On the Critical Behavior at the Lower Phase Transition of the Contact Process

Michael Aizenman and Paul Jung

Departments of Physics and Mathematics, Princeton University, Princeton, NJ 08544
E-mail address: `aizenman@princeton.edu`

Mathematics Department, Cornell University, 591 Malott Hall, Ithaca, NY 14850
E-mail address: `pjung@math.cornell.edu`

Abstract. We present general results for the contact process by a method which applies to all transitive graphs of bounded degree, including graphs of exponential growth. The model's infection rates are varied through a control parameter, for which two natural transition points are defined as: i. λ_T, the value up to which the infection dies out exponentially fast if introduced at a single site, and ii. λ_H, the threshold for the existence of an invariant measure with a non-vanishing density of infected sites. It is shown here that for all transitive graphs the two thresholds coincide. The method, which proceeds through partial differential inequalities for the infection density, yields also generally valid bounds on two related critical exponents. The main results discussed here were established by Bezuidenhout and Grimmett (1991) in an extension to the continuous-time process of the discrete-time analysis of Aizenman and Barsky (1987), and of the partially similar results of Menshikov (1986). The main novelty here is in the derivation of the partial differential inequalities on which the Aizenman and Barsky (1987) analysis is based by an argument which is formulated directly for the continuum.

1. Introduction and statement of the main results

Since its introduction by Harris (1974), the contact process has attracted interest as a model for the spread of "infection". The model undergoes a phase transition which is reached by varying the ratio of the infection rate to the healing rate, which in our notation is $\lambda : 1$. The small λ regime can be characterized by the finiteness of the "susceptibility", $\chi(\lambda)$, which is the total time lost to infection within the population if an infection is introduced at a single site. A duality argument allows to conclude that if $\chi(\lambda) < \infty$ then the infection dies out even if initially the entire

302 Michael Aizenman and Paul Jung

population was infected. As λ approaches the edge of the regime $\{\lambda : \chi(\lambda) < \infty\}$ the susceptibility diverges and the contact process exhibits critical behavior with characteristics similar to those observed in models of statistical mechanics.

Upon analysis, it turns out to be a generally valid statement that right past the point of divergence of $\chi(\lambda)$ a homogeneous contact process enters the phase at which there is a stationary measure with persistent infection. The technique presented below allows to establish this basic feature for the contact processes on the broad class of *transitive graphs*. This is the class of graphs which are invariant under the action of a symmetry group which acts transitively. Included in the collection are some graphs for which the contact process is known to exhibit more than one transition, in the sense explained below. In this generality, the basic properties of the model include the following:

(1) For λ with $\chi(\lambda) < \infty$ the probability that infection from a single site will persist in the population for time t decays exponentially in time. In models for which the rates for the direct transmission have suitable exponential decay, the probability that the infection would reach distance d away also exhibits, in that phase, exponential decay in the distance.

(2) At the edge of the above region $\chi(\lambda) \nearrow \infty$, i.e., the model exhibits criticality at the point

$$\lambda_T := \sup\{\lambda \,|\, \chi(\lambda) < \infty\}. \tag{1.1}$$

which is named for Temperley.

(3) For $\lambda > \lambda_T$ the model is in the phase at which the infection persists. The threshold for the latter condition has been recognized by the term λ_H, for Hammersley. Thus the above statement amounts to the coincidence of the two points: $\lambda_T = \lambda_H$.

(4) At the transition point, which can now be denoted simply λ_c, the model exhibits critical behavior with characteristic exponents which in general are bounded by their 'mean-field' values. The bounds are realized in certain situations.

In regard to the spread of infection, the contact process can be viewed as oriented percolation. That offers a helpful perspective, as the above characteristics are shared by transitive percolation models with or without orientation. It was in that context that the characteristics 1.-3. of the phase diagram were initially established for the discrete-time version of the models, in two different and independently derived methods, presented in the works of Menshikov (1986) and Aizenman and Barsky (1987). The argument of Menshikov (1986) was limited to graphs of subexponential growth, such as $\mathbb{Z}^d$. The method of Aizenman and Barsky (1987) readily extends to transitive graphs and yields also additional information on the critical exponents, which follow through partial differential inequalities on which more is said below. However, both analyses were initially presented for only the discrete-time version of the models. The extension to the continuous-time contact process was accomplished in the work of Bezuidenhout and Grimmett (1991), through a detailed control of the (1D) continuum limit. Our main goal here is to present a direct extension of the method of Aizenman and Barsky (1987) to the contact process in terms which are natural for the continuum, casting the argument in the generality of the transitive graphs of arbitrary growth rate.

1.1. *The model and its parameters.* We shall introduce the model in the context of transitive graphs. Before giving a formal description of the generator of the time evolution, let us set some notation. For $\mathcal{G} = (V, E)$ a connected transitive graph with vertex set V and edge collection E, the contact process $\{A_t\}_{t \in \mathbb{R}}$ is a random time-dependent collection of subsets of V describing a set of infected sites. If the initial set of infected sites is given by B at time T, then the corresponding measure on $\{A_t\}_{t \in \mathbb{R}}$ is given by $\mathbb{P}^{(B,T)}(\cdot)$. Let

$$c(y, A) = \begin{cases} 1 & \text{if } y \in A \\ \lambda \sum_{x \in A} J_{x,y} & \text{if } y \notin A \end{cases}.$$

where $J_{x,y}$ is a translation invariant kernel with

$$|J| = \sum_y J_{x,y} < \infty.$$

The generator of the contact process A_t is formally given by

$$Lf(A) = \sum_x c(x, A)[f(A \bullet \{x\}) - f(A)], \tag{1.2}$$

where $\bullet$ denotes the symmetric difference operation, i.e., for $x \in V$: $A \bullet \{x\} = A \cup \{x\}$ when $x \notin A$ and $A \bullet \{x\} = A \backslash \{x\}$ when $x \in A$.

Often times the kernel is 1 when x and y are neighbors and 0 otherwise; the finiteness of $|J|$ implies that the vertex degree is finite. A simple description of the process in this case is as follows. At time t, the set of infected vertices is denoted by A_t. A vertex heals independently with exponential rate 1, while uninfected vertices become infected at exponential rate λ times the number of infected neighbors.

Two significant quantities which reflect properties of the model are:

(1) the infection density of the upper invariant measure

$$\theta_+(\lambda) := \lim_{T \to -\infty} \mathbb{P}^{(V,T)}(o \in A_0) = \lim_{T \to -\infty} \mathbb{P}^{(\{o\},T)}(A_0 \neq \emptyset) \tag{1.3}$$

where the second equation is by duality and stationarity,

(2) the susceptibility

$$\chi(\lambda) := \int_0^\infty \mathbb{E}^{(\{o\},0)}(|A_t|)dt < \infty, \tag{1.4}$$

which, by Fubini, equals the expected value of the sum of the times lost to infection at the different sites. The contact process exhibits a number of different phases, depending on the control parameter λ. Some of the thresholds of interest are defined as follows.

Definition 1.1.

$$\begin{aligned} \lambda_T &:= \sup\{\lambda : \chi(\lambda) < \infty\} \\ \lambda_H &:= \sup\{\lambda : \theta_+(\lambda) = 0\} \\ \lambda_{GN} &:= \sup\{\lambda : \mathbb{P}^{(\{o\},0)}(o \in A_t) \xrightarrow[t \to \infty]{} 0\} \end{aligned}$$

Their general relation is:

$$\lambda_T \leq \lambda_H \leq \lambda_{GN}. \tag{1.5}$$

Remarks: The two first transition points were already mentioned above. The third has appeared in the work of Grimmett and Newman (1990) within the context of percolation models on products of regular trees and Euclidean lattices, where its

analog is the threshold for the uniqueness of the infinite cluster. The above work motivated (Pemantle, 1992), where it was shown that $\lambda_H < \lambda_{GN}$ for the contact process on regular trees of degree four or more; the proof was extended to all regular trees in (Liggett, 1996) and then more succinctly in (Stacey, 1996).

