$ to be close to 1, the critical mean for the GW branching process. (This intuition is of course well-known and was pointed by Penrose, among others.) Write $a(R) \sim b(R)$ iff the ratio approaches 1 as $R \to \infty$. Van der Hofstad and Sakai [7] have obtained finer asymptotics on $\lambda_c(R)$ for $d > 6$ using the lace expansion:

$$\lambda_c(R) - 1 \sim \frac{\theta_d}{R^d} \tag{1.2}$$

where θ_d has an explicit expression in terms of a random walk with uniform steps on $[-1, 1]^d$. The extension of (1.2) to $d \geq 4$ (with logarithmic corrections in $d = 4$) is the subject of ongoing work of one of us [EP] with Xinghua Zheng. In the next subsection we recall a parallel conjecture (Conjecture 1.3) for $d = 2, 3$ suggested in Lalley-Perkins-Zheng [10]. Our main result (Theorem 1.2 below) is a lower bound on $\lambda_c(R)$ which confirms the conjecture aside from the constant. The first step is to make the above connection between bond percolation and "interactive branching processes" more precise in the next subsection.

1.2 SIR epidemic models and bond percolation

We recall a well-known connection between bond percolation and an epidemic model which dates back at least to [12]. For the SIR epidemic model on $\mathbb{Z}_R^d$, each vertex $x \in \mathbb{Z}^d/R$ can be in one of three possible states: we denote $x \in \xi_n$ if x is susceptible at

time n; $x \in \eta_n$ if x is infected at time n; and $x \in \rho_n$ if x is recovered at time n. Each vertex x is in exactly one of the three states, so that for each time n,

$$\mathbb{Z}^d/R = \xi_n \dot\cup \eta_n \dot\cup \rho_n \ .$$

The epidemic starts from finite initial configurations of infected sites, η_0, and recovered sites, ρ_0. An infected site $x \in \eta_n$ infects a susceptible neighbour $y \in \xi_n,\ y \sim x$, with probability $p = p(R)$, where the infection events are conditionally independent given the current configuration of states. Infected sites are infected for unit time, after which they become recovered, and recovered sites are immune from infection, so that if $x \in \eta_n$, then $x \in \rho_k$ for all $k \geq n+1$. Note that

$$\rho_n = \rho_0 \dot\cup \eta_0 \dot\cup \ldots \dot\cup \eta_{n-1}, \text{ and is finite for all } n. \tag{1.3}$$

We let

$$\partial C_n := \{e = (x,y) \in E(\mathbb{Z}^d_R) : x \in \eta_n, y \in \xi_n\} \tag{1.4}$$

be the set of possible infection edges leading to η_{n+1}, where in describing edges (x,y) in ∂C_n we will use the convention that the first coordinate refers to the site in η_n. So that for any given edges $(x_i, y_i), i \leq m$, the event $\{\partial C_n = \{(x_i,y_i) : i \leq m\}\}$ will include this specification. (This amounts to choosing a particular representative for an equivalence class in our notation.) Let $\mathcal{F}_n = \sigma(\rho_0, \eta_k, k \leq n)$. Clearly the finite random set ∂C_n is $\sigma(\eta_n, \rho_n) \subset \mathcal{F}_n$-measurable because $\xi_n = (\eta_n \cup \rho_n)^c$. If y is a site in $\mathbb{Z}^d_R$, let

$$D_n(y) = \{x \in \eta_n : (x,y) \in \partial C_n\}. \tag{1.5}$$

If $S = \{(x_i,y_i) : i \leq m\}$ is a set of distinct edges in $\mathbb{Z}^d_R$ and $V \subset \{y_i : i \leq m\} := V_2$, then the above description implies that

$$\mathbb{P}(\eta_{n+1} = V | \mathcal{F}_n) = \prod_{y \in V_2 \setminus V} (1-p)^{|D_n(y)|} \prod_{y \in V} [1 - (1-p)^{|D_n(y)|}] \text{ a.s. on } \{\partial C_n = S\}. \tag{1.6}$$

Clearly (1.6) and the joint law of (η_0, ρ_0) uniquely determines the law of the SIR epidemic η. Since $\rho_{n+1} = \rho_n \cup \eta_n$, (1.6) also gives the (time-homogeneous) Markov property of (η_n, ρ_n):

$$\mathbb{P}((\eta_{n+1}, \rho_{n+1}) \in \cdot \,\big|\mathcal{F}_n) = \mathbb{P}((\eta_{n+1}, \rho_{n+1}) \in \cdot \,\big|\eta_n, \rho_n) = \mathbb{P}_{\eta_n, \rho_n}((\eta_1, \rho_1) \in \cdot) \ , \tag{1.7}$$

where $\mathbb{P}_{\eta_0,\rho_0}$ is the law of (η, ρ) starting at (η_0, ρ_0), whenever (η_0, ρ_0) are disjoint finite sets of sites. Unless otherwise indicated we will assume the SIR epidemic under $\mathbb{P}$ starts from $\eta_0 = \{0\}$ and $\rho_0 = \emptyset$, and assume $p = p(R)$, is such that

$$\lambda(R) := p(R) V(R) \geq 1.$$

We may use the edge percolation variables $\{B(e) : e \in E(\mathbb{Z}^d_R)\}$ from the last section to define the infection dynamics of η as follows: Every infected–susceptible pair $(x,y) \in \partial C_n$ has a successful infection iff $B(x,y) = 1$. *The dynamics of the SIR epidemic force each bond variable $B(x,y)$ to be used at most once in defining the epidemic process, precisely when x is infected and y is susceptible or conversely.* The fact that the above specification of the infection dynamics leads to (1.6) and hence defines an SIR epidemic, is immediate from Lemma 2.1 below.

We collect the dynamics here:

$$\begin{aligned} \eta_{n+1} &= \bigcup_{x \in \eta_n} \{y \in \xi_n : B(x,y) = 1\}, \\ \rho_{n+1} &= \rho_n \cup \eta_n, \\ \xi_{n+1} &= \xi_n \setminus \eta_{n+1}. \end{aligned} \tag{1.8}$$

Let $d_G(x,y)$ denotes the graph distance in the percolation graph, G, between $x, y \in \mathbb{Z}_R^d$ (it may be infinite), and similarly let $d_G(A,x)$ denote the distance in G between a set of vertices A and $x \in \mathbb{Z}_R^d$. An easy induction (see below) shows that if $(\eta_0, \rho_0) = (\{0\}, \emptyset)$, then for $n \in \mathbb{Z}_+$, $\eta_n = \{x \in \mathbb{Z}_R^d : d_G(0,x) = n\}$. More generally for a given pair of disjoint finite sets of sites in $\mathbb{Z}_R^d$, (η_0, ρ_0), let $G(\rho_0)$ be the percolation graph but where all edges containing a vertex in ρ_0 are closed. Then for (η, ρ) starting at (η_0, ρ_0),

$$\eta_n = \{x \in \mathbb{Z}_R^d : d_{G(\rho_0)}(\eta_0, x) = n\} := \eta_n^{\eta_0, \rho_0}. \tag{1.9}$$

We present the inductive argument for the above. If $n = 0$, this is obvious. Assume the result for n. Assume first $x \in \eta_{n+1}$. Then by (1.8), $x \in \xi_n \subset \rho_0^c$ and there is an $x' \in \eta_n$ such that $B(x', x) = 1$. By hypothesis, $d_{G(\rho_0)}(\eta_0, x') = n$. Since $x' \notin \rho_0$ (or else the above distance would be ∞) this implies that $d_{G(\rho_0)}(\eta_0, x) \le n+1$. As $x \notin \cup_{k \le n} \eta_k = \rho_n \cup \eta_n$ ($x \in \xi_n$ is disjoint from this union by (1.8)), the induction hypothesis imples $d_{G(\rho_0)}(\eta_0, x) > n$ and so $d_{G(\rho_0)}(\eta_0, x) = n+1$. Conversely assume that $d_{G(\rho_0)}(\eta_0, x) = n+1$. This again implies $x \notin \rho_0$. There is an $x' \notin \rho_0$ so that $d_{G(\rho_0)}(\eta_0, x') = n$, $x' \sim x$ and $B(x', x) = 1$. By hypothesis $x' \in \eta_n$. Also by the induction hypothesis $x \notin \cup_{k=1}^n \eta_k = \rho_n \cup \eta_n$ (or else $d_{G(\rho_0)}(\eta_0, x) \le n$). This means x must be in ξ_n and so by (1.8) we conclude that $x \in \eta_{n+1}$. This completes the induction.

Here is one simple consequence of the coupling of initial conditions that comes from this construction. It will be used in Section 3.

Lemma 1.1. *For any disjoint finite sets $\eta_0, \rho_0 \subset \mathbb{Z}_R^d$ and any $n \in \mathbb{Z}_+$,*
(a) $\mathbb{P}_{\eta_0,\rho_0}(|\cup_{i=0}^n \eta_i| \le r) \ge \mathbb{P}_{\eta_0,\emptyset}(|\cup_{i=0}^n \eta_i| \le r) \quad \forall r \ge 0.$
(b) $\mathbb{E}_{\eta_0,\emptyset}(|\rho_n|) \le |\eta_0| E_{\{0\},\emptyset}(|\rho_n|)$.

Proof. (a) Clearly $d_{G(\rho_0)}(\eta_0, x) \le n$ implies $d_G(\eta_0, x) \le n$, since the former gives the existence of a chain of at most n open bonds in G starting at a vertex in η_0 and ending at x with the additional property that each vertex is not in ρ_0, and this clearly implies the latter. The result is now immediate from the characterization of η_n given in (1.9).

(b) As we are dealing with coupled versions of ρ_n all starting at $\rho_0 = \emptyset$ but with different η_0's, we let $\rho_n^{\eta_0} = \cup_{k=0}^{n-1} \eta_k^{\eta_0,\emptyset}$. As $\rho_0 = \emptyset$ now, by (1.3) and (1.9),

$$\rho_n^{\eta_0} = \{y \in \mathbb{Z}_R^d : d_G(\eta_0, y) < n\} = \cup_{x \in \eta_0} \{y \in \mathbb{Z}_R^d : d_G(x,y) < n\},$$

the last by an elementary argument. This shows that $|\rho_n^{\eta_0}| \le \sum_{x \in \eta_0} |\rho_n^{\{x\}}|$, and so taking means we may conclude that

$$\mathbb{E}_{\eta_0,\emptyset}(|\rho_n|) \le \sum_{x \in \eta_0} \mathbb{E}_{\{x\},\emptyset}(|\rho_n|) = |\eta_0| \mathbb{E}_{\{0\},\emptyset}(|\rho_n|).$$

The last equality follows from translation invariance of the dynamics of η. □

We say that an epidemic *survives* if with positive probability, for all $n \in \mathbb{N}$, we have $\eta_n \neq \emptyset$. If with probability one, for some finite n we have $\eta_n = \emptyset$, we say the epidemic becomes *extinct*.

If we have survival of the epidemic, then with positive probability there is an infinite sequence of sites x_k such that $x_k \in \eta_k$, $x_k \sim x_{k-1}$, and x_{k-1} infected x_k at time k, and hence the edge $B(x_{k-1}, x_k)$ is open. This implies that with positive probability, we have percolation from $\eta_0 = \{0\}$ to infinity in range-R bond percolation. Likewise, percolation from η_0 to infinity in the percolation model induces an infinite sequence of infections and hence survival in the epidemic. In this way, percolation to infinity is equivalent to survival of the analogous SIR epidemic on the same graph.

