Decaying correlations for the supercritical contact process conditioned on survival

Fiocco, M.; Zwet, W.R. van

Bernoulli **9**(5), 2003, 763–781

Decaying correlations for the supercritical contact process conditioned on survival

MARTA FIOCCO[1] and WILLEM R. VAN ZWET[2]

[1]*Department of Medical Statistics, Leiden University Medical Centre, P.O. Box 9604, 2300 RC Leiden, The Netherlands. E-mail: m.fiocco@lumc.nl*
[2]*Mathematical Institute, University of Leiden, PO Box 9512, 2300 RA Leiden, The Netherlands. E-mail: vanzwet@math.leidenuniv.nl*

Contact processes – and, more generally, interacting particle processes – can serve as models for a large variety of statistical problems, especially if we allow some simple modifications that do not essentially complicate the mathematical treatment of these processes. In a forthcoming paper (Fiocco and van Zwet 2003) we shall begin a statistical study of the supercritical contact process $\xi_t^{\{0\}}$ that starts with a single infected site at the origin and is conditioned on survival. There we shall consider the simplest statistical problem imaginable, that is, to find an estimator of the parameter of the process based on observing the set of infected sites at a single time t. We shall show that this estimator is consistent as $t \to \infty$ and establish its limit distribution after proper normalization. First, however, we must push some known properties of the contact process a little further. The present paper is devoted to these matters. In particular, we study the convex hull of the set of infected sites for the conditional $\xi_t^{\{0\}}$ process as well as its spatial correlation. We find that under some restrictions this correlation decays faster than any negative power of the distance.

Keywords: contact process; coupling; decaying correlations; moment inequality; shape theorem; supercritical

1. Introduction

The contact process was introduced and studied by Harris (1974). It is a simple model for the spread of an infection or – more generally – of a biological population on the lattice $\mathbb{Z}^d$. At each time $t \geqslant 0$, each site can be in one of two possible states: infected or healthy. The state of the site $x \in \mathbb{Z}^d$ at time t will be indicated by a random variable $\xi_t(x)$, given by

$$\xi_t(x) = \begin{cases} 1 & \text{if } x \text{ is infected,} \\ 0 & \text{if } x \text{ is healthy.} \end{cases} \tag{1.1}$$

The function $\xi_t : \mathbb{Z}^d \to \{0, 1\}$ gives the state of the process at time t. It is a $\{0, 1\}$-valued random field over $\mathbb{Z}^d$.

The evolution of this random field in time is described by the following dynamics. A healthy site is infected at rate λ by each of its $2d$ immediate neighbours which is itself infected; an infected site recovers at rate 1. Given the configuration ξ_t at time t, the processes involved are independent until a change occurs.

1350–7265

It is sometimes convenient to represent the state of the contact process at time t by the set of infected sites rather than by the function $\xi_t : \mathbb{Z}^d \to \{0, 1\}$. Usually this set is also denoted by ξ_t. Thus, by an abuse of notation,

$$\xi_t = \{x \in \mathbb{Z}^d : \xi_t(x) = 1\}.$$

It remains to specify the initial state of the process at time $t = 0$. If this is deterministic it will be given by the set $A \subset \mathbb{Z}^d$ of infected sites at time $t = 0$, and we denote this contact process by $\{\xi_t^A : t \geqslant 0\}$. For example, $\{\xi_t^{\mathbb{Z}^d} : t \geqslant 0\}$ or $\{\xi_t^{\{0\}} : t \geqslant 0\}$ will denote the process starting with every site infected, or with infection only at the origin. Obviously $\xi_0^A = A$ for any A. The initial set of infected sites A may also be chosen at random according to a probability measure α, and in this case we indicate the contact process interchangeably by $\{\xi_t^A : t \geqslant 0\}$ or $\{\xi_t^\alpha : t \geqslant 0\}$. If we do not wish to specify the initial state of the process at all, we simply write $\{\xi_t : t \geqslant 0\}$.

The probability distribution of the state of the processes ξ_t^A and ξ_t^α at time t will be denoted by μ_t^A and μ_t^α, respectively. Obviously, $\mu_0^\alpha = \alpha$. Probability measures on the state space $H = \{0, 1\}^{\mathbb{Z}^d}$, such as μ_t^A and μ_t^α, are defined on the σ-algebra $\mathcal{B}$ generated by the 'rectangles' $\{\eta \in H : \eta(x) = 1\}$. This is also the σ-algebra of Borel sets if we equip the state space $H = \{0, 1\}^{\mathbb{Z}^d}$ with the product topology. Notice that this ensures that any function $g : H \to \mathbb{R}^1$ depending only on a finite number of coordinates of $\eta(x)$ is continuous. For a rigorous construction of the contact process we refer the reader to Liggett (1985).

When considering the contact process, the first question that comes to mind is whether the distribution μ_t^A of ξ_t^A will converge weakly to a limit measure μ^A as $t \to \infty$. Since we employ the product topology on the state space H,

$$\mu_t^A \xrightarrow{w} \mu^A \Leftrightarrow \mu_t^A\{B \subset \mathbb{Z}^d : B \supset F\} \xrightarrow{w} \mu^A\{B \subset \mathbb{Z}^d : B \supset F\}$$

for every finite set $F \subset \mathbb{Z}^d$. In terms of functions $\eta = I_B$, the set $\{B \subset \mathbb{Z}^d : B \supset F\}$ corresponds to the cylinder set $\{\eta \in \{0, 1\}^{\mathbb{Z}^d} : \eta(x) = 1, x \in F\}$. Thus weak convergence is equivalent to convergence in distribution of the finite-dimensional projections $\{\xi_t^A(x) : x \in F\}$.

To address the convergence of μ_t^A, we appeal to Part I, Section 2 of Liggett (1999). First of all, there exists a critical value λ_d such that for $\lambda \leqslant \lambda_d$, the contact process dies out with probability 1, regardless of its initial state at time $t = 0$ (subcritical case). In the supercritical case when $\lambda > \lambda_d$, the contact process ξ_t^A survives forever with positive probability for every non-empty set $A \subset \mathbb{Z}^d$. It survives forever with probability 1 if A is infinite. It is easy to show that the distribution $\mu_t^{\mathbb{Z}^d}$ of the process $\xi_t^{\mathbb{Z}^d}$ which starts with all sites infected, converges weakly to the so-called upper invariant measure $\nu = \nu_\lambda$. Here 'invariant' refers to the fact that the contact process $\{\xi_t^\nu : t \geqslant 0\}$ with ν as initial measure is stationary; in particular, the distribution μ_t^ν of ξ_t^ν is independent of t. Also, both $\{\xi_t^{\mathbb{Z}^d} : t \geqslant 0\}$ and $\{\xi_t^\nu : t \geqslant 0\}$ are spatially translation-invariant in the sense that the distribution of $\{c \oplus \xi_t : t \geqslant 0\}$ is independent of the shift $c \in \mathbb{Z}^d$. Here $\{c\} \oplus \xi_t = \{c + x : x \in \xi_t\}$ is the Minkowski sum. Finally, for $\lambda > \lambda_d$, ν_λ assigns probability 0 to the empty set.

For a general non-empty initial state A the convergence issue is decided by the *complete convergence theorem*. Define the random hitting time

$$\tau^A = \inf\{t : \xi_t^A = \varnothing\}, \qquad A \subset \mathbb{Z}^d, \tag{1.2}$$

with the convention that $\tau^A = \infty$ if $\xi_t^A \neq \varnothing$ for all $t \geqslant 0$. Let $\delta_\varnothing$ be the distribution on H that assigns probability 1 to the empty set.

Theorem 1.1. *Let $A \subset \mathbb{Z}^d$ and $\lambda > \lambda_d$. Then, as $t \to \infty$,*

$$\mu_t^A \xrightarrow{w} \mathbb{P}(\tau^A < \infty)\delta_\varnothing + \mathbb{P}(\tau^A = \infty)\nu_\lambda. \tag{1.3}$$

Thus, given that the process ξ_t^A survives, it tends in distribution to $\nu = \nu_\lambda$, the weight assigned to ν being the probability of survival starting from A. For a proof, see Liggett (1999, p. 55).

If $\lambda > \lambda_d$ and $A = \mathbb{Z}^d$, the process $\xi_t^{\mathbb{Z}^d}$ survives forever with probability 1 and converges exponentially to the limit process, that is, for positive C and γ and all $t \geqslant 0$,

$$0 \leqslant \mathbb{P}(\xi_t^{\mathbb{Z}^d}(x) = 1) - \mathbb{P}(\xi^\nu(x) = 1) \leqslant C\,\mathrm{e}^{-\gamma t} \tag{1.4}$$

(Liggett 1999, p. 57).

A second major result concerning the contact process is the so-called *shape theorem*. To formulate this result we first have to describe the graphical representation of contact processes due to Harris (1978). This is a particular coupling of all contact processes of a given dimension d and with a given value of λ, but with every possible initial state A or initial distribution α. Consider space-time $\mathbb{Z}^d \times [0, \infty)$. For every site $x \in \mathbb{Z}^d$ we define on the line $x \times [0, \infty)$ a Poisson process with rate 1; for every ordered pair (x, y) of neighbouring sites in $\mathbb{Z}^d$ we define a Poisson process with rate λ. All of these Poisson processes are independent.

