Survival of contact processes on the hierarchical group

Siva R. Athreya
Indian Statistical Institute
8th Mile, Mysore Road
RVCE Post
Bangalore 560 059
India
e-mail: athreya@isibang.ac.in

Jan M. Swart
Institute of Information
Theory and Automation
of the ASCR (ÚTIA)
Pod vodárenskou věží 4
18208 Praha 8
Czech Republic
e-mail: swart@utia.cas.cz

August 27, 2008

Abstract

We consider contact processes on the hierarchical group Ω_N with freedom N, where sites infect other sites at hierarchical distance k with rate $\alpha_k N^{-k}$, and sites become healthy with recovery rate δ. We show that the critical recovery rate δ_c is zero (i.e., the process dies out for any $\delta > 0$) if $\liminf_{k\to\infty} N^{-k}\log(\beta_k) = -\infty$, where $\beta_k := \sum_{n=k}^{\infty}\alpha_n$. On the other hand, in the special case that N is a power of two, we show that $\delta_c > 0$ provided that $\sum_k N^{-k}\log(\alpha_k) > -\infty$. The proof of this latter fact is based on a coupling argument that compares contact processes on Ω_2 with contact processes on a renormalized lattice.

MSC 2000. Primary: 82C22; Secondary: 60K35, 82C28.
Keywords. Contact process, survival, hierarchical group, coupling, renormalization group.
Acknowledgements. Work sponsored by GAČR grant 201/06/1323 and CSIR Grant 25(0164) /08/EMRII. Much of this work was carried out during a visit of J.M. Swart to the Indian Statistical Institute, Bangalore centre and Kolkata centre, during Nov. and Dec. 2007.

Contents

1 Introduction

1.1 Main result

Let Λ be a finite or countably infinite set, called lattice, let $(a(i,j))_{i,j\in\Lambda,\ i\neq j}$ be nonnegative constants, and $\delta \geq 0$. Then the *contact process* on Λ with *infection rates* $a(i,j)$ and *recovery rate* δ is the $\{0,1\}^\Lambda$-valued Markov process $X = (X_t)_{t\geq 0}$ with the following description. If $X_t(i) = 0$ (resp. $X_t(i) = 1$), then we say that the site $i \in \Lambda$ is *healthy* (resp. *infected*) at time $t \geq 0$. An infected site i infects a healthy site j with rate $a(i,j) \geq 0$, and infected sites become healthy with rate $\delta \geq 0$. It can be shown (see [Lig85, Prop. I.3.2]) that X is well-defined provided the infection rates are summable, in the sense that

$$|a| := \sup_{j\in\Lambda} \sum_{i\in\Lambda} a(i,j) < \infty. \tag{1.1}$$

A contact process may be used to model the spread of an infection in a spatially ordered population; see [Lig99] as a general reference. Historically, the main focus has been on nearest-neighbor contact processes. For such processes, Λ is a connected, undirected graph, and $a(i,j)$ equals some fixed constant $\lambda > 0$ if i and j are connected by an edge and is zero otherwise. We will be interested in a situation where there is no graph structure, but Λ is the hierarchical group. In the hierarchical group, sites are ordered into groups containing N sites each, groups are further ordered into 2-level groups containing N groups each, and so on. Sites are at distance one if they are in the same group, and more generally at distance k if they are in the same k-level group, but not in the same $(k-1)$-level group. Such a hierarchical structure may be a reasonable model for the spatial distribution of some species, including humans, which are organised in townships, cities, countries, and continents. Population dynamical models on the hierarchical group have been studied before in e.g. [SF83, DG93, Daw00].

A basic feature of the contact process is that it exhibits a phase transition between survival and extinction. Let $0 \in \Lambda$ be some fixed site, called *origin*. We say that a contact process on a lattice Λ with given infection rates $(a(i,j))_{i,j\in\Lambda,\ i\neq j}$ and recovery rate $\delta \geq 0$ *survives* if there is a positive probability that the process started with only the origin infected never recovers completely, i.e., if

$$\mathbb{P}^{\delta_0}[X_t \neq \underline{0}\ \forall t \geq 0] > 0, \tag{1.2}$$

where $\delta_i \in \{0,1\}^\Lambda$ is defined as $\delta_i(j) := 1$ if $i = j$ and $\delta_i(j) := 0$ otherwise, and $\underline{0} \in \{0,1\}^\Lambda$ denotes the configuration with only healthy sites. (In typical cases, e.g. when the infection rates are irreducible in an appropriate sense or if the process has some translation-invariant structure, this definition will not depend on the choice of the origin 0.)

For given infection rates, we let

$$\begin{aligned} \delta_{\rm c} := \sup\big\{\delta \geq 0 : &\text{the contact process with infection rates} \\ &(a(i,j))_{i,j\in\Lambda,\ i\neq j} \text{ and recovery rate } \delta \text{ survives}\big\} \end{aligned} \tag{1.3}$$

denote the *critical recovery rate*. A simple monotone coupling argument shows that X survives for $\delta < \delta_{\rm c}$ and dies out for $\delta > \delta_{\rm c}$.

By comparison with a subcritical branching process, it is not hard to show that $\delta_{\rm c} \leq |a|$, where $|a|$ is the constant defined in (1.1). For a large class of lattices, it is known that moreover $\delta_{\rm c} > 0$. For example, this is the case for nearest-neighbor processes on infinite graphs, or if Λ is a finitely generated, infinite group, and the infection rates are irreducible and invariant under the left action of the group [Swa07, Lemma 4.18]. On groups that are not finitely generated, the question whether $\delta_c > 0$ becomes more subtle. Inspired by a question that came up in [Swa08], the main aim of the present paper is to give sufficient conditions for $\delta_c > 0$ (resp. $\delta_c = 0$) when Λ is the hierarchical group.

By definition, the *hierarchical group with freedom* N is the set

$$\Omega_N := \big\{i = (i_0, i_1, \ldots) : i_k \in \{0, \ldots, N-1\},\ i_k \neq 0 \text{ for finitely many } k\big\}, \tag{1.4}$$

equipped with componentwise addition modulo N. We set

$$|i| := \inf\{k \geq 0 : i_m = 0\ \forall m \geq k\}, \tag{1.5}$$

and call $|i-j|$ the *hierarchical distance* between two elements $i, j \in \Omega_N$. We choose nonnegative constants $(\alpha_k)_{k\geq 1}$ and define infection rates on Ω_N by

$$a(i,j) := \alpha_{|i-j|} N^{-|i-j|} \qquad (i, j \in \Omega_N,\ i \neq j). \tag{1.6}$$

In order for the infection rates $a(i,j)$ to summable in the sense of (1.1) we must assume that $\sum_{k=1}^{\infty} \alpha_k < \infty$. Here is our main result:

Theorem 1 ((Non-) triviality of the critical recovery rate) *Let* $N \geq 2$, *let* $(\alpha_k)_{k\geq 1}$ *be nonnegative constants such that* $\sum_{k=1}^{\infty} \alpha_k < \infty$, *and let* δ_c *be the critical death rate of the contact process on* Ω_N *with infection rates as in (1.6).*

(a) *Assume that*

$$\liminf_{k\to\infty} N^{-k} \log(\beta_k) = -\infty, \quad \textit{where} \quad \beta_k := \sum_{n=k}^{\infty} \alpha_n \quad (k \geq 0). \tag{1.7}$$

Then $\delta_c = 0$.

(b) *Let* M *be the largest power of 2 such that* $M \leq N$. *Assume that*

$$\sum_{k=m}^{\infty} M^{-k} \log(\alpha_k) > -\infty \tag{1.8}$$

for some $m \geq 0$. *Then* $\delta_c > 0$.

The special role played by powers of 2 in part (b) is entirely due to our methods of proof and has no real significance. In fact, we will carry out most of our calculations for the case $N = 2$ and then generalize to the statement in part (b) by some simple comparison arguments. Note that if N is a power of 2 and

$$\alpha_k = e^{-\theta^k} \qquad (k \geq 1), \tag{1.9}$$

then our results show that $\delta_c > 0$ for $1 < \theta < N$ and $\delta_c = 0$ for $\theta > N$. There is a gap between the conditions (1.7) and (1.8). We guess that (1.8) is not necessary for $\delta_c > 0$, since this condition is violated when infinitely many of the α_k's are zero, while it seems unlikely that the latter should imply $\delta_c = 0$. We do not know if condition (1.7) is sharp.

1.2 Discussion and outline

Finding upper bounds on the critical recovery rate of a contact process is generally easier than finding lower bounds. In line with this, it turns out that the proof of Theorem 1 (a) is rather simple, but part (b) is much more involved.

Before we discuss our method for proving Theorem 1 (b), we recall how survival is proved for contact processes on other lattices. Since most of the literature deals with nearest-neighbor processes, for which there is just a single infection rate, it has become customary to fix the recovery rate to 1, consider the infection rate as a variable, and prove upper bounds on the critical infection rate. By a trivial rescaling of time, we may instead fix the infection rate and vary the recovery rate, hence any upper bound on the critical infection rate in the traditional setting can be translated into a lower bound on the critical recovery rate in our setting.

If Λ is an infinite (connected, undirected) graph, then it is always possible to embed a copy of $\mathbb{Z}$ in Λ, hence the problem can be reduced to proving survival of the nearest-neighbor contact process on $\mathbb{Z}$. (It is often possible to do better than just embedding copy of $\mathbb{Z}$ in Λ, see [Lig85, Thm VI.4.1].)

For the nearest-neighbor contact process on $\mathbb{Z}$, we are aware of two independent proofs that $\delta_{\rm c} > 0$. If the recovery rate δ is sufficiently small, then it is not hard to set up a comparison between the contact process on $\mathbb{Z}$ and oriented percolation on $\mathbb{Z} \times \mathbb{Z}_+$, with a percolation parameter p close to one. The problem can then be reduced to showing that $p_{\rm c} < 1$ for oriented percolation on $\mathbb{Z} \times \mathbb{Z}_+$, which is known to follow from a Peierls argument (see [Dur88, Chapter 5]).

An independent aproach for proving survival of the nearest-neighbor contact process on $\mathbb{Z}$, which gives a better bound on the critical value, is the method of Holley and Liggett [HL78] (see also [Lig85, Section IV.1]). Their basic observation is that if there exists a translation invariant probability law on $\{0,1\}^{\mathbb{Z}}$ such that the process X started in this initial law satisfies

$$\frac{\partial}{\partial t}\mathbb{P}[\exists i \in A \text{ s.t. } X_t(i) = 1]\big|_{t=0} \geq 0 \tag{1.10}$$

for all finite $A \subset \Lambda$, then by duality $\mathbb{P}[X_0(0) = 1]$ gives a lower bound on the survival probability of the process started with a single infected site. Holley and Liggett then explicitly construct a renewal measure that solves (1.10). Their method has been refined in [Lig95], leading to the best rigorous lower bound on $\delta_{\rm c}$ available to date.

For lattices different from $\mathbb{Z}$, there exist other, independent methods for obtaining lower bounds on $\delta_{\rm c}$. On $\mathbb{Z}^2$, one may use comparison with a stochastic Ising model. One $\mathbb{Z}^d$ with $d \geq 3$, one may use comparison with certain linear systems; this method gives the sharpest known bounds in high dimensions. (For both these techniques, see [Lig85, Section VI.4].) For processes on trees, there is a very simple lower bound on $\delta_{\rm c}$ resulting from a supermartingale argument (see [Lig99, Thm I.4.1]).

