Probab. Theory Relat. Fields (2010) 148:527–566
DOI 10.1007/s00440-009-0239-1

Spatial epidemics and local times for critical branching random walks in dimensions 2 and 3

Steven P. Lalley · Xinghua Zheng

Received: 9 January 2009 / Revised: 18 June 2009 / Published online: 27 August 2009

Abstract The behavior at criticality of spatial SIR epidemic models in dimensions two and three is investigated. In these models, finite populations of size N are situated at the sites of the integer lattice, and infectious contacts are limited to individuals at the same or at neighboring sites. Susceptible individuals, once infected, remain contagious for one unit of time and then recover, after which they are immune to further infection. It is shown that the measure-valued processes associated with these epidemics, suitably scaled, converge, in the large-N limit, either to a standard Dawson–Watanabe process (super-Brownian motion) or to a Dawson–Watanabe process with location-dependent killing, depending on the size of the the initially infected set. A key element of the argument is a proof of Adler's 1993 conjecture that the local time processes associated with branching random walks converge to the local time density process associated with the limiting super-Brownian motion.

Keywords Spatial epidemic · Branching random walk · Dawson–Watanabe process · Local times · Critical scaling

Mathematics Subject Classification (2000) Primary 60H30 · Secondary 60K35

S. P. Lalley (✉)
Department of Statistics, The University of Chicago, Chicago, 60637 IL, USA
e-mail: lalley@galton.uchicago.edu

X. Zheng (✉)
Department of ISOM, Hong Kong University of Science and Technology,
Clear Water Bay, Kowloon, Hong Kong
e-mail: xhzheng@ust.hk

1 Introduction

1.1 Spatial susceptible/infected/recovered epidemics

Simple spatial models of epidemics are known to exhibit *critical thresholds* in one dimension. Roughly, when the density of the initially infected set exceeds a certain level, the epidemic evolves in a markedly different fashion than its *branching envelope*. See [16] for a precise statement, and [2,20], and [7] for analogous results in the simpler setting of mean-field models. The main purpose of this article is to show that spatial susceptible/infected/recovered epidemics (SIR) in dimensions two and three also exhibit critical thresholds.

The epidemic models studied here take place in populations of size N located at the sites of the integer lattice $\mathbb{Z}^d$ in d dimensions. Each of the N individuals at a site $x \in \mathbb{Z}^d$ may at any time be either susceptible, infected, or recovered. Infected individuals remain infected for one unit of time, and then recover, after which they are immune to further infection. The rules of infections are as follows: at each time $t = 0, 1, 2, \ldots,$ for each pair (i_x, s_y) of an infected individual located at x and a susceptible individual at y, i_x infects s_y with probability $p_N(x, y)$. We shall only consider the case where the transmission probabilities $p_N(x, y)$ are spatially homogeneous, nearest-neighbor, and symmetric, and scale with the village size N in such a way that the expected number of infections by a contagious individual in an otherwise healthy population is 1 (so that the epidemic is critical), that is,

Assumption 1 $p_N(x; y) = 1/[(2d+1)N]$ if $|y - x| \le 1$; and $= 0$ otherwise.

(Strictly speaking, the expected number of infections is $(N-1)/N$ since an infected particle cannot infect itself—we thank a referee for pointing this out. One can also define $p_N(x; y) = 1/[(2d+1)N - 1]$ to make the expectation exactly 1; this would not change our main result in Theorem 2.)

Our main result, Theorem 2, asserts that under suitable hypotheses on the initial configurations of infected individuals, the critical spatial SIR-d epidemic can be rescaled so as to converge to a Dawson–Watanabe measure-valued diffusion in both $d = 2$ and $d = 3$. Depending on the size of the initially infected set, the limiting Dawson–Watanabe process has either a positive killing rate or no killing at all. The analogous result for $d = 1$ was proved in [16], using the fact that one-dimensional super-Brownian motion (the Dawson–Watanabe process with no killing) has sample paths in the space of absolutely continuous measures. In higher dimensions this is no longer true, so a different strategy is needed.

1.2 Branching envelope of a spatial epidemic

The spatial SIR epidemic in d dimensions is naturally coupled with a nearest neighbor branching random walk on the integer lattice $\mathbb{Z}^d$; this branching random walk is often referred to as the *branching envelope* of the epidemic. Particles of this branching random walk represent *infection attempts* in the coupled epidemic, some of which may fail to be realized in the epidemic because the targets of the attempts are either recovered or are targets of other simultaneous infection attempts. The branching envelope evolves as follows: any particle located at site x at time t lives for one unit of time and then reproduces, placing random numbers ξ_y of offspring at the sites y such that $|y-x| \leq 1$. The random variables ξ_y are i.i.d., with Binomial$(N, 1/[(2d+1)N])$ distributions. Denote this reproduction rule by $\mathcal{R}_N$, and denote by $\mathcal{R}_\infty$ the corresponding offspring law in which the Binomial distribution is replaced by the Poisson distribution with mean $1/(2d+1)$. Since offspring are placed independently at each of the $(2d+1)$ nearest neighbors, the expected total number of offspring of a particle is 1, i.e., the branching random walk is critical. Moreover, under either reproduction rule $\mathcal{R}_N$ or $\mathcal{R}_\infty$, the displacement of each offspring from its parent is governed by the law of the simple nearest neighbor random walk[1] on $\mathbb{Z}^d$ (with holding probability $1/(2d+1)$). In particular, given that a particle at site x has k offspring, each of these offspring independently chooses a neighboring site y according to the law

$$P_1(x, y) = 1/(2d+1) \quad \text{for} \quad |y-x| \leq 1. \tag{1}$$

Note that the covariance matrix of the increment has determinant σ^{2d}, where σ^2 is the *variance parameter* of the jump distribution, defined by

$$\sigma^2 := \left(\frac{2}{2d+1}\right). \tag{2}$$

The spatial SIR-d epidemic can be constructed together with its branching envelope on a common probability space in such a way that the branching envelope dominates the epidemic, that is, for each time n and each site x the number of infected individuals at site x at time n is no larger than the number of particles in the branching envelope. Similar dominance arguments were also used in, e.g. [3] and [8] where contact processes were viewed as coalescing branching random walks. The construction, in brief, is as follows (see [16]). Particles of the branching random walk will be colored either *red* or *blue* according to whether or not they represent infections that actually take place, with red particles representing actual infections. Initially, all particles are red. At each time $t = 0, 1, 2, \ldots$, particles produce offspring at the same or neighboring sites according to the law $\mathcal{R}_N$ described above. Offspring of blue particles are always blue, but offspring of red particles may be either red or blue, with the choices made according to the following procedure. All offspring of red particles at a location y choose numbers $j \in [N] := \{1, 2, \ldots, N\}$ at random, independently of all other particles. If a particle chooses a number j that was previously chosen by a particle of an earlier generation at the same site y, then it is assigned color blue. If $k > 1$ offspring of red particles choose the same number j at the same time, and if j was not chosen in an earlier generation, then 1 of the particles is assigned color red, while the remaining $k-1$ are assigned color blue. Under this rule, the subpopulation of red particles evolves as an SIR-d epidemic.

[1] Throughout the paper, the term *simple random walk* will mean simple random walk with holding probability $1/(2d+1)$.

It is apparent that when the numbers of infected and recovered individuals at a site and its nearest neighbors are small compared to N, then blue particles will be produced only infrequently, and so the epidemic process will closely track its branching envelope. Only when the sizes of the recovered and infected sets reach certain critical thresholds will blue particles start to be produced in large numbers, at which point the epidemic will begin to diverge significantly from the branching envelope. Our main result, Theorem 2, implies that the critical threshold for the number of initially infected individuals is on the order $N^{2/(6-d)}$.

The SIR-d epidemic is related to its branching envelope in a second—and for our purposes more important—way. The law of the epidemic, as a probability measure on the space of possible population trajectories, is absolutely continuous relative to the law of its branching envelope. The likelihood ratio can be expressed as a product over time and space, with each site/neighbor/generation contributing a factor (see Sect. 3.3). Each such factor involves the *total occupation time* $R_n^N(x)$ of the site, that is, the sum of the number of particles at site x over all times prior to n. Thus, the asymptotic behavior of the occupation time statistics for branching random walks will play a central role in the analysis of the large-N behavior of the SIR-d epidemic.

1.3 Watanabe's Theorem

A fundamental theorem of [26] asserts that, under suitable rescaling (the Feller–Watanabe scaling) the measure-valued processes naturally associated with critical branching random walks converge to a limit, the standard Dawson–Watanabe process, also known as super-Brownian motion. Let $M_F(\mathbb{R}^d)$ denote the space of finite measures on $\mathbb{R}^d$ with the topology of weak convergence, $D([0,\infty); M_F(\mathbb{R}^d))$ be the Skorokhod space of càdlàg $M_F(\mathbb{R}^d)$-valued paths.

Definition 1 A standard Dawson–Watanabe process (also called super-Brownian motion) $(X_t : t \geq 0) \in D([0,\infty); M_F(\mathbb{R}^d))$ with diffusion coefficient σ^2 is an a.s.-continuous $M_F(\mathbb{R}^d)$-valued process, and can be characterized by the following martingale problem: for each test function $\phi \in C_c^\infty(\mathbb{R}^d)$,

$$M_t(\phi) := \langle X_t, \phi\rangle - \langle X_0, \phi\rangle - \int_0^t \langle X_s, \frac{\sigma^2}{2}\Delta\phi\rangle\, ds \tag{3}$$

is a martingale with quadratic variation process

$$[M(\phi)]_t = \int_0^t \langle X_s, \phi^2\rangle\, ds, \tag{4}$$

where $C_c^\infty(\mathbb{R}^d)$ is the space of smooth functions on $\mathbb{R}^d$ with compact support.

Definition 2 The Feller–Watanabe scaling operator $\mathcal{F}_k$ scales mass by $1/k$ and space by $1/\sqrt{k}$, that is, for any finite Borel measure $\mu(dx)$ on $\mathbb{R}^d$ and any test function $\psi \in C_c^\infty(\mathbb{R}^d)$,

$$\langle \psi, \mathcal{F}_k \mu \rangle = k^{-1} \int \psi(\sqrt{k}x)\mu(dx). \tag{5}$$

Watanabe's Theorem *For each $k = 1, 2, \dots$, let X_t^k be a nearest neighbor branching random walk with Poisson (1) offspring distribution and initial particle configuration X_0^k. (In particular, $X_t^k(x)$ denotes the number of particles at site $x \in \mathbb{Z}^d$ in generation $[t]$, and X_t^k is the corresponding counting measure.) If the initial mass distributions converge, after rescaling, as $k \to \infty$, that is, if*

$$\mathcal{F}_k X_0^k \Rightarrow \mu = X_0 \tag{6}$$

for some finite Borel measure μ on $\mathbb{R}^d$, then the rescaled measure-valued processes $(\mathcal{F}_k X^k)_{kt}$ converge in law as $k \to \infty$:

$$(\mathcal{F}_k X^k)_{kt} \Rightarrow X_t,$$

where $\Rightarrow$ represents the weak convergence relative to the Skorokhod topology on $D([0,\infty); M_F(\mathbb{R}^d))$. The limit is the standard Dawson–Watanabe process X_t (super-Brownian motion) with diffusion coefficient σ^2.

See [9] for an in-depth study of the Dawson–Watanabe process and a detailed proof of Watanabe's Theorem. Because the process X_t has continuous sample paths in the space of finite Borel measures, it follows routinely from Watanabe's theorem that the occupation measures for branching random walks converge to those of super-Brownian motion.

Lemma 1 *The following joint convergence holds*:

$$\left((\mathcal{F}_k X^k)_{kt}, \left(\int_0^t (\mathcal{F}_k X^k)_{ks}\, ds \right) \right) \Rightarrow \left(X_t, \int_0^t X_s\, ds \right), \tag{7}$$

where $\Rightarrow$ represents weak convergence relative to the Skorokhod topology on $D([0,\infty); M_F(\mathbb{R}^d))^2$.

Proof The Dawson–Watanabe process X_t has continuous sample paths in $D([0,\infty); M_F(\mathbb{R}^d))$, see, e.g., Proposition 2.15 in [9]. The functional $(X_t) \mapsto (\int_0^t X_s\, ds)$ is continuous relative to the Skorokhod topology on the subspace of continuous measure-valued processes, so the result follows from Watanabe's theorem and the continuous mapping principle. □

1.4 Local times of critical branching random walks

In dimension $d = 1$ the super-Brownian motion has sample paths in the space of absolutely continuous measures, that is, for each $t > 0$ the random measure X_t is absolutely continuous relative to Lebesgue measure ([15,22]). Moreover, the Radon–Nikodym derivative $X(t, x)$ is jointly continuous in t, x (for $t > 0$). It is shown in [16] that if a sequence of branching random walks satisfy the assumptions in Watanabe's Theorem, then the density processes associated with those branching random walks, under some smoothness assumptions on the initial configurations and after suitable scaling, converge to the density process of the limiting super-Brownian motion.

In dimensions $d \geq 2$ the measure X_t is a.s. singular [5]. Therefore, one cannot expect the convergence of density processes as in [16]. We shall prove, however, that the *occupation measures* of critical branching random walks have discrete densities that converge weakly—see Theorem 1. The limit process is the *local time density* process associated with the occupation measure

$$L_t := \int_0^t X_s \, ds$$

of the super-Brownian motion. In dimensions $d = 2, 3$, the random measure L_t is, for each $t > 0$, absolutely continuous, despite the fact that X_t is singular—see [24], [13] and [11]. Moreover, under suitable hypotheses on the initial condition X_0, the density process $L_t(x)$ is jointly continuous for $t > 0$ and $x \in \mathbb{R}^d$: This is the content of *Sugitani's theorem*. For the reader's convenience, we state Sugitani's Theorem precisely here. For $t > 0$ and $x \in \mathbb{R}^d$, set

$$q_t(x) = \int_0^t \phi_s(x) \, ds, \quad \text{where} \quad \phi_t(x) = \frac{e^{-|x|^2/2t}}{(2\pi t)^{d/2}}$$

is the usual heat kernel.