We now draw a picture of $\mathbb{Z}^d \times [0, \infty)$ where, for each site $x \in \mathbb{Z}^d$, we remove the points of the corresponding Poisson process with rate 1 from the line $x \times [0, \infty)$; for each ordered pair of neighbouring sites (x, y) we draw an arrow going perpendicularly from the line $x \times [0, \infty)$ to the line $y \times [0, \infty)$ at the points of the Poisson process with rate λ corresponding to the pair (x, y). Let us say that for x', $x \in \mathbb{Z}^d$ and $t \geqslant 0$, there exists an active path from $(x', 0)$ to (x, t) if one can travel from site x' at time 0 to site x at time t along unbroken segments of lines $z \times [0, \infty)$ in the direction of increasing time, as well as along arrows. Clearly the active paths represent the paths along which the infection can travel from site x' to site x between times 0 and t. For any set $A \subset \mathbb{Z}^d$, define ξ_t^A to be the set of all sites x for which there exists an active path from $(x', 0)$ to (x, t) for some site $x' \in A$. Clearly $\{\xi_t^A : t \geqslant 0\}$ is distributed as a contact process with initial set A. By choosing the initial set at random with distribution α, we define $\{\xi_t^\alpha : t \geqslant 0\}$. The obvious beauty of this construction is that for two initial sets of infected points $A \subset B$, we have $\xi_t^A \subset \xi_t^B$ for all t. **Unless indicated otherwise, we shall assume that all contact processes are coupled according to the graphical construction.**

The contact process has the property of self-duality. If, in the graphical representation, time is run backwards and all arrows representing infection of one site by another are reversed, then the new graphical representation has precisely the same probabilistic structure as the original one. In particular,

$$\mathbb{P}(\xi_t^A \cap B \neq \varnothing) = \mathbb{P}(\xi_t^B \cap A \neq \varnothing), \quad \text{for all } A, B \subset \mathbb{Z}^d \text{ and } t \geqslant 0. \tag{1.5}$$

With $A = \{0\}$ and $B = \mathbb{Z}^d$ this yields

$$\mathbb{P}(\tau^{\{0\}} > t) = \mathbb{P}(\xi_t^{\mathbb{Z}^d}(0) = 1)$$

which, letting $t \to \infty$ in the supercritical case, reduces to

$$\mathbb{P}(\tau^{\{0\}} = \infty) = \mathbb{P}(\xi_t^{\nu}(0) = 1).$$

Combining this with (1.4), we see that if $\lambda > \lambda_d$, then

$$\mathbb{P}(t < \tau^{\{0\}} < \infty) \leqslant C\,\mathrm{e}^{-\gamma t} \tag{1.6}$$

(cf. Liggett 1999, p. 57).

Let $\|\cdot\|$ denote the L^∞ norm on $\mathbb{R}^d$ and define

$$H_t = \left\{ y \in \mathbb{R}^d : \exists\, x \in \bigcup_{0 \leqslant s \leqslant t} \xi_s^{\{0\}} \text{ with } \|x - y\| \leqslant \tfrac{1}{2} \right\} \tag{1.7}$$

and

$$K_t = \{ y \in \mathbb{R}^d : \exists\, x \in \mathbb{Z}^d \text{ with } \|x - y\| \leqslant \tfrac{1}{2} \text{ and } \xi_t^{\{0\}}(x) = \xi_t^{\mathbb{Z}^d}(x) \}. \tag{1.8}$$

H_t and K_t are the unions of the unit cubes centred respectively at sites that were infected at some time prior to t, or where the two processes $\xi_t^{\{0\}}$ and $\xi_t^{\mathbb{Z}^d}$ are equal at time t. Recalling that $\xi_t^{\{0\}}$ and $\xi_t^{\mathbb{Z}^d}$ are defined by the graphical construction, we can now formulate the shape theorem (cf. Durrett 1991; Bezuidenhout and Grimmett 1990).

Theorem 1.2. *There exists a bounded convex subset U of $\mathbb{R}^d$ with the origin as an interior point such that, for any $\epsilon \in (0, 1)$,*

$$(1 - \epsilon)tU \subset H_t \cap K_t \subset H_t \subset (1 + \epsilon)tU, \tag{1.9}$$

eventually almost surely on the event $\{\tau^{\{0\}} = \infty\}$ where $\xi_t^{\{0\}}$ survives forever.

Having described these well-known facts concerning the contact process, we now list the main results of the present paper. **At this point we should stress once more that we shall only be concerned with the supercritical case, that is to say, in the rest of this paper we shall tacitly assume that** $\lambda > \lambda_d$. First, we strengthen the lower inclusion in Theorem 1.2 as follows.

Theorem 1.3. *For any $\epsilon \in (0, 1)$ and $r > 0$, there exists a positive number $A_{r,\epsilon}$ such that, for every $t > 0$,*

$$\mathbb{P}((1 - \epsilon)tU \subset H_t \cap K_t | \tau^{\{0\}} = \infty) \geqslant 1 - A_{r,\epsilon} t^{-r}. \tag{1.10}$$

For a proof we refer to Theorem 4.2.1 in Fiocco (1997). Bounds for such exceptional probabilities for the contact process are typically exponential, that is, of the form $A\,\mathrm{e}^{-\gamma t}$ rather than $A_r t^{-r}$ for any $r > 0$, and one may conjecture that this is also true in Theorem 1.3. If true, this would still be a highly technical matter to prove and for our purposes it would not make any difference.

For the upper inclusion in Theorem 1.2 a much cruder probability bound will suffice. Let $B_{x,r} = \{y \in \mathbb{R}^d : |y - x| \leqslant r\}$ denote the L^1 ball with radius r and centred at x. Then there exist positive numbers c, C and γ such that, for all $t \geqslant 0$,

$$\mathbb{P}(H_t \subset B_{\{0,ct\}}) \geqslant 1 - C\,\mathrm{e}^{-\gamma t}. \tag{1.11}$$

This follows immediately by comparing the contact process with Richardson's growth process and applying the result of Durrett (1988, Chapter 1).

For our purposes, Theorems 1.2 and 1.3 have two drawbacks. The first is that in statistical applications the set U – and sometimes also the time t – are unknown and the experimenter only observes the set $\xi_t^{\{0\}}$. It is therefore of interest to show that on $\{\tau^{\{0\}} = \infty\}$ the convex hull $\mathcal{C}(\xi_t^{\{0\}})$ of the set of infected sites has the same asymptotic shape tU as $H_t \cap K_t$ and H_t.

Theorem 1.4. *For every* $\epsilon \in (0, 1)$,

$$(1 - \epsilon)tU \subset \mathcal{C}(\xi_t^{\{0\}}) \subset (1 + \epsilon)tU, \tag{1.12}$$

eventually a.s. on the set $\{\tau^{\{0\}} = \infty\}$.

We can also obtain a probability bound for $\mathcal{C}(\xi_t^{\{0\}})$ corresponding to Theorem 1.3 for $H_t \cap K_t$.

Theorem 1.5. *For any* $0 < \epsilon < 1$ *and* $r > 0$, *there exists a positive number* $A_{r,\epsilon}$ *such that, for every* $t > 0$,

$$\mathbb{P}((1 - \epsilon)tU \subset \mathcal{C}(\xi_t^{\{0\}})|\tau^{\{0\}} = \infty) \geqslant 1 - A_{r,\epsilon}t^{-r}.$$

A second problem with Theorems 1.2 and 1.3 is that for large t, these results allow us to approximate the conditional distribution of $\xi_t^{\{0\}} \cap (1 - \epsilon)tU$ given $\{\tau^{\{0\}} = \infty\}$ by that of $\xi_t^{\mathbb{Z}^d} \cap (1 - \epsilon)tU$ given $\{\tau^{\{0\}} = \infty\}$. The latter distribution, however, is hardly more manageable than the former. However, it is easy to show that for large t it can be approximated in turn by the unconditional distribution of $\xi_t^{\mathbb{Z}^d} \cap (1 - \epsilon)tU$. Let $\bar{\xi}_t^{\mathbb{Z}^d}$ denote the process $\xi_t^{\mathbb{Z}^d}$ conditioned on $\{\tau^{\{0\}} = \infty\}$. The following theorem asserts that we can couple the processes $\bar{\xi}_t^{\mathbb{Z}^d}$ and $\xi_t^{\mathbb{Z}^d}$ in such a way that they coincide on tU except on a set of exponentially small probability. We shall not explicitly describe this coupling, other than to note that it is not in accordance with the graphical representation as on that probability space the process $\bar{\xi}_t^{\mathbb{Z}^d}$ is defined only on the subset $\{\tau^{\{0\}} = \infty\}$, whereas $\xi_t^{\mathbb{Z}^d}$ is defined on the entire space.