In general, one can say that proving survival for contact processes is easier in higher dimensions. In this context, returning to the hierarchical group, the following observation is useful. Let $\xi = (\xi_t)_{t\geq 0}$ be a random walk on Ω_N that jumps from a point i to j with rates $a(i,j)$ as in (1.6), with

$$\alpha_k = N^{-k(2/d)} \qquad (k \geq 1), \tag{1.11}$$

where $d > 0$ is some real constant. Then

$$P^0[\xi_t = 0] \sim \phi(\log t)\, t^{-d/2} \qquad \text{as} \quad t \to \infty, \tag{1.12}$$

where $f(t) \sim g(t)$ means that $f(t)/g(t) \to 1$ and ϕ is a positive, periodic, continuous real function. Thus, if d is an integer, then such a random walk is similar to a usual short-range random walk on $\mathbb{Z}^d$. Also, ξ is recurrent if and only if $d \leq 2$.

These observations are relevant when we consider comparison with linear systems as a method to prove survival of contact processes on the hierarchical group. Indeed, since this technique depends on the transience of the underlying random walk, for processes with rates α_k as in (1.11), it seems this technique can only work if $d > 2$. Note that our Theorem 1 (b) shows that $\delta_{\rm c} > 0$ for any $d > 0$, and in fact for processes with much faster decaying rates.

If we forget about other 'high-dimensional' techniques, this leaves us with two known techniques for establishing lower bounds on the critical recovery rate that might be succesful on the hierarchical group: comparison with oriented percolation plus a Peierls argument, or the method of Holley and Liggett.

It is not hard to set up a comparison between a contact process on Ω_N and some form of oriented percolation on $\Omega_N \times \mathbb{Z}_+$ (with a whole set of percolation parameters p_k depending on the hierarchical distance), but this only moves the problem to proving that the latter percolates if the p_k are sufficiently large. Since long-range infections are essential for survival, it is not obvious, and seems rather difficult, to define suitable contours which could then be counted and estimated in a Peierls argument.

While we did not spend much time investigating oriented percolation on $\Omega_N \times \mathbb{Z}_+$, we did spend a considerable amount of effort trying to adapt the method of Holley and Liggett. As explained in [Lig95], the renewal measure of Holley and Liggett may be interpreted as a certain type of Gibbs measure with the property that in (1.10), equality holds if A is an interval. The difficult part of the proof is then to show that this equality for intervals extends to an inequality for general subsets $A \subset \Lambda$. For the hierarchical group Ω_N, it is not hard to dream up a good analogue of Liggett's Gibbs measures and to show that (1.10) may be satisfied with equality for certain special sets. (In fact, we used blocks of sites within a given hierarchical distance of each other.) We were not able, however, to carry out the difficult step in the argument, which is to extend the equality in (1.10) for special A (the blocks) to an inequality for general $A \subset \Omega_N$. It may be that this method can be carried out succesfully; our failure to do so is no proof that it cannot be done.

The method we finally used for proving Theorem 1 (b) is a coupling argument. The intuitive idea behind the argument is easily explained. For given $i = (i_0, i_1, \ldots) \in \Omega_N$, set

$$B_i := \big\{(j, i_0, i_1, \ldots) \in \Omega_N : j \in \{0, \ldots, N\}\big\} \tag{1.13}$$

Then $(B_i)_{i\in\Omega_N}$ is a collection of blocks $B_i \subset \Omega_N$, each B_i containing N sites at distance 1 from each other. We would like to consider B_i as a single site in a 'renormalized' lattice, such that B_i can be either infected or healthy. Indeed, if N is large and $\alpha_1 > \delta$, then it can be shown that there exists a 'metastable' state on B_i in which roughly a $(1 - \delta/\alpha_1)$-fraction of the sites is infected, and that transitions from this metastable state to the all-healthy state are fast and happen rarely. Thus, as long as δ/α_1 is sufficiently small, we expect our 'renormalized' blocks B_i to behave effectively as a single site, with an effective 'renormalized' recovery rate $\tilde{\delta}$ that is much smaller than the original δ. Iterating this procedure, we expect the system to be more and more stable as we move up the spatial scale, until, in the limit, we never get extinct.

The problem with making this intuition rigorous is that we are not allowed to send N to infinity. If one tries to make the heuristic picture above rigorous in a straightforward way, then one is forced to give exact bounds, for fixed N, on how stable the 'metastable' state on

B_i is, and how fast transitions between this state and the all-healthy state are. This soon becomes very messy and technical.

The solution we found for this problem is a technique the second author learned about from a talk by Tom Kurtz on the look-down construction for Fleming-Viot processes [DK96, DK99]. Basically, this is a technique for adding structure to a Markov process X, such that if in the enriched process (X, Y), one forgets the added structure Y, one obtains back the original process X. An interesting feature of this technique is that in the enriched process (X, Y), the process X is in general not an autonomous Markov process, i.e., the dynamics of X depend on Y. In practice, we will set up a coupling between a contact process X on the hierarchical group Ω_2 with freedom 2, and an 'added-on' process $\tilde{Y}$ that lives on a renormalized lattice and that is almost a contact process itself. In particular, $\tilde{Y}$ can be stochastically estimated from below by a contact process Y, which is sufficient for our purposes.

After proving Theorem 1 (a) in Section 2, we present and prove our coupling of contact processes on Ω_2 in Section 3 below. A more detailed discussion of our coupling can be found in Sections 3.1– 3.3 while Sections 3.4–3.8 contain proofs. The proof of Theorem 1 (b) is given in Section 4. Appendix A contains a simple, but rather tedious argument needed in Section 3.6.

2 Extinction

2.1 Some general notation

Fix $N \geq 2$ and let $\Omega = \Omega_N$ denote the hierarchical group with freedom N. We introduce contact processes whose state spaces are finite analogues of Ω. For $n \geq 1$, set

$$\Omega^n := \big\{ i = (i_0, \ldots, i_{n-1}) : i_k \in \{0, \ldots, N-1\} \big\} \tag{2.1}$$

and

$$\Omega^0 := \{(\emptyset)\}, \tag{2.2}$$

where $(\emptyset)$ denotes the empty sequence. We equip Ω^n with componentwise addition modulo 2. For $m, n \geq 0$, we define the concatenation $i \circ j \in \Omega^{m+n}$ of elements $i \in \Omega^m$ and $j \in \Omega^n$ by

$$i \circ j := (i_0, \ldots, i_{m-1}, j_0, \ldots, j_{n-1}). \tag{2.3}$$

Given $0 \leq m \leq n$, by definition, the m-block in Ω^n with index $j \in \Omega^{n-m}$ is the set

$$B_m(j) := \{ i \circ j : i \in \Omega^m \} \qquad (j \in \Omega^{n-m},\ 0 \leq m \leq n). \tag{2.4}$$

We define the set of spin configurations on Ω^n by

$$S_n := \{0,1\}^{\Omega^n} = \big\{ x = (x(i))_{i \in \Omega^n} : x(i) \in \{0,1\} \big\} \qquad (n \geq 0). \tag{2.5}$$

Note that Ω^0 is a set containing one element and therefore $S_0 = \{0,1\}$. For $0 \leq m \leq n$, $i \in \Omega^{n-m}$, and $x \in S_n$, we define $x_i \in S_m$ by

$$x_i(j) := x(j \circ i) \qquad (i \in \Omega^{n-m},\ j \in \Omega^m,\ x \in S_n,\ 0 \leq m \leq n). \tag{2.6}$$

Note that x_i describes what the spin configuration x looks like on the m-block with index i.

For $i \in \Omega^n$, we define $|i|$ as in (1.5) with $|i| := n$ if $i_{n-1} \neq 0$. For given $\delta > 0$ and nonnegative constants $\alpha_1, \ldots, \alpha_n$, we define infection rates $a(i,j)$ on Ω^n as in (1.6), and we call the contact process with these infection rates and with recovery rate δ the $(\delta, \alpha_1, \ldots, \alpha_n)$-*contact process*.

2.2 Extinction

Proof of Theorem 1 (a) For $n \geq 0$, let $X^{(n)}$ be the $(\delta, \alpha_1, \ldots, \alpha_n)$-contact process, and set

$$l(n) := \mathbb{E}^{\delta_0}\big[\inf\{t \geq 0 : X^{(n)}_t = \underline{0}\}\big] \qquad (n \geq 0). \tag{2.7}$$

We will estimate $l(n)$ by a very crude argument. By a simple rescaling of time, we may assume that the constant $|a|$ in (1.1) satisfies $|a| = 1$. By an obvious coupling, it follows that $X^{(n)}$ may be stochastically bounded from above by a process $\tilde{X}^{(n)}$ in $S_n = \{0,1\}^{\Omega^n}$ where sites jump independently of each other from 0 to 1 with rate 1 and from 1 to 0 with rate δ. Obviously, the process $\tilde{X}^{(n)}$ has a unique equilibrium law, which is of product form, and if $\tilde{X}^{(n)}_\infty$ denotes a random variable distributed according to this law, then

$$\mathbb{P}[\tilde{X}^{(n)}_\infty = \underline{0}] = \Big(\frac{\delta}{1+\delta}\Big)^{N^n}. \tag{2.8}$$

On the other hand, since the Markov process $\tilde{X}^{(n)}$ stays on average a time $(N^n)^{-1}$ in the state $\underline{0}$ every time it gets there, one has

$$\mathbb{P}[\tilde{X}^{(n)}_\infty = \underline{0}] = \frac{N^{-n}}{\tilde{l}(n) + N^{-n}} = \frac{1}{1 + N^n \tilde{l}(n)}, \tag{2.9}$$

where

$$\tilde{l}(n) := \mathbb{E}^{\delta_0}\big[\inf\{t \geq 0 : \tilde{X}^{(n)}_t = \underline{0}\}\big] \qquad (n \geq 0). \tag{2.10}$$

Solving $\tilde{l}(n)$ from (2.9) and (2.10) and comparing with $l(n)$, we find that

$$l(n) \leq \tilde{l}(n) = N^{-n}\big((1+\delta^{-1})^{N^n} - 1\big) \leq N^{-n}(1+\delta^{-1})^{N^n}. \tag{2.11}$$

Now consider our original contact process on the (infinite) hierarchical group Ω_N. We may stochastically estimate this process from above by a process where infections over a hierarchical distance $> n$ yield infections of a new type, in such a way that infections of different types do not interact with each other (in particular, sites may be multiple infected with infections of different types). Thus, in our new process, each type evolves as a $(\delta, \alpha_1, \ldots, \alpha_n)$-contact process in some n-block, and in addition, for each $k > n$, each site that is infected with this type establishes with rate $\alpha_k N^{-k}(N^k - N^{k-1})$ another type at a uniformly chosen site in some uniformly chosen n-block at hierarchical distance k. Since at any point in time there are at most N^n infected sites of a given type, and each type exists for an expected time of length $l(n)$, it follows that the expected number of new types created by a type during its lifetime is bounded from above by

$$N^n l(n)(1 - N^{-1}) \sum_{k=n+1}^{\infty} \alpha_k. \tag{2.12}$$

In view of (2.11) and the definition of β_n, we may estimate this quantity from above by

$$(1 - N^{-1})(1+\delta^{-1})^{N^n} \beta_{n+1}. \tag{2.13}$$

If, for some $n \geq 1$, this quantity is less than 1, then types create new types according to a subcritical branching process, hence a.s. at most finitely many types are created at all time, hence our contact process dies out. Taking logarithms and dividing by N^n, we see that for all $\delta > 0$ there exists an $n \geq 1$ such that the quantity in (2.13) is less than one, provided that

$$\liminf_{n\to\infty} N^{-n} \log\big((1-N^{-1})(1+\delta^{-1})^{N^n}\beta_{n+1}\big) < 0 \qquad \forall \delta > 0. \tag{2.14}$$

This is equivalent to

$$\log(1+\delta^{-1}) + \liminf_{n\to\infty} N^{-n}\log(\beta_{n+1}) < 0 \qquad \forall \delta > 0, \tag{2.15}$$

which is in turn equivalent to (1.7). ∎

3 Coupling of contact processes

3.1 A coupling

Throughout this section, we fix $N = 2$ and consider finite $(\delta, \alpha_1, \ldots, \alpha_n)$-contact processes on Ω^n as defined in Section 2.1. We will prove the following result.