Sugitani's Theorem *Assume that $d = 2$ or 3, and that the initial configuration $\mu := X_0$ of the super-Brownian motion X_t is such that the convolution*

$$(q_t * \mu)(x) \text{ is jointly continuous in } t \geq 0 \text{ and } x \in \mathbb{R}^d. \tag{8}$$

Then for each $t \geq 0$, the occupation measure L_t is absolutely continuous, and there is a jointly continuous version $L_t(x)$ of the density process.

We call $(L_t(x))_{t \geq 0,\ x \in \mathbb{R}^d}$ the *local time density process* associated with the super-Brownian motion. In view of Watanabe's and Sugitani's theorems, it is natural to conjecture (see Remark 1) that the local time density processes of branching random walks, suitably scaled, converge to the local time density process of the super-Brownian motion. Theorem 1 asserts that this conjecture is true. Let X^k be a sequence of branching random walks on $\mathbb{Z}^d$. Write

$$
\begin{aligned}
X_i^k(x) &:= \#\text{ particles at } x \text{ at time } i, \quad \text{and} \\
R_n^k(x) &:= \sum_{0 \le i < n} X_i^k(x).
\end{aligned}
\tag{9}
$$

(We use the notation R_n^k instead of L_n^k because in the corresponding spatial epidemic model, the quantity $R_n^k(x)$ represents the number of *recovered* individuals at site x and time n.) Denote by

$$
\mathbb{P}^n = (P_n(x, y))_{x,y\in\mathbb{Z}^d} = (P_n(y - x))_{x,y\in\mathbb{Z}^d} \tag{10}
$$

the transition probability kernel of the simple random walk on $\mathbb{Z}^d$, that is, $\mathbb{P}^n = \mathbb{P} * \mathbb{P}^{n-1}$ is the nth convolution power of the one-step transition probability kernel given by (1). Let $P_0(x) = \delta_0(x)$, and let $G_n(x, y)$ be the associated Green's function:

$$
G_n(x) := \sum_{0 \le i < n} P_i(x)
$$

For any finite measure μ on $\mathbb{Z}^d$ with finite support, set

$$
(\mu G_n)(x) := (\mu * G_n)(x) = \sum_y \mu(y) G_n(x - y),
$$

and denote by $(\mu G_t)(y)$ the continuous extension to $[0, \infty) \times \mathbb{R}^d$ by linear interpolation.

Theorem 1 *Assume that $d = 2$ or $d = 3$. For each $k = 1, 2, \ldots$, let X_t^k be a branching random walk whose offspring distribution is Poisson with mean 1. Assume that the initial configurations $\mu^k := X_0^k$ satisfy hypothesis* (6) *of Watanabe's theorem, where the limit measure μ has compact support and satisfies the hypothesis* (8) *of Sugitani's theorem. Assume further that*

$$
\frac{\mu^k G_{kt}(\sqrt{k}x)}{k^{2-d/2}} \Longrightarrow [(q_{\sigma^2 t} * \mu)/\sigma^2](x), \tag{11}
$$

where $\Rightarrow$ indicates weak convergence in the topology of $D([0, \infty), C_b(\mathbb{R}^d))$. Then as $k \to \infty$,

$$
\frac{R_{kt}^k(\sqrt{k}x)}{k^{2-d/2}} \Longrightarrow L_t(x), \tag{12}
$$

where $L_t(x)$ is the local time density process associated with the super-Brownian motion with diffusion coefficient σ^2 started in the initial configuration $X_0 = \mu$.

Theorem 1 will be proved in Sect. 2.

Remark 1 The analogous result for critical branching Brownian motions was conjectured by Adler in [1], who proved the marginal convergence for any fixed t and x.

Remark 2 The assumption that the offspring distribution is Poisson with mean 1 can be relaxed. All that is really needed is that the offspring distribution has an exponentially decaying tail. See Remark 8 for more explanations.

Remark 3 The hypothesis (6) does not by itself imply (11), even if the limit measure μ satisfies the hypothesis (8) of Sugitani's theorem. Sufficient conditions for (11) are given in Proposition 1. In particular, in dimension 2, if (6) holds and the maximal number of particles on a single site is bounded in k, then (11) is satisfied.

Remark 4 Let X_t be super-Brownian motion in dimension $d = 2$. For each $t > 0$ the random measure X_t is singular, so by Fubini's theorem, for almost every point $x \in \mathbb{R}^2$ the set of times $t > 0$ such that x is a point of density of X_t has Lebesgue measure 0. Under hypothesis (11) we can make an analogous quantitative statement for branching random walk. For any fixed $x \in \mathbb{Z}^2$ and all $t > 0$,

$$E \sum_{m=1}^{[kt]} I_{\{X_m^k(x)>0\}} = O(k/\log k). \tag{13}$$

Proof By Proposition 35 in [18], there exists $\delta > 0$ such that for all k and m sufficiently large,

$$E\left[X_m^k(x)|X_m^k(x) > 0\right] \geq \delta \log m.$$

But hypothesis (11) implies that

$$ER_{kt}^k(x) = (\mu^k G_{kt})(x) = O(k),$$

and

$$ER_{kt}^k(x) = \sum_{m<kt} EX_m^k(x) = \sum_{m<kt} E\left[X_m^k(x)|X_m^k(x) > 0\right] \cdot P\left[X_m^k(x) > 0\right].$$

□

1.5 Scaling limit of spatial SIR epidemic

Before stating our result, we first recall the definition of Dawson–Watanabe processes with variable-rate killing. The Dawson–Watanabe process X_t with killing rate $\theta = \theta(x, t, \omega)$ [assumed to be progressively measurable and jointly continuous in (t, x)] and diffusion coefficient σ^2 can be characterized by a martingale problem ([6], Section

6.2). For any test function $\psi \in C_c^\infty(\mathbb{R}^d)$,

$$\langle X_t, \psi\rangle - \langle X_0, \psi\rangle - \frac{\sigma^2}{2}\int_0^t \langle X_s, \Delta\psi\rangle\, ds + \int_0^t \langle X_s, \theta(\cdot, s)\psi\rangle\, ds$$

is a martingale with the same quadratic variation as in (4). The Dawson–Watanabe process with killing rate 0 (which we sometimes refer to as the *standard Dawson–Watanabe process*) is super-Brownian motion. Existence and distributional uniqueness of Dawson–Watanabe processes in general is asserted in [6] and proved, in various cases, in [4] and [10]. It is also proved in these articles that the law of a Dawson–Watanabe process with killing on any finite time interval $[0, t]$ is absolutely continuous with respect to that of a standard Dawson–Watanabe process with the same diffusion coefficient, and that the likelihood ratio (Radon–Nikodym derivative) is [6]

$$\exp\left\{-\int_0^t \theta(s, x)\, dM(s, x) - \frac{1}{2}\int_0^t \langle X_s, \theta(s, \cdot)^2\rangle\, ds\right\}, \tag{14}$$

where $dM(s, x)$ is the *orthogonal martingale measure* attached to the standard Dawson–Watanabe process (see [25]). Absolute continuity implies that sample path properties are inherited. In particular, when $d = 2, 3$, if X_t is a Dawson–Watanabe process with killing, then a.s. its occupation time process L_t is absolutely continuous, with local time density $L_t(x)$ jointly continuous in x and t.

It is shown in [16] that for the SIR-1 epidemic in $\mathbb{Z}$ with village size N, the particle density processes, suitably rescaled, converge as $N \to \infty$ to the density process of a standard Dawson–Watanabe process or a Dawson–Watanabe process with location-dependent killing, depending on whether the total number of initial infections is below a critical threshold or not. In dimensions $d \geq 2$, one cannot expect such a result to hold, because the Dawson–Watanabe process is a.s. singular with respect to the Lebesgue measure and therefore has no associated density process. However, as *measure-valued* processes, the SIR-d ($d = 2, 3$) epidemics, under suitable scaling, do converge, as the next theorem asserts. For the SIR-d model with village size N, define

$$\begin{aligned} X_i^N(x) &:= \#\text{ infected particles at } x \text{ at time } i; \quad \text{and} \\ R_n^N(x) &:= \#\text{ recovered particles at } x \text{ at time } n = \sum_{i<n} X_i^N(x). \end{aligned} \tag{15}$$

Theorem 2 *Assume that $d = 2$ or 3, and suppose that for some $\alpha \leq 2/(6-d)$ the initial configurations $\mu^N := X_0^N$ are such that*

$$\mathcal{F}_{N^\alpha}\mu^N \Rightarrow \mu \text{ with compact support, } \quad \text{and} \tag{16}$$

$$((\mu^N G_{N^\alpha t})(\sqrt{N^\alpha}x))/N^{\alpha(2-d/2)} \Rightarrow [(q_{\sigma^2 t} * \mu)/\sigma^2](x) \in C(\mathbb{R}^{1+d}), \tag{17}$$

where the second convergence is in $D([0,\infty); C_b(\mathbb{R}^d))$. *Then*

$$(\mathcal{F}_{N^\alpha} X^N)_{N^\alpha t} \Longrightarrow X_t, \tag{18}$$

where the limit process X_t *is a Dawson–Watanabe process with initial configuration* $X_0 = \mu$, *diffusion coefficient* σ^2, *and killing rate* θ. *The killing rate depends on the value of* α *as follows*:

(i) *if* $\alpha < 2/(6-d)$, *then* $\theta \equiv 0$; *and*
(ii) *if* $\alpha = 2/(6-d)$, *then* $\theta = L_t(x)$,

where $L_t(x)$ *is the local time density of the process* X_t. *The convergence* $\Rightarrow$ *in* (18) *is weak convergence relative to the Skorokhod topology on* $D([0,\infty); M_F(\mathbb{R}^d))$.

Theorem 2 will be proved in Sect. 3.

Remark 5 Theorem 2 asserts that there is a *critical threshold* for the SIR-*d* epidemic in dimensions $d = 2, 3$. Below the threshold (when the sizes of the initially infected populations are $\ll N^{\alpha_*}$, where $\alpha_* = 2/(6-d)$ is the *critical exponent*) the effect of finite population size is not felt, and the epidemic looks much like its branching envelope. At the critical threshold, the finite-population effects begin to show, and the epidemic now looks like a branching random walk with location-dependent killing.

Remark 6 If one replaces $p_N(x; y)$ in Assumption 1 with $(1+\beta/N^{\alpha_*})/([2d+1]N) \times \mathbf{1}_{|\mathbf{x}-\mathbf{y}|\leq 1}$ for some fixed constant β, then one can generalize the arguments in this paper to show that the convergence still holds and the limit is a Dawson–Watanabe process with initial configuration $X_0 = \mu$, diffusion coefficient σ^2, and killing rate $\theta - \beta$. This process will be studied in [17] in which we show that there exist critical values $\beta_c = \beta_c(d) > 0$ such that if $\beta > \beta_c$ then the process survives with positive probability while if $\beta < \beta_c$ then it dies out a.s.

Remark 7 The critical behavior of the SIR-*d* epidemics in dimensions $d \geq 4$ will be studied in [21] in which we connect SIR-*d* epidemics with bond percolations on $[N]^{\mathbb{Z}^d}$ and give sharp asymptotics for the critical probabilities.

1.6 Notational conventions

Since the proof of Theorem 2 is based on likelihood ratio calculations, we shall, at the risk of minor confusion, use the same letters X and R, with subscripts and/or superscripts, to denote particle counts and occupation counts for both branching random walks and the SIR-*d* epidemic processes (see Eqs. 14 and 15) and for their continuous limits. Throughout the paper, we use the notation $f \asymp g$ to mean that the ratio f/g remains bounded away from 0 and ∞. For any $a, b \in \mathbb{R}$, $a \wedge b := \min(a, b)$ and $a \vee b := \max(a, b)$. Also, C, C_1, etc. denote generic constants whose values may change from line to line. The notation $\delta_x(y)$ is reserved for the Kronecker delta function. The notation $Y_n = o_P(f(n))$ means that $Y_n/f(n) \to 0$ in probability; and $Y_n = O_P(f(n))$ means that the sequence $|Y_n|/f(n)$ is tight. Finally, we use a "local scoping rule" for notation. Any notation introduced in a proof is local to the proof, unless otherwise indicated.

2 Local time for branching random walk in $d = 2, 3$

2.1 Estimates on transition probabilities

Recall that $\mathbb{P}^n = (P_n(x - y))$ is the n-step transition probability kernel for the simple random walk on $\mathbb{Z}^d$ [with holding parameter $1/(2d + 1)$]. For critical branching random walk, $P_n(x, y)$ is the expected number of particles at site y at time n given that the branching random walk is initiated by a single particle at site x. For this reason, sharp estimates on these transition probabilities will be of crucial importance in the proof of Theorem 1. We collect several useful estimates here. As the proofs are somewhat technical, we relegate them to the Appendix (Sect. 4). Write

$$\Phi_n(x, y) = \phi_n(x) + \phi_n(y) \quad \text{where}$$
$$\phi_n(x) = \frac{1}{(2\pi n)^{d/2}} \exp\left(-\frac{|x|^2}{2n}\right)$$

is the Gauss kernel in $\mathbb{R}^d$. The first two results relate transition probabilities to the Gauss kernel.

Lemma 2 *For all sufficiently small $\beta > 0$ there exists constant $C = C(\beta) > 0$ such that for all integers $m, n \geq 1$ and all $x \in \mathbb{Z}^d$,*

$$P_n(x) \leq C\phi_n(\beta x) \quad \text{and} \tag{19}$$
$$\sum_y P_m(y)\phi_n(\beta(x - y)) \leq C\phi_{m+n}(\beta x/2). \tag{20}$$

Furthermore, for each $A > 0$ and each $T > 0$ there exists $C = C(\beta, A, T) > 0$ such that for all k sufficiently large and all $|x| \leq A\sqrt{k}$,

$$\sum_{n \leq kT} \phi_n(\beta x) \leq C \sum_{n \leq kT} P_n(x). \tag{21}$$

Lemma 3 *For all sufficiently small $\beta > 0$ there exists constant $C = C(\beta) > 0$ such that for all integers $n \geq 1$ and all $x, y \in \mathbb{Z}^d$,*

$$|P_n(x) - P_n(y)| \leq C\left(\frac{|x - y|}{\sqrt{n}} \wedge 1\right) \cdot \Phi_n(\beta x, \beta y). \tag{22}$$

In particular, for all $\gamma \leq 1$,

$$|P_n(x) - P_n(y)| \leq C\left(\frac{|x - y|}{\sqrt{n}}\right)^{\gamma} \cdot \Phi_n(\beta x, \beta y). \tag{23}$$

Our arguments will also require the following estimates on the discretized Green kernel.