The contact process can be viewed in terms of a graphical representation, whereby one traces the state of the infection over the 'space $\times$ time' graph, $\mathcal{G} \times \mathbb{R}$. Healing events are represented by Poisson processes of intensity 1 on the lines of $V \times \mathbb{R}$, and infection-transmission events are represented by Poisson processes of intensity $\lambda J_{x,y}$ on $(V \times V) \times \mathbb{R}$. For the latter, the set $V \times V$ appears as the collection of directed edges, and an event at (e_{xy}, t) represents a possible transmission from x to y at time t. The set of possible sources of infection for a site x at time t is the set of all points in $V \times \mathbb{R}$ from which there is a path which does not backtrack in time, reaching (x, t) without passing through any healing event. We refer to this set as $C(x, t)$. One may view it as the connected cluster of (x, t) in an oriented percolation model. For brevity we denote $C = C(o, 0)$. For a more detailed description of the graphical representation picture and its relation to the self-duality of the contact process we refer the reader to Liggett (1999).

The graphical representation highlights the strong relationship this process has with oriented percolation, and yields the following interpretation of the two transition points:

$$
\begin{aligned}
\lambda_T &= \inf\{\lambda : \mathbb{E}(|C|) = \infty\} \\
\lambda_H &= \inf\{\lambda : \mathbb{P}(|C| = \infty) > 0\},
\end{aligned}
\tag{1.6}
$$

where $|C|$ denotes the set's size. For a discrete set like A_t the size refers to the set's cardinality, whereas for a generic $S \subset V \times \mathbb{R}$, such as C, we denote by $|S|$ the total length of the set's vertical segments.

By known arguments, the small-λ phase has the following characteristics:

Proposition 1.2. *For any $\lambda < \lambda_T$ there exist some $c < \infty$ and $\tau > 0$ such that*

$$
\mathbb{E}^{(\{o\},0)}(|A_t|) < c\, e^{-t/\tau}.
\tag{1.7}
$$

Furthermore, if $\sum_x J_{o,x}\, e^{+\varepsilon|x|} < \infty$ for some $\varepsilon > 0$, then also

$$
\mathbb{P}^{(\{o\},0)}(A_t \cap B_r^c \neq \emptyset \quad \text{for some } t \geq 0) \leq k\, e^{-\mu r}
\tag{1.8}
$$

for some $k < \infty$ and $\mu > 0$, where $B_r^c \subset V$ represents the complement of a ball of radius r around the vertex o.

The proof of the above proposition follows from a subadditivity property of the contact process and can be found in several places in the literature. For completeness we include a proof of the above proposition in Appendix (A).

1.2. *Summary of the main results.* Among the key statements proven below is:

Theorem 1.3. *For any contact process on a transitive graph, with a translation-invariant infection-transmission rate $J_{x,y}$ and a constant healing rate 1,*

$$
\lambda_T = \lambda_H.
$$

As discussed above, for $\mathcal{G} = \mathbb{Z}^d$ the above result was established in Bezuidenhout and Grimmett (1991). The main method used there, based on Menshikov (1986), readily extends to transitive graphs of subexponential growth (such graphs are amenable, though the converse is not true). It allows to conclude that at $\lambda < \lambda_H$ the probability that the infection of one site will affect another decays exponentially

in the distance (and also in time). However, if $\mathcal{G}$ is non-amenable, e.g., a regular tree, exponential decay does not yet imply finiteness of $\chi(\lambda)$. Nevertheless, it is not difficult to extend the arguments given in Section 3.2 of Bezuidenhout and Grimmett (1991), which follows the approach of Aizenman and Barsky (1987), to prove the above theorem for the full class of transitive graphs.

In Aizenman and Barsky (1987), certain non-linear differential inequalities were derived within a somewhat natural extension of the model, for which one adds the possibility of spontaneous infection, at the rate h. Added insight is derived from the consideration of the model within the two parameter space of (λ, h). The original model is then recovered through the limit $h \to 0$. In terms of the graphical representation of the contact process, the spontaneous infection events are represented by a Poisson process on $V \times \mathbb{R}$ with density $h\,dt$.

It may be noted that while the extra parameter h has a very natural meaning for a contact process, in the original context of percolation it has appeared as a somewhat ad-hoc auxiliary "ghost field", whose introduction was motivated by an analogy with the external magnetic field of ferromagnetic Ising spin systems Aizenman et al. (1987).

Keeping the terminology used in the percolation discussion, the events of spontaneous infection (points in space $\times$ time) will be referred to as green sites, and their collection denoted by G. The function $\theta_+(\lambda)$ which referred to the limiting density of infection starting from the 'all infected' state, finds its extension to $h > 0$ in the function:

$$\theta(\lambda, h) = \mathbb{P}(C(o,0) \cap G \neq \emptyset). \tag{1.9}$$

Following are some of the relevant properties of this extension of the model.

Lemma 1.4. *For any contact process on a transitive graph:*
i. *At each $h > 0$ there is a unique stationary state, to which the state of the system converges for all asymptotic initial condition $((S_{-T}, -T)$ for $T \to \infty)$ with the infection density given by the above function $\theta(\lambda, h)$.*
ii. *For $h > 0$, the function $\theta(\lambda, h)$ is monotone in its arguments, continuous in λ, and continuously differentiable in h.*
iii. *In the limit $h \to 0+$, the function $\theta(\lambda, h)$ yields the quantities which were introduced above for the $h = 0$ model as follows:*

$$\lim_{h \searrow 0} \theta(\lambda, h) = \theta_+(\lambda) \tag{1.10}$$

$$\text{for } \lambda < \lambda_H: \quad \lim_{h \searrow 0} \frac{\partial \theta(\lambda, h)}{\partial h} = \chi(\lambda). \tag{1.11}$$

Since the main idea is rather standard, we relegate the proof of the above lemma to Appendix (B). As can be seen there, the graphical representation provides the following useful expressions for θ and its derivative $\chi(\lambda, h) := \partial\theta(\lambda, h)/\partial h$ is:

$$\theta(\lambda, h) = \mathbb{P}(C(o,0) \cap G \neq \emptyset) = \mathbb{E}\left(1 - e^{-h|C|}\right), \tag{1.12}$$

$$\text{for } \lambda < \lambda_H: \quad \chi(\lambda, h) = \mathbb{E}\left(|C(o,0)|;\, C(o,0) \cap G = \emptyset\right) = \mathbb{E}\left(|C|\,e^{-h|C|}\right) \tag{1.13}$$

The graphical representation enables the derivation of partial differential inequalities which, via integration through the two-parameter space prove Thm 1.3 and

provide also additional information about the behavior in the vicinity of the critical point, which can now be commonly denoted as $\lambda_c := \lambda_T = \lambda_H$.

Theorem 1.5. *For any transitive graph*
(i) for $\lambda > \lambda_c$:

$$\theta_+(\lambda) \geq Const. \, (\lambda - \lambda_c)^1 \tag{1.14}$$

(ii) at $\lambda = \lambda_c$:

$$\theta(\lambda_c, h) \geq Const. \, h^{1/2} \, . \tag{1.15}$$

The inequalities imply bounds for the associated critical exponents:

$$\beta \geq 1 \, , \qquad \qquad \delta \geq 2 \, . \tag{1.16}$$

As explained above, these results are known already for both the discrete-time contact process (Aizenman and Barsky, 1987), and the continuous-time model (Bezuidenhout and Grimmett, 1991). The main novelty here is in the derivation for the continuous-time process of the partial differential inequalities which are discussed next.

It should be noted that the critical exponent bounds (1.16) are saturated for the contact process on regular trees (of degree three or more) (Wu, 1995; Schonmann, 1998), and also on $\mathbb{Z}^d$ when d is very large or just $d > 4$ and the kernel is sufficiently 'spread-out' (Barsky and Wu, 1998; Sakai, 2001). The discrete-time version of this statement was proven earlier through the combination of the results of Barsky and Aizenman (1991) and Nguyen and Yang (1993).

1.3. *The key differential inequalities.* The derivation of the above results proceeds through certain non-linear partial differential inequalities (PDI). The simplest of these is:

$$\frac{\partial \chi}{\partial \lambda} \leq |J| \chi^2 \, . \tag{1.17}$$

This relation, which for percolation was presented in Aizenman and Newman (1984), is basically known in the generality considered here. It has been noted that (1.17) implies a critical exponent bound ($\gamma \leq 1$) which concerns the divergence rate for χ as $\lambda \nearrow \lambda_T$:

$$\chi(\lambda) \geq \frac{|J|^{-1}}{|\lambda_T - \lambda|_+} \, . \tag{1.18}$$

Next are partial differential inequalities which are similar to the PDI which were derived in Aizenman and Barsky (1987) for the discrete-time contact process, in the context of percolation model, extending an earlier differential inequality of Chayes and Chayes (1986), which has yielded a percolation analog of (1.14) with λ_c interpreted as λ_H.