Proposition 2 (Coupling of contact process) *Let $n \geq 1$, $\delta > 0$, and $\alpha_1, \ldots, \alpha_n \geq 0$. Let $X = (X_t)_{t\geq 0}$ be the $(\delta, \alpha_1, \ldots, \alpha_n)$-contact process started in any initial law. Set $\delta' := 2\xi\delta$ and $\alpha'_k := \frac{1}{2}\alpha_{k+1}$ $(k = 1, \ldots, n-1)$, where $\xi = f(\alpha_1/\delta)$ and f denotes the function*

$$f(r) := \gamma - \sqrt{\gamma^2 - \tfrac{1}{2}} \quad \textit{with} \quad \gamma := \tfrac{1}{4}\big(3 + \tfrac{1}{2}r\big) \qquad (r \geq 0). \tag{3.1}$$

Then X can be coupled to a process $(\tilde{Y}, Y)$ such that $(X_t, \tilde{Y}_t)_{t\geq 0}$ is a Markov process, $(Y_t)_{t\geq 0}$ is a $(\delta', \alpha'_1, \ldots, \alpha'_{n-1})$-contact process, $\tilde{Y}_0 = Y_0$, $\tilde{Y}_t \geq Y_t$ for all $t \geq 0$, and

$$\mathbb{P}\big[\tilde{Y}_t = y \,\big|\, (X_s)_{0\leq s\leq t}\big] = P(X_t, y) \quad \text{a.s.} \qquad (t \geq 0,\ y \in S_{n-1}), \tag{3.2}$$

where P is the probability kernel from S_n to S_{n-1} defined by (recall (2.6))

$$P(x, y) := \prod_{i\in\Omega^{n-1}} p(x_i, y(i)) \qquad (x \in S_n,\ y \in S_{n-1}), \tag{3.3}$$

with

$$\begin{pmatrix} p(00,0) & p(00,1) \\ p(01,0) & p(01,1) \\ p(10,0) & p(10,1) \\ p(11,0) & p(11,1) \end{pmatrix} := \begin{pmatrix} 1 & 0 \\ \xi & 1-\xi \\ \xi & 1-\xi \\ 0 & 1 \end{pmatrix}. \tag{3.4}$$

The coupling in Proposition 2 achieves the intuitive aim explained in Section 1.2, namely, to view blocks, consisting of two sites at distance one from each other, as single sites in a 'renormalized' lattice, which can either be infected or healthy. Indeed, (3.2) says that the conditional law of $\tilde{Y}_t$ given X_t has the following description. First, we group the sites of X_t into blocks, each consisting of two sites at distance one from each other. Then, independently for each block, if the configuration in such a block is 00 (resp. 11), then we let the corresponding single site in $\tilde{Y}_t$ be healthy (resp. infected), while if the configuration is 01 or 10, then we let the corresponding site in $\tilde{Y}_t$ be healthy with probability ξ and infected with probability $1-\xi$. This stochastic rule is demonstrated in Figure 1. The transition there has probability $\xi(1-\xi)$ and sites in Ω^3 and Ω^2 are depicted as leaves of a binary tree.

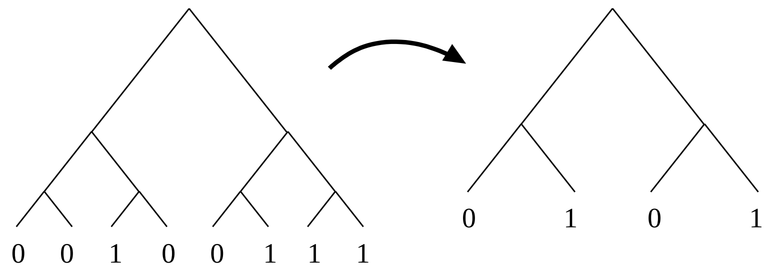

Figure 1: Coupling of X_t (left) and $\tilde{Y}_t$ (right). The conditional probability of the transition depicted here is $\xi(1-\xi)$.

It is interesting that a stochastic rule for deciding whether a block is healthy or infected seems to work better than a deterministic rule. We will choose the function $p(\,\cdot\,,1)$ in (3.4) in such a way that this is the leading eigenfunction of a one-level (δ,α_1)-contact process. Thus, our methods combine some elements of spectral analysis with probabilistic coupling tools.

The next lemma (which is proved in Section 3.5 below) lists some elementary properties of the function f defined in (3.1).

Lemma 3 (The function f**)** *The function f defined in (3.1) is decreasing on $[0,\infty)$ and satisfies $f(0)=\frac{1}{2}$ and*

$$f(r)=2r^{-1}+O(r^{-2})\qquad as\quad r\to\infty. \tag{3.5}$$

3.2 Markov processes with added structure

In Proposition 2, the coupling between the processes X and $\tilde{Y}$ is of a special kind. There exist general results that tell us how to construct processes with conditional probabilities as in (3.2), such that in addition $(X_t)_{t\geq 0}$, on its own, is a Markov process. In the present section, we formulate one such result, which will then be used to construct the coupling in Proposition 2.

Let S,S' be finite sets and set $\hat{S}:=S\times S'$. Let $(X,Y)=(X_t,Y_t)_{t\geq 0}$ be a Markov process with state space $\hat{S}$ and generator $\hat{G}$. For each $y\in S'$ (resp. $x\in S$), we define an operator $G_y:\mathbb{R}^S\to\mathbb{R}^S$ (resp. $G'_x:\mathbb{R}^{S'}\to\mathbb{R}^{S'}$) by

$$\begin{aligned} G_yf(x) &:= \hat{G}\overline{f}(x,y) \qquad \text{where} \quad \overline{f}(x,y):=f(x) \qquad (x\in S,\ y\in S',\ f\in\mathbb{R}^S),\\ G'_xf(y) &:= \hat{G}\overline{f}(x,y) \qquad \text{where} \quad \overline{f}(x,y):=f(y) \qquad (x\in S,\ y\in S',\ f\in\mathbb{R}^{S'}). \end{aligned} \tag{3.6}$$

We say that X *evolves according to the generator G_y while* $Y=y$ (resp. Y evolves according to the generator G'_x while $X=x$). In particular, if G_y does not depend on y, i.e., if $G_y=G$ $(y\in S')$ for some operator $G:\mathbb{R}^S\to\mathbb{R}^S$, then we say that X is an *autonomous Markov process* with generator G. This is equivalent to the statement that for *every* initial law of the joint process (X,Y), the process X, on its own, is the Markov process with generator G. The next proposition (which will be proved in Section 3.4) demonstrates that even when X is not autonomous, it may happen that there exists an operator G such that for certain *special* initial laws of the joint process (X,Y), the process X, on its own, is the Markov process with generator G. It seems that Rogers and Pitman [RP81] were the first who noticed this phenomenon. We will prove the proposition below by elaborating on their result.

Proposition 4 (Markov process with added structure) *Let X be a continuous-time Markov process with finite state space S and generator G. Let S' be a finite set, let P be a probability kernel from S to S', and let $(G'_x)_{x\in S}$ be a collection of generators of S'-valued Markov processes. Define an operator $\overline{G}:\mathbb{R}^{S'}\to\mathbb{R}^{S\times S'}$ by*

$$\overline{G}f(x,y) := G'_x f(y) \qquad (x\in S,\ y\in S',\ f\in\mathbb{R}^{S'}), \tag{3.7}$$

and define $P:\mathbb{R}^{S'}\to\mathbb{R}^S$ and $\overline{P}:\mathbb{R}^{S\times S'}\to\mathbb{R}^S$ by

$$Pf(x) := \sum_{y\in S'} P(x,y)f(y) \quad \text{and} \quad \overline{P}f(x) := \sum_{y\in S'} P(x,y)f(x,y). \tag{3.8}$$

Assume that

$$GPf = \overline{P}\,\overline{G}f \qquad (f\in\mathbb{R}^{S'}). \tag{3.9}$$

Then X can be coupled to an S'-valued process Y such that $(X,Y)=(X_t,Y_t)_{t\geq 0}$ is a Markov process with state space $S\times S'$, the process Y evolves according to the generator G'_x while $X=x$, and

$$\mathbb{P}\big[Y_t = y\,\big|\,(X_s)_{0\leq s\leq t}\big] = P(x_t,y) \quad \text{a.s.} \qquad (t\geq 0,\ y\in S'). \tag{3.10}$$

Remark 1 If X and Y are coupled as in Proposition 4, then it is typically not the case that X is an autonomous Markov process. Nevertheless, the joint Markov process (X,Y) has the property that if the initial law satisfies

$$\mathbb{P}\big[Y_0 = y\,\big|\,X_0 = x\big] = P(x,y) \qquad (t\geq 0), \tag{3.11}$$

then X, on its own, is the Markov process with generator G, and (3.10) holds.

Remark 2 If X and Y are coupled as in Proposition 4, then it may happen that Y is an autonomous Markov process. In this case, we will say that Y is an *averaged Markov process* associated with X. In the general case, we will say that Y is an *added-on* process.

3.3 Discussion

We mention a few open problems concerning our coupling.

1° Can one modify Proposition 2 such that $\tilde{Y}=Y$, i.e., (in terminology invented in the previous section), for a given $(\delta,\alpha_1,\ldots,\alpha_n)$-contact process X, can we find a $(\delta',\alpha'_1,\ldots,\alpha'_{n-1})$-contact process Y such that Y is an averaged Markov process of X? This would probably involve a kernel P and constants $\delta',\alpha'_1,\ldots,\alpha'_{n-1}$ that are more difficult to describe and less explicit than the ones in Proposition 2 but would have great theoretical value, since the resulting map $(\delta,\alpha_1,\ldots,\alpha_n)\mapsto(\delta',\alpha'_1,\ldots,\alpha'_{n-1})$ would represent a rigorous renormalization transformation.

2° Is it possible to construct a similar coupling as in Proposition 2, but with $\tilde{Y}_t\leq Y_t$? This could potentially be used to relax condition (1.7).

3° Can one use Proposition 2 to construct a probability law on $\{0,1\}^{\Omega_2}$ that satisfies condition (1.10) of Holley and Liggett? This would not add much in the line of proving survival (which is already achieved) but might add to our understanding of the method of Holley and Liggett, which is rather poor. In particular, in [Lig95] it is shown that this method may be used to calculate a sequence of approximations of the critical recovery rate, but beyond the second member of that sequence, there is no proof that these approximations are lower bounds on δ_c (though they are conjectured to be so).