Lemma 4 *For each $\gamma \in (0, 2-d/2)$, $\beta > 0$, $n \in \mathbb{N}$, and $x, y \in \mathbb{Z}^d$, define*

$$F_n(x, y; \beta) = F_{n;\gamma}(x, y; \beta) = \sum_{l<n} \frac{1}{l^{\gamma/2}} \Phi_l(\beta x, \beta y). \tag{24}$$

Then there exists $C = C(\gamma, \beta) < \infty$ such that for all $n \in \mathbb{N}$ and all $x, y \in \mathbb{Z}^d$, the following inequalities hold:

$$(F_n(x, y; \beta))^2 \le C n^{2-(d+\gamma)/2} F_n(x, y; \beta), \tag{25}$$

and

$$\begin{aligned}\sum_{i<n} \sum_{z} P_i(z) \cdot (F_{n-i}(x-z, y-z; \beta))^2 &\le C n^{2-(d+2\gamma)/2} \sum_{l<n} \Phi_l(\beta x/2, \beta y/2) \\ &\le C n^{2-(d+\gamma)/2} F_n(x, y; \beta/2). \end{aligned} \tag{26}$$

(*Note that in the last term the β parameter is changed to $\beta/2$.*)

Lemma 5 *For each $\beta > 0$, $m, n \ge 1$, and $x \in \mathbb{Z}^d$, define*

$$J_{m,n}(x; \beta) = \sum_{m \le l < m+n} \phi_l(\beta x). \tag{27}$$

Then there exists $C = C(\beta) > 0$ such that for all $m, n \ge 1$, and all $x \in \mathbb{Z}^d$, the following inequalities hold:

$$\left(J_{m,n}(x; \beta)\right)^2 \le C n^{2-d/2} J_{m,n}(x; \beta), \tag{28}$$

$$\sum_{i<n} \sum_{z} P_i(z) \cdot \left(J_{m,n-i}(x-z; \beta)\right)^2 \le C n^{2-d/2} J_{m,n}(x; \beta/2), \tag{29}$$

and

$$\sum_{i<m} \sum_{z} P_i(z) \cdot \left(J_{m-i,n}(x-z; \beta)\right)^2 \le C n^{2-d/2} J_{m,n}(x; \beta/2). \tag{30}$$

2.2 *Proof of Theorem* 1

For notational ease, we omit the superscript k in the arguments below: thus, we write $X_n(x)$ instead of $X_n^k(x)$, and $R_n(x)$ instead of $R_n^k(x)$. To prove the theorem it suffices to prove that (i) the sequence of random processes $(R_{kt}(\sqrt{k}x)/k^{2-d/2})$ is tight in the space $D([0, \infty); C_b(\mathbb{R}^d))$; and (ii) that the only possible weak limit is the local time density process $L_t(x)$. The second of these is easy, given Lemma 1. This implies that for any test function $\psi \in C_c(\mathbb{R}^d)$,

$$\frac{1}{k^2}\sum_x R_{kt}(\sqrt{k}x)\psi(x) = \frac{1}{k}\sum_{i\le kt}\sum_x X_i(\sqrt{k}x)\psi(x)/k \Rightarrow \int_0^t X_s(\psi)\,ds,$$

where (X_t) is the super-Brownian motion started in configuration $X_0 = \mu$ with diffusion coefficient σ^2. Hence, any weak limit of the sequence $(R_{kt}(\sqrt{k}x)/k^{2-d/2})$ must be a density of the occupation measure for super-Brownian motion. By Sugitani's Theorem (see Sect. 1.4),

$$\int_0^t X_s(\psi)\,ds = \int_x L_t(x)\psi(x)\,dx.$$

It follows that $L_t(x)$ is the only possible weak limit.

Thus, to prove Theorem 1 it suffices to prove that the sequence $(R_{kt}(\sqrt{k}x)/k^{2-d/2})$ is tight in the space $D([0,\infty); C_b(\mathbb{R}^d))$. In view of hypothesis (8), it is enough to prove the tightness of the re-centered sequence

$$Y_k(t,x) := \Big(R_{kt}(\sqrt{k}x) - (\mu^k G_{kt})(\sqrt{k}x)\Big)/k^{2-d/2}. \tag{31}$$

This we will accomplish by verifying a form of the Kolmogorov–Centsov criterion (see, e.g., Theorem 2.8 and Problem 2.9 in [14]). According to this criterion, to prove tightness it suffices to prove that for each compact subset K of $[0,\infty)\times\mathbb{R}^d$ there exist constants $C<\infty$, $\alpha>0$, and $\delta>d+1$ such that for all k and all pairs (s,a), $(t,b)\in K$,

$$E|Y_k(t,a) - Y_k(t,b)|^\alpha \le C|a-b|^\delta \quad \text{and} \tag{32}$$

$$E|Y_k(t,a) - Y_k(s,a)|^\alpha \le C|t-s|^\delta. \tag{33}$$

The trick is to not work with moments directly, but instead, following the strategy of [24], to work with cumulants.

Lemma 6 [Lemma 3.1 in [24]] *Let X be a random variable with moment generating function $E\exp(\theta X) = \exp(\sum_{n=1}^\infty \theta^n a_n)$. If for some integer m there exist $r, b > 0$ such that*

$$|a_n| \le br^n, \quad \textit{for all } n \le 2m,$$

then there exists a constant C depending only on b and m (and independent of X) such that

$$EX^{2m} \le Cr^{2m}.$$

Note that $R_1(x) \equiv \mu(x)$, and they cancel out in the right hand side of (31), hence we need to only work with $\sum_{1\le i<n} X_i(x)$, which, for notational ease, will still be denoted by R_n, that is, throughout this proof we redefine

$$R_n(x) := \sum_{1\le i<n} X_i(x), \text{ for all } x \in \mathbb{Z}^d,$$

and, accordingly, $G_n(x) := \sum_{1\le i<n} P_i(x)$. This modification eliminates some annoying computations involving $P_0 = \phi_0 = \delta_0$.

A. *Cumulants.* By the additivity and spatial homogeneity of the branching random walk, for any $\psi \in C_c(\mathbb{Z}^d)$ and for each $n \ge 1$ there exists a function $\nu_n = \nu_n^\psi \in C_c(\mathbb{Z}^d)$ such that for any (nonrandom) initial configuration μ,

$$E^\mu \exp(\langle R_n, \psi\rangle) = \exp(\langle \mu, \nu_n\rangle).$$

Note that $\nu_1 = 0$ (since $R_1(x) = \sum_{1\le i<1} X_i(x) \equiv 0$). The assignment $\psi \mapsto \nu_n^\psi$ is monotone in ψ, but not in general linear. Setting $\mu = \delta_x$ and conditioning on the first generation, we obtain

$$\exp(\nu_{n+1}(x)) = \sum_j Q_j \left(\frac{1}{2d+1}\sum_e \exp(\psi(x+e) + \nu_n(x+e))\right)^j,$$

where $\{Q_j\}$ is the offspring distribution (in the case of interest, the Poisson distribution with mean 1) and the inner sum is over the $2d+1$ nearest neighbors e of the origin in $\mathbb{Z}^d$ (recall that the origin is included in this collection, since particles of the branching random walk can stay at the same sites as their parents). Observe that if the offspring distribution is Poisson(1), then

$$\nu_{n+1}(x) = \frac{1}{2d+1}\sum_e \exp(\psi(x+e) + \nu_n(x+e)) - 1. \tag{34}$$

Define the *cumulants* $\kappa_{h,n}(x) = \kappa_{h,n}^\psi(x)$ in the usual way:

$$E^\mu \exp(\theta\langle R_n, \psi\rangle) = \exp\left(\left\langle \mu, \sum_{h\ge 1} \theta^h \kappa_{h,n}\right\rangle\right), \quad \text{for all } \theta \in \mathbb{R}.$$

By the arguments of the preceding paragraph, $\kappa_{h,1} = 0$ for all $h \ge 1$, and by (34),

$$\sum_{h\ge 1} \theta^h \kappa_{h,n+1}(x) = \frac{1}{2d+1}\sum_e \left(\exp\left\{\sum_{h\ge 1}\theta^h \kappa_{h,n}(x+e) + \theta\psi(x+e)\right\} - 1\right).$$

Consequently,

$$\kappa_{h,n+1}(x) = \frac{1}{2d+1} \sum_e \sum_{m=1}^{h} \frac{1}{m!} \sum_{\mathcal{P}_m(h)} \prod_{i=1}^{m} \{\kappa_{h_i,n}(x+e) + \delta_1(h_i) \cdot \psi(x+e)\}, \quad (35)$$

where $\mathcal{P}_m(h)$ denotes the set of m-tuples $(h_1, h_2, \ldots, h_m)$ of positive integers whose sum is h, and $\delta_1(\cdot)$ is the Kronecker delta function. When $h \geq 2$, the $m = 1$ summand in (35) equals $1/(2d+1) \cdot \sum_e \kappa_{h,n}(x+e) = (P_1 * \kappa_{h,n})(x)$, hence,

$$\kappa_{h,n+1}(x) = (P_1 * \kappa_{h,n})(x) + (P_1 * \Xi_n)(x),$$

where

$$\Xi_n(x) = \Xi_n(x; h) := \sum_{m=2}^{h} \frac{1}{m!} \sum_{\mathcal{P}_m(h)} \prod_{i=1}^{m} \{\kappa_{h_i,n}(x) + \delta_1(h_i)\psi(x)\}. \quad (36)$$

Since $\kappa_{h,1} = 0$, by iteration we get that for all $h \geq 2$,

$$\kappa_{h,n}(x) = \sum_{1 \leq l < n} (P_l * \Xi_{n-l})(x). \quad (37)$$

B. *Case* $\psi = \delta_a - \delta_b$: Consider now the special case $\psi = \psi_{a,b} := \delta_a - \delta_b$ where $a, b \in \mathbb{Z}^d$ and δ_x is the Kronecker delta function. Fix $0 < \gamma < 2 - d/2$ small, and let

$$\eta = \eta(\gamma) = 2 - (d+\gamma)/2 > 0.$$

Recall that in (24) we defined $F_n(x, y; \beta)$ for $\beta > 0, n \in \mathbb{N}$ and $x, y \in \mathbb{Z}^d$ as

$$F_n(x, y; \beta) = F_{n;\gamma}(x, y; \beta) = \sum_{1 \leq l < n} \frac{1}{l^{\gamma/2}} \Phi_l(\beta x, \beta y).$$

Claim For all sufficiently small $\beta > 0$, for each $h \geq 1$ there exists $C_h = C(h, \beta, \gamma) < \infty$ such that for all $n \in \mathbb{N}$ and all $x \in \mathbb{Z}^d$,

$$|\kappa_{h,n}(x)| \leq C_h |a-b|^{h\gamma} n^{\eta(h-1)} F_n(a-x, b-x; 2^{-(h-1)}\beta); \quad (38)$$

moreover, for all $h \geq 2$, all $n \in \mathbb{N}$ and all $x \in \mathbb{Z}^d$,

$$|\kappa_{h,n}(x)| \leq C_h |a-b|^{h\gamma} n^{\eta(h-1)-\gamma/2} \sum_{1 \leq l < n} \Phi_l(2^{-(h-1)}\beta(a-x), 2^{-(h-1)}\beta(b-x)). \quad (39)$$

In fact, when $h = 1$,

$$\begin{aligned}|\kappa_{1,n}(x)| &= |E^{\delta_x}\langle R_n, \psi\rangle| = |G_n(a-x) - G_n(b-x)| \\ &\le C|a-b|^{\gamma} \sum_{l<n} \frac{1}{l^{\gamma/2}} \Phi_l(\beta(a-x), \beta(b-x)) \\ &= C|a-b|^{\gamma} F_n(a-x, b-x; \beta),\end{aligned}$$

where in the middle inequality we used assertion (23) of Lemma 3. Furthermore, since $\psi(x) \ne 0$ if and only if $x = a$ or b, in which case $|\psi(x)| = 1$ and $\inf_n C|a-b|^{\gamma} F_n(a-x, b-x; \beta) > 0$, we get that for all n and all x,

$$|\kappa_{1,n}(x) + \psi(x)| \le C|a-b|^{\gamma} F_n(a-x, b-x; \beta). \tag{40}$$

Now suppose that the claim holds for $1, \dots, h-1$, and we want to prove the claim for h. First note that in the definition (36) of $\Xi_n(x)$, only κ_{h_i} for $h_i < h$ are involved, hence by induction, (40), relation (25), and the monotonicity of F in β, we get that for all n and x,

$$\Xi_n(x) \le C|a-b|^{h\gamma} n^{\eta(h-2)} \left(F_n(a-x, b-x; 2^{-(h-2)}\beta)\right)^2.$$

The claims then follows from (37) and (26).