Theorem 1.6. *There exist a coupling* $({}_c\xi_t^{\mathbb{Z}^d}, {}_c\bar{\xi}_t^{\mathbb{Z}^d})$ *of* $(\xi_t^{\mathbb{Z}^d}, \bar{\xi}_t^{\mathbb{Z}^d})$ *and positive constants* C *and* γ *such that, for all* $t > 0$,

$$\mathbb{P}\Big({}_c\xi_t^{\mathbb{Z}^d} \cap tU = {}_c\bar{\xi}_t^{\mathbb{Z}^d} \cap tU\Big) \geqslant 1 - C\,\mathrm{e}^{-\gamma t}.$$

Before formulating our results concerning the decaying correlations we need to introduce some notation. Recall that $H = \{0, 1\}^{\mathbb{Z}^d}$ denotes the state space for the contact process. For $f : H \to \mathbb{R}$ and $x \in \mathbb{Z}^d$, define

$$\Delta_f(x) = \sup\{|f(\eta) - f(\zeta)| : \eta, \zeta \in H \text{ and } \eta(y) = \zeta(y) \text{ for all } y \neq x\},$$

$$|||f||| = \sum_{x \in \mathbb{Z}^d} \Delta_f(x). \tag{1.13}$$

For $R_1, R_2 \subset \mathbb{Z}^d$, let $d(R_1, R_2)$ denote the L^1 distance of R_1 and R_2:

$$d(R_1, R_2) = \inf_{x \in R_1, y \in R_2} |x - y| = \inf_{x \in R_1, y \in R_2} \sum_{i=1}^{d} |x_i - y_i|.$$

Let

$$D_R = \{f : H \to \mathbb{R}, |||f||| < \infty, f(\eta) \text{ depends on } \eta \text{ only through } \eta \cap R\}, \tag{1.14}$$

that is to say, D_R is the class of functions f with $|||f||| < \infty$ such that $f(\eta)$ depends on η only through $\eta(x)$ with $x \in R$.

First, we show that the correlations between the states of sites for the process $\xi_t^{\mathbb{Z}^d}$ decay exponentially fast in the distance between them. This follows easily from known results.

Theorem 1.7. *There exist positive numbers* γ *and* C *such that, for every* $R_1, R_2 \subset \mathbb{Z}^d$, $f \in D_{R_1}$, $g \in D_{R_2}$, *and* $t \geqslant 0$,

$$|\mathrm{cov}\Big(f(\xi_t^{\mathbb{Z}^d}),\, g(\xi_t^{\mathbb{Z}^d})\Big)| \leqslant C|||f|||.|||g|||\mathrm{e}^{-\gamma d(R_1, R_2)}. \tag{1.15}$$

Let $\bar{\xi}_t^{\{0\}}$ denote the process $\xi_t^{\{0\}}$ conditioned on $\{\tau^{\{0\}} = \infty\}$. Our final result deals with decaying correlations for this process.

Theorem 1.8. *For every* $\epsilon \in (0, 1)$ *and* $r > 0$ *there exist a positive number* $A_{r,\epsilon}$, *as well as positive constants* C *and* γ, *such that, for all* $t > 0$ *and all* f *and* g *satisfying* $f \in D_{R_1}$ *with* $R_1 \subset (1 - \epsilon)tU \cap \mathbb{Z}^d$, and $g \in D_{R_2}$ *with* $R_2 \subset \mathbb{Z}^d$,

$$|\mathrm{cov}\Big(f(\bar{\xi}_t^{\{0\}}),\, g(\bar{\xi}_t^{\{0\}})\Big)| \leqslant |||f||| \cdot |||g|||\big(C\,\mathrm{e}^{-\gamma d(R_1, R_2)} + A_{r,\epsilon} t^{-r}\big). \tag{1.16}$$

It is perhaps of interest to compare Theorems 1.7 and 1.8. The term of order t^{-r} in (1.16) originates from Theorem 1.3 and can probably be shown to be exponentially small. This is relatively unimportant for our purposes. It is more interesting that a term depending on t occurs in Theorem 1.8 at all. But, in view of the tools at our disposal, this is to be expected. If $d(R_1, R_2)$ is of larger order than t, then at least one of the two sets R_1 or R_2 would be far outside the set tU and there would be no correlation unless H_t were to extend far beyond tU. All we know about this possibility is that it can occur with probability $\mathcal{O}(\mathrm{e}^{-\gamma t})$ by (1.11). It is therefore hardly surprising that the covariance bound in (1.16) should depend on t in that case.

For technical reasons these results will be proved in a different order than they are presented above. In Section 2 we consider the graphical representation a little more formally and compare a number of different processes in preparation for the proofs of Theorems 1.6, 1.7 and 1.8. These proofs are then given in Section 3. In Section 4 we prove a moment inequality which is needed in the next section as well as in Fiocco and van Zwet (2003). Theorems 1.4 and 1.5 concerning the behaviour of $\mathcal{C}(\bar{\xi}_t^{\{0\}})$ are proved in Section 5.

In Fiocco and van Zwet (2003) these probabilistic results will be used for a study of the estimation problem for the parameter λ of the supercritical contact process $\xi_t^{\{0\}}$. Based on an observation of $\xi_t^{\{0\}}$ at a single unknown time t, we obtain an estimator $\hat{\lambda}_t\{0\}$ of λ which is strongly consistent and asymptotically normal as $t \to \infty$. To establish these facts, we shall need a law of large numbers and a central limit theorem for $\hat{\lambda}_t^{\{0\}}$, and the results of the present paper will be needed to obtain these.

2. Comparison of processes

In this section we compare several processes, all of which are defined through the graphical representation. We begin by considering this representation somewhat more formally. The Poisson processes that underlie this construction serve only to define a Markov process $\{\zeta_t : t \geqslant 0\}$, with state space $\{0, 1\}^{\mathbb{Z}^{2d}}$, given by

$$\zeta_t(x', x) = \begin{cases} 1 & \text{if there is an active path from } (x', 0) \text{ to } (x, t), \\ 0 & \textit{otherwise}, \end{cases}$$

for every pair $x', x \in \mathbb{Z}^d$ and $t \geqslant 0$. This process evolves according to the following dynamics: for every $x \in \mathbb{Z}^d$, $\zeta_t(x', x)$ becomes 0 simultaneously for all $x' \in \mathbb{Z}^d$ at rate 1, and for every ordered pair of neighbouring sites x and y, $\zeta_t(x', y)$ becomes 1 for every x' with $\zeta_t(x', x) = 1$ at rate λ. The ζ_t process is all we need to describe the graphical construction, and contact processes are now defined as functions of ζ_t by

$$\xi_t^A(x) = \max_{x' \in A} \zeta_t(x', x). \tag{2.1}$$

There is a partial ordering on the states of ζ_t defined by $\zeta \leqslant \zeta'$ if and only if $\zeta(x', x) \leqslant \zeta'(x', x)$ for all $x', x \in \mathbb{Z}^d$. A real-valued function f of ζ is said to be non-decreasing if $f(\zeta) \leqslant f(\zeta')$ whenever $\zeta \leqslant \zeta'$. Since the process ζ_t can only jump between comparable states, it is easy to check the assumptions of Theorem 2.14 of Chapter II in Liggett (1985) and establish that, for every $t \geqslant 0$, the distribution of ζ_t has positive correlations, that is,

$$\mathrm{cov}(f(\zeta_t), g(\zeta_t)) \geqslant 0 \tag{2.2}$$

if f and g are non-decreasing and continuous. Obviously $g(\zeta_t) = \mathbb{P}(\tau^{\{0\}} = \infty | \zeta_t) = \mathbb{P}(\tau^{\{0\}} = \infty | \xi_t^{\{0\}})$ is a non-decreasing function of ζ_t which is continuous if the set $\xi_t^{\{0\}}$ is restricted to an arbitrarily large ball centred at the origin. In view of the shape theorem, this implies that we may apply (2.2) to obtain

$$\mathbb{E} f(\zeta_t) 1_{\{\tau^{\{0\}}=\infty\}} = \mathbb{E}(f(\zeta_t)\mathbb{P}(\tau^{\{0\}} = \infty | \zeta_t)) \geqslant \mathbb{E} f(\zeta_t)\mathbb{P}(\tau^{\{0\}} = \infty)$$

if f is non-decreasing and continuous.

A real-valued function f on the state space of the contact process (the subsets of $\mathbb{Z}^d$) is called non-decreasing if it respects the partial ordering of inclusion, that is, if $\xi \subseteq \xi'$ (or equivalently $\xi \leqslant \xi'$) implies $f(\xi) \leqslant f(\xi')$. By (2.1) a non-decreasing continuous function of a contact process is also a non-decreasing continuous function of ζ_t, and hence we have proved the intuitively obvious fact that for the processes defined by the graphical construction,

$$\mathbb{E}(f(\xi_t^A) | \tau^{\{0\}} = \infty) \geqslant \mathbb{E} f(\xi_t^A) \tag{2.3}$$

if f is a real-valued, non-decreasing and continuous function of ξ_t^A. In other words, the unconditional process ξ_t^A is stochastically smaller than the conditional process ξ_t^A given $\{\tau^{\{0\}} = \infty\}$.