To understand the large R behaviour of $p_c(R)$ we will fix a parameter $\theta > 0$, rescale our model by R^γ, choose $p = p(R)$ so that $\lambda := V(R)p(R) = 1 + \frac{\theta}{R^\gamma}$, and work with the empirical process of the rescaled epidemic model,

$$X_t^R = \frac{1}{R^\gamma} \sum_{x \in \eta_{\lfloor tR^\gamma \rfloor}} \delta_{x/(R\sqrt{R^\gamma})}.$$

We first have rescaled the range to size 1 and then applied the usual Brownian scaling due to the time rescaling by $R^\gamma t$. We will assume $\theta \leq 1$ which for $d > 1$, the case of interest, implies that $p = V(R)^{-1}[1 + \theta R^{-\gamma}] \in (0, 1)$ (recall R is a natural number). In the above $\gamma > 0$ must be chosen carefully if we want to obtain an interesting limit. For example, if γ is too small, then the range $R = (R^\gamma)^{1/\gamma}$ will be too large relative to the scaling parameter R^γ and the effect of suppressing infections onto recovered sites will become negligible. As a result we will recover the scaling limit of branching random walk, super-Brownian motion. The "correct" choice, taken from Lalley and Zheng [11] (see below), is $\gamma = \frac{2d}{6-d}$ (so that for $d = 2, 3$, $R^\gamma = R^{d-1} \in \mathbb{N}$ if $R \in \mathbb{N}$). Let $\Rightarrow$ denote weak convergence on the appropriate space, $M_F(\mathbb{R}^d)$ be the space of finite Borel measures on $\mathbb{R}^d$ with the topology of weak convergence, and $\mu(f)$ denote the integral $\int f d\mu$ of a function f with respect to a measure μ. C_K^2 is the space of C^2 functions from $\mathbb{R}^d$ to $\mathbb{R}$ with compact support, and $D([0, \infty), M_F(\mathbb{R}^d))$ is the Skorohod space of cadlag $M_F(\mathbb{R}^d)$-valued paths.

Conjecture 1.2. *Assume $d \leq 3$, $\gamma = \frac{2d}{6-d}$, and for appropriate $\{X_0^R\}$, $X_0^R \Rightarrow X_0$, for a compactly supported fixed finite measure X_0. Then $X^R \Rightarrow X$ in $D([0, \infty), M_F(\mathbb{R}^d))$ as $R \to \infty$. The limit X is the unique solution to the martingale problem*

$$X_t(\phi) = \int \phi dX_t = X_0(\phi) + M_t(\phi) + \int_0^t X_s\left(\frac{\Delta\phi}{6} + \theta\phi\right) ds - \int_0^t X_s(L_s\phi) ds \ , \ \forall \phi \in C_K^2, \tag{1.10}$$

where X is a continuous $M_F(\mathbb{R}^d)$-valued process, L_s is the local time of X, i.e., satisfies $\int_0^t X_s(\psi) ds = \int L_t(x)\psi(x) dx$, and $M(\phi)$ is a continuous martingale with $\langle M(\phi)\rangle_t = \int_0^t X_s(\phi^2) ds$.

The "appropriate" $\{X_0^R\}$ are those satisfying the regularity condition (17) in Theorem 2 of Lalley and Zheng [11]. That result establishes a weak convergence result which is very close to the above. Instead of $\mathbb{Z}_R^d$, they work with a "village model" on $\mathbb{Z}^d \times \{1, \ldots, N\}$ where sites (x, m) and (y, n) are neighbours iff x and y are nearest neighbours in $\mathbb{Z}^d$, or $x = y$ and $m \neq n$. The rescaling parameter in Theorem 2 of [11] is $N^{2/(6-d)}$. Equating N and R^d so that the number of neighbours in the two models are of the same order of magnitude, leads to the scaling parameter R^γ chosen above. Xinghua Zheng in fact has proved the above Conjecture for $d = 2$ [private communication].

The well-posedness of the above martingale problem is established, for example, in Theorem 2.2 of Lalley, Perkins, and Zheng[10] who also showed that for $d = 2, 3$, there is a phase transition in (1.10):

Theorem 1.1. Let $d = 2$ or $d = 3$. There exists a $\theta_c = \theta_c(d) > 0$ such that for all nonzero finite measures X_0,

a. for $\theta < \theta_c$, X becomes extinct.

b. for $\theta > \theta_c$, X survives (with positive probability).

If $d = 1$, then for all θ and all finite X_0, X becomes extinct.

(Although percolation fails trivially for any R and $p < 1$ in $d = 1$, there is some work to do to show this for the one-dimensional continuous model in the above.) For $d = 2, 3$, a

formal interchange of limits in t and R in the above results, and the equality $\frac{2d}{6-d} = d-1$ for $d = 2, 3$, leads us to the following refinement of M. Penrose's result (1.1):

Conjecture 1.3. *For $d = 2$ or 3, $V(R)p_c(R) - 1 \sim \frac{\theta_c}{R^{d-1}}$.*

(The analogue of the above for the "village model" described above is raised in Section 2.2 of Lalley, Perkins, and Zheng [10].) Simulations carried out by Deshin Finlay [5] suggest that

$$V(R)p_c(R) - 1 \sim \begin{cases} 1.2/R, & \text{if } d = 2, \\ .7/R^2, & \text{if } d = 3. \end{cases}$$

The conjecture really should be made with $\frac{2d}{6-d}$ in place of $d-1$ even though they are equal for the relevant values of d. For example, putting $d = 4$ into the former formula is consistent with the conjectured behaviour in the critical $d = 4$ case cited above (albeit with logarithmic corrections in R).

The similarity with the critical infection rate of the long range contact process is discussed in Section 1.5 of [10]. Our $d \leq 3$ setting corresponds to the $d = 1$ setting of the long range contact process where the conjecture corresponding to Conjecture 1.3 remains unresolved (see the Conjecture following Theorem 2 in Durrett and Perkins [4]). In the contact process setting, however, upper and lower bounds establishing the correct rate of convergence (if not the exact constant) were established in Theorem 1 of Bramson, Durrett and Swindle[2]. Our main result, which adapts some of the nice ideas in the 1-dimensional lower bound from [2], is a lower bound on $p_c(R)$, consistent with the rate in Conjecture 1.3:

Theorem 1.2. For $d = 2, 3$ there is a constant $\theta = \theta(d) > 0$, such that for all $R \in \mathbb{N}$

$$V(R)p_c(R) - 1 \geq \frac{\theta}{R^{d-1}}.$$

For the village model described above, independent work of E. Neuman and X. Zheng (in preparation) shows an analogue of the above result and a companion upper bound for p_c, using different methods.

In Section 2 we gather some results on the SIR epidemic η_n and an associated branching random walk (BRW) Z_n. In Subsection 2.1 we give a more careful derivation of the connection between bond percolation and η_n (Lemma 2.1), and use it to derive an expression on the conditional increments of $|\eta_n|$, Corollary 2.2. In Subsection 2.2 we couple η with a dominating BRW Z and in Subsection 2.3 we establish an upper bound on the probability that Z exits a large ball by generation n. The latter is classical (and goes back at least to Dawson-Iscoe-Perkins[3]) aside from the fact that the long range structure of the BRW means there is an additional parameter going to infinity. The changes needed to derive this result from the arguments in Section 4 of [3] are presented in an Appendix. In Section 3 we will use Corollary 2.2 and the above upper bound on the dominating BRW to show that for some fixed $\theta > 0$ and R satisfying $V(R)p(R) - 1 \geq \frac{\theta}{R^{d-1}}$ there is some k so that $\mathbb{E}(|\eta_k|) < 1$ (Proposition 3.1). We will then show, essentially by a comparison to subcritical branching (thanks in part to Lemma 1.1), that this gives a.s. extinction of η. In view of our assumption on $\lambda(R)$ this implies the lower bound on $\lambda_c(R)$.

2 Preliminary results

In this section and for the remainder of the paper, we will consider an SIR epidemic η on $\mathbb{Z}_R^d$ for $d = 2$ or 3, and choose the probability of infection $p = p(R)$, so that for a fixed $\theta \in (0, 1]$,

$$\lambda(R) = V(R)p(R) = 1 + \theta/R^{d-1} \ . \tag{2.1}$$

2.1 The increments of η_n

Throughout this subsection we assume our initial condition (η_0, ρ_0) is a fixed (deterministic) pair of disjoint finite sets of vertices and assume $\{\eta_n : n \geq 0\}$ is defined using the edge percolation variables $\{B(e) : e \in E(\mathbb{Z}^d_R)\}$. For each edge $(x, y) \in E(\mathbb{Z}^d_R)$, let

$$n(x, y) = \inf\{\ell \geq 1 : x \in \eta_{\ell-1} \text{ and } y \in \xi_{\ell-1}, \text{ or } x \in \xi_{\ell-1} \text{ and } y \in \eta_{\ell-1}\} .$$

(By convention, $\inf \emptyset = \infty$). Then $n(e)$ is the unique exploration time of the edge e when $B(e)$ is used to define η_n (and $n(e) = \infty$ means that the edge is never used in the definition of η). Clearly $n(e) = n + 1$ iff $e = (x, y) \in \partial C_n$, and so $n(e)$ is a predictable $\mathcal{F}_n$-stopping time, i.e., the events $\{n(e) = n + 1\}$ are $\mathcal{F}_n$-measurable.

For $e \in E(\mathbb{Z}^d_R)$, define

$$V_n(e) = \begin{cases} B(e), & \text{if } e \in \partial C_{n-1}, \\ 2, & \text{if } e \notin \partial C_{n-1}. \end{cases} \tag{2.2}$$

Then $x \in \eta_n$ if and only if there exists a sequence of points $x_0, \ \ldots \ , x_n = x$ such that $x_0 \in \eta_0$, $x_k \notin \rho_0$ for all $1 \leq k \leq n$, and $V_k(x_{k-1}, x_k) = 1$ for all $k = 1, \ldots, n$. In this way, $V_1, \ldots, V_n$ describe the epidemic in terms of the geometry of the percolation substructure. We define

$$\bar{\mathcal{F}}_n = \sigma(V_1, \ldots, V_n) .$$

It is clear from the above that $\mathcal{F}_n \subset \bar{\mathcal{F}}_n$, and it is not hard to see that this inclusion is strict (although this will not be needed). Here is the previously promised result from which it follows easily that η is an SIR epidemic process.