C. *Case* $\psi = \delta_a$: When $\psi = \delta_a$ for some $a \in \mathbb{Z}^d$, by (19),

$$\kappa_{1,n}(x) = E^{\delta_x}\langle R_n, \delta_a\rangle = G_n(a-x) \le C \sum_{1\le l<n} \phi_l(\beta(a-x)) = CJ_{1,n-1}(a-x; \beta),$$

where J is defined in (27). By a similar (in fact, slightly easier) argument as above and using relations (28) and (29) in Lemma 5, we get that for all $h \ge 1$, there exist $C_h = C(h, \beta, d) < \infty$ such that for all $n \in \mathbb{N}$ and all $x \in \mathbb{Z}^d$,

$$\kappa_{h,n}(x) \le C_h n^{(2-d/2)(h-1)} J_{1,n-1}(a-x; 2^{-(h-1)}\beta). \tag{41}$$

D. *Proof of* (32). Suppose the initial configurations μ^k satisfy the hypotheses of Theorem 1. For any $a, b \in \mathbb{R}^d$, we want to estimate $E^{\mu^k}|Y_k(t,a) - Y_k(t,b)|^h$. By Lemma 6, this can be done by setting $\psi = 1_{\sqrt{k}a} - 1_{\sqrt{k}b}$ and estimating $|\langle\mu^k, \kappa_{h,kt}\rangle|$. By (39), for all $h \ge 2$,

$$\begin{aligned}|\langle\mu^k, \kappa_{h,kt}\rangle| \le C_h|\sqrt{k}(a-b)|^{h\gamma} \cdot (kt)^{\eta(h-1)-\gamma/2} \cdot \\ \cdot \left\langle \mu^k, \sum_{l<kt} \Phi_l(2^{-(h-1)}\beta(\sqrt{k}a - \cdot), 2^{-(h-1)}\beta(\sqrt{k}b - \cdot)) \right\rangle.\end{aligned}$$

By assumption (6), we can find an $A > 0$ such that $\mathrm{Supp}(\mu^k) \subseteq B(0, A\sqrt{k})$ for all k, where $B(0, r)$ represents the ball of radius r around the origin. By relation (21), for any $T > 0$, there exists $C = C(\beta, h, T)$ such that for all $t \leq T$,

$$\begin{aligned}\max_x \left\langle \mu^k, \sum_{l<kt} \phi_l(2^{-(h-1)}\beta(x-\cdot)) \right\rangle &= \max_{|x|\leq A\sqrt{k}} \sum_y \mu_k(y) \sum_{l<kt} \phi_l(2^{-(h-1)}\beta(x-y)) \\ &\leq C \sum_y \mu_k(y) \sum_{l<kt} P_l(x-y) \\ &\leq Ck^{2-d/2}, \end{aligned} \tag{42}$$

where the last inequality is due to the relative compactness of $(\mu^k G_{kt})(\sqrt{k}\cdot)/k^{2-d/2}$ assumed in (11). Hence when $h \geq 2$,

$$|\langle \mu^k, \kappa_{h,kt}\rangle| \leq C_h k^{h\gamma/2+\eta(h-1)-\gamma/2+2-d/2} \cdot t^{\eta(h-1)-\gamma/2} \cdot |a-b|^{h\gamma}.$$

Plugging in $\eta = 2 - (d+\gamma)/2$ gives us

$$|\langle \mu^k, \kappa_{h,kt}\rangle| \leq C_h k^{(2-d/2)h} \cdot t^{(2-(d+\gamma))(h-1)-\gamma/2} \cdot |a-b|^{h\gamma}, \quad \text{for all } h \geq 2.$$

Noting that $E^{\mu^k}(Y_k(t,a) - Y_k(t,b)) = 0$, by Lemma 6 we get that there exists $C'_h = C'(T,h) > 0$ such that for all $t \leq T$,

$$E^{\mu^k} |Y_k(t,a) - Y_k(t,b)|^{2h} \leq C'_h |a-b|^{2h\gamma}.$$

By choosing h large such that $2h\gamma > d+1$ we obtain (32).

E. *Proof of* (33). By the additivity and spatial homogeneity of the branching random walk, for any $\psi \in C_c(\mathbb{Z}^d)$ and for all $m, n \in \mathbb{N}$ there exists a function $\nu_{(m,n)} = \nu^{\psi}_{(m,n)} \in C_c(\mathbb{Z}^d)$ such that for any (nonrandom) initial configuration μ,

$$E^{\mu} \exp(\langle R_{m+n} - R_m, \psi\rangle) = \exp(\langle \mu, \nu_{(m,n)}\rangle).$$

Letting $\mu = \delta_x$ and conditioning on the first generation, we get

$$\begin{aligned}\exp(\nu_{(m+1,n)}(x)) &= E^{\delta_x} \exp(\langle R_{m+1+n} - R_{m+1}, \phi\rangle) \\ &= \sum_j Q_j \left(\frac{1}{2d+1} \sum_e \exp\left(\nu_{(m,n)}(x+e)\right) \right)^j, \end{aligned}$$

where $\mathcal{Q} = \{Q_j\}_{j\geq 0}$ denotes the offspring distribution. In case where the offspring distribution is Poisson(1), the equation above implies

$$\nu_{(m+1,n)}(x) = \frac{1}{2d+1} \sum_e \exp\left(\nu_{(m,n)}(x+e)\right) - 1. \tag{43}$$

Define the cumulants $\kappa_{h,(m,n)}$ by

$$E^{\mu}\exp(\theta\langle R_{m+n}-R_m,\psi\rangle)=\exp\left(\langle\mu,\sum_{h\geq 1}\theta^h\kappa_{h,(m,n)}\rangle\right),\quad \text{for all } \theta\in\mathbb{R}.$$

Then by (43),

$$\sum_h\theta^h\kappa_{h,(m+1,n)}(x)=\frac{1}{2d+1}\sum_e\exp\left[\sum_h\theta^h\kappa_{h,(m,n)}(x+e)\right]-1.$$

Therefore

$$\kappa_{h,(m+1,n)}(x)=\frac{1}{2d+1}\sum_e\sum_{i=1}^{h}\frac{1}{i!}\sum_{\mathcal{P}_i(h)}\prod_{j=1}^{i}\kappa_{h_i,(m,n)}(x+e), \tag{44}$$

where $\mathcal{P}_i(h)$ denotes the set of i-tuples $(h_1,h_2,\ldots,h_i)$ of positive integers whose sum is h. The $i=1$ summand in (44) equals $1/(2d+1)\cdot\sum_e\kappa_{h,(m,n)}(x+e)=(P_1*\kappa_{h,(m,n)})(x)$, hence when $h\geq 2$,

$$\kappa_{h,(m+1,n)}(x)=\left(P_1*\kappa_{h,(m,n)}\right)(x)+\left(P_1*\tilde{\Xi}_{m,n}\right)(x),$$

where

$$\tilde{\Xi}_{m,n}(x)=\tilde{\Xi}_{m,n}(x;h):=\sum_{i=2}^{h}\frac{1}{i!}\sum_{\mathcal{P}_i(h)}\prod_{j=1}^{i}\kappa_{h_i,(m,n)}(x). \tag{45}$$

By iteration we then get that for all $h\geq 2$,

$$\kappa_{h,(m,n)}(x)=\sum_{i<m}(P_i*\tilde{\Xi}_{m-i,n})(x)+(P_m*\kappa_{h,(0,n)})(x), \tag{46}$$

where $\kappa_{h,(0,n)}$ is just $\kappa_{h,n}$ studied in Part A.

For $\psi=1_a$ for $a\in\mathbb{Z}^d$, by (19),

$$\kappa_{1,(m,n)}(x)=E^{\delta_x}\langle R_{m+n}-R_m,\psi\rangle=\sum_{m\leq l<m+n}P_l(a-x)\leq CJ_{m,n}(a-x;\beta),$$

where J is defined in (27). Similarly as in proving the claim in Part B, using relations (28) and (30) in Lemma 5 and induction to bound the first term in the right hand side of (46), and using (41) and (20) to bound the second term, we get that for all sufficiently small $\beta>0$ and for all h, there exist $C_h=C(h,\beta)>0$ such that

$$|\kappa_{h,(m,n)}(x)|\leq C_hn^{(2-d/2)(h-1)}J_{m,n}(a-x,2^{-h}\beta).$$

We are ready to verify (33). Setting $\psi = 1_{\sqrt{k}a}$ and using (42), we get that for any $T > 0$, there exists $C = C(C_h, T)$ such that for all $t + s \leq T$,

$$|\langle \mu^k, \kappa_{h,(kt,ks)} \rangle| \leq C k^{(2-d/2)(h-1)} s^{(2-d/2)(h-1)} \cdot k^{2-d/2} = C k^{h(2-d/2)} \cdot s^{(2-d/2)(h-1)}.$$

By Lemma 6 we then get that there exists $C''_h = C''(T, h) > 0$ such that for all $t + s \leq T$,

$$E^{\mu^k} |Y_k(t+s, a) - Y_k(t, a)|^{2h} \leq C''_h \cdot s^{(2-d/2)(2h-1)}.$$

So by choosing h large such that $(2 - d/2)(2h - 1) > d + 1$ we obtain (33). $\square$

Remark 8 For general offspring distributions $\mathcal{Q} = \{Q_j\}$, let $f(x) = \log(\sum_j Q_j x^j)$ where $x \geq 0$. If the offspring distribution $\mathcal{Q}$ has an exponentially decaying tail, then $f(x)$ can be expanded around $x = 1$ as $f(x) = \sum_{\ell=1}^{\infty} f^{(\ell)}(1)(x-1)^{\ell}/\ell!$. Thus (34) turns into

$$\nu_{n+1}(x) = \sum_{\ell=1}^{\infty} f^{(\ell)}(1) \left(\frac{1}{2d+1} \sum_e \exp\left(\psi(x+e) + \nu_n(x+e)\right) - 1 \right)^{\ell} \Big/ \ell!,$$

and

$$\sum_h \theta^h \kappa_{h,n+1}(x) = \sum_{\ell=1}^{\infty} f^{(\ell)}(1)$$
$$\times \left(\frac{1}{2d+1} \sum_e \exp\left(\theta\psi(x+e) + \sum_h \theta^h \kappa_{h,n}(x+e) \right) - 1 \right)^{\ell} \Big/ \ell!.$$

This enables us to express $\kappa_{h,n+1}(x)$ in terms of $\psi(x+e)$ and $\kappa_{h,n}(x+e)$ similarly as in (35) and in (37) (note $f^{(1)}(1) = 1$ because $\mathcal{Q}$ has mean 1), and prove the Kolmogorov–Centsov criterion for the spatial variable. Similarly one can verify the Kolmogorov–Centsov criterion for the time variable.

2.3 Sufficient conditions for Assumption (11)

Now we state some conditions that imply (11) and are easier to check.

Proposition 1 *Let $d = 2$ or 3. Suppose that the initial configurations μ^k are such that $\mathcal{F}_k \mu^k \Rightarrow \mu$, and satisfy*

$$\lim_{t \to 0} \sup_k \max_x (\mu^k G_{kt})(x)/k^{2-d/2} = 0. \tag{47}$$

Then (11) holds. In particular, if any of the following assumptions is satisfied, then (11) holds.

(i) *In dimension* 2*, the maximal number of particles on a single site is bounded in* k*, i.e.,* $\sup_k \max_y \mu^k(y) < \infty$.

(ii) *In dimension* 3*, there exist* $C_1, C_2 > 0$ *such that*

$$C_2 := \sup_k \max_x \sum_{y \in B(x,\ C_1 k^{1/6})} \mu^k(y) < \infty, \tag{48}$$

where $B(x, r)$ *denotes the ball of radius* r *around* x *for any* x *and* $r \geq 0$*; that is, the number of particles in any ball of radius* $C_1 k^{1/6}$ *is bounded in* k.

(iii) *In dimension* 2*,* μ^k *is such that* $\mu^k(y)$ *is a decreasing function in* $|y|$*, and there exists* $\alpha \in (0, 2)$ *such that*

$$\mu^k(y) \leq C \left(\sqrt{k/(|y|^2 + 1)} \right)^{\alpha}, \quad \textit{for all } y, k. \tag{49}$$

Remark 9 This proposition is a natural analogue of Proposition 1 in [24].

To prove Proposition 1, we will need the following two lemmas.

Lemma 7 *For any function* $\psi \in C_c(\mathbb{R}^d)$ *and each integer* $k \geq 1$*, define*

$$\Psi_t^k(x) = \sum_{y \in \mathbb{Z}^d} \psi(y/\sqrt{k}) G_{kt}(\sqrt{k}x - y)/k, \quad \textit{for } \ x \in \mathbb{Z}^d/\sqrt{k} \textit{ and } t \in \mathbb{Z}/k,$$

and extend by linear interpolation elsewhere. Then

$$\lim_{k \to \infty} \Psi_t^k(x) = [(q_{\sigma^2 t} * \psi)/\sigma^2](x), \tag{50}$$

and the convergence is locally uniform in t *and* x.

Proof Pointwise convergence (50) follows from the local central limit theorem. To prove that the convergence is locally uniform, it suffices to show that the sequence of functions $\left(\Psi_t^k(x)\right)$ is relatively compact in $C(\mathbb{R}^{1+d})$. For this, we use the Ascoli–Arzela criterion. First, we show that the functions $\Psi_t^k(x)$ are uniformly bounded on any compact set in $\mathbb{R}^{1+d}$. Denote by M the maximum of $|\psi(x)|$. Then

$$\begin{aligned} \left| \Psi_t^k(x) \right| &\leq \sum_{y \in \mathbb{Z}^d} \left| \psi(y/\sqrt{k}) \right| \cdot G_{kt}(\sqrt{k}x - y)/k \\ &\leq M \sum_{y \in \mathbb{Z}^d} G_{kt}(\sqrt{k}x - y)/k \\ &\leq Mt. \end{aligned} \tag{51}$$

Next, we show that the $\left(\Psi_t^k(x)\right)$ are equi-continuous. Fix $\varepsilon > 0$, and set $\delta = \varepsilon/M$. By (51), $|\Psi_t^k| \leq \varepsilon$ for all $t \leq \delta$; thus,

$$\left| \Psi_t^k(x) - \Psi_s^k(y) \right| \leq 2\varepsilon, \quad \text{for all } x, y \in \mathbb{R}^d \text{ and } s, t \leq \delta.$$

On the other hand, by (22), for all $t \geq \delta$ and $x \neq y \in \mathbb{R}^2$,

$$\begin{aligned}|\Psi_t^k(x) - \Psi_t^k(y)| &\leq 2\varepsilon + C\frac{\sqrt{k}|x-y|}{k} \sum_{k\delta \leq n \leq kt} \frac{1}{\sqrt{n}} \sum_z |\psi(z/\sqrt{k})| \\ &\quad \cdot \Phi_n(\beta(\sqrt{k}x - z), \beta(\sqrt{k}y - z)) \\ &\leq 2\varepsilon + C\frac{\sqrt{k}|x-y|}{k} \sum_{k\delta \leq n \leq kt} \frac{1}{\sqrt{n}^{1+d}} \cdot \sqrt{k}^d \\ &\leq 2\varepsilon + C\delta^{-(d-1)/2} \cdot |x-y|.\end{aligned}$$

(In the second inequality we used the fact that $\sum_z |\psi(z/\sqrt{k})| \leq C\sqrt{k}^d$; this holds because ψ is bounded and has compact support.) Finally, for all x and all $\delta \leq s < t$,

$$|\Psi_t^k(x) - \Psi_s^k(x)| \leq M \sum_{ks \leq n \leq kt} \sum_z P_n(\sqrt{k}x - z)/k \leq M(t-s).$$

□

Lemma 8 *Suppose that f and g are two nonnegative functions on $\mathbb{Z}^d$, and f has compact support. Suppose further that both $f(x)$ and $g(x)$ are decreasing functions in $|x|$, then*

$$\sum_y f(y)g(x-y) \leq \sum_y f(y)g(y), \quad \textit{for all } x. \tag{52}$$

Proof Since $f(x)$ has compact support, we can enumerate its positive values, say $a_1 \geq \cdots \geq a_n > 0$. We can also enumerate the values of g, say $b_1 \geq \cdots b_n \geq \cdots$. Moreover, since f and g are both decreasing functions in $|x|$, the enumerations can be made in such a way that for each $i \in \{1, \ldots, n\}$, a_i and b_i are the values of f and g at a same site, respectively. To show (52), it then suffices to show that

$$\sup_{i_1,\ldots,i_n} \sum_{k=1}^n a_k b_{i_k} = \sum_{k=1}^n a_k b_k.$$

But this is easily seen to be true. □

We now prove Proposition 1.