Theorem 1.6. *For any contact process on a transitive graph, at Lebesgue almost every $(\lambda, h) \in \mathbb{R}_+ \times \mathbb{R}_+$ (due to the monotonicity of θ, the derivatives exist in this sense):*

$$\frac{\partial \theta}{\partial \lambda} \leq \theta |J| \frac{\partial \theta}{\partial h} \tag{1.19}$$

and

$$\theta \leq h \frac{\partial \theta}{\partial h} + \left(2\lambda^2 |J|\theta + h\lambda\right) \frac{\partial \theta}{\partial \lambda} + \theta^2 \, . \tag{1.20}$$

In (Aizenman and Barsky, 1987), where the discrete version of the above theorem was established, it was envisioned that an extension to the continuum ought to be possible through a limiting argument, but the result may involve some more complicated coefficients. Nevertheless, as is shown here the inequalities are valid in a rather simple form, which is not that different from the discrete-time version. The proof proceeds through a finite-volume version of the statement, given in Theorem 3.2 in Section 3. Theorem 1.6 is proved in Section 4.

For the purpose of the derivation let us present some notions which are of general use when working with Poisson processes.

2. A Poisson process differentiation formula

The analysis is made clearer by recognizing a general expression for the derivatives of the probabilities of monotone events with respect to Poisson densities. Derivatives of functionals of Poisson processes have been useful in many contexts, see Borovkov (1996), Bacelli et al. (1995). In particular, one can apply such analysis to form a continuum analog of 'Russo's formula'. The formulae presented in this section are similar to those found in Zuev (1993).

Definition 2.1. *Let $\mathbb{X}$ be a measure space, and $\Psi(\rho)$ a monotone functional on the space $\mathcal{M}$ of non-negative measures $\rho(x)$ on $\mathbb{X}$. A function $K(x, \rho)$ on $\mathbb{X} \times \mathcal{M}$ is said to be the variational derivative of $\Psi(\cdot)$ at x if for all finite positive continuous measures α on $\mathbb{X}$*

$$\frac{d}{ds}\Psi(\rho + s\,\alpha)\Big|_{s=0^+} = \int K(x)\,d\alpha(x)\,. \tag{2.1}$$

It is easy to see that when it exists, the variational derivative is unique. We denote it

$$\frac{\delta\Psi}{\delta\rho(x)} = K(x)\,. \tag{2.2}$$

We shall now consider functionals of the form

$$\Psi(\rho) = \mathbb{P}_\rho(\omega \in F) \tag{2.3}$$

where F is an increasing event, i.e., one whose indicator function is a non-decreasing function of the configuration $\omega \subset \mathbb{X}$, and the subscript on $\mathbb{P}$ indicates that ω is distributed by the Poisson process with intensity measure ρ.

Definition 2.2. *Let F be an increasing event defined for the point process. A point $x \in \mathbb{X}$ is said to be* pivotal *for F in the configuration ω if $\omega\backslash\{x\} \notin F$ but $\omega \cup \{x\} \in F$. The set of pivotal points is denoted*

$$\Delta F(\omega) := \{x : \omega \cup \{x\} \in F \text{ and } \omega\backslash\{x\} \notin F\}\,. \tag{2.4}$$

Lemma 2.3. *For any Poisson process, the probability of any increasing event F has a variational derivative given by*

$$\frac{\delta\,\mathbb{P}_\rho(F)}{\delta\,\rho(x)} = \mathbb{P}_\rho(F^c; \{x \in \Delta F\})\,. \tag{2.5}$$

If the density ρ is non-atomic ($\rho(\{x\}) = 0$ for all $x \in \mathbb{X}$) then also

$$\frac{\delta\,\mathbb{P}_\rho(F)}{\delta\,\rho(x)} = \mathbb{P}_\rho(x \in \Delta F)\,. \tag{2.6}$$

Proof. Since the variational addition α is a continuous measure, a valid way to generate a random configuration distributed by the Poisson process at the density $\rho_s := \rho + s\,\alpha$ is to take the union of two configurations ω_0 and $\widetilde{\omega}_s$, drawn independently through a pair Poisson processes at intensities ρ and $s\alpha$, correspondingly. By this construction,

$$
\begin{aligned}
\mathbb{P}_{\rho_s}(F) - \mathbb{P}_{\rho_0}(F) &= \mathbb{P}\left(\omega_0 \notin F,\, \omega_0 \cup \widetilde{\omega}_s \in F\right) \\
&= \mathbb{P}\left(\omega_0 \notin F,\, \omega_0 \cup \widetilde{\omega}_s \in F,\, |\widetilde{\omega}_s| = 1\right) + \\
&\quad + \mathbb{P}\left(\omega_0 \notin F,\, \omega_0 \cup \widetilde{\omega}_s \in F,\, |\widetilde{\omega}_s| \geq 2\right)
\end{aligned}
\tag{2.7}
$$

where the first equality is due to monotonicity of F, and $|\cdot|$ denotes the cardinality of a set. The first of the two events in the last expression coincides with the event that *i.* $\omega_0 \notin F$, and *ii.* $\widetilde{\omega}_s$ is a one-point subset of $\Delta F(\omega_0)$. The second term is dominated by $[s\,\alpha(\mathbb{X})]^2$. Conditioning on ω_0, and using the explicit Poisson formula for $\widetilde{\omega}_s$, one gets:

$$
\begin{aligned}
\mathbb{P}_{\rho_s}(F) - \mathbb{P}_{\rho_0}(F) &= s\,\mathbb{E}\left(F^c; \alpha(\Delta F(\omega_0))\, e^{-s\alpha(\Delta F(\omega_0))}\right) + O(s^2) \\
&= s\int \mathbb{P}_\rho(F^c; x \in \Delta F)\, d\alpha(x) + O(s^2).
\end{aligned}
\tag{2.8}
$$

The first claim now readily follows.

If ρ is non-atomic then the probability that the site x seen on the right in eq. (2.5) is occupied, and thus F occurs, vanishes for each a-priori specified $x \in \mathbb{X}$. Hence the condition F^c can be omitted from eq. (2.5), which is thus reduced to eq. (2.6). $\square$

One may note that some auxiliary conditions are required for an extension of the differentiation formula (2.1) to apply also to the case where α is not a finite measure. E.g., for any set F which is measurable at infinity the pivotal set ΔF is a.s. empty, yet the probability of F need not be independent of ρ.

3. Derivation of the Partial Differential Inequalities

3.1. *A dictionary for the contact process.* We shall now translate Lemma 2.3 to the situation at hand. Recall that the space $\times$ time picture of the contact process is described in terms of three independent Poisson processes describing the random healing events, at constant rate 1, the spontaneous infection events, and the random infection-transmissions. In discussing the partial derivatives of the corresponding probability, we allow the latter two processes to be inhomogeneous, i.e., of densities given by functions rather than constants: $h_x(t)$ and $\lambda_{x,y}(t)\, J_{x,y}$. The corresponding probability measure is denoted by $\mathbb{P}_{\lambda,h}$.

Of particular interest will be the event $E = \{C \cap G \neq \emptyset\}$, where $C = C(o,0)$ is the infecting cluster for a particular site $(o,0)$. We apply in the natural way the terminology introduced in Definition 2.2 and say that in a given configuration ω a site $(x,t) \in V \times \mathbb{R}$ is *green pivotal* if a change of the green set G at (x,t) will have an affect on whether $\omega \in E$ or not. Likewise, we will say that an ordered bond e_{yx}^t, at time t, is *bond pivotal* for E if the presence of a transmission-event there will affect whether $\omega \in E$ or not.

Lemma 3.1. *For the contact process on an arbitrary graph,*

$$\frac{\delta\theta(\lambda,h)}{\delta\lambda_{y,x}(t)} \;=\; J_{y,x}\,\mathbb{P}_{\lambda,h}(e_{yx}^t \text{ is bond pivotal for } E) \qquad (3.1)$$

$$\frac{\delta\theta(\lambda,h)}{\delta h_x(t)} \;=\; \mathbb{P}_{\lambda,h}((x,t) \text{ is green pivotal for } E)\,, \qquad (3.2)$$

and for models with h constant:

$$h\,\frac{\partial}{\partial h}\theta(\lambda,h) \;=\; \mathbb{P}_{\lambda,h}(C \text{ has exactly one green site})\,. \qquad (3.3)$$

Proof. The first two assertions are direct consequences of Lemma 2.3.