Lemma 2.1. *(a) For any finite set of edges $S = \{(x_i, y_i) : i \leq m\}$ in $\mathbb{Z}^d_R$ and $\underline{S} \subset S$,*

$$\begin{aligned} &\mathbb{P}(B(e) = 1 \ \forall e \in \underline{S}, B(e') = 0 \ \forall e' \in S \setminus \underline{S} \,|\, \bar{\mathcal{F}}_n) \\ &\quad = p(R)^{|\underline{S}|}(1 - p(R))^{|S \setminus \underline{S}|} \text{ a. s. on the } \bar{\mathcal{F}}_n\text{-measurable set } \{\partial C_n = S\}. \end{aligned} \tag{2.3}$$

(b) If $x \sim y$, then

$$\mathbb{P}(B(x, y) = 1 \,|\, \bar{\mathcal{F}}_n) = p(R) \text{ on the } \bar{\mathcal{F}}_n\text{-measurable set } \{x \in \eta_n, y \in \xi_n\}. \tag{2.4}$$

Proof. (a) Let $A_k : E(\mathbb{Z}^d_R) \to \{0, 1, 2\}$, $k = 1, \ldots, n + 1$, be fixed functions such that each set

$$S_k := \{e \in E(\mathbb{Z}^d_R) : A_k(e) \neq 2\}, \quad k = 1, \ldots, n + 1,$$

is finite. (There are countably many such functions.) Since $\bar{\mathcal{F}}_n = \sigma(V_k, k \leq n)$, it suffices to show that

$$\begin{aligned} &\mathbb{P}(B(e) = 1 \ \forall e \in \underline{S}, B(e') = 0 \ \forall e' \in S \setminus \underline{S}, \partial C_n = S, (V_k)_{k \leq n} = (A_k)_{k \leq n}) \\ &= \mathbb{P}(B(e) = 1 \ \forall e \in \underline{S}, B(e') = 0 \ \forall e' \in S \setminus \underline{S}) \cdot \mathbb{P}(\partial C_n = S, (V_k)_{k \leq n} = (A_k)_{k \leq n}) \end{aligned} \tag{2.5}$$

The required result then follows by the independence of the $\{B(e) : e \in E(\mathbb{Z}^d/R)\}$ and the definition of conditional probability.

For simplicity of notation, let

$$C = \{\partial C_n = S, V_k = A_k, k \leq n\}.$$

Without loss of generality, we may assume that C is non-empty. On the event $\cap_{k=1}^n \{V_k = A_k\}$, we know that $\partial C_{k-1} = S_k$ for each $k = 1, \ldots, n$, since $V_k(e')$ is $\{0, 1\}$-valued only on edges $e' \in \partial C_{k-1}$, and is otherwise equal to 2. Moreover, on C we have $n(e) = n + 1$ for all $e \in S$, and hence the sets

$$S_1, \ \ldots, \ S_n, S$$

are a collection of $n+1$ mutually disjoint sets. Denote $S^{(n)} = \cup_{k\le n} S_k$. On $\cap_{k=1}^n \{V_k = A_k\}$, for all $1 \le k \le n$, $x \in \eta_k$ iff there is an $(x', x) \in S_k$ such that $A_k(x', x) = 1$ (since on this set $A_k(x', x) = B(x', x)$). Since $y \in \xi_n$ iff $y \notin \rho_0 \cup (\cup_{k=0}^n \eta_k)$, this means that on $\cap_{k=1}^n \{V_k = A_k\}$, we have $y \in \xi_n$ iff $y \notin \rho_0$ and for all $k \le n$ there is no $(y', y) \in S_k$ such that $A_k(y', y) = 1$. This shows that on $\cap_{k=1}^n \{V_k = A_k\}$,

$$\begin{aligned}\partial C_n = \{(x,y) : x \sim y, y \notin \rho_0, \exists (x', x) \in S_n \text{ s.t. } A_n(x', x) = 1,\\ \text{and } \forall k \le n \text{ there is no } (y', y) \in S_k \text{ s.t. } A_k(y', y) = 1\}\\ =: S'_{n+1}.\end{aligned}$$

Therefore

$$\cap_{k=1}^{n+1} \{V_k = A_k\} = \begin{cases} \cap_{k=1}^n \{V_k = A_k\} \cap \{A_{n+1}(e) = B(e)\ \forall e \in S_{n+1}\}, & \text{if } S_{n+1} = S'_{n+1},\\ \emptyset, & \text{if } S_{n+1} \ne S'_{n+1}.\end{cases}$$

The obvious induction now shows that $\cap_{k=1}^n \{V_k = A_k\} \in \sigma(B(e) : e \in S^{(n)})$. Also since $C \ne \emptyset$ we must have $S = S'_{n+1}$ and so $C = \cap_{k=1}^n \{V_k = A_k\} \in \sigma(B(e) : e \in S^{(n)})$. Since $S \cap S^{(n)}$ is empty, and distinct edges have independent Bernoulli variables, (2.5) is now immediate, and (a) is proved.

(b) Let S be as in (a) and containing the edge (x, y). Now sum the result in (a) over all $\bar{S}$ as in (a) with $(x, y) \in \bar{S}$ and S fixed. A simple application of the binomial theorem will give the conclusion of (b) on the set $\{\partial C_n = S\}$. Finally take the union over these events where $(x, y) \in S$ to derive (b). □

Corollary 2.2.

$$\mathbb{E}\Big[|\eta_{n+1}| - |\eta_n| \big| \bar{\mathcal{F}}_n\Big] = \left(\frac{\theta}{R^{d-1}}\right) |\eta_n| - \left(1 + \frac{\theta}{R^{d-1}}\right) \sum_{x \in \eta_n} \sum_{\substack{y \sim x \\ y \in \rho_n \cup \eta_n}} \frac{1}{V(R)}. \tag{2.6}$$

Proof. Conditioned on the history of infected sites up to time n, the sites that are infected at time $n+1$ are the susceptible neighbours $y \in \xi_n$ of infected sites $x \in \eta_n$ such that $B(x, y) = 1$. Thus,

$$\begin{aligned}\mathbb{E}[|\eta_{n+1}| \, \big| \bar{\mathcal{F}}_n] &= \mathbb{E}\left[\sum_{x \in \eta_n} \sum_{y \sim x} \mathbf{1}\{y \in \xi_n\} \mathbf{1}\{B(x, y) = 1\} \Big| \bar{\mathcal{F}}_n\right]\\ &= \sum_{x \in \eta_n} \sum_{y \sim x} (1 - \mathbf{1}\{y \in \rho_n \cup \eta_n\}) \mathbb{E}\left[\mathbf{1}\{B(x, y) = 1\} | \bar{\mathcal{F}}_n\right]\\ &= \sum_{x \in \eta_n} \sum_{y \sim x} (1 - \mathbf{1}\{y \in \rho_n \cup \eta_n\}) \cdot p(R) \qquad \text{(by Lemma 2.1}(b)\text{)}\\ &= p(R)V(R)|\eta_n| - p(R)V(R) \sum_{x \in \eta_n} \sum_{\substack{y \sim x \\ y \in \rho_n \cup \eta_n}} \frac{1}{V(R)},\end{aligned}$$

and so the result follows upon recalling $p(R)V(R) = \big(1 + \frac{\theta}{R^{d-1}}\big)$. □

The second term on the right-hand side of (2.6) is an *interference term* arising from the conditional expectation of those sites where an infection was attempted but suppressed as the site was already infected or recovered.

2.2 Coupling with a branching envelope

Throughout this subsection we assume our initial condition $(\eta_0, \rho_0) = (\{0\}, \emptyset)$. As in Lalley-Zheng[11], we may couple a copy of η with a dominating branching random walk

in our long range setting. To see this, label potential particles in the branching random walk by $I = \cup_{n=0}^{\infty}\mathcal{N}(0)^n$, where $\mathcal{N}(0)^0 = \{\phi\}$. It will be convenient to totally order $\mathcal{N}(0)$ as $\{e_1, e_2, \ldots, e_{V(R)}\}$ and then totally order each $\mathcal{N}(0)^n$ lexicographically by $<$. For $\alpha = (\alpha_1, \ldots, \alpha_n) \in \mathcal{N}(0)^n \subset I$, write $|\alpha| = n$, $\alpha|i = (\alpha_1, \ldots, \alpha_i)$ for $1 \le i \le n$ ($|\phi| = 0$), and $\pi\alpha = (\alpha_1, \ldots, \alpha_{n-1})$ be the parent of α, where if $n = 1$ the parent is the root index ϕ. Concatenation in I is denoted by $(\alpha_1, \ldots, \alpha_n) \vee (\beta_1, \ldots, \beta_m) = (\alpha_1, \ldots, \alpha_n, \beta_1, \ldots, \beta_m)$. Let $\{B^\alpha : \alpha \in I \setminus \{\phi\}\}$ be an iid collection of Bernoulli random variables with $\mathbb{P}(B^\alpha = 1) = 1 - \mathbb{P}(B^\alpha = 0) = p(R)$, and set $\mathcal{G}_n = \sigma(\{B^\alpha : |\alpha| \le n\})$. The intuition is that each $\alpha \in I$ with $|\alpha| = n$ labels a *potential* individual in generation n and $C^\alpha = \{e : B^{\alpha\vee e} = 1\}$ are the locations of the children of α relative to the position of the parent, α. Let $M^\alpha = |C^\alpha|$ be the number of children of α, so that $\{M^\alpha : \alpha \in I\}$ is a collection of independent binomial $(V(R), p(R))$ random variables. Moreover $\{M^\alpha : |\alpha| = n\}$ is jointly independent of $\mathcal{G}_n$. So for $n \ge 1$, each $\alpha \in I$ with $|\alpha| = n$ labels an individual alive in generation n iff $B^{\alpha|i} = 1$ for all $1 \le i \le n$, in which case we write $\alpha \approx n$. Note that $\phi \approx 0$ always and for $|\alpha| = n$,

$$\{\omega : \alpha \approx n\} \in \mathcal{G}_n.$$

If for each $\alpha \in I$,

$$Y^\alpha = \begin{cases} \sum_{i=1}^{|\alpha|} \alpha_i \in \mathbb{Z}^d/R & \text{if } \alpha \approx |\alpha| \\ \Delta & \text{otherwise,} \end{cases}$$

then Y^α is the location of the particle α if it is alive, and is set to the cemetery state, Δ, otherwise. In this way $Z_n = \sum_{\alpha\approx n} \delta_{Y^\alpha}$ ($n \ge 0$) defines the empirical distribution of a branching random walk (BRW) in which each individual in generation n produces a binomial $(V(R), p(R))$ number (M^α) of children whose positions, relative to their parent and given $\sigma(M^\alpha : |\alpha| = n) \vee \mathcal{G}_n$, are uniformly distributed over

$$\mathcal{N}(0)^{(M^\alpha)} = \{(e'_1, \ldots, e'_{M^\alpha}) : e'_i \in \mathcal{N}(0), e'_1, \ldots, e'_{M^\alpha} \text{ distinct}\}.$$

Note that conditional on $\sigma(M^\alpha : |\alpha| = n) \vee \mathcal{G}_n$, the steps of the siblings from their common parent in generation n are dependent but steps corresponding to distinct generation n parents are independent. We set $Z_n(x) = Z_n(\{x\})$ for $x \in \mathbb{Z}^d/R$. Clearly Z satisfies the natural Markov property with respect to $(\mathcal{G}_n)$ and $Z_0 = 1_{\{0\}}$.