For two configurations $\xi, \eta \in H = \{0, 1\}^{\mathbb{Z}^d}$ and $f : H \to \mathbb{R}^1$ we find by changing one coordinate at a time that

$$|f(\eta) - f(\xi)| \leqslant \sum_{x \in \mathbb{Z}^d} \Delta_f(x) I_{\eta(x) \neq \xi(x)},$$

with Δ_f as in (1.13). Now let ξ and η be random elements of H and suppose that $\xi \leqslant \eta$ (or equivalently $\xi \subset \eta$) a.s. If $f \in D_R$ for some $R \subseteq Z^d$, then

$$\begin{aligned} \mathbb{E}|f(\eta) - f(\xi)| &\leqslant \sum_{x \in \mathbb{R}} \Delta_f(x)[\mathbb{P}(\eta(x) \neq \xi(x)] \\ &= |||f||| \sup_{x \in \mathbb{R}} [\mathbb{P}(\eta(x) = 1) - \mathbb{P}(\xi(x) = 1)], \end{aligned} \tag{2.4}$$

where D_R and $|||f|||$ are defined in (1.14) and (1.13). If we assume only that ξ is stochastically smaller than $\eta (\xi \overset{st}{\leqslant} \eta)$, then there exists a coupling $({}_c\xi, {}_c\eta)$ of (ξ, η), such that ${}_c\xi \leqslant {}_c\eta$, a.s. (Liggett 1999, p. 6). Then, for $f \in D_R$, (2.4) implies

$$\begin{aligned} |\mathbb{E} f(\eta) - \mathbb{E} f(\xi)| &= |\mathbb{E} f({}_c\eta) - \mathbb{E} f({}_c\xi)| \leqslant \mathbb{E}|f({}_c\eta) - f({}_c\xi)| \\ &\leqslant |||f||| \sup_{x \in \mathbb{R}} [\mathbb{P}(\eta(x) = 1) - \mathbb{P}(\xi(x) = 1)]. \end{aligned} \tag{2.5}$$

Let us introduce a number of processes, all defined through the graphical representation. So far we have encountered the processes $\{\xi_t^A : t \geqslant 0\}$ for all $A \subseteq \mathbb{Z}^d$, in particular $\xi_t^{\{0\}}$ and $\xi_t^{\mathbb{Z}^d}$. We shall also have to consider the process $\{\xi_t' : t \geqslant s\} = \{\xi_{t-s}^{\mathbb{Z}^d} : t \geqslant s\}$ which starts at time s in the graphical representation with all sites infected. For $A \subseteq \mathbb{Z}^d$, let $\bar{\xi}_t^A$ denote a process which is distributed as ξ_t^A conditioned on $\{\tau^{\{0\}} = \infty\}$. It is defined through the graphical representation by restricting ξ_t^A to the set where $\{\tau^{\{0\}} = \infty\}$ and dividing all probabilities by $\mathbb{P}(\{\tau^{\{0\}} = \infty\}) > 0$. By (2.3) we have $\xi_t^A \overset{st}{\leqslant} \bar{\xi}_t^A$.

For fixed $s > 0$, we next define $\tilde{\xi}_t$ by

$$\tilde{\xi}_t = \begin{cases} \bar{\xi}_t^{\{0\}} & \text{for } t \in [0, s], \\ \xi_t^{\bar{\xi}_s^{\{0\}}} & \textit{for } t > s. \end{cases}$$

Thus $\tilde{\xi}_t$ starts out as $\xi_t^{\{0\}}$ conditioned on $\{\tau^{\{0\}} = \infty\}$ up to time s, and after this time this condition is dropped and if $\bar{\xi}_s^{\{0\}} = A$, the process continues as $\{\xi_{t-s}^A : t \geqslant s\}$. By (2.3) we have $\tilde{\xi}_t \overset{st}{\leqslant} \bar{\xi}_t^{\{0\}}$.

Finally, we shall have to consider a further modification of $\tilde{\xi}_t$ for any given $t > s$. Let $B_{\{x,c(t-s)\}} = \{z \in \mathbb{R}^d : |z - x| \leqslant c(t-s)\}$ denote the L^1 ball with centre x and radius $c(t-s)$ in $\mathbb{R}^d$, where c is the constant occurring in (1.11). In the graphical representation we have $\tilde{\xi}_t(x) = 1$ if a site in $\tilde{\xi}_s$ at time s is connected to site x at time t by an active path. We now construct ξ_t^* by defining $\xi_t^*(x)$ for each $x \in \mathbb{Z}^d$ in the same way as $\tilde{\xi}_t(x)$, but now ignoring all infections of sites in $B_{\{x,c(t-s)\}}$ by sites outside $B_{\{x,c(t-s)\}}$ at any time in the interval $[s, t]$. To put it another way, $\xi_t^*(x)$ is determined by a contact process restricted to the set $B_{\{x,c(t-s)\}}$, which starts at time s with the set of infected sites $\bar{\xi}_s^{\{0\}} \cap B_{\{x,c(t-s)\}}$. In the graphical representation we obviously have $\xi_t^* \leqslant \tilde{\xi}_t$ a.s., and we can only have $\xi_t^*(x) < \tilde{\xi}_t(x)$ if for the $\tilde{\xi}_t$ process there is an active path from a site in $(s, \tilde{\xi}_s)$ to (t, x) that is not entirely contained in $B_{\{x,c(t-s)\}}$ during the time interval $[s, t]$.

We now wish to compare these processes. To compare $\bar{\xi}_t^{\{0\}}$ and $\bar{\xi}_t^{\mathbb{Z}^d}$ we merely note that $\bar{\xi}_t^{\{0\}} \leqslant \bar{\xi}_t^{\mathbb{Z}^d}$ a.s. and hence Theorem 1.3 asserts the existence of a constant $A_{r,\epsilon}$ for every positive r and ϵ, such that, for every $x \in (1-\epsilon)tU$ and $t > 0$,

$$0 \leqslant \mathbb{P}\Big(\bar{\xi}_t^{\mathbb{Z}^d}(x) = 1\Big) - \mathbb{P}\Big(\bar{\xi}_t^{\{0\}}(x) = 1\Big) \leqslant A_{r,\epsilon} t^{-r}. \tag{2.6}$$

We have already noted that $\xi_t^{\mathbb{Z}^d} \overset{st}{\leqslant} \bar{\xi}_t^{\mathbb{Z}^d}$. On the one hand (1.4) ensures that, for $0 < s < t$,

$$\begin{aligned}\mathbb{P}\Big(\xi_t^{\mathbb{Z}^d}(x) = 1 | \tau^{\{0\}} > s\Big) &\leqslant \mathbb{P}\Big(\xi_{t-s}^{\mathbb{Z}^d}(x) = 1\Big) \leqslant \mathbb{P}(\xi^\nu(x) = 1) + C\,\mathrm{e}^{-\gamma(t-s)} \\ &\leqslant \mathbb{P}\Big(\xi_t^{\mathbb{Z}^d}(x) = 1\Big) + C\,\mathrm{e}^{-\gamma(t-s)};\end{aligned}$$

on the other hand (1.6) implies

$$\begin{aligned}\mathbb{P}\Big(\xi_t^{\mathbb{Z}^d}(x) = 1 | \tau^{\{0\}} > s\Big) &\geqslant \mathbb{P}\Big(\xi_t^{\mathbb{Z}^d}(x) = 1 \wedge \tau^{\{0\}} = \infty)/\mathbb{P}(\tau^{\{0\}} > s\Big) \\ &= \mathbb{P}\Big(\bar{\xi}_t^{\mathbb{Z}^d}(x) = 1\Big)\mathbb{P}\Big(\tau^{\{0\}} = \infty\Big)/\mathbb{P}\Big(\tau^{\{0\}} > s\Big) \\ &\geqslant \mathbb{P}\Big(\bar{\xi}_t^{\mathbb{Z}^d}(x) = 1\Big) - C'\,\mathrm{e}^{-\gamma s},\end{aligned}$$

with $C' = C/\mathbb{P}(\tau^{\{0\}} = \infty) < \infty$. Combining these facts for $s = t/2$, we find, for all $x \in \mathbb{Z}^d$,

$$0 \leqslant \mathbb{P}\left(\bar{\xi}_t^{\mathbb{Z}^d}(x) = 1\right) - \mathbb{P}\left(\xi_t^{\mathbb{Z}^d}(x) = 1\right) \leqslant (C + C')\mathrm{e}^{-\gamma t/2}. \tag{2.7}$$

Note that the same argument yields $\mathbb{P}(\xi_{t-s}^{\mathbb{Z}^d}(x) = 1) \leqslant \mathbb{P}(\xi_t^{\mathbb{Z}^d}(x) = 1) + C\,\mathrm{e}^{-\gamma(t-s)}$ and, as $\xi_t^{\mathbb{Z}^d} \leqslant \xi_{t-s}^{\mathbb{Z}^d} = \xi_t'$ a.s., we find that, for $0 < s < t$ and $x \in \mathbb{Z}^d$,

$$0 \leqslant \mathbb{P}(\xi_t'(x) = 1) - \mathbb{P}\left(\xi_t^{\mathbb{Z}^d}(x) = 1\right) \leqslant C\,\mathrm{e}^{-\gamma(t-s)}. \tag{2.8}$$

We pointed out earlier that $\tilde{\xi}_t \overset{st}{\leqslant} \bar{\xi}_t^{\{0\}}$. The two processes are defined on the set $\{\tau^{\{0\}} > s\}$ and are equal on $\{\tau^{\{0\}} = \infty\}$. Hence, again by (1.6),

$$\begin{aligned} 0 \leqslant \mathbb{P}(\bar{\xi}_t^{\{0\}}(x) = 1) - \mathbb{P}(\tilde{\xi}_t(x) = 1) &\leqslant \mathbb{P}(s < \tau^{\{0\}} < \infty)/\mathbb{P}(\tau^{\{0\}} > s) \\ &\leqslant C\,\mathrm{e}^{-\gamma s}/\mathbb{P}(\tau^{\{0\}} = \infty) = C'\,\mathrm{e}^{-\gamma s}, \end{aligned} \tag{2.9}$$

for C' as above, $0 < s < t$ and $x \in \mathbb{Z}^d$.