Proof of Proposition 1 For any $\psi \in C_c(\mathbb{R}^d)$, by Lemma 7, $\Psi_t^k(x)$ converge to $[(q_{\sigma^2 t} * \psi)/\sigma^2](x)$ in the local uniform topology. Therefore,

$$\begin{aligned}\sum_x (\mu^k G_{kt})(\sqrt{k}x)/k^2 \cdot \psi(x) &= \frac{1}{k}\sum_y \mu^k(\sqrt{k}y) \cdot \Psi_t^k(y) \\ &\to \langle \mu, [(q_{\sigma^2 t} * \psi)/\sigma^2] \rangle = \langle [(q_{\sigma^2 t} * \mu)/\sigma^2], \psi \rangle,\end{aligned}$$

where the second convergence holds because $\mu^k(\sqrt{k}\cdot)/k \Rightarrow \mu \in M_F(\mathbb{R}^d)$. On the other hand, if we can show that $(\mu^k G_{kt})(\sqrt{k}x)/k^{2-d/2}$ is relatively compact in $C(\mathbb{R}^{1+d})$, then for any possible limit $F(t, x)$,

$$\sum_x (\mu^k G_{kt})(\sqrt{k}x)/k^2 \cdot \psi(x) = \sum_x (\mu^k G_{kt})(\sqrt{k}x)/k^{2-d/2} \cdot \psi(x) \cdot 1/\sqrt{k}^d$$
$$\to \int_x F(t, x)\psi(x)dx.$$

Hence, $\langle[(q_{\sigma^2 t} * \mu)/\sigma^2], \psi\rangle = \int_x F(t, x)\psi(x)dx$, which implies that (i) the measure $[(q_{\sigma^2 t} * \mu)/\sigma^2]$ has density, and (ii) $(\mu^k G_{kt})(\sqrt{k}x)/k^{2-d/2}$ converge to $[(q_{\sigma^2 t} * \mu)/\sigma^2](x)$ in $C(\mathbb{R}^{1+d})$.

Now we show that (47) implies that $(\mu^k G_{kt})(\sqrt{k}x)/k^{2-d/2}$ is relatively compact in $C(\mathbb{R}^{1+d})$, by verifying the Ascoli–Arzela criterion. We first show that they are uniformly bounded on any compact set in $\mathbb{R}^{1+d}$. In fact, by (47), there exists $\delta > 0$ such that

$$\sup_k \max_x (\mu^k G_{k\delta})(\sqrt{k}x)/k^{2-d/2} \le 1;$$

moreover, for all $t \ge \delta$ and all x,

$$\begin{aligned}(\mu^k G_{kt})(\sqrt{k}x)/k^{2-d/2} &\le 1 + \sum_{k\delta \le n \le kt} \sum_z \mu^k(z) P_n(\sqrt{k}x - z)/k^{2-d/2} \\ &\le 1 + \sum_{k\delta \le n \le kt} C/n^{d/2} \cdot k/k^{2-d/2} \\ &\le C = C(t),\end{aligned}$$

where in the second inequality we used the facts that there exists $C > 0$ such that for all n and all $x \in \mathbb{Z}^d$, $P_n(x) \le C/n^{d/2}$ (cf. [23], Proposition 6 on p. 72), and that the total number of particles $\sum_z \mu^k(z) = O(k)$.

Next we show $(\mu^k G_{kt})(\sqrt{k}x)/k^{2-d/2}$ are equi-continuous. In fact, for any $\varepsilon > 0$, by (47), there exists $\delta > 0$ such that

$$\sup_k \max_x (\mu^k G_{k\delta})(\sqrt{k}x)/k^{2-d/2} \le \varepsilon;$$

therefore, for all $s, t \le \delta$ and all x, y,

$$\sup_k |(\mu^k G_{ks})(\sqrt{k}x)/k^{2-d/2} - (\mu^k G_{kt})(\sqrt{k}y)/k^{2-d/2}| \le 2\varepsilon;$$

moreover, for all $t \geq \delta$ and all $x \neq y$, by (22),

$$\begin{aligned}
&\left|(\mu^k G_{kt})(\sqrt{k}x)/k^{2-d/2} - (\mu^k G_{kt})(\sqrt{k}y)/k^{2-d/2}\right| \\
&\leq 2\varepsilon + \frac{1}{k^{2-d/2}} \sum_{k\delta \leq n \leq kt} \sum_z \mu^k(z) \left|P_n(\sqrt{k}x - z) - P_n(\sqrt{k}y - z)\right| \\
&\leq 2\varepsilon + \frac{C\sqrt{k}|x-y|}{k^{2-d/2}} \sum_{k\delta \leq n \leq kt} \sum_z \mu^k(z) \frac{1}{\sqrt{n}} \Phi_n(\beta(\sqrt{k}x - z), \beta(\sqrt{k}y - z)) \\
&\leq 2\varepsilon + \frac{C\sqrt{k}|x-y|}{k^{2-d/2}} \sum_{k\delta \leq n \leq kt} \frac{1}{\sqrt{n}^{1+d}} \cdot k \\
&\leq 2\varepsilon + C\delta^{-(d-1)/2}|x-y|;
\end{aligned}$$

and for all x and all $\delta \leq s < t$,

$$\begin{aligned}
&\left|(\mu^k G_{kt})(\sqrt{k}x)/k^{2-d/2} - (\mu^k G_{ks})(\sqrt{k}x)/k^{2-d/2}\right| \\
&= \frac{1}{k^{2-d/2}} \sum_{ks \leq n \leq kt} \sum_z \mu^k(z) P_n(\sqrt{k}x - z) \\
&\leq \frac{C}{k^{2-d/2}} \sum_{ks \leq n \leq kt} k/n^{d/2} \\
&\leq \begin{cases} C\log(t/s) \leq C(t-s)/\delta, & \text{if } d = 2; \\ C(1/\sqrt{s} - 1/\sqrt{t}) \leq C(t-s)/\delta^{3/2}, & \text{if } d = 3. \end{cases}
\end{aligned}$$

We have therefore proved that (47) implies the relative compactness of $(\mu^k G_{kt})(\sqrt{k}x)/k^{2-d/2}$. Next we show that any of the conditions in (i–iii) implies (47). Note that all three conditions imply that $\sup_k \max_x \mu^k(x)/k^{2-d/2} \to 0$, hence we need to only work with $((\mu^k G_{kt})(x) - \mu^k(x))/k^{2-d/2} = \sum_{1\leq n<kt} \sum_y \mu^k(y) P_n(x-y)/k^{2-d/2}$.

(i) For all $x \in \mathbb{R}^2$ and all $t \geq 0$,

$$(\mu^k G_{kt})(x)/k = \frac{1}{k} \sum_{n<kt} \sum_z \mu^k(z) P_n(x-z) \leq \frac{C}{k} \sum_{n<kt} 1 \leq Ct,$$

therefore (47) holds.

(ii) In order to verify (47), by (19), it suffices to show that

$$\lim_{t\to 0} \sup_k \max_x \sum_{1\leq n<kt} \sum_y \mu^k(y) \phi_n(\beta(x-y))/\sqrt{k} = 0. \tag{53}$$

Claim There exists $C_3 > 0$ such that for all k, n and all pairs $x, y \in \mathbb{Z}^d$ with $|x-y| \geq C_1 k^{1/6}$,

$$\sqrt{k}\phi_n(\beta(x-y)) \leq C_3 \sum_{z \in B(y, C_1 k^{1/6})} \phi_n(\beta(x-z)).$$

In fact, the above inequality is equivalent to

$$\sqrt{k} \leq C_3 \sum_{|z-y| \leq C_1 k^{1/6}} \exp\left(\frac{\beta^2(|x-y|^2 - |x-z|^2)}{n}\right), \quad \text{for all } |x-y| \geq C_1 k^{1/6}.$$

But this holds trivially since when $x \notin B(y, C_1 k^{1/6})$, there is a positive proportion of integer points z in the ball $B(y, C_1 k^{1/6})$ such that $|x-y| \geq |x-z|$, and the proportion does not depend on k, x and y. Now let us estimate $\sum_{n \leq kt} \sum_y \mu^k(y)\phi_n(\beta(x-y))/\sqrt{k}$. For any fixed k and x, this sum can be written as the sum of the following two terms:

$$I := \sum_{n<kt} \sum_{|y-x| \leq C_1 k^{1/6}} \mu^k(y)\phi_n(\beta(x-y))/\sqrt{k},$$

and

$$II := \sum_{n<kt} \sum_{|y-x| > C_1 k^{1/6}} \mu^k(y)\phi_n(\beta(x-y))/\sqrt{k}.$$

As to term I, we have

$$I \leq \sum_{n<kt} \sum_{|y-x| \leq C_1 k^{1/6}} \mu^k(y) \cdot C/n^{3/2}/\sqrt{k} \leq \sum_{n \leq kt} C_2 \cdot C/n^{3/2}/\sqrt{k} \leq C/\sqrt{k},$$

where in the second inequality we used (48). And by the claim and (48),

$$\begin{aligned} II &\leq \sum_{n<kt} \sum_{|y-x| > C_1 k^{1/6}} \mu^k(y)/\sqrt{k} \cdot C_3 \sum_{z \in B(y, C_1 k^{1/6})} \phi_n(\beta(x-z))/\sqrt{k} \\ &\leq C \sum_{n<kt} \sum_{z} \phi_n(\beta(x-z)) \cdot \sum_{y \in B(z, C_1 k^{1/6})} \mu^k(y)/k \\ &\leq \sum_{n \leq kt} C/k \leq Ct. \end{aligned}$$

Therefore (53) holds.

(iii) In order to verify (47), by (19), it suffices to show that

$$\lim_{t \to 0} \sup_k \max_x \sum_{1 \leq n < kt} \sum_y \mu^k(y)\phi_n(\beta(x-y))/k = 0. \tag{54}$$

By assumption, $\mu^k(y)$ is a decreasing function of $|y|$; so is $\phi_n(\beta y)$. Therefore, by Lemma 8, the last term is bounded by $\sum_{n\le kt}\sum_y \mu^k(y)\phi_n(\beta y)/k$, which, by assumption (49), can be further bounded by

$$C\sum_{1\le n<kt}\frac{1}{n\cdot k}\sum_y\left(\sqrt{\frac{k}{|y|^2+1}}\right)^{\alpha}e^{\frac{-\beta^2|y|^2/k}{2n/k}}\le C\sum_{n<kt}\frac{1}{n}\int\limits_{x\in\mathbb{R}^2}|x|^{-\alpha}e^{\frac{-\beta^2|x|^2}{2n/k}}dx$$
$$\le C\sum_{n<kt}\frac{1}{n}(n/k)^{-\alpha/2+1}\int\limits_{x\in\mathbb{R}^2}|x|^{-\alpha}e^{-\beta^2|x|^2/2}dx$$
$$\le C\frac{1}{k}\cdot\sum_{n<kt}\left(\frac{n}{k}\right)^{-\alpha/2}$$
$$\le C\int\limits_0^t s^{-\alpha/2}\,ds=O(t^{1-\alpha/2}),$$

where the third inequality and the last equation hold because $\alpha<2$ by assumption. □

Remark 10 In dimension 2, if the assumption in (iii) is satisfied, then the radius of the support of μ^k will be of order $\sqrt{k}$. This is because we need $\sum_y \mu^k(y)=O(k)$, hence for some $C>0$,

$$\sum_{y\in\mathrm{Supp}(\mu^k)}\left(\sqrt{k/(|y|^2+1)}\right)^{\alpha}\ge Ck,\quad\text{i.e.,}\quad\sum_{y\in\mathrm{Supp}(\mu^k)}\left(1/\sqrt{|y|^2+1}\right)^{\alpha}\ge Ck^{1-\alpha/2}.$$

But for any r,

$$\sum_{|y|\le r}\left(1/\sqrt{|y|^2+1}\right)^{\alpha}=O\left(\int\limits_{|y|\le r}\left(1/\sqrt{|y|^2+1}\right)^{\alpha}dy\right)$$
$$=O\left(\int\limits_0^r\left(1/\sqrt{s^2+1}\right)^{\alpha}\cdot s\,ds\right)$$
$$=O(r^{2-\alpha}).$$

In order that $O(r^{2-\alpha})\ge Ck^{1-\alpha/2}$, we need $r=O(k^{1/2})$.