We next define our coupled SIR epidemic (η_n, ξ_n, ρ_n) inductively in n so that $\eta_j(x) := 1(x \in \eta_j) \le Z_j(x)$ for all $x \in \mathbb{Z}^d_R$ and $j \le n$, $\mathcal{F}_n := \sigma(\rho_0, \eta_1, \eta_2, \ldots, \eta_n) \subset \mathcal{G}_n$, and $(\eta_j, \xi_j, \rho_j)_{j\le n}$ has the law of an SIR epidemic process. Set $\eta_0 = 1_{\{0\}} (= Z_0)$, assume the above for n, and consider $n+1$. Let $Y_n = \{y \in \mathbb{Z}^d/R : \exists x \in \eta_n \text{ s.t. } (x, y) \in \partial C_n\} \subset \xi_n$. If $x \in \eta_n$, then $1 \le Z_n(x)$ (by induction) and so we may choose a minimal α_n^x (with respect to our total order) in the non-empty set $\{\alpha \approx n : Y^\alpha = x\}$. One easily checks that α_n^x (set it equal to Δ if $x \notin \eta_n$) is $\mathcal{G}_n$-measurable. We define (recall that $D_n(y)$ is as in (1.5))

$$\eta_{n+1} = \{y \in Y_n : \exists x \in D_n(y) \text{ s.t. } B^{\alpha_n^x \vee (y-x)} = 1\},\ \xi_{n+1} = \xi_n \setminus \eta_{n+1},\ \rho_{n+1} = \xi_n \cup \eta_n. \quad (2.7)$$

In this way α_n^x labels the BRW representative at x in generation n for $x \in \eta_n$. The fact that $D_n(y)$ and Y_n are $\mathcal{F}_n$-measurable and α_n^x is $\mathcal{G}_n$-measurable shows that η_{n+1} is $\mathcal{G}_{n+1}$-measurable and so $\mathcal{F}_{n+1} \subset \mathcal{G}_{n+1}$. Assume next that $y \in \eta_{n+1}$. Therefore there is an $x \in D_n(y)$ such that $B^{\alpha_n^x \vee (y-x)} = 1$. As we have $\alpha_n^x \approx n$, the latter implies that $\alpha_n^x \vee (y - x) \approx n + 1$ and

$$Y^{\alpha_n^x \vee (y-x)} = Y^{\alpha_n^x} + y - x = y.$$

Therefore $Z_{n+1}(y) \ge 1$ and we have proved that $\eta_{n+1} \le Z_{n+1}$. To complete the induction it suffices to establish (1.6) in the stronger form

$$\mathbb{P}(\eta_{n+1} = V | \mathcal{G}_n) = \prod_{y \in V_2 \setminus V} (1-p)^{|D_n(y)|} \prod_{y \in V} [1 - (1-p)^{|D_n(y)|}] \text{ a.s. on } \{\partial C_n = S\}, \quad (2.8)$$

where S, V, and V_2 are as in (1.6). By (2.7) the left-hand side of the above equals

$$\mathbb{P}(\forall y \in V_2 \setminus V \ \forall x \in D_n(y)\ B^{\alpha_n^x \vee (y-x)} = 0, \\ \text{and } \forall y \in V \ \exists x \in D_n(y) \ \text{ s.t. } B^{\alpha_n^x \vee (y-x)} = 1 | \mathcal{G}_n).$$

One easily checks that the $\mathcal{G}_n$-measurable collection of superscripts on the Bernoulli variables are distinct and label particles in generation $n+1$. The fact that $\{B^\alpha : |\alpha| = n+1\}$ are jointly independent and independent of $\mathcal{G}_n$ now gives (2.8) from the above.

We restate what we have shown:

Proposition 2.3. *There is a BRW $(Z_n, n \in \mathbb{Z}_+)$ and an SIR epidemic $(\eta_n, n \in \mathbb{Z}_+)$ on a common probability space s.t. $\eta_n(x) \le Z_n(x)$ for all $x \in \mathbb{Z}^d/R$, $n \in \mathbb{Z}_+$, and $Z_0 = \eta_0 = 1_{\{0\}}$. Moreover Z is a BRW in which each parent independently gives birth to a binomial $(V(R), p(R))$ number of children (M), where conditional on M, the offspring locations relative to their parent is uniform over $\mathcal{N}(0)^{(M)}$. In addition, both Z and (η, ρ) satisfy their natural Markov properties with respect to a common filtration $(\mathcal{G}_n)$ (for (η, ρ) it is* (2.8) *above).*

2.3 Branching random walk bounds

We will need a pair of bounds on the BRW Z constructed above which are in the literature either explicitly or implicitly. The first is a bound on the survival probability for a sequence of Galton-Watson processes, which is almost immediate from Lemma 2.1(a) of Bramson et al [2]. As this latter result will also play a role in the Appendix we state it carefully (and note we have strengthened a uniform integrability hypothesis to a third moment bound).

Lemma 2.4. *For any $\alpha > 0$, let $\mathcal{K}_\alpha$ be a class of Galton-Watson processes each starting with a single individual with reproduction variances not less that α, and such that*

$$\sup_{p \in \mathcal{K}_\alpha} \sum_{j=0}^{\infty} j^3 p_j < \infty.$$

If p is a law in $\mathcal{K}_\alpha$, let $f(z) = \sum_{j=0}^\infty p_j z^j$, $\mu = f'(1)$, $\nu = f''(1)/2$, and

$$c_k = \begin{cases} \nu \frac{(1-\mu^{-k})}{\mu - 1} & \text{if } \mu \neq 1 \\ \nu k & \text{if } \mu = 1., \end{cases}$$

For any sequence of laws $p^k \in \mathcal{K}_\alpha$ with $\mu_k \to 1$ as $k \to \infty$,

(a) $c_k P(X_k^{(k)} > 0) \to 1$,

(b) $P(X_k^{(k)}/\mu_k^k c_k \le x | X_k^{(k)} > 0) \to 1 - e^{-x}$.

Here $X^{(k)}$ is the Galton Watson branching process with offspring law p^k, $c_k = c_k(p^k)$ is the constant at time k associated with p^k, and μ_k is the mean of p^k.

The first bound we will need is:

Lemma 2.5. *Let $\{X^{(k)}\}$ be a sequence of Galton-Watson branching processes each starting with a single particle. Assume $X^{(k)}$ has a Binomial (N_k, q_k) offspring law where for some $C > 0$, and large enough k,*
(i) $q_k \le 1/2$, *(ii)* $1 \le N_k q_k \le 1 + C/k$.
Then

$$\limsup_{k \to \infty} k P(X_k^{(k)} > 0) \le K_{2.5}(C) = \frac{4C}{1 - e^{-C}}. \tag{2.9}$$

Proof. If Y_k has a binomial distribution with parameters (N_k, q_k), then an easy calculation shows that for large enough n

$$E(Y_k^3) \le 2\left(1+\frac{C}{k}\right)^3 \le 2(1+C)^3,$$

and the third moment hypothesis in Lemma 2.4 holds. Note also that the variance of Y_k is $N_k q_k(1-q_k) \ge 1/2$ for large enough k, and so the variance bound holds with $\alpha = 1/2$. Finally the parameter $\nu_k := (1/2)E(Y_k(Y_k-1))$ in the definition of c_k satisfies $\liminf_k \nu_k \ge 1/4$. A bit of calculus allows one to apply Lemma 2.4(a) to get the above upper bound. □

The other result we will need concerns the range of the branching random walk Z constructed in Proposition 2.3. Recall that $Z_0(x) = 1(x=0)$ and the binomial offspring distribution has parameters $V(R)$ and $p(R)$ satisfying (2.1). Let $\mathcal{R}_n$ denote the range of Z up until n, that is,

$$\mathcal{R}_n = \{x \in \mathbb{Z}_R^d : \sum_{j=1}^n Z_j(x) > 0\}$$

.

Lemma 2.6. *For any $c > 0$ and $K \in \mathbb{N}$, there is an $A(c,K)$, non-decreasing in c and K, such that for all $R \in \mathbb{N}$, $n \le cR^{d-1}$, and $r \le K\sqrt{n}$,*

$$\mathbb{P}(\mathcal{R}_n \cap ([-r,r]^d)^c \ne \emptyset) \le A(r+1)^{-2}.$$

Finer results are available for nearest neighbour branching random walks (see, e.g., Theorem 7 of Le Gall and Lin [6]) but as our branching random walk has some nonstandard features (such as a "long range" random walk component) we outline the proof in the Appendix. It is a straightforward modification of the results in Section 4 (particularly Lemma 4.9) of Dawson, Iscoe, Perkins [3] from their branching Brownian motion setting to the long-range branching random walk setting here.

3 Proof of extinction

Assume that $\eta_0 = \{0\}$ and $\rho_0 = \emptyset$ throughout this section. Recall that $p(R)$ is chosen so that (2.1) holds. Our aim in this section is to establish Theorem 1.2 by showing η becomes extinct a.s. for some positive value of θ. The first and main step is the following:

Proposition 3.1. *There is a $\theta_0 > 0$ so that for all $R \in \mathbb{N}$, and all $0 < \theta \le \theta_0$, there is a $k \in \{1, 2, \ldots, R^{d-1}+1\}$ so that $\mathbb{E}(|\eta_k|) \le 1-\theta$.*

Proof. To shorten notation we set

$$\varepsilon = \frac{\theta}{R^{d-1}}, \tag{3.1}$$

where $\theta \in (0,1]$. We see that (2.6) implies

$$\mathbb{E}\Big[|\eta_{k+1}| - |\eta_k|\Big] \le \mathbb{E}\left[\varepsilon|\eta_k| - \sum_{x\in\eta_k}\sum_{\substack{y\sim x\\ y\in\eta_k\cup\rho_k}} \frac{1}{V(R)}\right].$$

Since $|\rho_{n+1}| = \sum_{k=0}^n |\eta_k|$ and $|\eta_0| = 1$, we can sum the above equation from $k=0$ to n to get a telescoping sum

$$\mathbb{E}\Big[|\eta_{n+1}| - 1\Big] \le \mathbb{E}\Big[\varepsilon|\rho_{n+1}| - \sum_{k=0}^n \sum_{x\in\eta_k}\sum_{\substack{y\sim x\\ y\in\eta_k\cup\rho_k}} \frac{1}{V(R)}\Big]. \tag{3.2}$$

Let us define the *infection time* $\tau(x)$ of a site $x \in \mathbb{Z}^d$ by

$$\tau(x) = \begin{cases} n, & \text{if } x \in \eta_n, \\ \infty, & \text{if } x \text{ is never infected.} \end{cases}$$

The second sum in (3.2) contains only $x, y \in \rho_{n+1}$ so that $\tau(y) \le \tau(x)$. The x's can appear at most once in this summation but the y's may appear multiple times corresponding to distinct values of k. Nonetheless we have the inequality

$$\sum_{k=0}^{n} \sum_{x \in \eta_k} \sum_{\substack{y \sim x \\ y \in \rho_k \cup \eta_k}} \frac{1}{V(R)} \ge \sum_{x \in \rho_{n+1}} \sum_{\substack{y \in \rho_{n+1} \\ y \sim x \\ \tau(y) \le \tau(x)}} \frac{1}{V(R)} \,. \tag{3.3}$$

By symmetry, we know that the sum on the right restricted to $\tau(y) < \tau(x)$ equals the same sum but now over $\tau(y) > \tau(x)$, and therefore,

$$\sum_{x \in \rho_{n+1}} \sum_{\substack{y \in \rho_{n+1} \\ y \sim x \\ \tau(y) \le \tau(x)}} \frac{1}{V(R)} \ge \frac{1}{2} \sum_{x \in \rho_{n+1}} \sum_{\substack{y \sim x \\ y \in \rho_{n+1}}} \frac{1}{V(R)} \,.$$

This, together with (3.2) and (3.3), implies

$$\mathbb{E}\Big[|\eta_{n+1}| - 1\Big] \le \mathbb{E}\Big[\varepsilon |\rho_{n+1}| - \frac{1}{2} \sum_{x \in \rho_{n+1}} \sum_{\substack{y \sim x \\ y \in \rho_{n+1}}} \frac{1}{V(R)}\Big] \,. \tag{3.4}$$

We can now proceed in a manner similar to that of Bramson *et al.* [2], and decompose the interference terms into regions of high density and low density. Fix $K \in \mathbb{N}$, and consider the following sets:

$$I_n = [-K\sqrt{n}, K\sqrt{n}]^d \,, \tag{3.5}$$

$$A_n = \{x \in \rho_{n+1} : |\{y \sim x : y \in \rho_{n+1}\}| \ge 6\varepsilon V(R)\} \,. \tag{3.6}$$