For finite graphs, equation (3.3) can be understood from (3.2), as we comment below. However a direct proof which is not limited by the finiteness condition can be obtained from the expression (1.12) for $\theta(\lambda,h)$ which readily yields:

$$\begin{aligned} h\,\frac{\partial}{\partial h}\theta(\lambda,h) &= \mathbb{E}\left(h|C|\,e^{-h|C|}\right) \\ &= \mathbb{P}_{\lambda,h}(C \text{ has exactly one green site})\,, \qquad (3.4) \end{aligned}$$

where the last step is an explicit Poisson process relation. $\square$

Remark: It is instructive to note that equation (3.3) can be explained by (3.2) through the following argument. Let $A = \{\omega : |C \cap G| = 1\}$ be the event that C has exactly one green site. Conditioned on A there is a uniquely defined site $Y(\omega)$, for which the event $Y(\omega) = (x,t)$ is characterized by:

i. $(x,t) \in G$, i.e., the site is an arrival point for the corresponding Poisson process,

ii. in the configuration ω, (x,t) is a pivotal site for $\{C \cap G \neq \emptyset\}$, i.e., for E.

The above two statements refer to independent conditions: *i.* referring to the status of the site itself (or arbitrarily small intervals including it), and *ii.* expressing a property of the configuration in the complement of this site. The probability of the former event (which is 0 for any a-priori specified t) has density h with respect to dt. A simple approximation argument can be used to show that event *ii.* is asymptotically independent of *i.* when the uncertainty interval is shrunk to a point. This yields the identity:

$$\mathbb{E}\left(1_A\,\delta(Y - (x,t))\right) \;=\; h\,\mathbb{P}\left((x,t) \text{ is green pivotal for } E\right)\,, \qquad (3.5)$$

which is to be interpreted in a distributional sense.

Thus,

$$\begin{aligned} \mathbb{P}\left(|C \cap G| = 1\right) &= \mathbb{E}\left(1_A\right) \;=\; \sum_x \int_{-\infty}^0 \mathbb{E}\left(1_A\,\delta(Y - (x,t))\right)\,dt \\ &= h\sum_x \int_{-\infty}^0 \mathbb{P}\left((x,t) \text{ is green pivotal for } E\right)\,dt \qquad (3.6) \\ &= h\sum_x \int_{-\infty}^0 \left.\frac{\delta\theta(\lambda,h)}{\delta h_x(t)}\right|_{h(\cdot)\equiv h}\,dt\,. \end{aligned}$$

where the last step is by (3.2). Now, in case the total time duration of the space $\times$ time graph is finite, the last expression yields $h\frac{\partial\theta}{\partial h}$, and thus we obtain (3.3). In this step we are applying the Definition 2.1 and Lemma 2.3 with α chosen to be

the Lebesgue measure dt. This argument is, however, limited by the restriction in Lemma 2.3 that the variational derivative α be finite.

3.2. *Differential inequalities at finite cutoffs.* As we just saw, certain technical issues need to be addressed in order to carry the analysis directly for an infinite graph, e.g., the decomposition of $\frac{\partial}{\partial h}\theta$ into the sum which appears at the last line of (3.6) is valid only if $\sum_x \int 1 dt < \infty$. We shall circumvent this problem through finite time and space cutoffs.

In order to apply arguments like the one seen above, we let $\theta_{T,L}$ the probability that infection is present at the origin o at time 0 due to a spontaneous infection event which has occurred within the finite time interval $(-T, 0]$ at some site within $V_L := \{x \in V : |x| \leq L\}$. Equivalently, $\theta_{T,L}$ is the infection probability at $(o, 0)$ for the finite subgraph $\mathcal{G}_L$ with the vertex set V_L, in the state which results from $A_{-T} = \emptyset$. More generally, the infection probability in this state at (x, t) is denoted by $\theta_{T,L}(x, t)$, and we let $\theta_{T,L}^{max} := \max_{x \in V_L} \theta_{T,L}(x, 0)$.

Due to the abundance of parameters, the dependence of the above quantities on (λ, h) will occasionally be suppressed in the notation. As a step towards Theorem 1.6 we first derive the following finite-volume version.

Theorem 3.2. *On the finite graph, $\mathcal{G}_L \times [-T, 0]$, the infection density introduced above $\theta_{T,L} \equiv \theta_{T,L}(\lambda, h)$, satisfies for $\lambda \geq 0$ and $h > 0$:*

$$\frac{\partial}{\partial\lambda}\theta_{T,L} \ \leq \ |J|\,\theta_{T,L}^{max}\,\frac{\partial}{\partial h}\theta_{T,L} \tag{3.7}$$

and

$$\theta_{T,L} \ \leq \ h\frac{\partial}{\partial h}\theta_{T,L} \ + \big(2\lambda^2|J|\theta_{T,L}^{max} + h\lambda\big)\,\frac{\partial}{\partial\lambda}\theta_{T,L} \ + \ [\theta_{T,L}^{max}]^2, \tag{3.8}$$

This statement is proven in the rest of this section. We start with the first inequality, using the dictionary provided by Lemma 3.1.

Proof of (3.7). Applying equation (3.1):

$$\frac{\partial}{\partial\lambda}\theta_{T,L} \ = \ \sum_{y,x \in V_L} \int_{-T}^0 \frac{\delta\theta_{T,L}(\lambda, h)}{\delta\lambda_{y,x}(t)}\bigg|_{\lambda(\cdot)\equiv\lambda} dt \tag{3.9}$$

$$= \ \sum_{y,x \in V_L} J_{y,x} \int_{-T}^0 \mathbb{P}_{\lambda,h}(e_{yx}^t \text{ is vacant and is bond pivotal for } E)\, dt.$$

Spelling out the condition on the right-hand side we get:

$$\frac{\partial}{\partial\lambda}\theta_{T,L} \ = \ \sum_{y,x \in V_L} J_{y,x} \int_{-T}^0 \mathbb{P}_{\lambda,h}((x, t) \in C; E^c; C(y, t) \cap G \neq \emptyset)\, dt \tag{3.10}$$

$$= \ \sum_{y,x \in V_L} J_{y,x} \int_{-T}^0 \mathbb{P}_{\lambda,h}(C(y, t) \cap G \neq \emptyset \mid (x, t) \in C; E^c) \ \times$$

$$\mathbb{P}_{\lambda,h}((x, t) \in C; E^c)\, dt \tag{3.11}$$

The conditional expectation is the average of the probability of the event that there is a connecting path from G to (y, t) in the complement of the cluster of sites which are reached from $(o, 0)$, moving back in time, without visiting (x, t). Conditioning on the exact extent of this cluster, we see that the conditional probability

is dominated by $\theta_{T,L}^{max}(\lambda, h)$. Thus,

$$
\frac{\partial}{\partial \lambda} \theta_{T,L}(\lambda, h) \;\leq\; \theta_{T,L}^{max}(\lambda, h) \sum_{y,x \in V_L} J_{y,x} \int_{-T}^{0} \mathbb{P}_{\lambda,h}((x,t) \in C; E^c) \, dt \quad (3.12)
$$

$$
=\; \theta_{T,L}^{max}(\lambda, h) \, |J| \sum_{x \in V_L} \int_{-T}^{0} \mathbb{P}_{\lambda,h}((x,t) \in C; E^c) \, dt \quad (3.13)
$$

$$
=\; \theta_{T,L}^{max}(\lambda, h) \, |J| \, \frac{\partial}{\partial h} \theta_{T,L}(\lambda, h), \quad (3.14)
$$

which is the statement we wanted to show. $\square$

   Inequality (3.8) is a bit more involved. Start by breaking the event $\{C \cap G \neq \emptyset\}$ into two cases:

$$
A \;=\; \{C \text{ has exactly one green site}\}
$$
$$
B \;=\; \{|C \cap G| \geq 2\}.
$$

By Lemma 3.1, we have $\mathbb{P}_{\lambda,h}(A) = h\frac{\partial}{\partial h}\theta_{T,L}$. To estimate the probability of B, we split it further. Let us define a "gate" as a space $\times$ time point which, in a given configuration, needs to be visited by all paths which connect G to $(o,0)$. It is easy to see that:

   (a) The collection of gates is well ordered.
   (b) The *last gate* is either a *green site* or a vertex of a *transmission bond*, "or" taken in a non-exclusive sense (though the probability that both occur is zero.) We denote the former event as B_s and the latter as B_b.