Remark 11 In dimension 3, if μ^k is such that $\mu^k(y)$ is a decreasing function in $|y|$, and there exists $\alpha\in(0,2)$ such that

$$\mu^k(y)\le C\left(\sqrt{k/(|y|^2+1)}\right)^{\alpha}/\sqrt{k},\ \text{ for all } y,k, \tag{55}$$

then using a similar proof as in (iii) we can show that (47) holds. But in fact, (55) can never be satisfied. The reason is that for any site y with at least one particle,

$$1 \le C\left(\sqrt{k/(|y|^2+1)}\right)^{\alpha} / \sqrt{k}, \quad \text{i.e.,} \quad |y| \le Ck^{1/2-1/(2\alpha)} = o(k^{1/4}).$$

On the other hand, we need $\sum_y \mu^k(y) = O(k)$, therefore for some $C > 0$,

$$\sum_{|y|=o(k^{1/4})} \left(\sqrt{k/(|y|^2+1)}\right)^{\alpha} \Big/ \sqrt{k} \ge Ck,$$

or

$$\sum_{|y|=o(k^{1/4})} \left(1/\sqrt{|y|^2+1}\right)^{\alpha} \ge Ck^{(3-\alpha)/2}. \tag{56}$$

However,

$$\begin{aligned}\sum_{|y|=o(k^{1/4})} \left(1/\sqrt{|y|^2+1}\right)^{\alpha} &= O\left(\int\limits_{|y|=o(k^{1/4})} \left(1/\sqrt{|y|^2+1}\right)^{\alpha} dy\right)\\ &= O\left(\int\limits_0^{o(k^{1/4})} \left(1/\sqrt{r^2+1}\right)^{\alpha} \cdot r^2\, dr\right)\\ &= o(k^{(3-\alpha)/4}) = o(k^{(3-\alpha)/2}),\end{aligned}$$

contradiction with (56).

3 Proof of Theorem 2: spatial epidemics in dimensions $d = 2, 3$

3.1 Strategy

The strategy is the same as that used by [16] in the one-dimensional case. Since the law of the SIR-d epidemic with village size N is absolutely continuous relative to that of its branching envelope, and since the branching envelopes converge weakly, after renormalization, to super-Brownian motion, it suffices to prove that the likelihood ratios converge weakly to the likelihood ratio (14) of the appropriate Dawson–Watanabe process relative to super-Brownian motion. The one- and higher-dimensional cases differ only in the behavior of the occupation statistics that enter into the likelihood ratios.

3.2 Modified SIR-*d* epidemic

As in the one-dimensional case, it is technically easier to work with the likelihood ratio for a modification of the SIR-*d* epidemic. Recall that (a) when an infected individual attempts to infect a recovered individual in an SIR epidemic, the attempt fails; and (b) when two (or more) infected individuals simultaneously attempt to infect the same susceptible individual, all but one of the attempts fail. Call an occurrence of type (a) an *errant attempt*, and an occurrence of type (b) a *collision*. In the modified SIR epidemic, collisions are not allowed, and there can be at most one errant attempt at any site/time. A formal specification of the modified SIR epidemic uses a variation of the standard coupling described in Sect. 1.2, as follows:

Modified Standard Coupling: particles are colored *red* or *blue*; red particles represent infected individuals in the modified SIR epidemic. Each particle produces a random number of offspring, according to the Poisson(1) distribution, which then randomly move to neighboring sites. Once situated, these offspring are assigned colors according to the following rules:

(A) Offspring of blue particles are blue; offspring of red particles may be either red or blue.
(B) At any site/time (x, t) there is at most one blue offspring of a red parent.
(C) Given that at site x and time t there are y offspring of red parents, the conditional probability $\kappa_N(y) = \kappa_{N,t,x}(y)$ that one of them is blue is

$$\kappa_N(y) = \{yR/N\} \wedge 1, \quad \text{where} \tag{57}$$

$$R = R_t^N(x) = \sum_{s<t} Y_s^N(x) \tag{58}$$

and $Y_t^N(x)$ is the number of *red* particles at site x in generation t. (Thus, $R = R_t^N(x)$ is the number of recovered individuals at site x at time t.) The red particle process is the *modified SIR epidemic*.

Proposition 2 *For each $N \geq 1$, versions of the SIR epidemic and the* modified *SIR epidemic can be constructed on a common probability space in such a way that* (*i*) *the initial configurations μ^N of infected individuals are identical, and satisfy the hypothesis* (16) *of Theorem* 2*; and* (*ii*) *the discrepancy $D_t(x)$ between the two processes at site x and time t* (*that is, the absolute difference in number of infected individuals*) *satisfies*

$$\max_t \sum_x D_t(x) = o_P(N^\alpha). \tag{59}$$

This implies that after the Feller–Watanabe scaling $(\mathcal{F}_{N^\alpha} X^N)_{N^\alpha t} = X^N_{N^\alpha t}(\sqrt{N^\alpha}\cdot)/N^\alpha$, the *SIR-d* epidemic and the modified *SIR-d* epidemic are indistinguishable. Consequently, to prove Theorem 2 it suffices to prove the corresponding result for the modified epidemic.

Proposition 2 is an easy consequence of Lemma 9.

Lemma 9 *For each pair $(n, x) \in \mathbb{N} \times \mathbb{Z}^d$, let $\Gamma_n^N(x)$ and $A_n^N(x)$ be the number of collisions and the number of errant infection attempts, respectively, at site x and time n in the SIR-d epidemic with village size N. Assume that the hypotheses (16–17) of Theorem 2 are satisfied, for some $\alpha \le 2/(6-d)$. Then*

$$\sum_n \sum_x \left\{ \Gamma_n^N(x) + \left(A_n^N(x) - 1 \right)_+ \right\} = o_P(N^\alpha). \tag{60}$$

The proof of this lemma makes use of the following result.

Lemma 10 [Proposition 28 in [18]] *Denote by $U_n(x)$ the number of particles at x at time n of a critical branching random walk started by one particle at the origin, then*

$$EU_n(x)^2 = P_n(x) + \sigma^2 \sum_{i=0}^{n-1} \sum_z P_i(z) P_{n-i}^2(x-z), \tag{61}$$

where σ^2 is the variance of the offspring distribution.

Proof of Lemma 9 Since the life length of the process is $O_p(N^\alpha)$, it suffices to show that for any $t > 0$,

$$\sum_{n \le N^\alpha t} \sum_x \left\{ \Gamma_n^N(x) + (A_n^N(x) - 1)_+ \right\} = o_p(N^\alpha).$$

Consider first the number $\Gamma_n^N(x)$ of collisions at site x and time n. For any susceptible individual η, a collision occurs at η if and only if there is some pair ξ, ζ of infected individuals at neighboring sites that simultaneously attempt to infect η. Therefore given the evolution up to time n, the conditional expectation of $\Gamma_{n+1}^N(x)$ is bounded by $C(\sum_e X_n^N(x+e))^2/N$. We want to show that

$$\sum_{n \le N^\alpha t} \sum_x (X_n^N(x))^2/N = o_P(N^\alpha). \tag{62}$$

By the dominance of the branching envelope over SIR epidemic, if we denote by $U_n(x)$ the number of particles at x at time n of a branching random walk with Poisson(1) offspring distribution started by one particle at the origin, and x_i $(i = 1, 2, \ldots)$ the positions of the initial particles of our epidemic model, then

$$E(X_n^N(x))^2 \le \sum_i EU_n(x - x_i)^2 + 2 \sum_{i \ne j} P_n(x - x_i) P_n(x - x_j),$$

which, in dimension 3, by Lemma 10, can be bounded by $C \sum_i P_n(x - x_i) + 2 \sum_{i \ne j} P_n(x - x_i) P_n(x - x_j)$. Therefore

$$\sum_{n \le N^\alpha t} \sum_x E(X_n^N(x))^2 \le C \sum_{n \le N^\alpha t} (CN^\alpha + CN^{2\alpha}/\sqrt{n}^3) = O(N^{2\alpha}), \tag{63}$$

which is $o(N \times N^{\alpha})$ since $\alpha \leq 2/3$. In dimension 2, again by Lemma 10, $EU_n(x)^2 \leq C(1 + \log n)P_n(x)$, therefore

$$\sum_{n \leq N^{\alpha}t} \sum_{x} E(X_n^N(x))^2 \leq C \sum_{n \leq N^{\alpha}t} (C(1 + \log n)N^{\alpha} + CN^{2\alpha}/n)$$
$$= O(N^{2\alpha} \log N), \tag{64}$$

which is also $o(N \times N^{\alpha})$ since $\alpha \leq 1/2$.

Now consider the number $A_n^N(x)$ of errant infection attempts at site x and time n. In order that there be more than one errant attempt, either (i) two or more infected individuals must simultaneously try to infect a recovered individual, or (ii) infected individuals must attempt to infect more than one recovered individual. The number of occurrences of type (i) during the course of the epidemic is $o_p(N^{\alpha})$, by the same argument that proved (62). Thus, it suffices to bound the number of errant attempts of type (ii). This is bounded by the number $B_n^N(x)$ of pairs ϱ, ϱ' of recovered individuals at site x and time n that are subject to simultaneous infection attempts. Clearly,

$$B_n^N(x) \leq \sum_{\xi,\varrho} \sum_{\zeta,\varrho'} Z_{\xi,\varrho} Z_{\zeta,\varrho'}$$

where the sums are over all pairs $((\xi, \varrho), (\zeta, \varrho'))$ in which ϱ, ϱ' are recovered individuals at site x and time n and ξ, ζ are infected individuals at neighboring sites, and $Z_{\xi,\varrho}$ and $Z_{\zeta,\varrho'}$ are independent Bernoulli $(1/((2d + 1)N))$. Hence,

$$E(B_{n+1}^N(x) \,|\, \mathcal{G}_n) \leq C \left(\sum_{e} X_n^N(x + e) \right)^2 (R_n^N(x)/N)^2.$$

By Theorem 1 and the dominance of the branching envelope over SIR epidemic, for all $\epsilon > 0$, there exists $C > 0$ such that with probability $\geq 1 - \epsilon$,

$$\max_{x} R_{N^{\alpha}t}^N(x) \leq CN^{\alpha(2-d/2)}.$$

Note further that

$$\sum_{n \leq N^{\alpha}t} \sum_{x} E\left(\sum_{e} X_n^N(x + e) \right)^2 \leq C \sum_{n \leq N^{\alpha}t} \sum_{x} E(X_n^N(x))^2,$$

which, by (64) and (63), is bounded by $CN^{2\alpha} \log N$ in dimension 2 and $CN^{2\alpha}$ in dimension 3. Therefore, by enlarging C if necessary we have that with probability $\geq 1 - 2\epsilon$, the following holds:

$$\sum_{n\le N^\alpha t}\sum_x\left(\sum_e X_n^N(x+e)\right)^2\left(R_n^N(x)/N\right)^2\le CN^{2\alpha}\log N\cdot N^{2\alpha(2-d/2)}/N^2=o(N^\alpha).$$

□

3.3 Convergence of likelihood ratios

In view of Proposition 2, to prove Theorem 2 it suffices to prove the corresponding result for the modified SIR epidemic defined in the previous section. For this, we shall analyze likelihood ratios. Denote by Q^N the law of the modified SIR epidemic, and by P^N the law of the branching envelope. Recall (cf. the *modified standard coupling*) that in the modified SIR process there can be at most one errant infection attempt, and no collisions, at any site/time x,t. Given the evolution of the process up to time $t-1$, infection *attempts* at site x and time t are made according to the same law as are offspring in the branching envelope; the conditional probability that one of the attempts is errant is $\kappa_N(y)$ (see Eq. 57). Consequently, the likelihood ratio dQ^N/dP^N at the sample evolution $X^N:=\{X_t^N(x)\}_{x,t}$ is

$$\frac{dQ^N}{dP^N}=\prod_{t\ge1}\prod_{x\in\mathbb{Z}^d}\frac{p(y|\lambda)(1-\kappa_N(y))+p(y+1|\lambda)\kappa_N(y+1)}{p(y|\lambda)},\tag{65}$$

where

$$\begin{aligned}y&=X_t^N(x),\\ \lambda=\lambda_t^N(x)&=\sum_e X_{t-1}^N(x+e)/(2d+1),\quad\text{and}\\ p(k\,|\,\lambda)&=\lambda^k e^{-\lambda}/k!.\end{aligned}$$

By the same calculation as that for Eq. 53 in [16], this can be rewritten as

$$\frac{dQ^N}{dP^N}=(1+\varepsilon_N)\exp\left\{-\sum_t\sum_x\Delta_t^N(x)\varrho_t^N(x)-\frac12\sum_t\sum_x\Delta_t^N(x)^2\varrho_t^N(x)^2\right\},\tag{66}$$

where

$$\begin{aligned}\Delta_t^N(x)&:=(X_t^N(x)-\lambda_t^N(x))/N^\alpha,\\ \varrho_t^N(x)&:=R_t^N(x)/N^{1-\alpha};\quad\text{and}\\ \varepsilon_N&=o_P(1)\quad\text{under }P^N.\end{aligned}$$

That the error term ε_N is $o_P(1)$ follows by an argument nearly identical to the proof of Lemma 9.