4° Is it possible to make the methods of the present paper work on $\mathbb{Z}$ instead of Ω_2? At first sight, it seems that the hierarchical structure of Ω_2 is essential to Proposition 2. However, when we think of the latter as 'forgetting the fast modes of the spectrum', something may be possible. Any link between Proposition 2 and the method of Holley and Liggett might also provide a clue.

3.4 Added-on processes

Proof of Proposition 4 We adopt the convention that sums over x, x', x'' always run over S and sums over y, y', y'' always run over S'. Write

$$\begin{aligned} Gf(x) &= \sum_{x'} r(x,x')\big(f(x') - f(x)\big), \\ G'_x f(y) &= \sum_{y'} r'_x(y,y')\big(f(y') - f(y)\big), \end{aligned} \tag{3.12}$$

where $r(x, x')$ (resp. $r'_x(y, y')$) denotes the rate at which the Markov process with generator G (resp. G'_y) jumps from a state x to a state x' (resp. from y to y').

Set $\hat{S} := S \times S'$. We let $(X, Y) = (X_t, Y_t)_{t\geq 0}$ be the Markov process in $\hat{S}$ started in an initial law satisfying (3.11), with generator $\hat{G}$ defined by

$$\begin{aligned} \hat{G}f(x,y) := &\sum_{x'} t_y(x,x')\big(f(x',y) - f(x,y)\big) \\ &+ \sum_{y' :\, P(x,y')>0} r'_x(y,y')\big(f(x,y') - f(x,y)\big) \\ &+ \sum_{y' :\, P(x,y')=0} r'_x(y,y') \sum_{x'} q_{y'}(x,x')\big(f(x',y') - f(x,y)\big) \end{aligned} \tag{3.13}$$

$((x, y) \in \hat{S},\ f \in \mathbb{R}^{\hat{S}})$, where

$$t_y(x,x') := \frac{r(x,x')P(x',y)}{P(x,y)} \quad \text{and} \quad q_y(x,x') := \frac{r(x,x')P(x',y)}{\sum_{x''} r(x,x'')P(x'',y)}. \tag{3.14}$$

These formulas are not defined if $P(x, y) = 0$ resp. $\sum_{x''} r(x, x'')P(x'', y) = 0$, so in the first case we define $t_y(x, x')$, in some arbitrary way, while in the second case we choose for $q_{y'}(x, \cdot\,)$ some arbitrary probability distribution on S. In any case, it will be true that

$$\Big(\sum_{x''} r(x,x'')P(x'',y)\Big) q_y(x,x') = r(x,x')P(x',y), \tag{3.15}$$

since the right-hand side of this equation is zero if $\sum_{x''} r(x, x'')P(x'', y) = 0$.

Formula (3.13) says that the process (X, Y) jumps from a state (x, y) to a state (x', y) with rate $t_y(x, x')$. In addition, while X is in the state x, the process Y jumps from the state y to the state y' with rate $r'_x(y, y')$. During such a jump, if $P(x, y') = 0$, then the process X does nothing but if $P(x, y') = 0$, then the process X jumps at the same time to a state x' chosen according to the probability kernel $q_{y'}(x, x')$. In particular, these rules say that the process Y evolves according to the generator G'_x while $X = x$.

It is known that (3.10) holds for the process (X, Y) started in any initial law satisfying (3.11), provided that

$$G\overline{P}f = \overline{P}\hat{G}f \qquad (f \in \mathbb{R}^{\hat{S}}). \tag{3.16}$$

The sufficiency of (3.16) follows, for example, from [Kur98, Corollary 3.5], which is a rather technical statement about martingale problems. A much less technical version of this result can be found in [RP81].

We may rewrite $\hat{G}$ in the form

$$\begin{aligned}\hat{G}f(x,y) := &\sum_{x'} t_y(x,x')\big(f(x',y) - f(x,y)\big)\\ &+ \sum_{y'} r'_x(y,y')\big(f(x,y') - f(x,y)\big)\\ &+ \sum_{y':P(x,y')=0} r'_x(y,y') \sum_{x'} q_{y'}(x,x')\big(f(x',y') - f(x,y')\big) \qquad (x \in S,\ y \in R).\end{aligned} \tag{3.17}$$

We calculate, remembering the definition of $t_y(x,x')$, and letting $G(x,x')$ denote the matrix associated with the operator G,

$$\begin{aligned}G\overline{P}f(x) &= \sum_{x'} r(x,x')\Big(\sum_y P(x',y)f(x',y) - \sum_y P(x,y)f(x,y)\Big)\\ &= \sum_y \sum_{x'} r(x,x')\big[P(x',y)\big(f(x',y) - f(x,y)\big) + \big(P(x',y) - P(x,y)\big)f(x,y)\big]\\ &= \sum_y P(x,y) \sum_{x'} t_y(x,x')\big(f(x',y) - f(x,y)\big)\\ &\quad + \sum_{y:\,P(x,y)=0} \sum_{x'} r(x,x')P(x',y)\big(f(x',y) - f(x,y)\big)\\ &\quad + \sum_y \Big(\sum_{x'} G(x,x')P(x',y)\Big) f(x,y),\end{aligned} \tag{3.18}$$

and

$$\begin{aligned}\overline{P}\hat{G}f(x) &= \sum_y P(x,y) \sum_{x'} t_y(x,x')\big(f(x',y) - f(x,y)\big)\\ &\quad + \sum_{yy'} P(x,y) r'_x(y,y')\big(f(x,y') - f(x,y)\big)\\ &\quad + \sum_y P(x,y) \sum_{y':\,P(x,y')=0} r'_x(y,y') \sum_{x'} q_{y'}(x,x')\big(f(x',y') - f(x,y')\big)\\ &= \sum_y P(x,y) \sum_{x'} t_y(x,x')\big(f(x',y) - f(x,y)\big)\\ &\quad + \sum_{y:\,P(x,y)=0} \sum_{x'} \Big(\sum_{y'} P(x,y') r'_x(y',y)\Big) q_y(x,x')\big(f(x',y) - f(x,y)\big)\\ &\quad + \sum_y \Big(\sum_{y'} P(x,y') G_x(y',y)\Big) f(x,y),\end{aligned} \tag{3.19}$$

where to get the second equality we have reordered our terms and relabelled indices. The first terms on the right-hand sides of (3.18) and (3.19) are equal while the third terms agree by

(3.9). Since by (3.9), for each x, y such that $P(x,y) = 0$, one has

$$\sum_{y'} P(x,y')r'_x(y',y) = \sum_{y'} P(x,y')G_x(y',y) = \sum_{x''} G(x,x'')P(x'',y) = \sum_{x''} r(x,x'')P(x'',y), \tag{3.20}$$

we see by (3.15) that also the second terms on the right-hand sides of (3.18) and (3.19) agree, hence (3.16) holds. ∎

Remark For fixed $y \in S'$, set $S_y := \{x \in S : P(x,y) > 0\}$ and consider the operator $\tilde{G}_y$ defined by (compare (3.13)–(3.14))

$$\tilde{G}_y f(x) := \sum_{x' \in S_y} t_y(x,x')\big(f(x') - f(x)\big) \qquad (x \in S_y,\ f \in \mathbb{R}^{S_y}). \tag{3.21}$$

Then $\tilde{G}_y$ is a 'compensated h-transform' of the operator G, with the function $h(x) := P(x,y)$. Here, if G is the generator of a Markov process on S and h is a nonnegative function on S, then

$$G^h f := h^{-1}G(hf) - h^{-1}(Gh)f \tag{3.22}$$

defines a generator of a Markov process on the space $S_h := \{x : h(x) > 0\}$. This sort of transformation has been called a *compensated h-transform* in [FS04]. In particular, if h is harmonic, i.e., $Gh = 0$, then G^h is the usual h-transform of G.

3.5 Definition of the added-on process

In this section, we prove Proposition 2. Our proof depends on some calculations that will be done in the next three sections. We wish to construct an S^{n-1}-valued added-on process $\tilde{Y}$ on X, such that $\tilde{Y}$ can be stochastically estimated from below by a $(\delta', \alpha'_1, \ldots, \alpha'_{n-1})$-contact process. We introduce the notation

$$x(i,j) := \big(x(i), x(j)\big) \qquad (x \in S_n,\ i,j \in \Omega^n). \tag{3.23}$$

With this notation, the generator of X can be written as

$$\begin{aligned} Gf(x) = {} & \delta \sum_{i \in \Omega^n} 1_{\{x(i)=1\}}\big(f(x - \delta_i) - f(x)\big) \\ & + \sum_{k=1}^{n} \alpha_k\, 2^{-k} \sum_{\substack{i,j \in \Omega^n \\ |i-j|=k}} 1_{\{x(i,j)=(0,1)\}}\big(f(x + \delta_i) - f(x)\big). \end{aligned} \tag{3.24}$$

For any $x \in S_1 = \{0,1\}^2$, we write

$$\overline{x} := \begin{cases} 00 & \text{if } x = (0,0), \\ 01 & \text{if } x = (0,1) \text{ or } (1,0), \\ 11 & \text{if } x = (1,1), \end{cases} \tag{3.25}$$

For each $x \in S_n$, we define a generator G'_x of an S_{n-1}-valued Markov process by (recall (2.6))

$$\begin{aligned} G'_x f(y) = {} & \delta' \sum_{i \in \Omega^{n-1}} 1_{\{y(i)=1\}}\big(f(y - \delta_i) - f(y)\big) \\ & + \sum_{k=1}^{n-1} \alpha_{k+1}\, 2^{-k} \sum_{\substack{i,j \in \Omega^{n-1} \\ |i-j|=k}} \big[a\big(\overline{x}_i, \overline{x}_j\big) 1_{\{y(i,j)=(0,1)\}} \\ & \qquad + b\big(\overline{x}_i, \overline{x}_j\big) 1_{\{y(i,j)=(0,0)\}}\big]\big(f(y + \delta_i) - f(y)\big), \end{aligned} \tag{3.26}$$

where δ' is defined as in Proposition 2 and a, b are the functions

$$\begin{pmatrix} a(00,00) & a(00,01) & a(00,11) \\ a(01,00) & a(01,01) & a(01,11) \\ a(11,00) & a(11,01) & a(11,11) \end{pmatrix} = \begin{pmatrix} * & 1-\xi & 2(1-\xi) \\ * & \frac{1}{2} & 1 \\ * & * & * \end{pmatrix} \tag{3.27}$$

and

$$\begin{pmatrix} b(00,00) & b(00,01) & b(00,11) \\ b(01,00) & b(01,01) & b(01,11) \\ b(11,00) & b(11,01) & b(11,11) \end{pmatrix} = \begin{pmatrix} 0 & 1-\xi & * \\ 0 & \frac{1}{2} & * \\ * & * & * \end{pmatrix}. \tag{3.28}$$

Here ξ is defined as in Proposition 2 and the symbol $*$ indicates that the definition of a and b in these points is irrelevant. Indeed, in the next three sections, we will prove the following fact.