Lemma 3.3.
$$
\mathbb{P}_{\lambda,h}(B_s) \;\leq\; \frac{\theta_{T,L}^{max}}{1 - \theta_{T,L}^{max}} \times h \, \frac{\partial}{\partial h} \theta_{T,L} \quad (3.15)
$$

Proof. As was discussed already, the last factor in (3.3) coincides with $\mathbb{P}_{\lambda,h}(A)$. The probabilities of B_s and A will be compared here through the probability densities for the uniquely defined 'markers' for the two events. For A, that role is played by the unique green site in C, and the corresponding decomposition of its probability is given by equations (3.5) and (3.6). For B_s we note that conditioned on it there is a unique site $W(\omega) \in V_L \times [-T, 0]$ for which the event $W(\omega) = (x,t)$ has the following characteristics

   (W1) $(x,t) \in G$
   (W2) $(x,t) \in C$ and there is no green site within the cluster of sites from which $(o,0)$ can be reached without visiting (x,t)
   (W3) there is a green site connected to (x,t) by a path in the complement of the above cluster.

The condition (W1) has an infinitesimal probabiliy, of density h with respect to dt. Let $\Phi_{x,t}$ denote the cluster described in (W2) and $\theta(x,t)_{\Phi^c}$ be the probability of the event (W3) conditioned on that cluster. We have the following analog of equation (3.5):

$$
\mathbb{E}\left(1_{B_s}\,\delta(W - (x,t))\right) \;=\; h\,\mathbb{E}\left(1_{[\Phi_{x,t}\cap G=\emptyset]}\,\theta(x,t)_{\Phi^c}\right)
$$
$$
\leq\; h\,\mathbb{E}\left(1_{[\Phi_{x,t}\cap G=\emptyset]}\,(1 - \theta(x,t)_{\Phi^c})\right) \times \frac{\theta_{max}}{1 - \theta_{max}} \quad (3.16)
$$

where we used the fact that $0 \leq \theta(x,t)_{\Phi^c} \leq \theta_{max} \leq 1$ and hence

$$\frac{\theta(x,t)_{\Phi^c}}{1-\theta(x,t)_{\Phi^c}} \leq \frac{\theta_{max}}{1-\theta_{max}} \tag{3.17}$$

Now, as is easily seen,

$$\mathbb{E}\left(1_{[(x,t)\in C]}\, 1_{[C\cap G=\emptyset]}\left[1-\theta(x,t)_{\Phi_{\widetilde{W}}^c}\right]\right) = \mathbb{P}((x,t)\in C;\; C\cap G=\emptyset)$$
$$= \mathbb{P}((x,t)\text{ is green pivotal for }E) \tag{3.18}$$

Putting it together, we get the following analog of (3.6)

$$\mathbb{P}(B_s) = \sum_{x\in V_L}\int_{-T}^{0}\mathbb{E}\left(1_{B_s}\,\delta(W-(x,t))\right)\,dt$$
$$\leq h\sum_{x\in V_L}\int_{-T}^{0}\mathbb{P}((x,t)\text{ is green pivotal for }E)\,dt \times \frac{\theta_{max}}{1-\theta_{max}}$$
$$= \frac{\theta_{max}}{1-\theta_{max}} \times h\frac{\partial}{\partial h}\theta_{T,L} \tag{3.19}$$

which proves the lemma, through a comparison with (3.6). $\square$

Now we come to the trickiest estimate:

Lemma 3.4.

$$\mathbb{P}_{\lambda,h}(B_b)(\lambda,h) \leq \left[2\lambda^2|J|\theta_{T,L}^{max}(\lambda,h)+h\lambda\right]\frac{\partial}{\partial\lambda}\theta_{T,L}(\lambda,h) \tag{3.20}$$

Proof. As in the last proof, we shall compare $\mathbb{P}(B_b)$ with $\frac{\partial}{\partial\lambda}\theta_{T,L}$ by expressing each of the quantities as integrals, with simple bounds relating the two integrands. The probability of B_b would be decomposed similarly to that of B_s there, except that we shall also integrate over the specifics of the last event occurring at x before the time t.

For a site (x,t), let $\tau_{x,t}$ be the time of the last event at x preceding t, which can be either healing, spontaneous infection, or an infection-transmission event into x. By properties of the Poisson distribution, for (x,t) specified: $\mathbb{P}(t-\tau_{x,t}\geq u) = e^{-(1+h+\lambda|J|)u}$, and conditioned on the value of $\tau_{x,t}$ the probabilities of the three possibilities for the event, have the ratios $1:h:\lambda|J|$.

If $\Phi_{x,t}$ is as in the proof of the previous lemma, then let $K_{x,t}^{(1)}$, $K_{x,t}^{(2)}$, and $K_{x,t}^{(3)}$ be the events that: $\{(x,t)\in C \text{ and } \Phi_{x,t}\cap G=\emptyset\}$ and the last event at x preceding t is correspondingly: healing, spontaneous infection, or an infection-transmission event. Due to the independence of future from the past events, we have:

$$\mathbb{P}\left(K_{x,t}^1\right):\mathbb{P}\left(K_{x,t}^2\right):\mathbb{P}\left(K_{x,t}^3\right) = 1:h:\lambda|J| \tag{3.21}$$

Now, if the event B_b occurs there is a unique bond $\widetilde{W}$ for which the event $\widetilde{W}=e_{yx}^t$ is characterized by the conditions:

($\widetilde{W}1$) the bond e_{yx}^t is realized as an infection-transmission event,

($\widetilde{W}2$) there is no green site which connects to $(o,0)$ without visiting (x,t)

($\widetilde{W}3$) the sites (x,t) and (y,t) are reached by a pair of disjoint paths from distinct green sites g_x and g_y, both in the complement of $\Phi_{x,t}$.

The existence of a site with the above characteristics is in fact equivalent to the event B_b. Thus $\mathbb{P}(B_b)$ can be written, in a form similar to (3.6), as a sum of integrals of $\mathbb{E}\left(\delta(\widetilde{W} - e_{yx}^t)\right)$. Splitting that further according to the characteristics of the last event at x preceding t, for which $K_{x,t}^{(1)}$ is not an option, we get:

$$\mathbb{P}(B_b) = \sum_{\substack{x,y \in V_L \\ k=2,3}} \lambda J_{y,x} \times \tag{3.22}$$

$$\int_{-T}^{0} \mathbb{P}\left(K_{x,t}^{(k)}\right) \mathbb{P}\left((x,\tau_{x,t}) \text{ and } (y,t) \text{ are disjointly connected to } G \mid K_{x,t}^{(k)}\right) dt .$$

In the statement that $(x,\tau_{x,t})$ and (y,t) are disjointly connected to G, it is possible that one of the paths has trivial length, e. g. the event $K_{x,t}^{(2)}$.

For $k = 2$ the conditional probability in the last expression satisfies:

$$\mathbb{P}\left((x,\tau_{x,t}) \text{ and } (y,t) \text{ are disjointly connected to } G \mid K_{x,t}^{(2)}\right) =$$

$$= \mathbb{P}\left((y,t) \text{ is connected to } G \text{ by a path avoiding } (x,\tau_{x,t}) \mid K_{x,t}^{(2)}\right) \tag{3.23}$$

$$= \mathbb{P}\left((y,t) \text{ is connected to } G \text{ by a path avoiding } (x,\tau_{x,t}) \mid K_{x,t}^{(1)}\right) .$$

The first equality holds since under the condition $K_{x,t}^{(2)}$ the site $(x,\tau_{x,t})$ is itself green, and the second equality expresses the fact that the conditional probability is not affected by the type of event which occurs at $(x,\tau_{x,t})$.