Then, as

$$\begin{aligned} \sum_{x \in \rho_{n+1}} \sum_{\substack{y \sim x \\ y \in \rho_{n+1}}} \frac{1}{V(R)} &\ge \sum_{x \in A_n} \sum_{\substack{y \sim x \\ y \in \rho_{n+1}}} \frac{1}{V(R)} \\ &\ge |A_n| 6\varepsilon \,, \end{aligned}$$

we have from (3.4) and an elementary argument,

$$\mathbb{E}\Big[|\eta_{n+1}| - 1\Big] \le \varepsilon \mathbb{E}\Big[|\rho_{n+1} \cap I_n^c| + |(\rho_{n+1} \setminus A_n) \cap I_n| - 2|A_n \cap I_n|\Big]. \tag{3.7}$$

For the remainder of this section, we will work on each of the terms in (3.7) to show that the right-hand side of (3.7) is negative some n for small enough θ. Henceforth we assume that

$$n \le 4R^{d-1}. \tag{3.8}$$

To prove that the term $|\rho_{n+1} \cap I_n^c|$ in (3.7) is small for large K, we will compare the range of the epidemic with the range of the branching random walk Z_n in Proposition 2.3. Recall that Z_n is constructed on the same probability space as the epidemic such that $|\eta_k \cap A| \le Z_k(A)$ for any set $A \subset \mathbb{Z}^d/R$. A routine calculation for the branching random

walk shows that if $X_k = X_k^R$ is a simple random walk, starting at the origin, taking steps uniformly in $\{x : x \in \mathbb{Z}_R^d, 0 < \|x\|_\infty \le 1\}$, we have

$$\mathbb{E}(Z_k(A)) = \left(1 + \frac{\theta}{R^{d-1}}\right)^k \mathbb{P}(X_k \in A) \, . \tag{3.9}$$

(Recall that the offspring steps at distinct times are independent in the BRW.) Therefore, we have

$$\begin{aligned}
\mathbb{E}(|\rho_{n+1} \cap I_n^c|) &= \sum_{k=0}^{n} \mathbb{E}(|\eta_k \cap I_n^c|) \\
&\le \sum_{k=0}^{n} \mathbb{E}(Z_k(I_n^c)|) \\
&= \sum_{k=0}^{n} \left(1 + \frac{\theta}{R^{d-1}}\right)^k \mathbb{P}(X_k \in I_n^c) \\
&\le e^4 \sum_{k=1}^{n} \mathbb{P}(X_k \in I_n^c) \, ,
\end{aligned} \tag{3.10}$$

the last by our choice of n in (3.8) and the fact that $X_0 = 0$ so that we can drop $k = 0$. We will get the desired bound, independent of R, by an application of the Azuma–Hoeffding inequality (see, for example Theorem 2 of Hoeffding [8] and the comments at the end of Section 2 of that reference).

Lemma 3.2. *(Azuma-Hoeffding). Let M_n be a martingale with increments satisfying $|M_k - M_{k-1}| \le c_k$ a.s. for $k \in \mathbb{N}$. Then we have for any $N \in \mathbb{N}$ and $D > 0$,*

$$\mathbb{P}(|M_N - M_0| \ge D) \le 2 \exp\left(\frac{-D^2}{2\sum_{k=1}^{N} c_k^2}\right) \, .$$

Applying the above to the martingale $(X_n^i)_n$, the i-th component of the random walk (X_n), we get that for any $1 \le k \le n$,

$$\mathbb{P}(|X_k^i| \ge K\sqrt{n}) \le 2 \exp\left(\frac{-K^2 n}{2k}\right) \le 2 \exp(-K^2/2) \, .$$

Therefore, we have for $1 \le k \le n$,

$$\mathbb{P}(\|X_k\|_\infty \ge K\sqrt{n}) \le 2d \exp(-K^2/2), \tag{3.11}$$

which by (3.10) yields

$$\mathbb{E}(|\rho_{n+1} \cap I_n^c|) \le e^4 2dn \exp(-K^2/2) \, . \tag{3.12}$$

Now consider the second term in (3.7) and define $\zeta_n = (\rho_{n+1} \setminus A_n) \cap I_n$. Let $C(r)$ denote a closed cube of ℓ^∞ diameter 1 that is ℓ^∞ distance $r \ge 0$ from the origin in $\mathbb{Z}^d/R$. There are at most $6\varepsilon V(R)$ elements inside $C(r) \cap \zeta_n$, since otherwise an element in $C(r) \cap \zeta_n$ would have more than $6\varepsilon V(R)$ neighbours in ρ_{n+1}, contradicting the definition of ζ_n. Comparing the epidemic with the coupled branching random walk Z_n with range $\mathcal{R}_n$ up to time n, we can use Lemma 2.6 to conclude (recall (3.8) and $\varepsilon = \theta/R^{d-1}$) that for $r \le 2K\sqrt{n}$,

$$\begin{aligned}
\mathbb{E}(|C(r) \cap \zeta_n|) \le \mathbb{E}(6\varepsilon V(R) 1(C(r) \cap \mathcal{R}_n \ne \emptyset)) &\le 6\varepsilon V(R) A(4, 2K)(r+1)^{-2} \\
&\le c_1(K)\theta R(r+1)^{-2} \, .
\end{aligned} \tag{3.13}$$

The number of such boxes $C(r)$ with "integer corners" at ℓ^∞ distance $r \in [m, m+1]$ ($m \in \mathbb{Z}_+$) from the origin is bounded by $C'(m+1)^{d-1}$. We can cover $\zeta_n \subset I_n$ by the collections of these boxes for $m = 0, \ldots, \lfloor K\sqrt{n} \rfloor$ and conclude that for a constant $c_2 = c_2(K)$,

$$\begin{aligned}\mathbb{E}(|\zeta_n|) &\le \sum_{m=0}^{\lfloor K\sqrt{n}\rfloor} C'(m+1)^{d-1} c_1(K)\theta R(m+1)^{-2} \\ &\le c_2(K)\theta R \sum_{m=0}^{\lfloor K\sqrt{n}\rfloor} (m+1)^{d-3}. \end{aligned} \tag{3.14}$$

We strengthen (3.8) and now assume

$$R^{d-1} \le n \le 4R^{d-1}. \tag{3.15}$$

Assume first that $d = 3$. Continuing from (3.14), we use $n^{-1/2} \le R^{-1}$ (from (3.15)) to see that for n as above,

$$\mathbb{E}(|\zeta_n|) \le c_2(K)(K\sqrt{n}+1)\theta R \le 2c_2(K)Knn^{-1/2}\theta R \le 2c_2(K)K\theta n\ . \tag{3.16}$$

Substitute (3.16) and (3.12) into (3.7) to get

$$\mathbb{E}\Big[|\eta_n| - 1\Big] \le \varepsilon n \left[2e^4 d\exp(-K^2/2) + 2c_2(K)K\theta - \frac{2}{n}\mathbb{E}(|A_n \cap I_n|)\right] \text{ for } R^2 \le n \le 4R^2\ . \tag{3.17}$$

Now set $n = R^2$ and suppose that $\mathbb{E}(|\eta_k|) \ge 0.9$ for $k = 1, \ldots, n$. Then, as $|\rho_{n+1}| = \sum_{k=0}^n |\eta_k|$, we have

$$\mathbb{E}(|\rho_{n+1}|) \ge 0.9n. \tag{3.18}$$

First choose K large enough so that $2e^4 d\exp(-K^2/2) < 0.01$, and then $\theta > 0$ small enough so that $2c_2(K)K\theta < 0.01$. This and the bounds (3.16) and (3.12) show that

$$\begin{aligned} 0.9n \le \mathbb{E}(|\rho_{n+1}|) &\le \mathbb{E}(|\rho_{n+1} \cap I_n^c| + |(\rho_{n+1} \setminus A_n) \cap I_n| + |A_n \cap I_n|) \\ &\le n[e^4 2de^{-K^2/2} + 2c_2(K)\theta K + \frac{1}{n}\mathbb{E}(|A_n \cap I_n|)] \\ &\le n \times .02 + \mathbb{E}(|A_n \cap I_n|), \end{aligned}$$

and so $\mathbb{E}(|A_n \cap I_n|) \ge .88n$. Therefore inserting the above and our choices of n, K and $\theta > 0$ into (3.17) we arrive at

$$\mathbb{E}(|\eta_n| - 1) \le \theta[.02 - 1.76] \le -\theta.$$

Recall we had assumed that $\mathbb{E}(|\eta_k|) \ge 0.9$ for $k = 1, \ldots, n = R^2$, and so we may conclude that in any case for small enough $\theta > 0$ as above, for all $R \in \mathbb{N}$,

$$\text{for some } 1 \le k \le R^2 + 1,\ \mathbb{E}(|\eta_k| - 1) \le -(.1 \wedge \theta) < 0. \tag{3.19}$$

This completes the proof for $d = 3$ as we may take $\theta \le .1$.

Assume next that $d = 2$. If one proceeds in the same manner as in the $d = 3$ case above, the sum in (3.14) will lead to an extra logarithmic factor in the lower bound of λ_c. In this case we will need to improve the bound (3.13) for smaller values of r by an appeal to the local central limit theorem.

Continue to assume (3.15), now with $d = 2$. Let $\{U_i^R : i \in \mathbb{N}\}$ be iid random vectors which are uniformly distributed over the $V(R)$ points in $\mathcal{N}(0)$. We can couple this sequence an iid sequence, $\{U_i : i \in \mathbb{N}\}$ of random vectors which are uniformly distributed

over $[-1,1]^2$ and such that $\|U_i^R - U_i\|_\infty \le R^{-1}$ (the precise assignment of regions of $[-1,1]^2$ to points in $\mathcal{N}(0)$ is not important). If $S_k = \sum_{i=1}^k U_i$ and $X_k^R = \sum_{i=1}^k U_i^R$, then by the triangle inequality,

$$\text{for } k \le n,\ \|S_k - X_k^R\|_\infty \le k/R \le n/R \le 4. \tag{3.20}$$

If $C(r)$ is as above and $\bar{C}(r)$ is a box with the same centre but with ℓ^∞ diameter 9, then by our comparison result with BRW and (3.9) we have

$$\begin{aligned}
\mathbb{E}(|\zeta_n \cap C(r)|) &\le \mathbb{E}\Big(\sum_{k=0}^n Z_k(C(r))\Big) \\
&\le e^{\theta n/R} \sum_{k=0}^n \mathbb{P}(X_k^R \in C(r)) \\
&\le e^{4\theta}\left[\sum_{k=1}^n \mathbb{P}(S_k k^{-1/2} \in k^{-1/2}\bar{C}(r)) + 1_{C(r)}(0)\right],
\end{aligned} \tag{3.21}$$

the last by (3.20) and (3.15). In the last line $k^{-1/2}\bar{C}(r) = \{k^{-1/2}x : x \in \bar{C}(r)\}$. Let q_k be the density of $S_k/\sqrt{k}$, $\sigma^2 = 1/3$ (the variance of the uniform law on $[-1,1]$) and $\|(x_1,x_2)\|_2^2 = x_1^2 + x_2^2$. The local central limit theorem (e.g. (19.26) in Bhattacharya and Rao [1]) implies

$$\lim_{n\to\infty} \sup_{x\in\mathbb{R}^d} (1 + \|x\|_\infty^2)|q_n(x) - (2\pi\sigma^2)^{-1}\exp(-\|x\|_2^2/(2\sigma^2))| = 0. \tag{3.22}$$