Lemma 5 (Added-on process) *Let G be the generator of the $(\delta, \alpha_1, \ldots, \alpha_n)$-contact process on S_n, let P be the probability kernel from S_n to S_{n-1} defined in Proposition 2, and let $(G'_x)_{x \in S_n}$ be the generators defined in (3.26), where the functions a and b are defined as in (3.27)–(3.28). Then, no matter how we define a and b in points indicated with the symbol $*$, one has*

$$GPf = \overline{P}\,\overline{G}f \qquad (f \in \mathbb{R}^{S_{n-1}}), \tag{3.29}$$

where $\overline{G}, P$, and $\overline{P}$ are defined as in (3.7)–(3.8).

Based on Lemma 5, we can now prove Proposition 2.

Proof of Proposition 2 By Proposition 4 and Lemma 5, we can couple X to an S_{n-1}-valued process $\tilde{Y}$ such that $(X_t, \tilde{Y}_t)_{t \geq 0}$ is a Markov process, $\tilde{Y}$ evolves according to the generator G'_x while $X = x$, and (3.2) holds.

By Lemma 3, $0 < \xi \leq \frac{1}{2}$. It follows that the functions a and b in (3.27)–(3.28) satisfy $a \geq \frac{1}{2}$ and $b \geq 0$. From this and (3.26), it is easy to see that $(X, \tilde{Y})$ can be coupled to a $(\delta', \alpha'_1, \ldots, \alpha'_{n-1})$-contact process Y, such that $\tilde{Y}_0 = Y_0$ and $\tilde{Y}_t \geq Y_t$ for all $t \geq 0$. ∎

For completeness, we give here the:

Proof of Lemma 3 Set $\xi(\gamma) := \gamma - \sqrt{\gamma^2 - \frac{1}{2}}$. Then it is straightforward to check that $\xi(\frac{3}{4}) = \frac{1}{2}$. Moreover, $\frac{\partial}{\partial\gamma}\xi(\gamma) = 1 - \gamma(\gamma^2 - \frac{1}{2})^{-1/2} = 1 - (1 - \frac{1}{2}\gamma^{-2})^{-1/2} < 0$ on $[\frac{3}{4}, \infty)$, so $\gamma \mapsto \xi(\gamma)$ is decreasing on $[\frac{3}{4}, \infty)$. Set $\varepsilon := \gamma^{-1}$. Then $\xi(\varepsilon^{-1}) = \varepsilon^{-1}\big(1 - \sqrt{1 - \frac{1}{2}\varepsilon^2}\big)$. We observe that

$$\begin{aligned} &\big(1 - \sqrt{1 - \tfrac{1}{2}\varepsilon^2}\big)\big|_{\varepsilon=0} = 0, \\ &\tfrac{\partial}{\partial\varepsilon}\big(1 - \sqrt{1 - \tfrac{1}{2}\varepsilon^2}\big)\big|_{\varepsilon=0} = \tfrac{1}{2}\varepsilon(1 - \tfrac{1}{2}\varepsilon^2)^{-1/2}\big|_{\varepsilon=0} = 0, \\ &\tfrac{\partial^2}{\partial\varepsilon^2}\big(1 - \sqrt{1 - \tfrac{1}{2}\varepsilon^2}\big)\big|_{\varepsilon=0} = \tfrac{\partial}{\partial\varepsilon}\tfrac{1}{2}\varepsilon(1 - \tfrac{1}{2}\varepsilon^2)^{-1/2}\big|_{\varepsilon=0} \\ &\quad = \tfrac{\partial}{\partial\varepsilon}\Big(\tfrac{1}{2}(1 - \tfrac{1}{2}\varepsilon^2)^{-1/2} + \tfrac{1}{4}\varepsilon^2(1 - \tfrac{1}{2}\varepsilon^2)^{-3/2}\Big)\big|_{\varepsilon=0} = \tfrac{1}{2}, \end{aligned} \tag{3.30}$$

hence $\xi(\varepsilon^{-1}) = \varepsilon^{-1}\big(\frac{1}{4}\varepsilon^2 + O(\varepsilon^3)\big) = \frac{1}{4}\varepsilon + O(\varepsilon^2)$, i.e.,

$$\xi(\gamma) = \tfrac{1}{4}\gamma^{-1} + O(\gamma^{-2}) \qquad \text{as} \quad \gamma \to \infty. \tag{3.31}$$

To translate this to the statements in Lemma 3, it suffices to note that the function $r \mapsto \gamma(r) := \frac{1}{4}(2 + \frac{1}{2}r)$ is increasing on $[0, \infty)$, satisfies $\gamma(0) = \frac{3}{4}$, and $\gamma(r) = \frac{1}{8}r + O(1)$ as $r \to \infty$. ∎

3.6 Reduction to a one- and two-level system

In this section, we prove Lemma 5. Our proof is based on two lemmas which will be proved in the next two sections.

Proof of Lemma 5 We start by rewriting the generator in (3.24) as follows:

$$
\begin{aligned}
Gf(x) = {} & \delta \sum_{i\in\Omega^{n-1}} \sum_{i'\in\Omega^1} 1_{\{x(i'\circ i)=1\}} \big(f(x-\delta_{i'\circ i}) - f(x)\big) \\
& + \alpha_1 2^{-1} \sum_{i\in\Omega^{n-1}} \sum_{\substack{i',i''\in\Omega^1 \\ |i'-i''|=1}} 1_{\{x(i'\circ i)=0,\ x(i''\circ i)=1\}} \big(f(x+\delta_{i'\circ i}) - f(x)\big) \\
& + \sum_{k=2}^{n} \alpha_k\, 2^{-k} \sum_{\substack{i,j\in\Omega^{n-1} \\ |i-j|=k-1}} \sum_{i',j'\in\Omega^1} 1_{\{x(i'\circ i)=0,\ x(j'\circ j)=1\}} \big(f(x+\delta_{i'\circ i}) - f(x)\big) \\
= {} & \sum_{i\in\Omega^{n-1}} R_i f(x) + \sum_{k=1}^{n-1} \alpha_{k+1}\, 2^{-k} \sum_{\substack{i,j\in\Omega^{n-1} \\ |i-j|=k}} I_{ij} f(x),
\end{aligned}
\tag{3.32}
$$

where

$$
\begin{aligned}
R_i f(x) := {} & \delta \sum_{i'\in\{0,1\}} 1_{\{x(i'\circ i)=1\}} \big(f(x-\delta_{i'\circ i}) - f(x)\big) \\
& + \alpha_1 \sum_{\substack{i',i''\in\{0,1\} \\ i'\neq i''}} 1_{\{x(i'\circ i)=0,\ x(i''\circ i)=1\}} \big(f(x+\delta_{i'\circ i}) - f(x)\big), \\
I_{ij} f(x) := {} & 2^{-1} \sum_{i',j'\in\{0,1\}} 1_{\{x(i'\circ i)=0,\ x(j'\circ j)=1\}} \big(f(x+\delta_{i'\circ i}) - f(x)\big).
\end{aligned}
\tag{3.33}
$$

Likewise, we may write the operator in (3.26) as

$$
G'_x f(y) = \sum_{i\in\Omega^{n-1}} R'_i f(y) + \sum_{k=1}^{n-1} \alpha_{k+1}\, 2^{-k} \sum_{\substack{i,j\in\Omega^{n-1} \\ |i-j|=k}} I'^{x}_{ij} f(y), \tag{3.34}
$$

where

$$
\begin{aligned}
R'_i f(y) &:= \delta' 1_{\{y(i)=1\}} \big(f(y-\delta_i) - f(y)\big), \\
I'^{x}_{ij} f(y) &:= \big[a\big(\overline{x}_i, \overline{x}_j\big) 1_{\{y(i)=0,\ y(j)=1\}} + b\big(\overline{x}_i, \overline{x}_j\big) 1_{\{y(i)=0,\ y(j)=0\}}\big] \big(f(y+\delta_i) - f(y)\big).
\end{aligned}
\tag{3.35}
$$

In view of (3.32) and (3.34), in order to prove (3.29), it suffices to show that

$$
\left.\begin{aligned}
&\text{(i)} \quad R_i P f = P R'_i f \\
&\text{(ii)} \quad I_{ij} P f = \overline{P}\, \overline{I}_{ij} f
\end{aligned}\right\} \quad (f \in \mathbb{R}^{S_{n-1}},\ i,j \in \Omega^{n-1},\ i \neq j), \tag{3.36}
$$

where $\overline{I}_{ij} : \mathbb{R}^{S_{n-1}} \to \mathbb{R}^{S_n \times S_{n-1}}$ is defined as $\overline{I}_{ij} f(x,y) := I'^{x}_{ij} f(y)$. Note that since R'_i does not depend on x, there is no need to define $\overline{R}_i$.

The operators R_i, R'_i, I_{ij}, and $\overline{I}_{ij}$ act only on certain coordinates. In view of this, our problem reduces to a lower-dimensional one, and (3.36) follows from Lemmas 6 and 7 stated below. It is not difficult, but notationally cumbersome, to give a formal derivation of (3.36) from Lemmas 6 and 7. For completeness, we give this derivation in Appendix A. ∎

Recall that $S_1 = \{0,1\}^{\Omega^1}$ and $S_0 = \{0,1\}^{\Omega^0} = \{0,1\}$. Let $\delta > 0$, $\alpha_1 \geq 0$, and let δ' be defined as in Proposition 2. Let P be the probability kernel from S_1 to S_0 defined in (3.3)–(3.4) and let $P : \mathbb{R}^{S_0} \to \mathbb{R}^{S_1}$ be defined as in (3.8). Let R be the generator of the (δ, α_1)-contact process X on S_1 and let R' be the generator of the δ'-contact process Y on S_0. The latter is just the Markov process with state space $\{0,1\}$ that jumps from 1 to 0 with rate δ'. The next lemma implies that Y is an averaged Markov process associated with X, i.e., X and Y can be coupled such that (3.10) holds.

Lemma 6 (One-level system) *One has*

$$RPf = PR'f \qquad (f \in \mathbb{R}^{S_0}). \tag{3.37}$$

Formula (3.37) implies (3.36) (i). We next formulate a lemma that implies (3.36) (ii).

Let $\delta > 0$, $\alpha_1 \geq 0$, let P be the probability kernel from S_2 to S_1 defined in (3.3)–(3.4), and let a, b be the functions in (3.27)–(3.28). We define a generator I of a Markov process in S_2 and generators $(I'_x)_{x \in S_2}$ of Markov processes in S_1 by

$$\begin{aligned} If(x) &:= \tfrac{1}{2} \sum_{i,j \in \{0,1\}} 1_{\{x(i,0)=0,\ x(j,1)=1\}} \big(f(x + \delta_{(i,0)}) - f(x) \big), \\ I'_x f(y) &:= \big[a(\overline{x}_0, \overline{x}_1) 1_{\{y(0,1)=(0,1)\}} + b(\overline{x}_0, \overline{x}_1) 1_{\{y(0,1)=(0,0)\}} \big] \big(f(y + \delta_0) - f(y) \big). \end{aligned} \tag{3.38}$$

Our next lemma says that the $(I'_x)_{x \in S_2}$ define an added-on Markov process associated with the process with generator I.

Lemma 7 (Two-level system) *One has*

$$IPf = \overline{P}\,\overline{I}f \qquad (f \in \mathbb{R}^{S_1}), \tag{3.39}$$

where $P, \overline{P}$ are defined as in (3.8) and $\overline{I}f(x,y) := I'_x f(y)$ $(x \in S_2,\ y \in S_1)$.