For $k = 3$ the condition that $(x,\tau_{x,t})$ is infected can be met in two ways, since the site is at the end of an infection transmitting bond. By the van den Berg - Kesten inequality van den Berg and Kesten (1985), which applies to independent systems, the probability of the disjoint occurrence of two events is dominated by the product of their separate probabilities. Peeling off one of the factors, and then switching the value of k, we obtain:

$$\mathbb{P}\left((x,\tau_{x,t}) \text{ and } (y,t) \text{ are disjointly connected to } G \mid K_{x,t}^{(3)}\right) \leq$$

$$\leq 2\theta_{T,L}^{max} \times \mathbb{P}\left((y,t) \text{ is connected to } G \text{ by a path avoiding } (x,\tau_{x,t}) \mid K_{x,t}^{(1)}\right) \tag{3.24}$$

After the above bounds are inserted in (3.22), we use (3.21) to change also the value of k in the factor $\mathbb{P}\left(K_{x,t}^{(k)}\right)$ appearing there. This results in:

$$\mathbb{P}(B_b) \leq [h + 2\theta_{T,L}^{max} \lambda |J|] \sum_{x,y \in V_L} \lambda J_{y,x} \times$$

$$\int_{-T}^{0} \mathbb{P}\left(K_{x,t}^{(1)}; \{(y,t) \text{ is connected to } G \text{ by a path avoiding } (x,\tau_{x,t})\}\right) dt$$

$$\leq [h + 2\theta_{T,L}^{max} \lambda |J|] \sum_{x,y \in V_L} \lambda J_{y,x} \int_{-T}^{0} \mathbb{P}\left(e_{(x,t)}^t \text{ is bond pivotal for } E\right) dt$$

$$= [h + 2\theta_{T,L}^{max} \lambda |J|] \lambda \frac{\partial}{\partial \lambda}\theta_{T,L} . \tag{3.25}$$

where the last equation is by (3.9). $\qquad\qquad\square$

Proof of (3.8). Putting the above together, we have:

$$
\begin{aligned}
\theta_{T,L} \;&\le\; \mathbb{P}_{\lambda,h}(A) \;+\; \mathbb{P}_{\lambda,h}(B_s) \;+\; \mathbb{P}_{\lambda,h}(B_b) \\
&\le\; h\frac{\partial \theta_{T,L}}{\partial h} \;+\; h\frac{\theta_{T,L}^{max}}{1-\theta_{T,L}^{max}}\frac{\partial \theta_{T,L}}{\partial h} \;+\; \left(2\lambda^2|J|\theta_{T,L}^{max}+h\lambda\right)\frac{\partial \theta_{T,L}}{\partial \lambda}\,. \quad (3.26)
\end{aligned}
$$

Collecting the first two terms and multiplying through by $(1-\theta_{T,L}^{max})$, one gets (3.8). $\qquad\square$

This concludes the proof of Theorem 3.2.

4. Analysis: from the PDI to the critical behavior

We shall now extend the inequalities of Thm 3.2 to the infinite-volume (Theorem 1.6), and then explain how they yield the main results stated in the introduction.

Proof of Theorem 1.6. By the monotonicity of the contact process,

$$
\theta_{T,L}^{max}(\lambda,h) \;\le\; \theta_{T,2L}(\lambda,h) \;\le\; \lim_{T',L'\to\infty}\theta_{T',L'}(\lambda,h) \;=\; \theta(\lambda,h)\,. \qquad (4.1)
$$

This relation permits us to simplify (linearize) the problem of passage to the limit, by replacing $\theta_{T,L}^{max}$ in the inequalities (3.7) and (3.8) by the limiting function θ. We get

$$
\frac{\partial}{\partial \lambda}\theta_{T,L}(\lambda,h) \;\le\; \theta(\lambda,h)\,|J|\,\frac{\partial}{\partial h}\theta_{T,L}(\lambda,h) \qquad (4.2)
$$

and

$$
\theta_{T,L} \;\le\; h\frac{\partial}{\partial h}\theta_{T,L} \;+\; \left(2\lambda^2|J|\,\theta+h\lambda\right)\frac{\partial}{\partial \lambda}\theta_{T,L} \;+\; \theta^2\,. \qquad (4.3)
$$

The finite-volume quantities are differentiable for $h>0$ (in fact analytic in $(\lambda,\cdot)$). In order to take the limit, we shall interpret the inequalities in a weaker sense, as indicators of the corresponding relations for integrals of the quantities over $d\lambda\,dh$ against suitable test functions. General arguments permit to conclude that in this sense the inequalities remain valid also in the limit.

More explicitly, through integration by parts (4.2) can be expressed as the relation of the following Stieltjes integrals (each over $\mathbb{R}_+$) with positive, compactly supported, test functions $g\in C_o(\mathbb{R}_+,\mathbb{R}_+)$

$$
-\int\left[\int\theta_{T,L}(\lambda,h)\,dg\right]dh \;\le\; -\int\left[\int|J|\,\theta_{T,L}(\lambda,h)\,d[g\theta]\right]d\lambda\,, \qquad (4.4)
$$

where dg on the left is a Stieltjes integral at fixed h and $d[g\theta]$ on the right is a Stieltjes integral at fixed λ. By the bounded convergence theorem, as $T,L\to\infty$, the integrals converge to those of the limit. Since the limiting function is also monotone in its arguments (λ,h), the integration by parts can be reversed in the limit. The ultimate conclusion, allowed since the derivatives of monotone functions are locally absolutely integrable, is that the limiting inequality (1.19) holds in the sense of a relation holding at Lebesgue almost every (λ,h). A similar argument permits to deduce (1.20) from (4.3), thereby proving Theorem 1.6. $\qquad\square$

The inequalities which are established in Theorem 1.6 are very close to what was proven in (Aizenman and Barsky, 1987) for the model's discrete-time version on $\mathbb{Z}^d$. From this point on, the analysis of the PDI is identical, and it is covered by the general results of Lemma 4.1 and Lemma 5.1 of (Aizenman and Barsky, 1987),

which yield the following statement (formulated here in the notation of (Aizenman and Barsky, 1987)).

Proposition 4.1. *Let $M(\beta, h) : \mathbb{R}^2 \mapsto \mathbb{R}$ be a positive function which for $h = 0$ is continuous from above and for $h > 0$ is continuous, increasing in each of its arguments, and satisfies (in the a.e. sense):*

$$\frac{\partial M}{\partial \beta} \quad \leq \quad \phi \, M \, \frac{\partial M}{\partial h} \tag{4.5}$$

$$M \quad \leq \quad h \frac{\partial M}{\partial h} \; + \psi \, M^a \, \frac{\partial M}{\partial \beta} \; + \; M^2 \,, \tag{4.6}$$

with some $0 < a < \infty$ and some $\phi(\beta, h)$ and $\psi(\beta, h)$ which are finite on compact subsets of $\mathbb{R}_+ \times \mathbb{R}_+$. If there exists a value β_0 for which

$$\lim_{h \searrow 0} M(\beta_0, h)/h \; = \; \infty \tag{4.7}$$

then for $h \searrow 0$

$$M(\beta_0, h) \; \geq \; c_1 \, h^{1/(1+a)} \tag{4.8}$$

and for $\beta \geq \beta_0$

$$M(\beta, 0) \; \geq \; c_2 \, |\beta - \beta_0|_+^{1/a} \,, \tag{4.9}$$

with some $c_1, c_2 < \infty$.

Remark: Since at first glance it may appear surprising that hard information about the critical behavior can be obtained from "soft " inequalities like (4.5) and (4.6), let us outline here the heuristics behind Theorem 4.1.

First, combining (4.5) and (4.6), one gets:

$$M \; \leq \; h \frac{\partial M}{\partial h} \; + \phi(\beta, h) \, \psi(\beta, h) \, M^{(1+a)} \, \frac{\partial M}{\partial h} \; + \; M^2 \,. \tag{4.10}$$

We shall apply this relation to study the h dependence at small h in the vicinity of β_0, which is analogous to our λ_T.

It may be noted that the inequality (4.10) does not add much information about the regime where M is linear in h since there $h \frac{\partial}{\partial h} M \approx M$, and thus already the first term on the right accounts for the left side. However, at β_0 the dependence of M on h is singular, and may be given by a power law: $M(\beta_0, h) \approx h^{1/\delta}$, with some $\delta > 1$ (the physicists convention for the corresponding exponent). For a shortcut, which is of course not made in the actual proof, let us allow such an assumption – taken in the literal sense that $h \frac{\partial}{\partial h} M \approx \frac{1}{\delta} M$. We now see that at β_0 and h small, (4.10) holds not because of the first term on the right, but due to the presence of the second:

$$(1 - \frac{1}{\delta})M \; \leq \; \phi(\beta, h) \, \psi(\beta, h) \, M^{(1+a)} \, \frac{\partial M}{\partial h} \; + \; o(M) \,, \tag{4.11}$$

Dividing by M and integrating from $h = 0$ up, one gets (4.8), which leads to the interesting conclusion that there is a gap in the allowed values of the exponent by which M may vanish: it either vanishes linearly in h or at a slower power, $1/\delta$, with $\delta \geq 1 + a$ (in our case $a = 1$).