Integrate the above bound to conclude that for some C, whose exact value may change from line to line,

$$\begin{aligned}
\sum_{k=1}^n \mathbb{P}(S_k k^{-1/2} \in k^{-1/2}\bar{C}(r)) &\le C\sum_{k=1}^n \int_{\bar{C}(r)k^{-1/2}} (1+\|x\|_\infty^2)^{-1}\,dx + \int_{\bar{C}(r)k^{-1/2}} \exp(-\|x\|_2^2)\,dx \\
&\le C\sum_{k=1}^n (9k^{-1/2})^2\Big[\frac{1}{1+((r-4)^+)^2/k} + \exp(-((r-4)^+)^2/k)\Big] \\
&\le C\sum_{k=1}^n (k+r^2)^{-1} \\
&\le C\left[\log\Big(\frac{n+r^2}{1+r^2}\Big) + \frac{1}{1+r^2}\right] \\
&\le C\left(\log\Big(1+\frac{2n}{(r+1)^2}\Big) + \frac{2}{(r+1)^2}\right) \le C\log\Big(1+\frac{2n}{(1+r)^2}\Big),
\end{aligned}$$

the last by some calculus. Use this in (3.21) and recall that $\theta \le 1$ to conclude that

$$\mathbb{E}(|\zeta_n \cap C(r)|) \le C\log\Big(1+\frac{2n}{(1+r)^2}\Big). \tag{3.23}$$

We consider cubes of the form $C(r) = [j_1, j_1+1] \times [j_2, j_2+1]$ for $j = (j_1, j_2) \in \mathbb{Z}^2$ such that $C(r) \cap I_n \ne \emptyset$. Recalling that r is the $\ell^\infty-$distance of $C(r)$ from 0 we see that $r \le K\sqrt{n}$. If $m \in \{1, 2\ldots, \lceil\sqrt{n}K\rceil\}$, the number of such cubes $C(r)$ with $r \in [m-1, m]$ is bounded by $c_0 m$ for some c_0. As $m \le r+1$, (3.23) and (3.13) imply that

$$\mathbb{E}(|\zeta_n \cap C(r)|) \le \min\left(\frac{c_1(K)\theta R}{m^2}, C\log\Big(1+\frac{2n}{m^2}\Big)\right). \tag{3.24}$$

Let $\delta = \theta \log(1/\theta)$ where we now take $0 < \theta < e^{-1}$, and note that $\delta \le e^{-1}$. Set

$$M_1 = \{m \in \mathbb{N} : m \le \lceil \sqrt{n}K \rceil : nm^{-2} \le \delta^{-1}\} \text{ and}$$
$$M_2 = \{m \in \mathbb{N} : m \le \lceil \sqrt{n}K \rceil : nm^{-2} > \delta^{-1}\}.$$

The obvious covering argument and (3.24) show that

$$\begin{aligned}
\mathbb{E}(|\zeta_n|) &\le \sum_{m=1}^{\lceil \sqrt{n}K \rceil} c_0 m \min\left(\frac{c_1(K)\theta R}{m^2}, C \log\Big(1 + \frac{2n}{m^2}\Big) \right) \\
&\le c_3(K) \left[\sum_{m \in M_1} \theta R m^{-1} + \sum_{m \in M_2} m \log\Big(1 + \frac{2n}{m^2}\Big) \right] \\
&\le c_3(K) \left[\theta n (1 + \log((K+1)/\sqrt{\delta})) + n \sum_{m \in M_2} \frac{m}{\sqrt{n}} \log\Big(1 + \frac{2n}{m^2}\Big) \frac{1}{\sqrt{n}} \right] \text{ (recall } R \le n) \\
&\le c_3(K) n \left[\theta \log((K+1)/\sqrt{\delta})) + \int_0^{\sqrt{\delta}} u \log(1 + 2u^{-2})\, du \right].
\end{aligned}$$

The last line follows by a bit of calculus, and a bit more gives the bound

$$\begin{aligned}
\mathbb{E}(|\zeta_n|) &\le c_3(K) n \Big[\theta \log((K+1)/\sqrt{\delta})) + \delta \log(1/\delta) \Big] \\
&\le c_3(K) \theta n \Big[\log((K+1)/\sqrt{\delta})) + \log(1/\theta) \log(1/\delta) \Big] \\
&\le c_3(K) \theta n [\log((K+1)/\theta)]^2.
\end{aligned}$$

Use the above bound in place of (3.16), so that instead of (3.17) we get

$$E(|\eta_n| - 1) \le \varepsilon n [2e^4 \exp(-K^2/2) + C(K)\theta [\log((K+1)/\theta)]^2 - \frac{2}{n} \mathbb{E}(|A_n \cap I_n|)] \quad \text{for } R \le n \le 4R.$$

Since $\theta[\log((K+1)/\theta)]^2$ decreases to 0 as $\theta \downarrow 0$, we may now proceed just as for $d = 3$ to conclude (in place of (3.19)) that for small enough $\theta > 0$, for all $R \in \mathbb{N}$,

$$\text{for some } 1 \le k \le R+1,\ \mathbb{E}(|\eta_k| - 1) \le -(.1 \wedge \theta) < 0. \tag{3.25}$$

This completes the proof for $d = 2$. □

For $n \in \mathbb{Z}_+ \cup \{\infty\}$, let

$$L_n = \sum_{j=0}^{n} |\eta_j| = |\cup_{j=0}^{n} \eta_j|.$$

Clearly if $\rho_0 = \emptyset$, then $L_n = |\rho_{n+1}|$.

Proof of Theorem 1.2. Let $\theta = \theta_0 > 0$ and k be as in Proposition 3.1. Let $c_k = \mathbb{E}(L_k)$, where it is understood that $(\eta_0, \rho_0) = (\{0\}, \emptyset)$ under $\mathbb{P}$, as usual. By Proposition 2.3 and (2.1),

$$c_k \le \mathbb{E}\Big[\sum_{i=0}^{k} Z_i(1) \Big] \le \sum_{i=0}^{k} \Big(1 + \frac{\theta}{R^{d-1}}\Big)^i < \infty.$$

Let $\{\beta_n\}$ be a Galton–Watson branching process with offspring law $\mathbb{P}(|\eta_k| \in \cdot)$ and initial state $\beta_0 = 1$. We claim

$$\forall m \in \mathbb{Z}_+,\ \forall n \ge mk,\ \mathbb{E}(L_n) \le c_k \sum_{j=0}^{m-1} \mathbb{E}(\beta_j) + \mathbb{E}(\beta_m)\mathbb{E}(L_{n-km})\ . \tag{3.26}$$

First, we show that the claim would complete the proof. Let $n = (m+1)k$ in (3.26) and then let $m \to \infty$ to conclude that

$$\mathbb{E}(L_\infty) \le c_k \sum_{j=0}^{\infty} \mathbb{E}(\beta_j) + \lim_{m\to\infty} \mathbb{E}(\beta_m) c_k = c_k \sum_{j=0}^{\infty} \mathbb{E}(|\eta_k|)^j < \infty.$$

This implies that $\cup_{j=0}^{\infty} \eta_j$ is finite a.s. and so $\eta_j = \emptyset$ for j large enough. Recalling that extinction of η starting at $\{0\}$ occurs iff the percolation cluster $\mathcal{C}_0$ is finite, we can conclude that $p_c(R)V(R) \ge 1 + \frac{\theta_0}{R^{d-1}}$, and the proof is complete.

We prove (3.26) by induction on m. The result is trivial for $m = 0$. Assume the result for m. Let $\{\eta_\cdot^i : i \in \mathbb{N}\}$ be independent and identically distributed copies of η under $\mathbb{P}$, independent of the branching process β. Let $L_n^i = \sum_{j\le n} |\eta_j^i|$, and $\mathcal{F}_n^i = \sigma(\eta_k^i, k \le n)$ be the generated filtrations. If $n \ge (m+1)k$, we can rewrite the last term in (3.26) as

$$\begin{aligned}
\sum_{i=1}^{\infty} \mathbb{P}(i \le \beta_m)\mathbb{E}(L_{n-mk}^i) &= \sum_{i=1}^{\infty} \mathbb{P}(i \le \beta_m)[\mathbb{E}(\mathbb{E}(L_{n-mk}^i - L_k^i | \mathcal{F}_k^i) + L_k^i)] \\
&= \mathbb{E}\left[\sum_{i=1}^{\beta_m} \mathbb{E}\left[L_{n-km}^i - L_k^i | \mathcal{F}_k^i\right]\right] + c_k\mathbb{E}(\beta_m) \\
&= \mathbb{E}\left[\sum_{i=1}^{\beta_m} \mathbb{E}_{\eta_k^i,\rho_k^i}\left[L_{n-(m+1)k}^i\right]\right] \\
&\quad + c_k\mathbb{E}(\beta_m) \qquad \text{(by the Markov property (1.7))} \\
&\le \mathbb{E}\left[\sum_{i=1}^{\beta_m} \mathbb{E}_{\eta_k^i,\emptyset}\left[L_{n-(m+1)k}^i\right]\right] \\
&\quad + c_k\mathbb{E}(\beta_m) \qquad \text{(by Lemma 1.1}(a)\text{)}.
\end{aligned}$$

Lemma 1.1(b) implies that

$$\mathbb{E}_{\eta_k^i,\emptyset}\left[L_{n-(m+1)k}^i\right] \le |\eta_k^i|\, \mathbb{E}(L_{n-(m+1)k}).$$

So substituting this into the previous display, we conclude that

$$\begin{aligned}
\mathbb{E}(\beta_m)\mathbb{E}(L_{n-km}) &\le \mathbb{E}\left[\sum_{i=1}^{\beta_m} |\eta_k^i|\right] \mathbb{E}(L_{n-(m+1)k}) + c_k\mathbb{E}(\beta_m) \\
&= \mathbb{E}(\beta_{m+1})\mathbb{E}(L_{n-(m+1)k}) + c_k\mathbb{E}(\beta_m).
\end{aligned}$$

Put this into (3.26) (the induction hypothesis), to see that (3.26) holds for $m+1$, completing the induction, and hence the proof of the Theorem. □

4 Appendix: proof of Lemma 2.6

Consider μ particles starting at the origin in $\mathbb{R}^d$. For each $i \in \mathbb{Z}^+$, on $[\frac{i}{\mu}, \frac{i+1}{\mu})$ each particle follows an independent d-dimensional Brownian motion, and at time $\frac{i+1}{\mu}$ the particle is replaced by 0 or 2 offspring at the parent's location, each with probability $\frac{1}{2}$. Let $\hat{X}_t^\mu$ be the random measure which puts mass μ^{-1} at the location of each particle at time t. (See Section 2 of [3] for a more detailed description of this branching Brownian motion.) Let $\hat{X}$ be the super-Brownian motion which is the unique in law solution of the following martingale problem:

$$\hat{X}_t(\phi) = \phi(0) + \hat{M}_t(\phi) + \int_0^t \hat{X}_s\left(\frac{\sigma^2}{2}\Delta\phi + \theta'\phi\right) ds\ ,\ \forall \phi \in C_K^2, \qquad (MP)_{\sigma^2,\theta'}$$

where $\hat{X}$ is a continuous $M_F(\mathbb{R}^d)$-valued process, and $\hat{M}(\phi)$ is a continuous martingale with $\langle \hat{M}(\phi)\rangle_t = \int_0^t \hat{X}_s(\phi^2)ds$.