3.7 The one-level system

Proof of Lemma 6 We may write (3.37) in the matrix form

$$\sum_{x' \in S_1} \sum_{y \in S_0} R(x,x') P(x',y) f(y) = \sum_{y' \in S_0} \sum_{y \in S_0} P(x,y') R'(y',y) f(y) \qquad (x \in S_1,\ f \in \mathbb{R}^{S_0}), \tag{3.40}$$

which is equivalent to

$$\sum_{x' \in S_1} R(x,x') p(x',y) = \sum_{y' \in S_0} p(x,y') R'(y',y) \qquad (x \in S_1,\ y \in S_0), \tag{3.41}$$

where p is the function in (3.4). Here

$$\sum_{x' \in S_1} R(x,x') p(x',y) = Rp(\,\cdot\,, y)(x) \qquad (x \in \{0,1\}^2,\ y \in \{0,1\}). \tag{3.42}$$

Thus, (3.41) says that $Rp(\,\cdot\,,0)$ and $Rp(\,\cdot\,,1)$ can be written as a linear combination of the functions $p(\,\cdot\,,0)$ and $p(\,\cdot\,,1)$. It follows that $\mathcal{F} := \mathrm{span}\{p(\,\cdot\,,0), p(\,\cdot\,,1)\}$ is an invariant subspace of the operator R. Since $p(\,\cdot\,,0)+p(\,\cdot\,,1) = 1$, the space $\mathcal{F}$ contains the constant function 1. We will show that $\mathcal{F}$ is in fact the span of 1 and one nontrivial eigenfunction of R.

We start by noting that by symmetry, the space $\mathcal{H} := \{f \in \mathbb{R}^{S_1} : f(0,1) = f(1,0)\}$ is invariant under R. Since $(0,0)$ is a trap of the (δ,α_1)-contact process, the space $\mathcal{H}_0 := \{f \in \mathcal{H} : f(0,0) = 0\}$ is also invariant under R; in fact, $\mathcal{H}$ is the span of $\mathcal{H}_0$ and the trivial eigenfunction 1. In view of this, we look for eigenfunctions of R in $\mathcal{H}_0$. We observe that for $f \in \mathcal{H}_0$,

$$\begin{aligned}\begin{pmatrix} Rf(0,1) \\ Rf(1,1) \end{pmatrix} &= \begin{pmatrix} \delta\big(0 - f(0,1)\big) + \frac{1}{2}\alpha_1\big(f(1,1) - f(0,1)\big) \\ 2\delta\big(f(0,1) - f(1,1)\big) \end{pmatrix} \\ &= \begin{pmatrix} -(\delta + \frac{1}{2}\alpha_1) & \frac{1}{2}\alpha_1 \\ 2\delta & -2\delta \end{pmatrix} \begin{pmatrix} f(0,1) \\ f(1,1) \end{pmatrix}.\end{aligned} \tag{3.43}$$

To find the eigenvalues, we must solve

$$\det\begin{pmatrix} -(\delta + \frac{1}{2}\alpha_1) - \lambda & \frac{1}{2}\alpha_1 \\ 2\delta & -2\delta - \lambda \end{pmatrix} = 0, \tag{3.44}$$

which gives

$$\begin{aligned} &(\delta + \tfrac{1}{2}\alpha_1 + \lambda)(2\delta + \lambda) = \delta\alpha_1 \\ \Leftrightarrow\quad & \lambda^2 + (3\delta + \tfrac{1}{2}\alpha_1)\lambda + 2\delta^2 = 0 \\ \Leftrightarrow\quad & \Big(\lambda + \frac{3\delta + \frac{1}{2}\alpha_1}{2}\Big)^2 = \Big(\frac{3\delta + \frac{1}{2}\alpha_1}{2}\Big)^2 - 2\delta^2 \\ \Leftrightarrow\quad & \lambda = -\Big(\frac{3\delta + \frac{1}{2}\alpha_1}{2}\Big) \pm \sqrt{\Big(\frac{3\delta + \frac{1}{2}\alpha_1}{2}\Big)^2 - 2\delta^2} \\ \Leftrightarrow\quad & \lambda = -2\delta\big(\gamma \pm \sqrt{\gamma^2 - \tfrac{1}{2}}\big), \end{aligned} \tag{3.45}$$

where $\gamma := \frac{1}{4}(3 + \frac{1}{2}\frac{\alpha_1}{\delta})$ (compare (3.1)). In particular, the leading eigenvalue is $\lambda = -2\delta\xi = -\delta'$, where ξ and δ' are defined as in Proposition 2. To find the corresponding eigenfunction, we need to solve

$$\begin{aligned} &2\delta\big(f(0,1) - f(1,1)\big) = \lambda f(11) \\ \Leftrightarrow\quad & 2\delta f(0,1) = (2\delta + \lambda) f(1,1) = 2\delta(1 - \xi) f(1,1) \\ \Leftrightarrow\quad & f(0,1) = (1 - \xi) f(1,1), \end{aligned} \tag{3.46}$$

which yields the eigenfunction

$$\begin{pmatrix} f(0,0) \\ f(0,1) \\ f(1,0) \\ f(1,1) \end{pmatrix} = \begin{pmatrix} 0 \\ 1 - \xi \\ 1 - \xi \\ 1 \end{pmatrix} = p(\,\cdot\,,1). \tag{3.47}$$

Our calculations so far show that $\mathcal{F} := \mathrm{span}\{1, p(\,\cdot\,,1)\} = \mathrm{span}\{p(\,\cdot\,,0), p(\,\cdot\,,1)\}$ is an invariant subspace of the operator R. It follows that there exist constants $(R'(y',y))_{y,y' \in \{0,1\}}$ such that

$$Rp(\,\cdot\,,y) = \sum_{y' \in \{0,1\}} p(\,\cdot\,,y') R'(y',y). \tag{3.48}$$

In fact

$$\begin{pmatrix} Rp(\,\cdot\,,0) \\ Rp(\,\cdot\,,1) \end{pmatrix} = \begin{pmatrix} -\delta' p(\,\cdot\,,1) \\ \delta' p(\,\cdot\,,1) \end{pmatrix} = \begin{pmatrix} p(\,\cdot\,,0) & p(\,\cdot\,,1) \end{pmatrix} \begin{pmatrix} 0 & 0 \\ -\delta' & \delta' \end{pmatrix}, \tag{3.49}$$

hence

$$\begin{pmatrix} R'(0,0) & R'(0,1) \\ R'(1,0) & R'(1,1) \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ -\delta' & \delta' \end{pmatrix}, \tag{3.50}$$

which we reckognize as the generator of a Markov process on $\{0,1\}$ that jumps from 1 to 0 with rate δ'. ∎

3.8 The two-level system

Proof of Lemma 7 We may write (3.39) in the matrix form

$$\sum_{x'\in S_2}\sum_{y\in S_1} I(x,x')P(x',y)f(y) = \sum_{y'\in S_1}\sum_{y\in S_1} P(x,y')I'_x(y',y)f(y) \qquad (x\in S_2,\ f\in\mathbb{R}^{S_1}), \tag{3.51}$$

which is equivalent to

$$\sum_{x'\in S_2} I(x,x')P(x',y) = \sum_{y'\in S_1} P(x,y')I'_x(y',y) \qquad (x\in S_2,\ y\in S_1). \tag{3.52}$$

Here

$$\sum_{x'\in S_2} I(x,x')P(x',y) = IP(\,\cdot\,,y)(x) \qquad (x\in S_2,\ y\in S_1). \tag{3.53}$$

Note that $S_n = \{0,1\}^{\Omega^n} = \{0,1\}^{\{0,1\}^n}$ has 2^{2^n} elements, so $|S_1| = 2^2 = 4$ and $|S_2| = 2^4 = 16$, hence $(IP(\,\cdot\,,y)(x))_{x\in S_1,\ y\in S_2}$ is a matrix with $4\cdot 16 = 64$ entries. Luckily, using symmetry, we can reduce the size of our problem quite a bit. We start by calculating

$$P(x,y) = P_y(x_0,x_1) = p(x_0,y(0))p(x_1,y(1)) \tag{3.54}$$

for $x_0,x_1,y\in\{(0,0),(0,1),(1,1)\}$. For brevity, we write $00=(0,0)$, $01=(0,0)$, and $11=(1,1)$. We have

$$\begin{pmatrix} P_{00}(00,00) & P_{00}(00,01) & P_{00}(00,11) \\ P_{00}(01,00) & P_{00}(01,01) & P_{00}(01,11) \\ P_{00}(11,00) & P_{00}(11,01) & P_{00}(11,11) \end{pmatrix} = \begin{pmatrix} 1 & \xi & 0 \\ \xi & \xi^2 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \tag{3.55}$$

$$\begin{pmatrix} P_{01}(00,00) & P_{01}(00,01) & P_{01}(00,11) \\ P_{01}(01,00) & P_{01}(01,01) & P_{01}(01,11) \\ P_{01}(11,00) & P_{01}(11,01) & P_{01}(11,11) \end{pmatrix} = \begin{pmatrix} 0 & 1-\xi & 1 \\ 0 & \xi(1-\xi) & \xi \\ 0 & 0 & 0 \end{pmatrix}, \tag{3.56}$$

and

$$\begin{pmatrix} P_{11}(00,00) & P_{11}(00,01) & P_{11}(00,11) \\ P_{11}(01,00) & P_{11}(01,01) & P_{11}(01,11) \\ P_{11}(11,00) & P_{11}(11,01) & P_{11}(11,11) \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & (1-\xi)^2 & 1-\xi \\ 0 & 1-\xi & 1 \end{pmatrix}. \tag{3.57}$$

Recall the definition of $\overline{x}$ from (3.25). If $(X(t))_{t\geq 0} = (X_0(t), X_1(t))_{t\geq 0}$ is a Markov process in $S_2 = S_1 \times S_1$ with generator I, then $(\overline{X}_0(t), \overline{X}_1(t))_{t\geq 0}$ is a Markov process that jumps with the following rates:

$$
\begin{array}{ccc}
(00,00) & (00,01) & (00,11) \\
 & \Big\downarrow 1 & \Big\downarrow 2 \\
(01,00) & (01,01) & (01,11) \\
 & \Big\downarrow \tfrac{1}{2} & \Big\downarrow 1 \\
(11,00) & (11,01) & (11,11).
\end{array}
\tag{3.58}
$$

From this, we see that the functions $IP(\,\cdot\,,y)(x) = IP_y(x)$ are given by

$$
\begin{pmatrix} IP_{00}(00,00) & IP_{00}(00,01) & IP_{00}(00,11) \\ IP_{00}(01,00) & IP_{00}(01,01) & IP_{00}(01,11) \\ IP_{00}(11,00) & IP_{00}(11,01) & IP_{00}(11,11) \end{pmatrix} = \begin{pmatrix} 0 & -\xi(1-\xi) & 0 \\ 0 & -\frac{1}{2}\xi^2 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \tag{3.59}
$$

$$
\begin{pmatrix} IP_{01}(00,00) & IP_{01}(00,01) & IP_{01}(00,11) \\ IP_{01}(01,00) & IP_{01}(01,01) & IP_{01}(01,11) \\ IP_{01}(11,00) & IP_{01}(11,01) & IP_{01}(11,11) \end{pmatrix} = \begin{pmatrix} 0 & -(1-\xi)^2 & -2(1-\xi) \\ 0 & -\frac{1}{2}\xi(1-\xi) & -\xi \\ 0 & 0 & 0 \end{pmatrix}, \tag{3.60}
$$

and

$$
\begin{pmatrix} IP_{11}00,00) & IP_{11}00,01) & IP_{11}00,11) \\ IP_{11}01,00) & IP_{11}01,01) & IP_{11}01,11) \\ IP_{11}11,00) & IP_{11}11,01) & IP_{11}11,11) \end{pmatrix} = \begin{pmatrix} 0 & (1-\xi)^2 & 2(1-\xi) \\ 0 & \frac{1}{2}\xi(1-\xi) & \xi \\ 0 & 0 & 0 \end{pmatrix}. \tag{3.61}
$$