It remains to compare $\tilde{\xi}_t$ and ξ_t^*. We have already noted that $\xi_t^* \leqslant \tilde{\xi}_t$ a.s., and that $\xi_t^*(x) < \tilde{\xi}_t(x)$ implies that for the $\tilde{\xi}_t$-process there is an active path from a site in $(s, \tilde{\xi}_s)$ to (t, x) that is not entirely contained in $B_{\{x,c(t-s)\}}$ during the time interval $[s, t]$. Reversing time and the direction of infection arrows, the self-duality of the graphical construction and (1.11) yield

$$\begin{aligned} \mathbb{P}(\xi_t^*(x) \neq \tilde{\xi}_t(x)) &\leqslant \mathbb{P}\left(\tau^{\{x\}} \geqslant t-s, \bigcap_{0\leqslant u\leqslant t-s} \xi_u^{\{x\}} \not\subset B_{\{x,c(t-s)\}}\right) \\ &= \mathbb{P}\left(\bigcap_{0\leqslant u\leqslant t-s} \bar{\xi}_u^{\{0\}} \not\subset B_{\{0,c(t-s)\}}\right) \leqslant \mathbb{P}(H_{t-s} \not\subset B_{\{0,c(t-s)\}}) \leqslant C\,\mathrm{e}^{-\gamma(t-s)}. \end{aligned}$$

It follows that

$$0 \leqslant \mathbb{P}(\tilde{\xi}_t(x) = 1) - \mathbb{P}(\xi_t^*(x) = 1) \leqslant C\,\mathrm{e}^{-\gamma(t-s)}. \tag{2.10}$$

3. Proof of Theorems 1.6–1.8

We are now in a position to prove Theorems 1.6, 1.7 and 1.8.

Proof of Theorem 1.6. Since $\xi_t^{\mathbb{Z}^d} \overset{st}{\leqslant} \bar{\xi}_t^{\mathbb{Z}^d}$, there exists a coupling $({}_c\xi_t^{\mathbb{Z}^d}, {}_c\bar{\xi}_t^{\mathbb{Z}^d})$ such that ${}_c\xi_t^{\mathbb{Z}^d} \leqslant {}_c\bar{\xi}_t^{\mathbb{Z}^d}$ a.s. The cardinality of $tU \cap \mathbb{Z}^d$ is at most ct^d for some constant $c > 0$ and hence, by (2.7),

$$\begin{aligned} \mathbb{P}\left({}_c\xi_t^{\mathbb{Z}^d} \cap tU \neq {}_c\bar{\xi}_t^{\mathbb{Z}^d} \cap tU\right) &\leqslant ct^d \sup_{x\in tU}\left[\mathbb{P}({}_c\bar{\xi}_t^{\mathbb{Z}^d}(x) = 1) - \mathbb{P}({}_c\xi_t^{\mathbb{Z}^d}(x) = 1)\right] \\ &\leqslant ct^d c'\,\mathrm{e}^{-\gamma' t/2} \leqslant C\,\mathrm{e}^{-\gamma t}, \end{aligned}$$

for appropriate positive c', γ', C and γ. This proves the theorem. □

Proof of Theorem 1.7. We note that $\xi^\nu \overset{st}{\leqslant} \xi_t^{\mathbb{Z}^d}$ and hence, by (2.5) and (1.4),

$$\begin{aligned}|\mathbb{E}f(\xi_t^{\mathbb{Z}^d}) - \mathbb{E}f(\xi^\nu)| &\leqslant |||f|||\sup_{x\in\mathbb{R}}[\mathbb{P}(\xi_t^{\mathbb{Z}^d}(x) = 1) - \mathbb{P}(\xi^\nu(x) = 1)] \\ &\leqslant C|||f|||\mathrm{e}^{-\gamma t}.\end{aligned}$$

Now (1.15) follows by applying Theorem 4.20 in Chapter 1 of Liggett (1985). □

Proof of Theorem 1.8. Let $\epsilon \in (0, 1)$, $r > 0$, $f \in D_{R_1}$ and $g \in D_{R_2}$, with $R_1 \subset (1-\epsilon)tU \cap \mathbb{Z}^d$ and $R_2 \subset \mathbb{Z}^d$. Write $d^* = d(R_1, R_2)$ and choose $s \in (0, t)$ with

$$t - s \leqslant \frac{d^*}{3c}, \tag{3.1}$$

where c is the constant occurring in (1.11). It follows that the L^1 ball $B_{\{x,c(t-s)\}}$ in the definition of ξ_t^* is contained in $B_{\{x,d^*/3\}}$. The construction of ξ_t^* therefore implies that, conditional on $\tilde{\xi}_s$, $f(\xi_t^*)$ depends only on those Poisson processes in the graphical representation associated with sites or pairs of neighbouring sites in $R_1 \oplus B_{\{0,d^*/3\}}$. The same is true for $g(\xi_t^*)$ and $R_2 \oplus B_{\{0,d^*/3\}}$. As $d(R_1, R_2) = d^*$, $R_1 \oplus B_{\{0,d^*/3\}}$ and $R_2 \oplus B_{\{0,d^*/3\}}$ are disjoint and hence $f(\xi_t^*)$ and $g(\xi_t^*)$ are conditionally independent given $\tilde{\xi}_s$.

Without loss of generality, we replace g and f by $g - g(\eta)$ and $f - f(\eta)$ for some fixed $\eta \in H$. The effect of this is that $\|g\| \leqslant |||g|||$, $\|f\| \leqslant |||f|||$, and hence $|||fg||| \leqslant \|f\| \cdot |||g||| + \|g\| \cdot |||f||| \leqslant 2|||f||| \cdot |||g|||$. As $\xi_t^* \overset{st}{\leqslant} \bar{\xi}_t^{\{0\}}$, (2.5), (2.9) and (2.10) imply that there exist $C > 0$ and $\gamma > 0$ such that

$$|\mathbb{E}f(\bar{\xi}_t^{\{0\}})g(\bar{\xi}_t^{\{0\}}) - \mathbb{E}f(\xi_t^*)g(\xi_t^*)| \leqslant C|||f||| \cdot |||g|||\left(\mathrm{e}^{-\gamma s} + \mathrm{e}^{-\gamma(t-s)}\right). \tag{3.2}$$

The conditional independence of $f(\xi_t^*)$ and $g(\xi_t^*)$ yields

$$\mathbb{E}\left[f(\xi_t^*)g(\xi_t^*)|\tilde{\xi}_s\right] = \mathbb{E}\left[f(\xi_t^*)|\tilde{\xi}_s\right] \cdot \mathbb{E}\left[g(\xi_t^*)|\tilde{\xi}_s\right]. \tag{3.3}$$

We have

$$\begin{aligned}\mathbb{E}\Big(\mathbb{E}\left[f(\xi_t^*)|\tilde{\xi}_s\right].\mathbb{E}\left[g(\xi_t^*)|\tilde{\xi}_s\right]\Big) &= \mathbb{E}\Big(\mathbb{E}\left[f(\xi_t')|\tilde{\xi}_s\right].\mathbb{E}\left[g(\xi_t^*)|\tilde{\xi}_s\right]\Big) \\ &\quad + \mathbb{E}\Big(\mathbb{E}\left[f(\xi_t^*) - f(\xi_t')|\tilde{\xi}_s\right].\mathbb{E}\left[g(\xi_t^*)|\tilde{\xi}_s\right]\Big),\end{aligned} \tag{3.4}$$

where we recall that ξ_t' starts at time s in the graphical representation with all sites infected and is therefore independent of $\tilde{\xi}_s = \bar{\xi}_s^{\{0\}}$. Hence the first term on the right in (3.4) equals

$$\mathbb{E}f(\xi_t')\mathbb{E}\Big(\mathbb{E}\left[g(\xi_t^*)|\tilde{\xi}_s\right]\Big) = \mathbb{E}f(\xi_t')\mathbb{E}g(\xi_t^*). \tag{3.5}$$