It is well-known that $\hat{X}^\mu$ converges weakly to $\hat{X}$ (with $\sigma^2 = 1$ and $\theta' = 0$) on $D([0,\infty), M_F(\mathbb{R}^d))$ but we will need a result on the convergence of the ranges which does not follow from this alone. Let

$$\hat{\mathcal{R}}_t^\mu = \{x \in \mathbb{R}^d : \exists s \le t \text{ s. t. } \hat{X}_s^\mu(\{x\}) > 0\},$$

and let $\hat{\mathcal{R}}_t$ denote the closed support of $\int_0^t \hat{X}_s(\cdot)\, ds$.

Lemma 4.1. *If $\mathbb{P}_{\delta_0}$ denotes the law of $\hat{X}$ as above with $\sigma^2 = 1$ and $\theta' = 0$, then*

$$\limsup_{\mu\to\infty} \mathbb{P}(\hat{\mathcal{R}}_1^\mu \cap ((-1,1)^d)^c \neq \emptyset) \le \mathbb{P}_{\delta_0}(\hat{\mathcal{R}}_1 \cap ((-1,1)^d)^c \neq \emptyset). \tag{4.1}$$

Proof. This is immediate from Lemma 4.9 and Theorem 4.7(a) of Dawson, Iscoe, and Perkins [3] and the Transfer Principle of nonstandard analysis (to translate into standard terms). More specifically the first two results imply that for μ infinite (fixed), if $x \in \hat{\mathcal{R}}_1^\mu \cap ((-1,1)^d)^c$, then there is a sequence x_n in $\hat{\mathcal{R}}_1$ converging to the standard part of x, $st(x)$. This shows the latter must be in $\hat{\mathcal{R}}_1$ as this set is closed. Since $st(x)$ is also in the complement of $(-1,1)^d$, we have shown that $\hat{\mathcal{R}}_1 \cap ((-1,1)^d)^c$ is non-empty. An immediate application of the Transfer Principle now gives the required result. □

Remark 4.2. Just using the above weak convergence and elementary properties of $\hat{X}$ (it never charges the boundary of $(-1,1)^d$) one can easily show that

$$\liminf_{\mu\to\infty} \mathbb{P}(\hat{\mathcal{R}}_1^\mu \cap ((-1,1)^d)^c \neq \emptyset) \ge \mathbb{P}_{\delta_0}(\hat{\mathcal{R}}_1 \cap ((-1,1)^d)^c \neq \emptyset),$$

but it is the upper bound that will be of interest.

Our immediate goal is to extend Lemma 4.1 from the above branching Brownian motion to the context of the BRW Z constructed in Section 2.2. We shall see that Lemma 2.6 then follows easily.

To more closely parallel the setting in [3] we modify the setup for the branching envelope Z in Section 2.2 while constructing a BRW $\tilde{Z}$ with the same branching dynamics as Z. Let $\tilde{I} = \cup_{n=0}^\infty \mathbb{N} \times \{1,\dots,V(R)\}^n$ and for $\beta, \beta' \in \tilde{I}$, we define $|\beta|$, $\beta|i$, $\pi\beta$ and $\beta < \beta'$ as for I in Section 2.2. Assume $\{\tilde{M}^\beta : \beta \in \tilde{I}\}$ are iid Binomial $(V(R), p)$ random variables, denoting the number of offspring of particle β, where we assume $pV(R) \ge 1$ and, as always, $R \in \mathbb{N}$. Fix an initial number of particles $\mu \in \mathbb{N}$. Write $\beta \approx n$ iff $|\beta| = n$, $\beta_0 \le \mu$, and $\beta_{i+1} \le \tilde{M}^{\beta|i}$, for all $0 \le i < n$, meaning that β labels a particle which is alive in the nth generation. Next, let $(d^{\beta\vee i}, i \le V(R))_{\beta\in\tilde{I}}$ be a collection of iid random vectors, each uniformly distributed over $\mathcal{N}(0)^{(V(R))} = \{(e_1,\dots,e_{V(R)}) : \{e_i\} \text{ all distinct}\}$. For each β these are the displacements of the *potential* children from the parent β. Therefore the historical path followed by the ancestors of a particle $\beta \in \tilde{I}$ is

$$\tilde{Y}_t^{\beta,\mu} = \tilde{Y}_t^\beta = \sum_{i=1}^{|\beta|} 1(i \le \lfloor \mu t \rfloor) d^{\beta|i},$$

and its current location is

$$\tilde{Y}^\beta = \sum_{i=1}^{|\beta|} d^{\beta|i} \in \mathbb{Z}^d/R \quad (\text{so if } |\beta| = 0, \text{ then } \tilde{Y}^\beta = 0).$$

Note that for each β, $\tilde{Y}_\cdot^\beta$ is a random walk which jumps at times i/μ for $i \le |\beta|$, and whose step distribution is uniform over $\mathcal{N}(0)$. Let $\tilde{\mathcal{F}}_n = \sigma(\tilde{M}^\beta : |\beta| < n) \vee \sigma(d^\beta : 1 \le |\beta| \le n)$.

Note for each fixed $|\beta| = n$, the event $\{\beta \approx n\}$ is in $\tilde{\mathcal{F}}_n$ and $\tilde{Y}^\beta$ is $\tilde{\mathcal{F}}_n$-measurable. Therefore

$$\tilde{Z}_n = \sum_{\beta \approx n} \delta_{\tilde{Y}^\beta} \text{ is an } \tilde{\mathcal{F}}_n\text{-measurable random measure.}$$

Conditional on $\tilde{\mathcal{F}}_n$, $(\tilde{M}^\beta : \beta \approx n\}$ are iid binomial $(V(R), p)$ random variables, and conditional on $\tilde{\mathcal{F}}_n \vee \sigma(\tilde{M}^\beta : \beta \approx n)$, $(\tilde{Y}^{\beta\vee i} - \tilde{Y}^\beta : i \le \tilde{M}^\beta)_{\beta\approx n} = (d^{\beta\vee i} : i \le \tilde{M}^\beta)_{\beta\approx n}$ are independent random vectors which for each $\beta \approx n$ are uniformly distributed over $\mathcal{N}(0)^{(\tilde{M}^\beta)}$. This shows $\tilde{Z}$ is a BRW with the same offspring law as that of the branching envelope Z, and hence:

$$\begin{gathered}\text{if } \mu = 1, \text{ the laws of } \tilde{Z} \text{ and } Z \text{ (from Proposition 2.3) are identical,} \\ \text{and in general } \tilde{Z} \text{ is equal in law to a sum of } \mu \text{ iid copies of } Z.\end{gathered} \tag{4.2}$$

Consider the rescaled random measures given by

$$\tilde{X}^\mu_t(A) = \frac{1}{\mu}\tilde{Z}_{\lfloor \mu t\rfloor}(\sqrt{\mu}A) = \Big(\frac{1}{\mu}\sum_{\beta\approx\lfloor\mu t\rfloor} \delta_{\tilde{Y}^\beta/\sqrt{\mu}}\Big)(A).$$

Proposition 4.3. *If $\mu_n \to \infty$ and we choose $R_n \to \infty$ and p_n so that $p_nV(R_n) \ge 1$ and $\lim_n \mu_n(p_nV(R_n) - 1) = \theta' \ge 0$, then $\tilde{X}^{\mu_n} \Rightarrow \hat{X}$ in $D([0,\infty), M_F(\mathbb{R}^d))$, where $\hat{X}$ is the super-Brownian motion satisfying $(MP)_{\sigma^2,\theta'}$, with $\sigma^2 = 1/3$.*

This is a minor modification of the classical convergence theorem and may be proved by making minor changes in the proof, for example, in Chapter II of [14]. The value $\sigma^2 = 1/3$ arises as the variance of the marginals of the uniform distributions over $[-1,1]^d$ and the drift θ' arises since the mean number of offspring is $p_nV(R_n) \sim 1 + \frac{\theta'}{\mu_n}$. Note there is dependence between particle steps only if the particles are siblings and even here the steps are uncorrelated. This leads only to very minor alterations to the usual proof in the setting of completely independent displacements.

We assume in the rest of this section that (μ_n, p_n, R_n), $\tilde{X}^{\mu_n}$, and $\hat{X}$ are as in Proposition 4.3, $\mathbb{P}_{\delta_0}$ is the law of $\hat{X}$ and $\hat{R}_t$ is the closed support of $\int_0^t \hat{X}_s\,ds$. Let

$$\tilde{\mathcal{R}}^{\mu_n}_t = \{x \in \mathbb{R}^d : \exists s \le t \text{ s. t. } \tilde{X}^{\mu_n}_s(\{x\}) > 0\}.$$

Here is the version of Lemma 4.1 we will need.

Lemma 4.4.

$$\limsup_{n\to\infty} \mathbb{P}(\tilde{\mathcal{R}}^{\mu_n}_1 \cap ((-1,1)^d)^c \ne \emptyset) \le \mathbb{P}_{\delta_0}(\hat{\mathcal{R}}_1 \cap ((-1,1)^d)^c \ne \emptyset). \tag{4.3}$$

The result will follow just as in the proof of Lemma 4.1, once the analogues of Lemma 4.9 and Theorem 4.7 of [3] are established. The analogue of Theorem 4.7 will be immediate from the following uniform modulus of continuity for the historical paths of all particles in the BRW (just as Theorem 4.7 of [3] follows from Theorem 4.5 of that reference).

Lemma 4.5. *For each $L \in \mathbb{N}$ there are positive constants $c_i(L)$, $i = 1,2,3$ and non-negative random variables, $\delta(L,\mu_n)$, such that for all n,*

$$\mathbb{P}(\delta(L,\mu_n) \le \rho) \le c_1(L)\rho^{c_2(L)} \text{ for } 0 \le \rho \le c_3(L),$$

and if $s,t \in [0,L]$ satisfy $\mu_n^{-1} \le t - s \le \delta(L,\mu_n)$, then for all $\beta \approx \lfloor \mu_n t\rfloor$,

$$|Y^{\beta,\mu_n}_t - Y^{\beta,\mu_n}_s| < (t-s)^{1/8}.$$

In Theorem 4.5 of [3] the analogous result is stated for the branching Brownian motion with $c\sqrt{(t-s)\log(1/(t-s))}$ ($c>2$) in place of $(t-s)^{1/8}$ but any modulus function will do for our purposes. The proof of the above lemma is very similar. In place of the Gaussian bounds for the Brownian paths one uses Lemma 3.2 as the coordinates of Y_t^β are martingales with bounded jumps. The above cruder modulus helps handle the very small values of $t-s$. Note that the restriction $t-s\geq 1/\mu_n$ is natural as the modulus is really only needed for $s,t\in\{i/\mu_n : i\in\mathbb{Z}_+\}$. The slight super-criticality of our BRW also leads to some minor changes including the time cut-off L and the use of the survival bound (2.9) for our BRW (see the calculation in (4.4) below). We omit the proof.