Observe that under P^N, the increments (in t) of the first sum in the exponential constitute a martingale difference sequence. Furthermore, the quantities $\Delta_t^N(x)$ in Eq. 66 are the atoms of the *orthogonal martingale measures* M^N associated with the branching random walks X^N. See [16] for the analogous representation in the one-dimensional case, and [25] for background on stochastic integration against orthogonal martingale measures. The martingale measures M^N can be defined by their actions on test functions $\psi \in C_c^\infty(\mathbb{R}^d)$. Write $\langle \mu, \psi \rangle$ for the integral of ψ against a finite Borel measure μ on $\mathbb{R}^d$, and $\mathcal{F}_k$ for the Feller–Watanabe rescaling operator (5); then

$$M_t^N(\psi) = \langle \mathcal{F}_{N^\alpha} X^N_{N^\alpha t}, \psi \rangle - \langle \mathcal{F}_{N^\alpha} X^N_0, \psi \rangle - \int_0^t \langle \mathcal{F}_{N^\alpha} X^N_{N^\alpha s}, A_{N^\alpha} \psi \rangle \, ds,$$

where A_k is the difference operator

$$A_k \psi(x) = \left(\sum_e \psi(x + e/\sqrt{k}) - (2d+1)\psi(x) \right) \Big/ \left[(2d+1)k^{-1} \right].$$

The first sum in the exponential of equation (66) can be expressed as a stochastic integral against the orthogonal martingale measure M^N:

$$\sum_{t \geq 1} \sum_{x \in \mathbb{Z}^d} \Delta_t^N(x) \varrho_t^N(x) = \int \int \theta^N(t, x) M^N(dt, dx), \tag{67}$$

where

$$\theta^N(t, x) = R^N_{N^\alpha t}(\sqrt{N^\alpha} x)/N^{1-\alpha}.$$

Proposition 3 *Let X be the Dawson–Watanabe process with initial configuration μ and diffusion coefficient σ^2, and let $M(dt, dx)$ and $L_t(x)$ be the associated orthogonal martingale measure and local time density process. Then under P^N, given the hypotheses of Theorem 2, as $N \to \infty$,*

$$(\mathcal{F}_{N^\alpha} X^N, \theta^N, M^N) \Longrightarrow (X, 0, M) \quad \text{if } \alpha < 2/(6-d) \quad \text{and}$$
$$(\mathcal{F}_{N^\alpha} X^N, \theta^N, M^N) \Longrightarrow (X, L, M) \quad \text{if } \alpha = 2/(6-d).$$

Proof Given the weak convergence of the second marginal θ^N, the joint convergence of the triple follows by the same argument as in Proposition 4 of [16]. The asymptotic behavior of the processes θ^N follows from Theorem 1. □

Corollary 1 *If $\alpha < 2/(6-d)$ then under P^N, as $N \to \infty$,*

$$\frac{dQ^N}{dP^N} \longrightarrow 1 \quad \text{in probability} \tag{68}$$

provided that the hypotheses of Theorem 2 on the initial configurations are satisfied.

Proof Proposition 3 implies that the sums (67) converge to zero in probability as $N \to \infty$. That the second sum in the likelihood ratio (66) also converges to zero in probability follows by a simple mean calculation (see 72). □

Proof of Theorem 2 Corollary 1 implies that the modified SIR epidemics have the same scaling limit as their branching envelopes when $\alpha < 2/(6-d)$. Thus, to complete the proof of Theorem 2, it suffices to prove the assertion (18) when $\alpha = 2/(6-d)$. For this, it suffices to show that the two sums in the exponential of equation (66) converge to the corresponding integrals in the exponential of equation (14). The convergence of the first sum follows from Proposition 3 and the representation (67), and can be proved by the same argument as in the proof of Corollary 4 of [16]. Basically, by the Skorhod embedding we can assume that $(\mathcal{F}_{N^\alpha} X^N, \theta^N, M^N)$ converges to (X, L, M) a.s.. In particular, since $L_t(x) \in C(\mathbb{R}^{1+d})$ and has compact support a.s., we have that for any $t > 0$ and $\epsilon > 0$, there exists N_0 such that for all $N \geq N_0$,

$$P\left(\max_{s\leq t,\ x} |\theta^N(s,x) - L_s(x)| \geq \epsilon\right) \leq \epsilon. \tag{69}$$

By the convergence of M^N, $\int\int L_s(x) M^N(ds\,dx) \Rightarrow \int\int L_s(x) M(ds\,dx)$. Furthermore, the local martingale

$$\int\int_{s\leq t} (\theta^N(s,x) - L_s(x)) M^N(ds\,dx) := \int\int_{s\leq t} D^N(s,x) M^N(ds\,dx) = (D^N \cdot M^N)_t$$

has quadratic variation

$$\begin{aligned}[D^N \cdot M^N]_t &= \sum\sum \mathbf{1}_{[0,t]}(s) D^N(s,x)^2 \lambda_s^N(x)/N^{2\alpha} \\ &\leq C \max_{s\leq t,\ x} D^N(s,x)^2 \cdot \sum\sum X_s^N(x)/N^{2\alpha}.\end{aligned}$$

But by Feller's Theorem, the total rescaled mass $\sum\sum X_s^N(x)/N^{2\alpha}$ of the branching process converges to the total mass of the limiting super-Brownian motion, hence by (69), with high probability, $[D^N \cdot M^N]_t$ will be small. That the approximation error $(D^N \cdot M^N)_t$ will be small with high probability uniformly in N follows by standard martingale arguments. To sum up, we get

$$\sum_n \sum_x \Delta_n^N(x) \varrho_n^N(x) = \int\int \theta^N(t,x) M^N(dt\,dx) \Rightarrow \int\int L_t(x) M(dt\,dx). \tag{70}$$

The convergence of the second sum

$$A^N := \sum_n \sum_x \Delta_n^N(x)^2 \varrho_n^N(x)^2 \Rightarrow \int \langle X_t, (L_t)^2\rangle\, dt \tag{71}$$

follows by an argument similar to the proof of equation (60) in [16]. The idea is that if one substitutes the conditional expectation $\lambda_n^N(x)/N^{2\alpha} = E(\Delta_n^N(x)^2 \,|\, \mathcal{G}_{n-1})$ for the quantity $\Delta_n^N(x)^2$ in the sum (71) then the modified sum converges; in particular, by Theorem 1 and Watanabe's theorem,

$$\begin{aligned}
B^N &:= \sum_n \sum_x \lambda_n^N(x)/N^{2\alpha} \times [R_n^N(x)/N^{1-\alpha}]^2 \\
&= \frac{1}{N^\alpha} \sum_n \sum_x \left[\sum_e X_{n-1}^N(x+e)/(2d+1) \right] /N^\alpha \times \left[R_n^N(x)/N^{\alpha(2-d/2)} \right]^2 \\
&\Longrightarrow \int \langle X_t, (L_t)^2 \rangle \, dt,
\end{aligned} \tag{72}$$

where the second equation holds because $\alpha = 2/(6-d)$. Therefore, it suffices to show that replacing $\Delta_n^N(x)^2$ by its conditional expectation has an asymptotically negligible effect on the sum, that is,

$$A^N - B^N = o_P(1).$$

By a simple variance calculation (see [16] for the one-dimensional case), this reduces to proving that

$$\sum_n \sum_x \left(\lambda_n^N(x) \right)^2 /N^{4\alpha} \times \left[R_n^N(x)/N^{\alpha(2-d/2)} \right]^4 = o_P(1). \tag{73}$$

In fact, by Theorem 1, for all $\epsilon > 0$, there exists $C > 0$ such that with probability $\geq 1 - \epsilon$,

$$\max_x R_{N^\alpha t}^N(x) \leq C N^{\alpha(2-d/2)}.$$

Note further that $\sum_{n \leq N^\alpha t} \sum_x E\left[\sum_e X_n^N(x+e)\right]^2 \leq C \sum_{n \leq N^\alpha t} \sum_x E(X_n^N(x))^2$, which, by (64) and (63), is bounded by $CN^{2\alpha} \log N$ in dimension 2 and $CN^{2\alpha}$ in dimension 3. Therefore, by enlarging C if necessary we have that with probability $\geq 1 - 2\epsilon$, the following holds:

$$\sum_{n \leq N^\alpha t} \sum_x (\lambda_n^N(x))^2/N^{4\alpha} \times [R_n^N(x)/N^{\alpha(2-d/2)}]^4 \leq CN^{2\alpha} \log N/N^{4\alpha} = o(1).$$

□

Acknowledgments S. P. Lalley supported by NSF grant DMS-0805755; Xinghua Zheng partially supported by NSERC (Canada). Xinghua Zheng is also very grateful to the support from the University of Chicago where most of the research was carried out. The authors thank an editor and two referees for careful reading of the paper and for their quick feedbacks that help improve this article.

4 Appendix: Proofs of Lemmas 2–5

4.1 Proofs of Lemmas 2–3

The strategy is to consider the regions $|x| \le (2Ln\log n)^{1/2}$ and $|x| \ge (Ln\log n)^{1/2}$ separately. We begin with the unbounded region. Recall that Hoeffding's inequality (Theorem 2 in [12]) asserts that if $Y_1, \ldots, Y_n$ are independent random variables taking values in bounded intervals $[a_i, b_i]$ respectively, then for the sum $A_n = X_1 + \cdots + X_n$, for any $t > 0$,

$$P(|A_n - EA_n| \ge t) \le 2\exp\left(-\frac{2t^2}{\sum_{i=1}^n (b_i - a_i)^2}\right).$$

In our case, since the increments of S_n are no larger than 1 in modulus,

$$P_n(x) \le P(|S_n| \ge |x|) \le d\sum_{i=1}^d P(|S_n^i| \ge |x|/\sqrt{d}) \le 2d\exp(-|x|^2/(2dn)),$$

where S_n^i is the ith coordinate of S_n. When L is sufficiently large, for any $0 < \beta < 1/(2\sqrt{d})$ we have that

$$\exp(-|x|^2/(2dn)) \le \exp(-\beta^2|x|^2/(2n))/n^{(d+1)/2}, \quad \text{for all } |x| \ge \sqrt{Ln\log n}.$$

Thus,

$$P_n(x) \le C\exp(-\beta^2|x|^2/(2n))/n^{(d+1)/2} = C\phi_n(\beta x)/\sqrt{n}, \qquad \text{for all } |x| \ge \sqrt{Ln\log n}, \tag{74}$$

and

$$\begin{aligned} |P_n(x) - P_n(y)| &\le C\Phi_n(\beta x, \beta y)/\sqrt{n} \\ &\le C\left(\frac{|x-y|}{\sqrt{n}} \wedge 1\right)\Phi_n(\beta x, \beta y), \quad \text{for all } |x|, |y| \ge \sqrt{Ln\log n}. \end{aligned} \tag{75}$$

This proves inequalities (19) and (22) for x and y outside the ball of radius $(Ln\log n)^{1/2}$.

To deal with the region $|x| \le (2Ln\log n)^{1/2}$ we shall use the following crude estimate, valid for all points $x \in \mathbb{Z}^d$ (Theorem 2.3.5 in [19]):

$$|P_n(x) - \sigma^{-d}\phi_n(x/\sigma)| \le C/(\sqrt{n}^d \cdot n).$$

For $\beta = \beta(L) > 0$ sufficiently small,

$$\phi_n(\beta x) \ge 1/(\sqrt{n}^d \cdot \sqrt{n}), \quad \text{for all } |x| \le \sqrt{2Ln\log n};$$

consequently,

$$|P_n(x) - \sigma^{-d}\phi_n(x/\sigma)| \le C\phi_n(\beta x)/\sqrt{n}, \quad \text{for all } |x| \le \sqrt{2Ln\log n}. \tag{76}$$

This obviously implies (19) for x in the region $|x| \le (2Ln\log n)^{1/2}$, and hence, together with the argument of the preceding paragraph, completes the proof of (19).

Similar arguments can be used to establish inequality (22) for points x and y in the ball of radius $(2Ln\log n)^{1/2}$ centered at the origin. First, it is easily seen that for sufficiently small $\beta > 0$,

$$|\phi_n(x) - \phi_n(y)| \le C\left((|x-y|/\sqrt{n}) \wedge 1\right)\Phi_n(\beta x, \beta y), \quad \text{for all } x, y \in \mathbb{R}^d.$$

Hence, by (76), (22) holds for for x and y in the ball of radius $(2Ln\log n)^{1/2}$. Therefore, to complete the proof of (22) it suffices to consider the case where $|x| \le (Ln\log n)^{1/2}$ and $|y| \ge (2Ln\log n)^{1/2}$. In this case, choose a point z in the annulus $|z| \in ((Ln\log n)^{1/2}, (2Ln\log n)^{1/2})$ such that $|x-z| + |z-y| \le 2|x-y|$. Using the fact that (22) holds for each of the pairs x, z and z, y, we have

$$\begin{aligned} |P_n(x) - P_n(z)| &\le C\left((|x-z|/\sqrt{n}) \wedge 1\right)\Phi_n(\beta x, \beta z) \\ &\le 2C\left((|x-z|/\sqrt{n}) \wedge 1\right)\Phi_n(\beta x, \beta y) \end{aligned}$$

and

$$\begin{aligned} |P_n(z) - P_n(y)| &\le C\left(|z-y|/\sqrt{n} \wedge 1\right)\Phi_n(\beta z, \beta y) \\ &\le C\left((|z-y|/\sqrt{n}) \wedge 1\right)\Phi_n(\beta x, \beta y). \end{aligned}$$

Consequently,

$$|P_n(x) - P_n(y)| \le C\left((|x-y|/\sqrt{n}) \wedge 1\right)\cdot\Phi_n(\beta x, \beta y).$$

This completes the proof of (22). Inequality (23) obviously follows from (22).

Proof of (21) *when* $d = 3$ The following argument works for all $d \ge 3$. First, $\sum_{n\le kT}\phi_n(\beta x)$ is bounded by $\sum_{n=1}^{\infty}\phi_n(\beta x)$. This is a decreasing function in $|x|$; moreover, by Lemma 4.3.2 in [19], it equals $C_1/|x|^{d-2} + O(1/|x|^{d+2})$ as $|x| \to \infty$ for some $C_1 > 0$. Second, for all k sufficiently large and all $|x| \le A\sqrt{k}$,

$$\sum_{n\le kT} P_n(x) \ge \sum_{kT/2\le n\le kT} P_n(x) \ge \sum_{kT/2\le n\le kT} n^{-d/2}C \ge Ck^{1-d/2};$$

note further that

$$\sum_{n>kT} P_n(x) \le C\sum_{n>kT} n^{-d/2} \le Ck^{1-d/2};$$

therefore there exists $\delta > 0$ such that all k sufficiently large and all $|x| \le A\sqrt{k}$,

$$\sum_{n \le kT} P_n(x) \ge \delta \sum_{n=1}^{\infty} P_n(x).$$

For the nearest neighbor random walk, by Theorem 4.3.1 in [19], $\sum_{n=1}^{\infty} P_n(x)$ equals $C_2/|x|^{d-2} + O(1/|x|^d)$ as $|x| \to \infty$ for some $C_2 > 0$. Relation (21) follows.
Proof of (21) *when* $d = 2$ In this case, one can deduce from the proof of Theorem 4.4.4 in [19] that there exist $C_i > 0$ such that for all $|x| \le A\sqrt{k}$,

$$\sum_{n \le kT} \phi_n(\beta x) \asymp C_1 + C_2 \log(kT/|x|^2),$$

and

$$\sum_{n \le kT} P_n(x) \asymp C_3 + C_4 \log(kT/|x|^2).$$

(21) follows.