Once it is known that for $\beta \geq \beta_0$: $M(\beta, h) \geq c \, h^{1/(1+a)}$, a similar treatment of (4.6) yields for that regime

$$(1 - \frac{1}{1 + a})M \; \leq \; \psi(\beta, 0) \, M \, \frac{\partial M}{\partial \beta} \; + \; o(M) \,. \tag{4.12}$$

Dividing by M, and integrating from β_0 upward, one gets (4.9). In particular, one learns that

$M(\beta, 0+) > 0$ for any $\beta > \beta_0$!

The complete proof of Theorem 4.1, which does not rely on the power law assumption, can be obtained through the integration of the inequalities (4.5) and (4.6) through suitable regimes in the (β, h) plane, as is done in Lemmas 4.1 and 5.1 of Aizenman and Barsky (1987).

Proofs of Theorem 1.3 and Theorem 1.5. The two statements follow now by applying the principle expressed in Theorem 4.1, to the inequalities of Theorem 1.6, with the correspondence:

$(\theta, \lambda, h) \mapsto (M, \beta, h)$. For this purpose we note that by a simple estimate of the contact process on a graph with only one vertex, $\theta(\lambda, h) \geq h(1+h)$ or equivalently $h \leq \theta/(1-\theta)$, and hence inequality (1.20) can be brought to the form (4.6) with $\psi = 2\lambda^2 |J| + \lambda/(1-\theta)$ and $a = 1$. $\square$

5. Remarks

1. The results presented here can be extended also to graphs which are only quasi-transitive in the following sense. The analysis can be adapted as long as it can be shown that for each bounded region in the (λ, h) plane, there are $0 < c_1 \leq c_2 < \infty$ such that

$$\mathbb{P}_{\lambda, h}(C(x, t) \cap G \neq \emptyset) \in [c_1 \theta(\lambda, h), c_2 \theta(\lambda, h)] \qquad (5.1)$$

uniformly in (x, t). In particular, the conclusions of Theorem 1.3 hold under this 'weak inhomogeneity' condition.

2. Among the cases for which Theorem 1.3 applies to are the many graphs of exponential growth which are the subject of current research. These include hyperbolic tessellations, Cayley graphs of non-amenable groups, and exponentially growing amenable graphs such as the lamplighter group and the Diestel-Leader graph (see Lyons and Peres (2007) and references therein.) An example for which the discrete-time version of Theorem 1.3 was recently applied is the thermodynamic limit of the small-world graphs, see Durrett and Jung (2006).

3. The method and results presented here apply also to unoriented percolation models on transitive graphs similar to those considered here, i.e. $\mathcal{G} \times \mathbb{R}$ with one continuum dimension. Similar independence of the argument from the presence of orientation was noted in the previous related results on the contact process (Aizenman and Barsky, 1987; Bezuidenhout and Grimmett, 1991).

4. A topic which our discussion did not address is whether in addition to the general Properties 1.-4. it is also true, for contact process in the generality considered here, that the upper stationary infection density vanishes at the critical point. An equivalent formulation is that at $\lambda = \lambda_c$ infection from a single site will almost surely die out. Such a statement was established for $\mathcal{G} = \mathbb{Z}^d$ in the celebrated work of Bezuidenhout and Grimmett (1990), and their arguments can most likely be extended to all graphs of subexponential growth satisfying a certain homogeneity involving block structures. The only related results known to the authors for contact processes on graphs of exponential growth are those of (Morrow et al., 1994) – where the corresponding statement is proven for regular trees, and the corresponding statement for regular percolation on Cayley graphs of non-amenable groups, of (Benjamini et al., 1999).

5. Finally, we note that Theorem 1.3 allows to sharpen a statement which was derived in Morrow et al. (1994). As a step towards the proof that infection from a single site dies out almost surely at λ_H, it is shown there, for contact processes on tree graphs, that

$$\exp(\xi(\lambda)\,t) \leq \mathbb{E}^{(\{o\},0)}(|A_t|) \leq c\exp(\xi(\lambda)\,t)\,. \tag{5.2}$$

at some continuous $\xi(\lambda)$, and $c < \infty$. It is not difficult to see that $\xi(\lambda) > 0$ for $\lambda > \lambda_H$, and by Proposition 1.2 $\xi(\lambda) < 0$ for $\lambda < \lambda_T$. Thus, Theorem 1.3 $(\lambda_H = \lambda_T)$ allows to conclude that $\xi(\cdot)$ actually changes sign at the transition point.

Appendix A. Exponential decay in the subcritical regime

For completeness, we provide here a proof that throughout the regime $\lambda < \lambda_T$, which is characterized by $\chi(\lambda) < \infty$, the probability that the infection, if is introduced at a single site, would persist for time t and/or spread over distance L decays exponentially in t and L. The proof uses generally known arguments.

Proof of Proposition 1.2. By the additivity of the contact process and the transitivity of $\mathcal{G}$,

$$\mathbb{E}^{(\{o\},0)}(|A_{t+s}|) \leq \mathbb{E}^{(\{o\},0)}(|A_t|)\mathbb{E}^{(\{o\},0)}(|A_s|)\,. \tag{A.1}$$

Subadditivity arguments permit to conclude that

$$\lim_{t\to\infty} \frac{1}{t}\log\mathbb{E}^{(\{o\},0)}(|A_t|) = \inf_{t>0}\frac{1}{t}\log\mathbb{E}^{(\{o\},0)}(|A_t|) = \eta \tag{A.2}$$

exists so that $\exp(\eta t) \leq \mathbb{E}^{(\{o\},0)}(|A_t|)$. Since $\lambda < \lambda_T$, it must be that $\eta < 0$. For $0 < \delta < -\eta$ we can find $\bar{t}$ so that $\mathbb{E}^{(\{o\},0)}(|A_t|) < \exp((\eta+\delta)t)$ for all $t > \bar{t}$. Letting $1/\tau = -(\eta+\delta)$ and choosing c large enough completes the proof of (1.7).

For (1.8), we consider a process $F_t^{\{o\}}$ which is defined to be the contact process ignoring all healing events. In particular $A_s^{\{o\}} \subset F_s^{\{o\}} \subset F_t^{\{o\}}$ for all $s < t$. We have that the left-hand side of (1.8) is bounded by

$$\mathbb{P}^{(\{o\},0)}(x \in F_t \text{ for some } |x| > r) + \mathbb{P}^{(\{o\},0)}(A_t \neq \emptyset)\,.$$

Coupling $F_t^{\{o\}}$ with a branching random walk starting from one particle at the origin gives the bound

$$\mathbb{P}^{(\{o\},0)}(x \in F_t) < e^{Kt\lambda}p_t(o,x) \tag{A.3}$$

for some constant K where $p_t(o,x)$ is the transition probability of a random walk on $\mathcal{G}$ (note that this is a general bound which does not require $\lambda < \lambda_c$). A standard large deviations result which holds for $\sum_x J_{o,x}\,e^{+\varepsilon|x|} < \infty$, says that for all $\xi > 0$ there is a $c > 0$ such that $\sum_{|x|>ut} p_t(o,x) \leq ce^{-\xi t}$. This together with (A.3) gives an exponentially decaying bound on $\mathbb{P}^{(\{o\},0)}(F_{r/u} \cap B_r^c \neq \emptyset)$, whereas (1.7) implies that $\mathbb{P}^{(\{o\},0)}(A_{r/u} \neq \emptyset)$ decays exponentially in r. $\qquad\square$

Appendix B. Uniqueness of the invariant measure in the presence of spontaneous infection

In this appendix we prove the basic regularity properties which were asserted for the process at $h > 0$ and its relation with the standard $h = 0$ version of the model.