We wish to express the functions $(IP(\,\cdot\,,y))_{y\in S_1}$ in the functions $(P(\,\cdot\,,y))_{y\in S_1}$. Unlike in the previous section, the span of the functions $(P(\,\cdot\,,y))_{y\in S_1}$ is not invariant under the operator I, so we cannot express the functions $(IP(\,\cdot\,,y))_{y\in S_1}$ as a linear combination of the functions $(P(\,\cdot\,,y))_{y\in S_1}$. However, we can find expressions of the form (compare (3.52))

$$
IP(\,\cdot\,,y)(x) = \sum_{y'\in S_1} P(x,y')I'_x(y',y) \qquad (x\in S_2,\ y\in S_1), \tag{3.62}
$$

where the coefficients $I'_x(y',y)$ do not depend too strongly on x. Solutions to this problem are not unique. The claim of Lemma 7 is that we can choose

$$
\begin{aligned}
&\begin{pmatrix} I'_x(00,00) & I'_x(00,01) & I'_x(00,10) & I'_x(00,11) \\ I'_x(01,00) & I'_x(01,01) & I'_x(01,10) & I'_x(01,11) \\ I'_x(10,00) & I'_x(10,01) & I'_x(10,10) & I'_x(10,11) \\ I'_x(11,00) & I'_x(11,01) & I'_x(11,10) & I'_x(11,11) \end{pmatrix} \\
&\qquad = \begin{pmatrix} -b(\overline{x}_0,\overline{x}_1) & 0 & b(\overline{x}_0,\overline{x}_1) & 0 \\ 0 & -a(\overline{x}_0,\overline{x}_1) & 0 & a(\overline{x}_0,\overline{x}_1) \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix},
\end{aligned}
\tag{3.63}
$$

where a, b are the functions in (3.27)–(3.28). Thus, we need to check that

$$
\begin{array}{rl}
\text{(i)} & IP(\,\cdot\,,00)(x) = -b(\overline{x}_0,\overline{x}_1)P(x,00), \\
\text{(ii)} & IP(\,\cdot\,,01)(x) = -a(\overline{x}_0,\overline{x}_1)P(x,01), \\
\text{(iii)} & IP(\,\cdot\,,10)(x) = b(\overline{x}_0,\overline{x}_1)P(x,00), \\
\text{(iv)} & IP(\,\cdot\,,11)(x) = a(\overline{x}_0,\overline{x}_1)P(x,01).
\end{array}
\tag{3.64}
$$

Since $\sum_{y\in S_1} IP(\,\cdot\,,y) = I1 = 0$, it suffices to check only three of these equations, say (i), (ii), and (iv). We observe from (3.60)–(3.61) that $IP(\,\cdot\,,01) = -IP(\,\cdot\,,11)$. In view of this, it suffices to check only (i) and (ii). By (3.27)–(3.28), (3.55)–(3.56), and (3.59)–(3.60), we need to check that

$$\left(\begin{array}{ccc} 0 & -\xi(1-\xi) & 0 \\ 0 & -\frac{1}{2}\xi^2 & 0 \\ 0 & 0 & 0 \end{array}\right) = -\left(\begin{array}{ccc} 0 & 1-\xi & * \\ 0 & \frac{1}{2} & * \\ * & * & * \end{array}\right) \bullet \left(\begin{array}{ccc} 1 & \xi & 0 \\ \xi & \xi^2 & 0 \\ 0 & 0 & 0 \end{array}\right) \tag{3.65}$$

and

$$\left(\begin{array}{ccc} 0 & -(1-\xi)^2 & -2(1-\xi) \\ 0 & -\frac{1}{2}\xi(1-\xi) & -\xi \\ 0 & 0 & 0 \end{array}\right) = -\left(\begin{array}{ccc} * & 1-\xi & 2(1-\xi) \\ * & \frac{1}{2} & 1 \\ * & * & * \end{array}\right) \bullet \left(\begin{array}{ccc} 0 & 1-\xi & 1 \\ 0 & \xi(1-\xi) & \xi \\ 0 & 0 & 0 \end{array}\right), \tag{3.66}$$

where $\bullet$ denotes the componentwise product of functions and the symbol $*$ indicates that the value of the functions a and b in these points is irrelevant. We see by inspection that (3.65) and (3.66) are satisfied. ∎

4 Survival

4.1 Survival bounds

Until further notice, we continue to study the contact process on the hierarchical group Ω_N with $N = 2$ and its finite analogues defined in Section 3.1. Our proof of Theorem 1 (b) is based on the following basic estimate.

Proposition 8 (Survival bound for finite systems) *Let $\delta > 0$ and let $(\alpha_k)_{k\geq 1}$ be non-negative constants. Let $X^{(n)}$ be the $(\delta,\alpha_1,\ldots,\alpha_n)$-contact process started in $X^{(n)}_0 = \delta_0$. Then*

$$\mathbb{P}^{\delta_0}[X^{(n)}_t \neq \underline{0}] \geq \Big(\prod_{k=0}^{n-1}(1-\xi(k))\Big) e^{-\delta(n)t} \qquad (t\geq 0), \tag{4.1}$$

where $\delta(0) := \delta$, $\alpha_k(0) := \alpha_k$ $(k \geq 1)$, and we define inductively, for $n \geq 0$,

$$\begin{aligned} \delta(n+1) &:= 2\xi(n)\delta(n), \\ \alpha_k(n+1) &:= \tfrac{1}{2}\alpha_{k+1}(n) \qquad (k\geq 1), \end{aligned} \tag{4.2}$$

where $\xi(n) := f(\alpha_1(n)/\delta(n))$ with f as in (3.1).

Proof By Lemma 3, one has $0 < \xi(n) \leq \frac{1}{2}$ for all $n \geq 0$. For $0 < \xi \leq \frac{1}{2}$ and $k \geq 1$, let $P_{k,\xi}$ denote the probability kernel from S_k to S_{k-1} defined in (3.3)–(3.4). Let $X^{(n)}$ be the $(\delta,\alpha_1,\ldots,\alpha_n)$-contact process. Applying Proposition 2 inductively, we can couple $X^{(n)}$ to processes

$$\tilde X^{(n-1)}, X^{(n-1)}, \ldots, \tilde X^{(0)}, X^{(0)}$$

such that $\tilde X^{(n-m)}, X^{(n-m)}$ take values in S_{n-m}, one has $\tilde X^{(n-m)}_0 = X^{(n-m)}_0$, $\tilde X^{(n-m)}_t \geq X^{(n-m)}_t$ for all $t \geq 0$,

$$\mathbb{P}\big[\tilde X^{(n-m-1)}_t = y \,\big|\, X^{(n-m)}_t = x\big] = P_{n-m,\xi(m)}(x,y), \tag{4.3}$$

and the process $X^{(n-m)}$ is a $(\delta(m), \alpha_1(m), \ldots, \alpha_n(m))$-contact process. A little thinking convinces us that this coupling can be done in a Markovian way, i.e., in such a way that $(\tilde{X}^{(n-m-1)}, X^{(n-m-1)})$ is conditionally independent of

$$X^{(n)}, (\tilde{X}^{(n-1)}, X^{(n-1)}), \ldots, (\tilde{X}^{(n-m+1)}, X^{(n-m+1)})$$

given $(\tilde{X}^{(n-m)}, X^{(n-m)})$. By this Markovian property and the definition of $P_{k,\xi}(x,y)$, if we start $X^{(n)}$ in the initial state $X_0^{(n)} = \delta_0$, then

$$\mathbb{P}[X_0^{(n-m)} = \delta_0] = \prod_{k=0}^{m-1} (1 - \xi(k)), \tag{4.4}$$

and $X_0^{(n-m)} = \underline{0}$ with the remaining probability. Since $\mathbb{P}[X_t^{(n-m-1)} = \underline{0} \,|\, X_t^{(n-m)} = \underline{0}] = 1$ for each m, we have

$$\mathbb{P}[X_t^{(n)} \neq \underline{0}] \geq \mathbb{P}[X_t^{(n-m)} \neq \underline{0}] \qquad (0 \leq m \leq n). \tag{4.5}$$

In particular, since $X^{(0)}$ is a Markov process in $S_0 = \{0, 1\}$ that jumps from 1 to 0 with rate $\delta(n)$, we observe that

$$\mathbb{P}[X_t^{(n)} \neq \underline{0}] \geq \mathbb{P}[X_t^{(0)} \neq \underline{0}] = e^{-\delta(n)t}\, \mathbb{P}[X_0^{(0)} = \delta_0] = e^{-\delta(n)t} \prod_{k=0}^{n-1} (1 - \xi(k)) \qquad (t \geq 0), \tag{4.6}$$

which proves (4.1). ∎

As an immediate corollary to Proposition 8, we obtain:

Proposition 9 (Survival bound for infinite systems) *Let $\delta > 0$ and let $(\alpha_k)_{k \geq 1}$ be nonnegative constants satisfying $\sum_{k=1}^{\infty} \alpha_k < \infty$. Let $(\xi(k))_{k \geq 0}$ be defined as in Proposition 8. Let X be the contact process on Ω_2 with infection rates as in (1.6) and recovery rate δ. Then the process started in $X_0 = \delta_0$ satisfies*

$$\mathbb{P}^{\delta_0}[X_t \neq \underline{0} \ \forall t \geq 0] \geq \prod_{k=0}^{\infty} (1 - \xi(k)). \tag{4.7}$$

Proof It is easy to see that the process $X^{(n)}$ in Proposition 8 and X can be coupled such that $X_t^{(n)} \leq X_t$ for all $t \geq 0$. Therefore (4.7) follows from (4.1), provided we show that $\delta(k) \to 0$ as $k \to \infty$. In fact, it suffices to prove this under the assumption that $\prod_{k=0}^{\infty}(1 - \xi(k)) > 0$, for otherwise (4.7) is trivial. Indeed, $\prod_{k=0}^{\infty}(1 - \xi(k)) > 0$ implies that $\xi(k) \to 0$ as $k \to \infty$, which by the fact that

$$\delta(n) = \delta \prod_{k=0}^{n-1} \big(2\xi(k)\big) \tag{4.8}$$

implies that $\delta(k) \to 0$ as $k \to \infty$. ∎

4.2 The critical recovery rate

In view of Proposition 9, we wish to find sufficient conditions for $\prod_{k=0}^{\infty}(1-\xi(k)) > 0$. The next lemma shows that this is equivalent to showing that the $\xi(k)$'s are summable.