The second term is bounded in absolute value by $|||g||| \cdot \mathbb{E}|f(\xi_t^*) - f(\xi_t'|$. Since $\xi_t^* \leqslant \tilde{\xi}_t \leqslant \xi_t'$ a.s. for $0 < s < t$ and $f \in D_{R_1}$ with $R_1 \subset (1-\epsilon)tU \cap \mathbb{Z}^d$, (2.4) ensures that this is bounded in turn by

$$\|\!|f\|\!| \cdot \|\!|g\|\!| \cdot \sup_{x\in R_1}\big[\mathbb{P}(\xi'_t(x)=1)-\mathbb{P}(\xi^*_t(x)=1)\big]$$

$$\leqslant \|\!|f\|\!| \cdot \|\!|g\|\!| \cdot \sup_{x\in R_1}\Big\{\Big[\mathbb{P}(\tilde{\xi}_t(x)=1)-\mathbb{P}(\xi^*_t(x)=1)\Big]+\Big[\mathbb{P}(\bar{\xi}^{\{0\}}_t(x)=1)-\mathbb{P}(\tilde{\xi}_t(x)=1)\Big]$$

$$+\Big[\mathbb{P}(\bar{\xi}^{\mathbb{Z}^d}_t(x)=1)-\mathbb{P}(\bar{\xi}^{\{0\}}_t(x)=1)\Big]-\Big[\mathbb{P}(\bar{\xi}^{\mathbb{Z}^d}_t(x)=1)-\mathbb{P}(\xi^{\mathbb{Z}^d}_t(x)=1)\Big]$$

$$+\Big[\mathbb{P}(\xi'_t(x)=1)-\mathbb{P}(\xi^{\mathbb{Z}^d}_t(x)=1)\Big]\Big\},$$

which by (2.6)–(2.10) is bounded by $\|\!|f\|\!| \cdot \|\!|g\|\!| \cdot [C(\mathrm{e}^{-\gamma s}+\mathrm{e}^{-\gamma(t-s)})+A_{r,\epsilon}t^{-r}]$. Hence, by (3.2)–(3.5),

$$|\mathbb{E}f(\bar{\xi}^{\{0\}}_t)g(\bar{\xi}^{\{0\}}_t)-\mathbb{E}f(\xi'_t)\mathbb{E}g(\xi^*_t)| \leqslant \|\!|f\|\!| \cdot \|\!|g\|\!| \cdot [C(\mathrm{e}^{-\gamma s}+\mathrm{e}^{-\gamma(t-s)})+A_{r,\epsilon}t^{-r}] \qquad (3.6)$$

for appropriate positive C, γ and $A_{r,\epsilon}$.

Because $R_1 \subset (1-\epsilon)tU$, (2.5)–(2.8) yield

$$|\mathbb{E}f(\xi'_t)-\mathbb{E}f(\bar{\xi}^{\{0\}}_t)| \leqslant |\mathbb{E}f(\bar{\xi}^{\mathbb{Z}^d}_t)-\mathbb{E}f(\bar{\xi}^{\{0\}}_t)|+|\mathbb{E}f(\xi^{\mathbb{Z}^d}_t)-\mathbb{E}f(\bar{\xi}^{\mathbb{Z}^d}_t)|$$

$$+|\mathbb{E}f(\xi'_t)-\mathbb{E}f(\xi^{\mathbb{Z}^d}_t)| \leqslant \|\!|f\|\!| \cdot [C\,\mathrm{e}^{-\gamma(t-s)}+A_{r,s}t^{-r}],$$

and by (2.5), (2.9) and (2.10),

$$|\mathbb{E}g(\xi^*_t)-\mathbb{E}g(\bar{\xi}^{\{0\}}_t)| \leqslant \|\!|g\|\!| \cdot C(\mathrm{e}^{-\gamma s}+\mathrm{e}^{-\gamma(t-s)}).$$

Combining this and (3.6), we find that

$$|\mathrm{cov}\Big(f(\bar{\xi}^{\{0\}}_t),\ g(\bar{\xi}^{\{0\}}_t)\Big)| \leqslant \|\!|f\|\!| \cdot \|\!|g\|\!| \cdot [C(\mathrm{e}^{-\gamma s}+\mathrm{e}^{-\gamma(t-s)})+A_{r,\epsilon}t^{-r}],$$

for appropriate positive C, γ and $A_{r,\epsilon}$. Choosing $(t-s)=d^*/(3c)\wedge(t/2)$, we arrive at

$$|\mathrm{cov}\Big(f(\bar{\xi}^{\{0\}}_t),\ g(\bar{\xi}^{\{0\}}_t)\Big)| \leqslant \|\!|f\|\!| \cdot \|\!|g\|\!| \cdot [C\,\mathrm{e}^{-\gamma d^*}+A_{r,\epsilon}t^{-r}],$$

for positive C, γ and $A_{r,\epsilon}$. □

4. A moment inequality

In this section we prove an inequality for the central moments of certain functions of $\xi^{\mathbb{Z}^d}_t$ that will be needed in Section 5 as well as in Fiocco and van Zwet (2003). For $A \subset \mathbb{Z}^d$, define the total number of infected sites in the set A at time t as

$$n^{\mathbb{Z}^d}_t(A)=\sum_{x\in A}\xi^{\mathbb{Z}^d}_t(x). \qquad (4.1)$$

The cardinality of a set $A \subset \mathbb{Z}^d$ will be denoted by $|A|$.

Lemma 4.1. *For any $k = 1, 2, \ldots$, there exists a number $C_k > 0$ such that, for every $A \subset \mathbb{Z}^d$ and $t \geqslant 0$,*

$$\mu_{2k} = \mathbb{E}\Big(n_t^{\mathbb{Z}^d}(A) - \mathbb{E}n_t^{\mathbb{Z}^d}(A)\Big)^{2k} \leqslant C_k|A|^k. \tag{4.2}$$

Proof. Given any $x_1, x_2, \ldots, x_{2k} \in \mathbb{Z}^d$, let $R = \max_j d(x_j, \{x_1, \ldots, x_{j-1}, x_{j+1}, \ldots, x_{2k}\})$ and write μ_{2k} as

$$\mu_{2k} = \mathbb{E}\left(\sum_{x\in A}(\xi_t^{\mathbb{Z}^d}(x) - \mathbb{E}\xi_t^{\mathbb{Z}^d}(x))\right)^{2k} = \sum_{x_1\in A}\cdots\sum_{x_{2k}\in A}\mathbb{E}\prod_{i=1}^{2k}(\xi_t^{\mathbb{Z}^d}(x_i) - \mathbb{E}\xi_t^{\mathbb{Z}^d}(x_i)).$$

For every $j = 1, 2, \ldots, 2k$, we have

$$\left|\left|\left|\prod_{i=1, i\neq j}^{2k}\Big(\xi_t^{\mathbb{Z}^d}(x_i) - \mathbb{E}\xi_t^{\mathbb{Z}^d}(x_i)\Big)\right|\right|\right| \leqslant 2k, \qquad \left|\left|\left|\xi_t^{\mathbb{Z}^d}(x_j) - \mathbb{E}\xi_t^{\mathbb{Z}^d}(x_j)\right|\right|\right| = 1,$$

and hence Theorem 1.7 and the definition of R imply that

$$\mu_{2k} = 2Ck\sum_{x_1\in A}\cdots\sum_{x_{2k}\in A}\mathrm{e}^{-\gamma R}.$$

Notice that the distance $d(x_j, \{x_1, \ldots, x_{j-1}, x_{j+1}, \ldots, x_{2k}\}) = 0$ unless x_j occurs only once in the sequence $x_1, \ldots, x_{2k}$.

Let m_j denote the number of sites that occur j times in the sequence of sites $x_1, \ldots, x_{2k}$. Define, for $m_1, \ldots, m_{2k}$ with $\sum_{i=1}^{2k} jm_j = 2k$ and $r \geqslant 0$, $F_{m_1,\ldots,m_{2k}}(r)$ as the number of sequences $x_1, \ldots, x_{2k} \in A$ with given $m_1, \ldots, m_{2k}$ and $R < r$. Then $F_{m_1,\ldots,m_{2k}}(0) = 0$ and, for $r > 0$,

$$F_{m_1,\ldots,m_{2k}}(r) \leqslant \begin{cases} C_k'|A|^{m_2+m_3+\cdots}|A|^{m_1/2}r^{dm_1/2} & \text{if } m_1 \text{ even,} \\ C_k'|A|^{m_2+m_3+\cdots}|A|^{(m_1-1)/2}r^{d(m_1+1)/2} & \text{if } m_1 \text{ odd,} \end{cases}$$

where d is the dimension of the contact process and C_k' is an appropriate positive constant. This bound for $F_{m_1,\ldots,m_{2k}}(r)$ is computed by noting that the $m_2 + m_3 + \ldots$ sites which occur more than once can be chosen anywhere in A without contributing to R, which accounts for the factor $|A|^{m_2+m_3+\cdots}$. The m_1 sites which occur only once, however, have to be chosen within a distance r from another member of $x_1, \ldots, x_{2k}$, and this gives rise to a factor $|A|^{m_1/2}r^{dm_1/2}$ or $|A|^{(m_1-1)/2}r^{d(m_1+1)/2}$ for even or odd m_1, respectively. Finally, the combinatorics of the situation refers to ordering $2k$ sites and can therefore be bounded by C_k'.