Proof of Lemma 1.4. i.   In the presence of spontaneous infection, i.e., for $h > 0$, the introduction of initial infection at time $-T < 0$, on a set S, will have negligible effect on the infection at time 0 as $T \to \infty$. To see this, we split the probability of infection at a site at time 0 into two cases, *(a)* when it can be accounted for by a spontaneous infection event, and *(b)* when it can be present only due to the initial conditions:

$$\mathbb{P}^{(B,T)}(o \in A_0) \ = \ \mathbb{P}\left(C(o,0)\right) \cap G \cap V \times [-T,0] \neq \emptyset) + \tag{B.1}$$
$$+ \mathbb{P}\left(C(o,0)\right) \cap G \cap V \times [-T,0] = \emptyset; \ C(o,0) \cap B \times \{-T\} \neq \emptyset)$$

The second term is negligibly small in the limit $T \to \infty$, since

$$\mathbb{P}\left(C(o,0) \cap G \cap V \times [-T,0] = \emptyset; \ C(o,0) \cap S \times \{-T\} \neq \emptyset\right) \ =$$
$$= \ \mathbb{E}\left(e^{-h|C(o,0)\cap G\cap V\times[-T,0]|}; \ C(o,0) \cap S \times \{-T\} \neq \emptyset\right) \leq \ e^{-hT} \tag{B.2}$$

which is obtained by first conditioning on the percolation structure, i.e., the bond variables and the healing events. Thus:

$$\lim_{T\to\infty} \mathbb{P}^{(B,-T)}(o \in A_0) \quad = \quad \mathbb{P}\left(C(o,0) \cap G \neq \emptyset\right) \tag{B.3}$$
$$= \quad \mathbb{E}\left(1 - e^{-h|C(o,0)|}\right) \ \equiv \ \theta(\lambda,h).$$

This implies assertion *i*.

ii. The monotonicity of $\theta(\lambda,h)$ is a standard observation (and is valid also for the approximating functions), and the continuous differentiability of $\theta(\lambda,h)$ in h, for $h > 0$ is an easy consequence (B.3). We turn our attention to the continuity of $\theta(\lambda,h)$ in λ.

As explained above, the probability that events occurring earlier that T ago are of relevance is bounded by e^{-hT}. Restricting to times $[-T,0]$: the probability that the cluster $C \cap V \times [-T,0]$ reaches a site with $|x| \geq uT$ can in turn be bounded by the estimates which are used in Appendix A. These show that, under the assumption which is made on $\{J_{x,y}\}$, there exists $0 < u$ such that the probability that the infection reaches the origin from a site at distance greater than uT is dominated by $ce^{-\xi T}$. Hence, for $L = uT$:

$$|\theta(\lambda,h) - \theta_{T,L}(\lambda,h)| \ \leq \ e^{-hT} \ + ce^{-\xi T}. \tag{B.4}$$

The continuity of $\theta_{T,L}(\lambda,h)$ in λ is obvious (for a detailed argument see (Liggett, 1999)). Since (B.4) shows that $\theta_{T,L}(\cdot,h)$ converges uniformly to $\theta(\cdot,h)$ on $\mathbb{R}_+$, it must be that θ is continuous in λ.

iii. The representation (B.3) readily implies, via the monotone convergence theorem, that

$$\theta(\lambda,0+) \ := \ \lim_{h\searrow 0} \theta(\lambda,h) \ = \ \mathbb{P}\left(|C(o,0)| = \infty\right) \tag{B.5}$$

and for $\lambda < \lambda_H$,

$$\chi(\lambda) = \lim_{h\searrow 0} \frac{\partial\theta(\lambda,h)}{\partial h} \ = \ \mathbb{E}\left(|C(o,0)|1_{[|C(o,0)|\neq\infty]}\right). \tag{B.6}$$

The event $\{|C(o,0)| = \infty\}$ does not coincide with $\cap_{T<0}\{o \in A_0^{(V,T)}\}$ however we claim that the difference is of probability zero. More explicitly:

$$0 \ \leq \ \theta(\lambda,0+) - \theta_+(\lambda) \tag{B.7}$$
$$\leq \ \mathbb{P}\left(|C(o,0)| = \infty, \ \exists T \in (-\infty,0) : C(o,0) \cap V \times \{T\} = \emptyset\right) \ = \ 0$$

since it is easy to see that if $|C(o, 0) \cap V \times [T_1, 0]| = \infty$, for some $T_1 < 0$, then with probability 1: $C(o, 0) \cap V \times \{T\} \neq \emptyset$ for all $T < 0$. $\qquad \square$

The imbedding of the contact process within the its extended two-parameter version through the relations established in this Appendix plays a fundamental role in our analysis. In effect, it allows to relate the regimes of $\lambda < \lambda_T$ and $\lambda > \lambda_H$, at $h = 0$, by exploring the model along contours in the half plane $\{(\lambda, h) \in \mathbb{R}_+^2 : h > 0\}$.

References

M. Aizenman and D. J. Barsky. Sharpness of the phase transition in percolation models. *Comm. Math. Phys.* **108**, 489–526 (1987).

M. Aizenman, D. J. Barsky and R. Fernández. The phase transition in a general class of Ising-type models is sharp. *J. Statist. Phys.* **47**, 343–374 (1987).

M. Aizenman and C. M. Newman. Tree graph inequalities and critical behavior in percolation models. *J. Statist. Phys.* **36**, 107–143 (1984).

F. Bacelli, M. Klein and S. A. Zuev. Perturbation analysis of functionals of random measures. *Adv. Appl. Probab.* **27**, 306–325 (1995).

D. J. Barsky and M. Aizenman. Percolation critical exponents under the triangle condition. *Ann. Probab.* **19**, 1520–1536 (1991).

D. J. Barsky and C. C. Wu. Critical exponents for the contact process under the triangle condition. *J. Statist. Phys.* **91**, 95–124 (1998).

I. Benjamini, R. Lyons, Y. Peres and O. Schramm. Critical percolation on any nonamenable group has no infinite clusters. *Ann. Probab.* **27**, 1347–1356 (1999).

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

C. Bezuidenhout and G. Grimmett. Exponential decay for subcritical contact and percolation processes. *Ann. Probab.* **19**, 984–1009 (1991).

A. A. Borovkov. Asymptotic expansions for functionals of dilation of point processes. *J. Appl. Probab.* **33**, 573–591 (1996).

J.T. Chayes and L. Chayes. An inequality for the infinite cluster density in bernoulli percolation. *Phys. Rev. Lett.* **56**, 1619–1622 (1986).

R. Durrett and P. Jung. Two phase transitions for the contact process on small worlds (2006). Submitted for publication.

G. Grimmett and C. Newman. Percolation in $\infty + 1$ dimensions. In *Disorder in Physical Systems*, pages 167–190. Oxford University Press, New York (1990).

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

T. M. Liggett. Multiple transition points for the contact process on the binary tree. *Ann. Probab.* **24**, 1675–1710 (1996).

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

R. Lyons and Y. Peres. Probability on Trees and Networks (2007). Available at `http://mypage.iu.edu/~rdlyons/prbtree/prbtree.html`.

M. V. Menshikov. Coincidence of the critical points in percolation problems. *Soviet Math. Dokl.* **33**, 856–859 (1986).

G. J. Morrow, R. B. Schinazi and Y. Zhang. The critical contact process on a homogeneous tree. *J. Appl. Probab.* **31**, 250–255 (1994).

B. G. Nguyen and W. S. Yang. Triangle condition for oriented percolation in high dimensions. *Ann. Probab.* **21**, 1809–1844 (1993).

R. Pemantle. The contact process on trees. *Ann. Probab.* **20**, 2089–2116 (1992).

A. Sakai. Mean-field critical behavior for the contact process. *J. Statist. Phys.* **104**, 111–143 (2001).

R. H. Schonmann. The triangle condition for contact processes on homogeneous trees. *J. Stat. Phys.* **90**, 1429–1440 (1998).

A. M. Stacey. The existence of an intermediate phase for the contact process on trees. *Ann. Probab.* **24**, 1711–1726 (1996).

J. van den Berg and H. Kesten. Inequalities with applications to percolation and reliability. *J. Appl. Probab.* **22**, 556–569 (1985).

R. van der Hofstad and A. Sakai. Gaussian scaling for the critical spread-out contact process above the upper critical dimension. *Electron. J. Probab.* **9**, 710–769 (2004).

C. C. Wu. The contact process on a tree: behavior near the first phase transition. *Stoch. Proc. Appl.* **57**, 99–112 (1995).

C. C. Wu. Inhomogeneous contact process on trees. *J. Stat. Phys.* **88**, 1399–1408 (1997).

S. A. Zuev. Russo's formula for Poisson point fields and its applications. *Discrete Math. Appl.* **3**, 63–73 (1993).