To complete the proof of Lemma 2, it remains to prove inequality (20).
Proof of (20) By (19), it suffices to show that there exists $C > 0$ such that for all $x \in \mathbb{Z}^d$ and all $i, j \in \mathbb{N}$,

$$\sum_y \phi_i(\beta y)\phi_j(\beta(x-y)) \le C\phi_{i+j}(\beta x/2). \tag{77}$$

For all $y \in \mathbb{Z}^d$, Let Q_y be the cube centered at y with side length 1, and define

$$\tilde{\phi}_i(y) = \int\limits_{z \in Q(y)} (2\pi i/\beta^2)^{-d/2} \exp(-\beta^2|z|^2/(2i))dz.$$

Then there exists $C > 0$ such that for all i and all x,

$$\phi_i(\beta x) \le C\tilde{\phi}_i(x).$$

Therefore to show (77), it suffices to show that there exists $C > 0$ such that for all $x \in \mathbb{Z}^d$ and all $i, j \in \mathbb{N}$,

$$\sum_y \tilde{\phi}_i(y)\tilde{\phi}_j(x-y) \le C\phi_{i+j}(\beta x/2). \tag{78}$$

Note that $(\tilde{\phi}_i(\cdot))$ is the probability mass function of the random variable $[\Lambda_i]$, where $\Lambda_i \sim N(0, i/\beta^2 \cdot I_d)$, and for any $z \in \mathbb{R}^d \setminus \cup_y \partial Q(y)$, $[z]$ is the unique y such that $z \in Q(y)$ (Λ_i takes values on $\cup_y \partial Q(y)$ with probability 0, so $[\Lambda_i]$ is well defined a.s.). Hence $\sum_y \tilde{\phi}_i(y)\tilde{\phi}_j(\cdot - y)$ is the probability mass function of $[\Lambda_i] + [\Lambda_j]$ with Λ_i and Λ_j being independent. Since $[\Lambda_i] + [\Lambda_j]$ differs from $\Lambda_i + \Lambda_j$ by distance at most 2,

$$\sum_y \tilde{\phi}_i(y)\tilde{\phi}_j(x - y) \le \int\limits_{|z-x|\le 2} (2\pi(i + j)/\beta^2)^{-d/2} \exp(-\beta^2|z|^2/(2(i + j)))\, dz.$$

It is easy to see that the last term can be bounded by $C\phi_{i+j}(\beta x/2)$ for some C independent of i, j and x. □

4.2 Proof of Lemma 4

By the monotonicity of $\phi_n(x)$ in $|x|$, for all integers $m, l \ge 1$ and all $x, y \in \mathbb{R}^d$ we have

$$\begin{aligned}\phi_m(x)\phi_l(y) &\le \phi_m(x)\phi_l(x) + \phi_m(y)\phi_l(y)\\ &\le C(ml)^{-d/4}(\phi_{ml/(m+l)}(x) + \phi_{ml/(m+l)}(y)).\end{aligned}$$

Now note that for any $t > 0$ and any x,

$$\phi_t(x) = (2\pi t)^{-d/2} \exp(-|x|^2/(2t)) \le 2^{d/2} \cdot \phi_{2t}(x),$$

and when $t \ge 1$,

$$\phi_t(x) \le \phi_{\lceil t\rceil}(x) \cdot (\lceil t\rceil/t)^{d/2} \le 2^{d/2}\phi_{\lceil t\rceil}(x),$$

where $\lceil t\rceil$ stands for the smallest integer bigger than or equal to t. Further note that when $m, l \ge 1$, $ml/(m + l) \ge 1/2$. Using the three inequalities above we then get

$$\phi_m(x)\phi_l(y) \le C(ml)^{-d/4} \left(\phi_{\lceil 2ml/(m+l)\rceil}(x) + \phi_{\lceil 2ml/(m+l)\rceil}(y)\right). \tag{79}$$

and

$$\begin{aligned}\Phi_m(x, y)\Phi_l(x, y) &= \phi_m(x)\phi_l(x) + \phi_m(x)\phi_l(y) + \phi_m(y)\phi_l(x) + \phi_m(y)\phi_l(y)\\ &\le C(ml)^{-d/4}\Phi_{\lceil 2ml/(m+l)\rceil}(x, y).\end{aligned}$$

Therefore

$$\begin{aligned}(F_n(x, y; \beta))^2 &= \sum_{m<n}\sum_{l<n}(ml)^{-\gamma/2}\Phi_m(\beta x, \beta y) \cdot \Phi_l(\beta x, \beta y)\\ &\le C\sum_{m<n}\sum_{l<n}(ml)^{-d/4-\gamma/2}\Phi_{\lceil 2ml/(m+l)\rceil}(\beta x, \beta y). \end{aligned}\tag{80}$$

Observe that when $m, l \in [1, n)$, $ml/(m+l) \in [1/2, n/2)$, hence the last term is bounded by

$$C\left(\sum_{m<n} m^{-d/4-\gamma/2}\right)^2 \sum_{l<n} \Phi_l(\beta x, \beta y) \leq C n^{2-(d+\gamma)/2} \sum_{l<n} l^{-\gamma/2} \Phi_l(\beta x, \beta y)$$
$$= C n^{2-(d+\gamma)/2} F_n(x, y; \beta),$$

i.e., (25) holds.

We now prove (26). By (80),

$$\sum_{i<n} \sum_{z} P_i(z) \cdot (F_{n-i}(x-z, y-z; \beta))^2$$
$$\leq C \sum_{i<n} \sum_{z} P_i(z) \sum_{m<n-i} \sum_{l<n-i} (ml)^{-d/4-\gamma/2} \Phi_{\lceil 2ml/(m+l) \rceil}(\beta(x-z), \beta(y-z))$$
$$\leq C \sum_{m<n} \sum_{l<n} (ml)^{-d/4-\gamma/2} \sum_{i<(n-m)\wedge(n-l)} \sum_{z} P_i(z) \Phi_{\lceil 2ml/(m+l) \rceil}(\beta(x-z), \beta(y-z)).$$

Using relation (20) and noting that $\lceil 2ml/(m+l) \rceil \leq m \vee l$, we can further bound the last term by

$$C\left(\sum_{m<n} m^{-d/4-\gamma/2}\right)^2 \sum_{i<n} \Phi_i(\beta x/2, \beta y/2) \leq C n^{2-(d+2\gamma)/2} \sum_{i<n} \Phi_i(\beta x/2, \beta y/2)$$
$$\leq C n^{2-(d+\gamma)/2} F_n(x, y; \beta/2).$$

□

4.3 Proof of Lemma 5

For all $x \in \mathbb{Z}^d$ and all integers $m, n \geq 1$, by (79),

$$\begin{aligned}(J_{m,n}(x; \beta))^2 &= \sum_{m \leq l_1, l_2 < m+n} \phi_{l_1}(\beta x) \phi_{l_2}(\beta x) \\ &\leq C \sum_{m \leq l_1, l_2 < m+n} (l_1 l_2)^{-d/4} \phi_{\lceil 2l_1 l_2/(l_1+l_2) \rceil}(\beta x).\end{aligned} \tag{81}$$

Note that when $l_1, l_2 \in [m, m+n)$, $\lceil 2l_1 l_2/(l_1+l_2) \rceil \in [m, m+n)$, hence the last term is bounded by

$$C\left(\sum_{m \leq l_1 < m+n} l_1^{-d/4}\right)^2 \sum_{m \leq l < m+n} \phi_l(\beta x) \leq C n^{2-d/2} J_{m,n}(x; \beta),$$

i.e., (28) holds.

As to (29), by (81),

$$\begin{aligned}
&\sum_{i<n}\sum_{z} P_i(z)\left(J_{m,n-i}(x-z;\beta)\right)^2 \\
&\quad\le C\sum_{i<n}\sum_{z} P_i(z)\sum_{m\le l_1,l_2<m+n-i}(l_1l_2)^{-d/4}\phi_{\lceil 2l_1l_2/(l_1+l_2)\rceil}(\beta(x-z)) \\
&\quad\le C\sum_{m\le l_1,l_2<m+n}(l_1l_2)^{-d/4}\sum_{i<(m+n-l_1)\wedge(m+n-l_2)}\sum_{z} P_i(z)\phi_{\lceil 2l_1l_2/(l_1+l_2)\rceil}(\beta(x-z)).
\end{aligned}$$

Using relation (20) and noting that $\lceil 2l_1l_2/(l_1+l_2)\rceil\in[l_1\wedge l_2, l_1\vee l_2)$, we can further bound the last term by

$$C\left(\sum_{m\le l_1<m+n} l_1^{-d/4}\right)^2\cdot\sum_{m\le i<m+n}\phi_i(\beta x/2)\le Cn^{2-d/2}J_{m,n}(x;\beta/2).$$

Relation (30) can be proved similarly: by (81),

$$\begin{aligned}
&\sum_{i<m}\sum_{z} P_i(z)\left(J_{m-i,n}(x-z;\beta)\right)^2 \\
&\quad\le C\sum_{i<m}\sum_{z} P_i(z)\sum_{m-i\le l_1,l_2<m-i+n}(l_1l_2)^{-d/4}\phi_{\lceil 2l_1l_2/(l_1+l_2)\rceil}(\beta(x-z)).
\end{aligned}$$

By (20), the last term is bounded by

$$C\left(\sum_{l_1<n} l_1^{-d/4}\right)^2\cdot\sum_{m\le i<m+n}\phi_i(\beta x/2)\le Cn^{2-d/2}J_{m,n}(x;\beta/2).$$

□

References

1. Adler, R.J.: Superprocess local and intersection local times and their corresponding particle pictures. In: Seminar on Stochastic Processes, 1992 (Seattle, WA, 1992), vol 33 of Progr. Probab., pp. 1–42. Birkhäuser Boston (1993)
2. Aldous, D.: Brownian excursions, critical random graphs and the multiplicative coalescent. Ann. Probab. **25**(2), 812–854 (1997)
3. Bramson, M., Durrett, R., Swindle, G.: Statistical mechanics of crabgrass. Ann. Probab. **17**(2), 444–481 (1989)
4. Dawson, D.A.: Geostochastic calculus. Can. J. Stat. **6**(2), 143–168 (1978)
5. Dawson, D.A., Hochberg, K.J.: The carrying dimension of a stochastic measure diffusion. Ann. Probab. **7**(4), 693–703 (1979)
6. Dawson, D.A., Perkins, E.A.: Measure-valued processes and renormalization of branching particle systems. In: Stochastic Partial Differential Equations: Six Perspectives, vol 64 of Math. Surveys Monogr., pp. 45–106. Amer. Math. Soc., Providence (1999)

7. Dolgoarshinnykh, R.G., Lalley, S.P.: Critical scaling for the SIS stochastic epidemic. J. Appl. Probab. **43**(3), 892–898 (2006)
8. Durrett, R., Perkins, E.A.: Rescaled contact processes converge to super-Brownian motion in two or more dimensions. Probab. Theory Relat. Fields **114**(3), 309–399 (1999)
9. Etheridge, A.M.: An introduction to superprocesses, volume 20 of University Lecture Series. American Mathematical Society, Providence, RI (2000)
10. Evans, S.N., Perkins, E.A.: Explicit stochastic integral representations for historical functionals. Ann. Probab. **23**(4), 1772–1815 (1995)
11. Fleischmann, K.: Critical behavior of some measure-valued processes. Math. Nachr. **135**, 131–147 (1988)
12. Hoeffding, W.: Probability inequalities for sums of bounded random variables. J. Am. Stat. Assoc. **58**, 13–30 (1963)
13. Iscoe, I.: Ergodic theory and a local occupation time for measure-valued critical branching Brownian motion. Stochastics **18**(3–4), 197–243 (1986)
14. Karatzas, I., Shreve, S.E.: Brownian Motion and Stochastic Calculus, vol. 113 of Graduate Texts in Mathematics, 2nd edn. Springer, New York (1991)
15. Konno, N., Shiga, T.: Stochastic partial differential equations for some measure-valued diffusions. Probab. Theory Relat. Fields **79**(2), 201–225 (1988)
16. Lalley, S.P.: Spatial epidemics: critical behavior in one dimension. Probab. Theory Relat. Fields **144**(3–4), 429–469 (2009)
17. Lalley, S.P., Perkins, E., Zheng, X.: A phase transition in spatial sir epidemics (in preparation, 2009)
18. Lalley, S.P., Zheng, X.: Occupation statistics of critical branching random walks. Ann. Probab., arXiv:0707.3829v2 (in revision 2007)
19. Lawler, G.F., Limic, V.: Random Walk: A Modern Introduction. http://www.math.uchicago.edu/~lawler/srwbook.pdf (2007)
20. Martin-Löf, A.: The final size of a nearly critical epidemic, and the first passage time of a Wiener process to a parabolic barrier. J. Appl. Probab. **35**(3), 671–682 (1998)
21. Perkins, E., Zheng, X.: Spatial epidemics and percolations in dimensions four and higher (in preparation, 2009)
22. Reimers, M.: One-dimensional stochastic partial differential equations and the branching measure diffusion. Probab. Theory Relat. Fields **81**(3), 319–340 (1989)
23. Spitzer, F.: Principles of Random Walks, 2nd edn. Graduate Texts in Mathematics, vol. 34. Springer, New York (1976)
24. Sugitani, S.: Some properties for the measure-valued branching diffusion processes. J. Math. Soc. Jpn **41**(3), 437–462 (1989)
25. Walsh, J.B.: An introduction to stochastic partial differential equations. In: École d'été de probabilités de Saint-Flour, XIV—1984, volume 1180 of Lecture Notes in Math., pp. 265–439. Springer, Berlin (1986)
26. Watanabe, S.: A limit theorem of branching processes and continuous state branching processes. J. Math. Kyoto Univ. **8**, 141–167 (1968)