If m_1 is odd, then $m_3 + m_4 + \ldots > 0$ as $\sum_j jm_j = 2k$ and, as a result,

$$\frac{m_1 - 1}{2} + m_2 + m_3 + \ldots \leqslant k - 1,$$

while if m_1 is even,

$$\frac{m_1}{2} + m_2 + m_3 + \ldots \leqslant k.$$

Hence, if we define $F_m(r)$ as the number of sequences $x_1, \ldots, x_{2k} \in A$ with $m_1 = m$ and $R < r$, then

$$F_m(r) \leqslant. \begin{cases} C_k''|A|^k r^{dm/2} & \text{if } m \text{ even,} \\ C_k''|A|^{k-1} r^{d(m+1)/2} & \text{if } m \text{ odd.} \end{cases}$$

For $r > 0$, let $F(r)$ be the number of sequences $x_1, \ldots, x_{2k} \in A$ with $R < r$, so that

$$F(r) = \sum_{m=0}^{2k} F_m(r).$$

Summing the terms with even or odd values of m separately we obtain, for $r > 0$,

$$F(r) = \sum_{s=0}^{k} F_{2s}(r) + \sum_{s=1}^{k} F_{2s-1}(r) \leqslant C_k'''|A|^k r^{dk},$$

where C_k''' is an appropriately chosen constant. As $F(0) = 0$,

$$\begin{aligned} \mu_{2k} &\leqslant 2Ck \sum_{x_1 \in A} \cdots \sum_{x_{2k} \in A} \mathrm{e}^{-\gamma R} \\ &\leqslant 2Ck \sum_{r=0}^{\infty} \mathrm{e}^{-\gamma r}(F(r+1) - F(r)) \leqslant C_k|A|^k \end{aligned}$$

for an appropriate $C_k > 0$. □

A more general version of Lemma 4.1 may be formulated as follows. Let $g : H \to \mathbb{R}$ satisfy $g \in D_{B_{\{0,r\}}}$, where $B_{\{0,r\}}$ is an L^1 ball centred at the origin with radius r. Hence $g(\eta)$ depends on η only through $\eta(x)$ with x in a fixed set $B_{\{0,r\}} \subset \mathbb{R}^d$. For $a \in \mathbb{Z}^d$, let ${}_a\eta$ denote a shifted version of η with ${}_a\eta(x) = \eta(a + x)$ for all $x \in \mathbb{Z}^d$, and define $g_a : H \to R$ by $g_a(\eta) = g({}_a\eta)$ for $\eta \in H$. Note that $g_a \in D_{B_{\{a,r\}}}$. Lemma 4.1 can now easily be generalized as follows.

Theorem 4.1. *For any $k = 1, 2, \ldots$ and $r > 0$, there exists a number $C_{k,r} > 0$ such that, for every $A \subset \mathbb{Z}^d$, $g \in D_{B_{\{0,r\}}}$ and $t \geqslant 0$,*

$$\mathbb{E}\left(\sum_{a \in A}(g_a(\xi_t^{\mathbb{Z}^d}) - \mathbb{E}g_a(\xi_t^{\mathbb{Z}^d}))\right)^{2k} \leqslant C_{k,r}|||g|||^{2k}|A|^k. \tag{4.3}$$

Proof. Without loss of generality we may replace g by $g - g(\zeta)$ and g_a by $g_a - g_a(\zeta)$ for some fixed $\zeta \in H$, so that $\|g_a\| = \|g\| \leqslant |||g||| = |||g_a|||$. Then, for some $c_{k,r} > 0$,

$$\left|\left|\left| \prod_{i=1, i\neq j}^{2k} \left(g_{a_i}(\xi_t^{\mathbb{Z}^d}) - \mathbb{E} g_{a_i}(\xi_t^{\mathbb{Z}^d}) \right) \right|\right|\right| \leqslant c_{k,r} |||g|||^{2k-1},$$

$$\left|\left|\left| \left(g_{a_j}(\xi_t^{\mathbb{Z}^d}) - \mathbb{E} g_{a_j}(\xi_t^{\mathbb{Z}^d}) \right) \right|\right|\right| = |||g|||,$$

for some $c_{k,r} > 0$. Using the fact that $d(B_{\{a,r\}}, B_{\{b,r\}}) \geqslant d(a, b) - 2r$, the proof goes through the same counting argument employed to prove Lemma 4.1. $\square$

Recall the definition of the random hitting time for the process ξ_t^A,

$$\tau^A = \inf\{t : \xi_t^A = \varnothing\}, \qquad A \subset \mathbb{Z}^d.$$

We shall need the following corollary.

Corollary 4.1. *For any $r > 0$ there exists a number $C_r > 0$ such that, for every $A \subset \mathbb{Z}^d$ and $t \geqslant 0$,*

$$\mathbb{P}(n_t^{\mathbb{Z}^d}(A)) \leqslant \tfrac{1}{2} \mathbb{E} n_t^{\mathbb{Z}^d}(A)) \leqslant C_r |A|^{-r}, \tag{4.4}$$

$$\mathbb{P}(\tau^A < \infty) \leqslant C_r |A|^{-r}. \tag{4.5}$$

Proof. It is obviously enough to prove the corollary for integer $k = 1, 2, \ldots$ instead of real $r > 0$. Applying the self-duality property (1.5) for $B = \mathbb{Z}^d$, we find

$$\mathbb{P}(\tau^A \leqslant t) = \mathbb{P}(\xi_t^{\mathbb{Z}^d} \cap A = \varnothing) = \mathbb{P}(n_t^{\mathbb{Z}^d}(A) = 0).$$

Since the process $\xi_t^{\mathbb{Z}^d}$ is translation-invariant, the graphical construction yields

$$\mathbb{E} n_t^{\mathbb{Z}^d}(A) = \sum_{x \in A} \mathbb{E} \xi_t^{\mathbb{Z}^d}(0) = |A| \mathbb{E} \xi_t^{\mathbb{Z}^d}(0) \geqslant |A| \mathbb{E} \xi^\nu(0),$$

where the right-hand side is independent of t. Therefore for every $t \geqslant 0$ and $k = 1, 2, \ldots,$

$$\begin{aligned}
\mathbb{P}(\tau^A \leqslant t) &= \mathbb{P}\left(n_t^{\mathbb{Z}^d}(A) = 0 \right) \leqslant \mathbb{P}\left(n_t^{\mathbb{Z}^d}(A) \leqslant \tfrac{1}{2} \mathbb{E} n_t^{\mathbb{Z}^d}(A) \right) \\
&\leqslant \mathbb{P}\left(|n_t^{\mathbb{Z}^d}(A) - \mathbb{E} n_t^{\mathbb{Z}^d}(A)| \geqslant \tfrac{1}{2} \mathbb{E} n_t^{\mathbb{Z}^d}(A) \right) \\
&\leqslant \frac{2^{2k} \mu_{2k}}{[\mathbb{E} n_t^{\mathbb{Z}^d}(A)]^{2k}} \leqslant \frac{2^{2k} \mu_{2k}}{[\mathbb{E} \xi^\nu(0)]^{2k} |A|^{2k}} \leqslant C_k |A|^{-k}
\end{aligned}$$

by Lemma 4.1. This implies that

$$\mathbb{P}(\tau^A < \infty) \leqslant C_k |A|^{-k}$$

because the bound does not depend on t. $\square$

We note that (4.5) may be improved by a more delicate argument to yield

$$\mathbb{P}(\tau^A < \infty) \leqslant C\,\mathrm{e}^{-\gamma|A|} \tag{4.6}$$

for positive C and γ (cf. Liggett 1999, p. 57).

5. The asymptotic shape of the convex hull

In this section we prove a shape theorem for the convex hull $\mathcal{C}(\xi_t^{\{0\}})$ of the set of infected sites $\xi_t^{\{0\}}$, as well as a corresponding probability bound (cf. Theorems 1.4 and 1.5).

Definition 5.1. *A convex polytope is a set which is the convex hull of a finite number of points.*

Lemma 5.1. *For every $0 < \epsilon < \frac{1}{2}$, there exists a convex polytope $P \subset \mathbb{R}^d$ such that*

$$(1-2\epsilon)U \subset P \subset \left(1 - \frac{3\epsilon}{2}\right)U. \tag{5.1}$$

Proof. By Theorem 33 in Chapter 4 of Eggleston (1958) we have, for every $\delta > 0$, a convex polytope P containing $(1-2\epsilon)U$ and contained in a δ-neighbourhood $\{x : d(x, (1-2\epsilon)U) \leqslant \delta\}$ of $(1-2\epsilon)U$. Here d is L^1 distance. Since 0 is an interior point of U, this δ-neighbourhood of $(1-2\epsilon)U$ is contained in $(1-3\epsilon/2)U$ for sufficiently small δ. □

Let $x_1, x_2, \ldots, x_k$ be the extreme points of a convex polytope P satisfying (5.1). For each of these points x_i we define a set

$$A_i = \{x : \exists\, \eta > 0,\, x_i - \eta(x - x_i) \in P\} \cap (1-\epsilon)U. \tag{5.2}$$

The set A_i is the intersection of $(1-\epsilon)U$ and the exterior cone of P at x_i, and as $P \subset (1 - 3\epsilon/2)U$, we see that A_i contains an open set in $\mathbb{R}^d$. For any $B \subset \mathbb{R}^d$, let $\mathcal{C}(B)$ denote the convex hull of B. We have:

Lemma 5.2. *If $x_i' \in A_i$, for $i = 1, \ldots, k$, then*

$$P \subset \mathcal{C}(\{x_1', \ldots, x_k'\}). \tag{5.3}$$

Proof. Suppose that for some $m = 1, 2, \ldots, k$, $P \subset \mathcal{C}(\{x_1', \ldots, x_{m-1}', x_m, \ldots, x_k\})$, where the latter set is interpreted as $\mathcal{C}(\{x_1, \ldots, x_k\})$ for $m = 1$. As $x_m \in A_m$,

$$x_m - \eta(x_m' - x_m) \in P \subset \mathcal{C}(\{x_1', \ldots, x_{m-1}', x_m \ldots, x_k\}),$$

or equivalently,

$$x_m = \frac{\eta}{(1+\eta-\lambda_m)}x_m' + \sum_{j=1}^{m-1}\frac{\lambda_j}{1+\eta-\lambda_m}x_j' + \sum_{j=m+1}^{k}\frac{\lambda_j}{1+\eta-\lambda_m}x_j,$$

for some $\eta > 0$, $\lambda_j \geqslant 0$ for each j, and $\sum_{j=1}^k \lambda_j = 1$. This implies that $x_m \in \mathcal{C}(\{x_1', \ldots, x_m', x_{m+1}, \ldots, x_k\})$ and as a result $P \subset \mathcal{C}(\{x_1', \ldots, x_m', x_{m+1}, \ldots, x_k\})$. Induction yields $P \subset \mathcal{C}(\{x_1', \ldots, x_k'\})$. □

Proof of Theorem 1.4. On the set where $\xi_t^{\{0\}}$ survives forever, $H_t \subset (1+\epsilon)tU$ eventually a.s. by Theorem 1.2. Since U is convex, this implies that $\mathcal{C}(H_t) \subset (1+\epsilon)tU$. In view of the definition of H_t in (1.7), $\xi_t^{\{0\}} \subset H_t$ and hence $\mathcal{C}(\xi_t^{\{0\}}) \subset \mathcal{C}(H_t)$. Combining these inclusions we arrive at

$$\mathcal{C}(\xi_t^{\{0\}}) \subset \mathcal{C}(H_t) \subset (1+\epsilon)tU \tag{5.4}$$

eventually a.s. on the set where $\xi_t^{\{0\}}$ survives forever. This establishes the almost sure upper bound for $\mathcal{C}(\xi_t^{\{0\}})$ in (1.12).

To obtain the lower bound in (1.12), we begin by noting that (4.4) ensures that, for every $r > 0$ and $i = 1, 2, \ldots, k$,

$$\mathbb{P}(n_t^{\mathbb{Z}^d}(tA_i) \leqslant \tfrac{1}{2}\mathbb{E}n_t^{\mathbb{Z}^d}(tA_i)) \leqslant C_r|tA_i|_D^{-r} \leqslant C_{r,\epsilon}t^{-dr}.$$

Here

$$|A|_D = |A \cap \mathbb{Z}^d|$$

denotes the discrete cardinality of a set $A \subset \mathbb{R}^d$ and the final inequality follows from the fact that for fixed ϵ, $A_1, \ldots, A_k$ are fixed subsets of $\mathbb{R}^d$ with non-empty interiors. In view of the graphical representation, we have for $i = 1, \ldots, k$ and $t \geqslant m_\epsilon$,

$$\tfrac{1}{2}\mathbb{E}n_t^{\mathbb{Z}^d}(tA_i) \geqslant \tfrac{1}{2}\mathbb{E}n_t^{\nu}(tA_i) = \tfrac{1}{2}|tA_i|_D\mathbb{E}\xi^{\nu}(0) \geqslant c_\epsilon t^d.$$

It follows that, for $t \geqslant m_\epsilon$,

$$\mathbb{P}(n_t^{\mathbb{Z}^d}(tA_i) \leqslant c_\epsilon t^d) \leqslant C_{r,\epsilon}t^{-dr} \tag{5.5}$$

for appropriately chosen positive c_ϵ and $C_{r,\epsilon}$ and integer $m_\epsilon > 0$. Hence, for $m \geqslant m_\epsilon$,

$$\begin{aligned}&\mathbb{P}(n_t^{\mathbb{Z}^d}(tA_i) = 0 \text{ for some } t \in [m, m+1)) \\ &\qquad \leqslant C_{r,\epsilon}m^{-dr} + \mathbb{P}(n_t^{\mathbb{Z}^d}(tA_i) = 0 \text{ for some } t \in [m, m+1) | n_m^{\mathbb{Z}^d}(mA_i) > c_\epsilon m^d).\end{aligned}$$

The latter conditional probability is bounded by the probability that the maximum of $[c_\epsilon m^d] + 1$ independent standard exponential waiting times does not exceed 1, which is $\mathcal{O}(m^{-dr})$ as $m \to \infty$ for any $r > 0$. As a result we have, for $m = m_\epsilon, m_{\epsilon+1}, \ldots,$

$$\mathbb{P}(n_t^{\mathbb{Z}^d}(tA_i) = 0 \text{ for some } t \in [m, m+1)) \leqslant C'_{r,\epsilon}m^{-dr},$$

and, choosing $r \geqslant 2/d$, we see that the Borel–Cantelli lemma implies that, for $i = 1, \ldots, k$,

$$n_t^{\mathbb{Z}^d}(tA_i) \neq 0 \text{ eventually a.s.} \tag{5.6}$$

Obviously this also holds for $i = 1, 2, \ldots, k$ simultaneously.

By (5.2), $tA_i \subset (1-\epsilon)tU$ for $i = 1, \ldots, k$, and hence Theorem 1.2 implies that on the set $\{\tau^{\{0\}} = \infty\}$, $\xi_t^{\{0\}}(x) = \xi_t^{\mathbb{Z}^d}(x)$ eventually a.s., for all $x \in tA_i$ for $i = 1, \ldots, k$. Hence, (5.6) ensures that on the set where $\xi_t^{\{0\}}$ survives forever,

$$n_t^{\{0\}}(tA_i) \neq 0 \qquad \text{for } i = 1, \ldots, k, \text{ eventually a.s.} \tag{5.7}$$

If $n_t^{\{0\}}(tA_i) \neq 0$ for $i = 1, \ldots, k$, then each of the sets tA_i contains a point of $\xi_t^{\{0\}}$, and by (5.1)–(5.3) this implies that

$$(1-2\epsilon)tU \subset tP \subset \mathcal{C}(\xi_t^{\{0\}}). \tag{5.8}$$

In view of (5.7), (5.8) holds eventually a.s. on the set where $\xi_t^{\{0\}}$ survives forever. Since ϵ is arbitrary, the theorem is proved. $\square$

Proof of Theorem 1.5. In the proof of Theorem 1.4 we note that (2.3) and (5.5) imply that

$$\mathbb{P}(n_t^{\mathbb{Z}^d}(tA_i) = 0 | \tau^{\{0\}} = \infty) \leqslant \mathbb{P}(n_t^{\mathbb{Z}^d}(tA_i) = 0) \leqslant C_{r,\epsilon} t^{-dr}.$$

Invoking Theorem 1.3, we arrive at

$$\mathbb{P}(n_t^{\{0\}}(tA_i) \neq 0 \text{ for } i = 1, \ldots, k | \tau^{\{0\}} = \infty) \geqslant 1 - C_{r,\epsilon} k t^{-dr} - A_{r,\epsilon} t^{-r}.$$

By the argument leading to (5.8) we obtain a probability bound for $\mathcal{C}(\xi_t^{\{0\}})$, which is the statement of Theorem 1.5. $\square$

Acknowledgement

The authors are endebted to an associate editor and three referees for constructive criticism. They also wish to thank the past editor, Ole Barndorff-Nielsen, for his willingness to continue handling this paper well after the end of his tenure.

References

Bezuidenhout, C. and Grimmett, G. (1990) The critical contact process dies out. *Ann. Probab.*, **18**, 1462–1482.

Durrett, R. (1988) *Lectures Notes on Particle Systems and Percolation*. Pacific Grove, CA: Wadsworth.

Durrett, R. (1991) The contact process, 1974–1989. In W.E. Kohler and B.S. White (eds), *Mathematics of Random Media*, Lectures in Appl. Math. 27, pp. 1–18. Providence, RI: American Mathematical Society.

Fiocco, M. (1997) Statistical estimation for the supercritical contact process. Doctoral thesis, University of Leiden, The Netherlands.

Fiocco, M. and van Zwet, W.R. (2003) Parameter estimation for the supercritical contact process. *Bernoulli*, **9**. To appear.

Eggleston, H.G. (1958) *Convexity.* Cambridge: Cambridge University Press.

Harris, T.E. (1974) Contact interaction on a lattice. *Ann. Probab.*, **2**, 969–988.

Harris, T.E. (1978) Additive set-valued Markov processes and graphical methods. *Ann. Probab.*, **6**, 355–378.

Liggett, T. (1985) *Interacting Particle Systems*. New York: Springer-Verlag.

Liggett, T. (1999) *Stochastic Interacting Systems: Contact, Voter and Exclusion Processes*. New York: Springer-Verlag.

Received December 1998 and revised April 2003