Proof of Lemma 4.4. As noted above, given the previous lemma, it suffices to establish the analogue of Lemma 4.9 in [3] for our slightly supercritical rescaled BRW's $\tilde{X}^{\mu_n}$. The proof for branching Brownian motion goes through unchanged using Lemma 4.5 once the analogue of Lemma 4.8 of [3] is established so we now consider this result. It is a nonstandard formulation of the fact that if $t>0$ and $t_n\to t$ are fixed, then the sets $S^{\mu_n}_{t_n}=\{x\in\mathbb{Z}^d/R_n : \hat{X}^{\mu_n}_{t_n}(\{x\})>0\}$ converge weakly to the closed support, $S(\hat{X}_t)$, of $\hat{X}_t$ w.r.t. the Hausdorff metric as $n\to\infty$. Fix $t_n\to t>0$ and assume without loss of generality that $\inf_n t_n>0$. If $\beta\in\tilde{I}$ and $0\leq\varepsilon\leq s$, let

$$I_n(s,\varepsilon)=\{\gamma\in\tilde{I}:\gamma\approx\lfloor(s-\varepsilon)\mu_n\rfloor,\ \exists\beta\approx\lfloor s\mu_n\rfloor\ s.t.\ \gamma<\beta\},$$

that is, $I_n(s,\varepsilon)$ is the set of individuals in population $\tilde{X}^{\mu_n}_{s-\varepsilon}$ which have descendants alive in $\tilde{X}^{\mu_n}_s$. In what follows we consider m large enough so that $2^{-m}\leq\inf_n t_n$. For $\gamma\in I_n(t_n,2^{-m})$, let $N_n(\gamma,t_n)=|\{\beta\approx\lfloor\mu_n t_n\rfloor:\beta>\gamma\}|$ be the number of descendants of γ alive in the population $X^{\mu_n}_{t_n}$. A branching process argument (with offspring law binomial $(V(R_n),p_n)$) using Lemma 2.4(b) , shows that for some $C>0$,

$$\limsup_{n\to\infty}\mathbb{P}(N_n(\gamma,t_n)\leq 8^{-m}\mu_n|\gamma\in I_n(t_n,2^{-m}))\leq 1-e^{-C4^{-m}}\leq C4^{-m}.$$

Therefore

$$\begin{aligned}
&\limsup_{n\to\infty}\mathbb{P}(N_n(\gamma,t_n)\leq 8^{-m}\mu_n \text{ for some } \gamma\in I_n(t_n,2^{-m}))\\
&\quad\leq\limsup_{n\to\infty}\sum_{\gamma_0\leq\mu_n,|\gamma|=\lfloor\mu_n(t_n-2^{-m})\rfloor}\mathbb{P}(N_n(\gamma,t_n)\leq 8^{-m}\mu_n|\gamma\in I_n(t_n,2^{-m}))\mathbb{P}(\gamma\in I_n(t_n,2^{-m}))\\
&\quad\leq C4^{-m}\limsup_{n\to\infty}\sum_{\gamma_0\leq\mu_n,|\gamma|=\lfloor\mu_n(t_n-2^{-m})\rfloor}\mathbb{P}(\gamma\approx\lfloor(t_n-2^{-m})\mu_n\rfloor)\mathbb{P}(\tilde{Z}_{\lfloor\mu_n 2^{-m}\rfloor}(\mathbb{R}^d)\\
&\qquad\qquad>0|\tilde{Z}_0=1_{\{0\}})\\
&\quad\leq C4^{-m}\limsup_{n\to\infty}2K_{2.5}(2\theta'2^{-m})2^m\mu_n^{-1}\mathbb{E}(|\{\gamma\in\tilde{I}:\gamma\approx\lfloor(t_n-2^{-m})\mu_n\rfloor\}|)\\
&\quad\leq C2^{-m}\limsup_{n\to\infty}(V(R_n)p_n)^{\lfloor(t_n-2^{-m})\mu_n\rfloor}\\
&\quad\leq Ce^{Ct}2^{-m},
\end{aligned}\tag{4.4}$$

where C may depend on θ'. In the last line we have used the growth condition on $(V(R_n),p_n)$, in the fourth inequality we have used the definition of $K_{2.5}$ from Lemma 2.5 to absorb it into C, and in the third inequality we have used the survival probability bound in (2.9) with $k=\lfloor\mu_n 2^{-m}\rfloor$, for m fixed. To check that the hypotheses of Lemma 2.5 are in force, note that for large enough n,

$$1\leq V(R_n)p(R_n)\leq 1+\frac{2\theta'}{\mu_n}\leq 1+\frac{2\theta'2^{-m}}{\lfloor\mu_n 2^{-m}\rfloor}.$$

(4.4) gives the inequality in the display just before (4.27) in [3] and the rest of the proof of Lemma 4.8 of [3] now proceeds as for branching Brownian motion in that reference, again using our modulus of continuity in Lemma 4.5. The idea is that the above bound shows that any point in the support of $X^{\mu_n}_{t_n}$ will have enough mass nearby from its ancestor at time $t_n - 2^{-m}$ for m large enough that it will be arbitrarily close to a point in the support of the limiting $\hat{X}_t$. □

Proof of Lemma 2.6. It clearly suffices to consider $r \in \mathbb{N}$. Let $T_0 = \min\{n : Z_n(\mathbb{R}^d) = 0\}$. For $n, r \in \mathbb{N}$ as in the statement of the lemma we have

$$r^2 \leq cK^2 R^{d-1}, \tag{4.5}$$

which implies that

$$1 \leq V(R)p(R) = 1 + \frac{\theta}{R^{d-1}} \leq 1 + \frac{\theta c K^2}{r^2}. \tag{4.6}$$

Clearly we have

$$\mathbb{P}(\mathcal{R}_n \cap ([-r,r]^d)^c \neq \emptyset) \leq \mathbb{P}(T_0 > r^2) + \mathbb{P}(\mathcal{R}_{r^2} \cap ([-r,r]^d)^c \neq \emptyset,\, T_0 \leq r^2).$$

By the extinction bound (2.9) we have for sufficiently large r, $\mathbb{P}(T_0 > r^2) \leq 2K_{2.5}(\theta c K^2) r^{-2}$ (if not, choose sequences $\{r_n\}$ and $\{R_n\}$ both going to ∞ so that we can contradict the conclusion of Lemma 2.5.) Therefore for all $r \in \mathbb{N}$ we have $\mathbb{P}(T_0 > r^2) \leq Br^{-2}$ for some $B = B(cK^2)$. Hence, it suffices to show

$$\mathbb{P}(\mathcal{R}_{r^2} \cap ([-r,r]^d)^c \neq \emptyset) \leq Ar^{-2} \text{ for } r \in \mathbb{N} \text{ as in (4.5) and } A = A(c,K).$$

Assume that this is not the case. Then there are sequences of natural numbers $r_n \to \infty$ and $R_n \to \infty$ such that

$$r_n^2 \leq cK^2 R_n^{d-1}, \tag{4.7}$$

and

$$\lim_{n\to\infty} r_n^2 \mathbb{P}(\mathcal{R}_{r_n} \cap ([-r_n, r_n]^d)^c \neq \emptyset) = \infty. \tag{4.8}$$

Recall that we have chosen $p(R_n)$ so that $V(R_n)p(R_n) = 1 + \theta/R_n^{d-1}$ for some $\theta > 0$. This and (4.7) show that $V(R_n)p(R_n) \leq 1 + \frac{\theta K^2 c}{r_n^2}$. The probability in (4.8) will only increase if we raise $p(R_n)$ to p_n so that $V(R_n)p_n = 1 + \frac{\theta K^2 c}{r_n^2}$ and so we may use this modified Bernoulli probability for which (4.8) holds, and if $\mu_n = r_n^2$, then

$$\mu_n(V(R_n)p_n - 1) = \theta K^2 c =: \theta' > 0. \tag{4.9}$$

Now consider $\tilde{X}^{\mu_n}$ as above, and recalling (4.2), we have

$$\mathbb{P}(\tilde{\mathcal{R}}_1^{\mu_n} \cap ((-1,1)^d)^c = \emptyset) = (1 - \mathbb{P}(\mathcal{R}_{r_n} \cap ((-r_n, r_n)^d)^c \neq \emptyset))^{r_n^2} \to 0 \text{ as } n \to \infty, \tag{4.10}$$

by (4.8). On the other hand if $B(0,1)$ is the Euclidean open unit ball, then by Lemma 4.4 (recall (4.9))

$$\begin{aligned}
\liminf_{n\to\infty} \mathbb{P}(\tilde{\mathcal{R}}_1^{\mu_n} \cap ((-1,1)^d)^c = \emptyset) &\geq \mathbb{P}_{\delta_0}(\hat{\mathcal{R}}_1 \cap ((-1,1)^d)^c = \emptyset) \\
&\geq \mathbb{P}_{\delta_0}(\hat{X}_s(\overline{B(0,1)}^c) = 0 \ \ \forall s \geq 0) \\
&= e^{-u(0)} > 0,
\end{aligned} \tag{4.11}$$

where u is the unique radial solution of $\Delta u = u^2$ on $B(0,1)$ and $u(x) \to \infty$ as $|x| \uparrow 1$ (Theorem 1 of [9]). Together (4.10) and (4.11) give us the contradiction which completes the proof. □

Acknowledgments. The second author thanks Xinghua Zheng for a number of helpful comments.

References

[1] R. Bhattacharya and R. Rao. *Normal Approximation and Asymptotic Expansions*. Society for Industrial and Applied Mathematics, 2010. MR-3396213

[2] M. Bramson, R. Durrett, and G. Swindle. Statistical mechanics of crabgrass. *Annals of Probability*, 17: 444–481, 1989. MR-MR0985373

[3] D. A. Dawson, I. Iscoe, and E. A. Perkins. Super-Brownian motion: path properties and hitting probabilities. *Probability Theory Related Fields*, 83(1–2): 135–205, 1989. MR-1012498

[4] R. Durrett and E. Perkins. Rescaled contact processes converge to super-Brownian motion for $d \geq 2$. *Prob. Theory Rel. Fields*, 114: 309–399, 1999. MR-1705115

[5] D. Finlay. Percolation and contact processes through invasion methods. *USRA Report, UBC*, 2013.

[6] J.F. Le Gall and S. Lin. The range of tree-indexed random walk in low dimensions. *Annals of Probabiity*, 43: 2701–2728, 2015. MR-3395472

[7] R. van der Hofstad and A. Sakai. Critical points for spread-out self-avoiding walk, percolation and the contact process above the the upper critical dimensions. *Probability Theory and Related Fields*, 132: 438–470, 2005. MR-2197108

[8] W. Hoeffding. Probability inequalities for sums of bounded random variables. *Journal of American Statistical Association*, 158: 13–30, 1963. MR-0144363

[9] I. Iscoe. On the supports of measure-valued critical branching brownian motion. *Annals of Probabiity*, 16: 200–221, 1988. MR-0920265

[10] S. Lalley, E. A. Perkins, and X. Zheng. A phase transition for measure-valued SIR epidemic processes. *Annals of Probability*, 42(1): 237–310, 2014. MR-3161486

[11] S. Lalley and X. Zheng. Spatial epidemics and local times for critical branching random walks in dimensions 2 and 3. *Probability Theory and Related Fields*, 148(3–4): 527–566, 2010. MR-2678898

[12] D. Mollison. Spatial contact models for ecological and epidemic spread. *J. Roy. Statist. Soc. B*, 39: 283–326, 1977. MR-0496851

[13] M. D. Penrose. On the spread-out limit for bond and continuum percolation. *Annals of Applied Probability*, 3: 253–276, 1993. MR-1202526

[14] E. Perkins. Dawson-Watanabe Superprocesses and Measure-valued Diffusions, in *Lectures on Probability and Statistics*, 1781, Springer, Berlin, 2002. MR-1915445