Lemma 10 (Infinite products) *Let $(\xi(k))_{k\geq 0}$ be constants, $0 < \xi(k) < 1$. Then one has $\prod_{k=0}^{\infty}\big(1-\xi(k)\big) > 0$ if and only if $\sum_{k=0}^{\infty}\xi(k) < \infty$.*

Proof Since $1-\xi-\xi' \leq (1-\xi)(1-\xi')$ and $\log(1-\xi) \leq -\xi$ for all $0 < \xi, \xi' < 1$, we see that

$$\Big(\prod_{k=0}^{m-1}\big(1-\xi(k)\big)\Big)\Big(1-\sum_{k=m}^{\infty}\xi(k)\Big) \leq \prod_{k=0}^{\infty}\big(1-\xi(k)\big) = e^{\sum_{k=0}^{\infty}\log(1-\xi(k))} \leq e^{-\sum_{k=0}^{\infty}\xi(k)}. \tag{4.9}$$

Here the left-hand side is positive for m large enough if and only if $\sum_{k=0}^{\infty}\xi(k) < \infty$, while obviously the right-hand side is positive if and only if $\sum_{k=0}^{\infty}\xi(k) < \infty$. ∎

The next lemma casts the inductive formula (4.2) in a more tractable form.

Lemma 11 (Inductive formula) *Let $\delta(n), \alpha_k(n)$, and $\xi(n)$ be defined as in Proposition 9 and assume that the constants $(\alpha_k)_{k\geq 1}$ are positive. Set $\varepsilon(k) := \delta(k)/\alpha_1(k)$ $(k \geq 0)$. Then $\xi(k) = f(1/\varepsilon(k))$ and*

$$\varepsilon(k+1) = \tfrac{\alpha_{k+1}}{\alpha_{k+2}}g(\varepsilon(k)) \qquad (k \geq 0), \tag{4.10}$$

where

$$g(\varepsilon) := 4\varepsilon f(1/\varepsilon) \quad (\varepsilon > 0), \tag{4.11}$$

and f is the function defined in (3.1).

Proof It is clear from (4.2) that $\xi(n) = f(1/\varepsilon(n))$ and

$$\alpha_k(n) = 2^{-n}\alpha_{k+n} \qquad (k \geq 1,\ n \geq 0). \tag{4.12}$$

Using (4.2) once more, it follows that

$$\varepsilon(n+1) = \frac{2\xi(n)\delta(n)}{\frac{1}{2}\alpha_2(n)} = \frac{\alpha_1(n)}{\alpha_2(n)}\frac{\delta(n)}{\alpha_1(n)}4f(\alpha_1(n)/\delta(n)) = \frac{2^{-n}\alpha_{n+1}}{2^{-n}\alpha_{n+2}}4\varepsilon(n)f(1/\varepsilon(n)). \tag{4.13}$$

∎

The next lemma collects some elementary facts about the function g from Lemma 11.

Lemma 12 (The function g) *The function g defined in (4.11) is increasing on $(0,\infty)$ and satisfies*

$$g(\varepsilon) = 8\varepsilon^2 + O(\varepsilon^3) \qquad \text{as} \quad \varepsilon \to 0. \tag{4.14}$$

Proof This follows from the fact that, by Lemma 3, the function $\varepsilon \mapsto f(1/\varepsilon)$ is increasing and satisfies

$$f(1/\varepsilon) = 2\varepsilon + O(\varepsilon^2) \qquad \text{as} \quad \varepsilon \to 0. \tag{4.15}$$

∎

The next proposition answers the question when the infinite product in (4.7) is positive for δ small enough.

Proposition 13 (Nontrivial survival bound) *Let* $(\alpha_k)_{k\geq 0}$ *be nonnegative constants. For given* $\delta > 0$*, set* $\Pi(\delta) := \prod_{k=0}^{\infty}(1-\xi(k))$*, where the* $(\xi(k))_{k\geq 0}$ *are defined as in Proposition 8. Then* $\Pi(\delta)$ *is nonincreasing in* δ*. Moreover,* $\Pi(\delta) > 0$ *for* δ *suffiently small if and only if*

$$\sum_{k=m}^{\infty} 2^{-k}\log(\alpha_k) > -\infty \quad \textit{for some} \quad m \geq 0. \tag{4.16}$$

Proof We start by showing that $\Pi(\delta)$ is nonincreasing in δ. By continuity, it suffices to prove this under the additional assumption that the α_k's are all positive. In this case, we observe from Lemma 11 and the monotonicity of g that the $\varepsilon(k)$'s are nondecreasing in δ. Since $\xi(k) = f(1/\varepsilon(k))$ and f is decreasing, it follows that the $\xi(k)$'s are nondecreasing in δ, hence $\Pi(\delta)$ is nonincreasing in δ.

We next show that $\Pi(\delta) > 0$ for $\delta > 0$ suffiently small if and only if (4.16) holds. If $\alpha_k = 0$ for some $k \geq 1$, then $\xi(k-1) = f(0) = \frac{1}{2}$, hence if infinitely many of the α_k's are zero then $\Pi(\delta) = 0$ for all $\delta > 0$, while (4.16) is obviously violated. If finitely many of the α_k's are zero, then we may start our inductive formulas after the first m iterations, where we observe that $\delta(m)$ can be made arbitrarily small by choosing δ small enough. Thus, without loss of generality, we may assume that the α_k's are all positive, and under this assumption we need to show that $\Pi(\delta) > 0$ for δ suffiently small if and only if

$$\sum_{k=0}^{\infty} 2^{-k}\log(\alpha_k) > -\infty. \tag{4.17}$$

By Lemma 10, $\Pi(\delta) > 0$ if and only if $\sum_{k=0}^{\infty}\xi(k) < \infty$. Using (4.15) and the fact that $\xi(k) = f(1/\varepsilon(k))$, it is easy to see that this is equivalent to $\sum_{k=0}^{\infty}\varepsilon(k) < \infty$.

Now assume that (4.17) holds, and, in view of (4.14), define $(\tilde\varepsilon(k))_{k\geq 0}$ by

$$\tilde\varepsilon(0) := \frac{\delta}{\alpha_1} \quad \text{and} \quad \tilde\varepsilon(k+1) := 9\tfrac{\alpha_{k+1}}{\alpha_{k+2}}(\tilde\varepsilon(k))^2 \qquad (k \geq 0). \tag{4.18}$$

Then

$$\begin{aligned}
\tilde\varepsilon(0) &= \tfrac{\delta}{\alpha_1},\\
\tilde\varepsilon(1) &= 9\tfrac{\alpha_1}{\alpha_2}(\tfrac{\delta}{\alpha_1})^2,\\
\tilde\varepsilon(2) &= 9\tfrac{\alpha_2}{\alpha_3}(9\tfrac{\alpha_1}{\alpha_2})^2(\tfrac{\delta}{\alpha_1})^4,\\
\tilde\varepsilon(3) &= 9\tfrac{\alpha_3}{\alpha_4}(9\tfrac{\alpha_2}{\alpha_3})^2(9\tfrac{\alpha_1}{\alpha_2})^4(\tfrac{\delta}{\alpha_1})^8 = \frac{\frac{1}{9}(9\delta)^8}{\alpha_4\alpha_3\alpha_2^2\alpha_1^4}.
\end{aligned} \tag{4.19}$$

More generally, it is not hard to see that

$$\tilde\varepsilon(n) = \frac{\frac{1}{9}(9\delta)^{2^n}}{\alpha_{n+1}\prod_{k=1}^{n}(\alpha_k)^{2^{n-k}}} \tag{4.20}$$

By Lemma 14 below, we can choose δ sufficiently small such that $\sum_{n=0}^{\infty}\tilde\varepsilon(n) < \infty$. By (4.14), there exists a $c > 0$ such that $g(\varepsilon) \leq 9\varepsilon^2$ for all $\varepsilon \leq c$. By making δ smaller if necessary, we can arrange that $\tilde\varepsilon(n) \leq c$ for all n, hence $\sum_{n=0}^{\infty}\varepsilon(n) \leq \sum_{n=0}^{\infty}\tilde\varepsilon(n) < \infty$.

On the other hand, assume that $\sum_{n=0}^{\infty}\varepsilon(n) < \infty$ for some $\delta > 0$ while (4.17) does not hold. Define $(\tilde\varepsilon(k))_{k\geq 0}$ as in (4.18) but with the factor 9 replaced by 7. By (4.14), there exists a $c > 0$ such that $7\varepsilon^2 \leq g(\varepsilon)$ for all $\varepsilon \leq c$. Making δ smaller if necessary, we can arrange that $\varepsilon(n) \leq c$ for all n, hence $\varepsilon(n) \geq \tilde\varepsilon(n)$ for all n. Since $\tilde\varepsilon(n) \to \infty$ by Lemma 14 below, this leads to a contradiction. ∎

Lemma 14 (Summability) *For $\eta > 0$, set*

$$F_\eta(n) := \frac{\eta^{2^n}}{\alpha_{n+1}\prod_{k=1}^{n}(\alpha_k)^{2^{n-k}}} \qquad (n \geq 0). \tag{4.21}$$

If (4.17) holds, then $\sum_{n=0}^{\infty} F_\eta(n) < \infty$ for η sufficiently small. On the other hand, if (4.17) does not hold, then $\lim_{n\to\infty} F_\eta(n) = \infty$ for all $\eta > 0$.

Proof We start by observing that

$$\begin{aligned}
&\limsup_{n\to\infty} F_\eta(n) \leq 1 \\
&\quad\Leftrightarrow\quad \limsup_{n\to\infty} \Big(2^n \log(\eta) - \log(\alpha_{n+1}) - \sum_{k=1}^{n} 2^{n-k}\log(\alpha_k)\Big) \leq 0 \\
&\quad\Leftrightarrow\quad \limsup_{n\to\infty} \Big(\log(\eta) - 2^{-n}\log(\alpha_{n+1}) - \sum_{k=1}^{n} 2^{-k}\log(\alpha_k)\Big) \leq 0 \\
&\quad\Leftrightarrow\quad \log(\eta) - \sum_{k=1}^{\infty} 2^{-k}\log(\alpha_k) \leq 0,
\end{aligned} \tag{4.22}$$

which is satisfied for η sufficiently small if (4.17) holds. In this case, we may choose $\eta > 0$ such that $K := \sup_{n\geq 0} F_\eta(n) < \infty$ and observe that for any $\eta' < \eta$

$$\sum_{n=0}^{\infty} F_\eta(n) \leq K \sum_{n=0}^{\infty} \big(\frac{\eta'}{\eta}\big)^{2^n} < \infty. \tag{4.23}$$

On the other hand, if (4.17) does not hold, then a calculation as in (4.22) shows that $\liminf_{n\to\infty} F_\eta(n) \geq 1$ for any $\eta > 0$, and therefore, for any $0 < \eta < \eta'$,

$$\liminf_{n\to\infty} F_{\eta'}(n) \geq \liminf_{n\to\infty} \big(\frac{\eta'}{\eta}\big)^{2^n} F_\eta(n) = \infty. \tag{4.24}$$

∎

4.3 Comparison arguments

Proof of Theorem 1 (b) For $N = 2$, Theorem 1 (b) follows from Propositions 9 and 13. We next consider the case that $N = 2^n$ for some integer $n \geq 1$. Let X be our contact process of interest and let $\tilde{X}$ be the contact process on Ω_2 with the same recovery rate and infection rates

$$a(i,j) = \tilde{\alpha}_{|i-j|} 2^{-|i-j|}, \tag{4.25}$$

where

$$\begin{aligned}
\tilde{\alpha}_1 2^{-1} = \tilde{\alpha}_2 2^{-2} = \cdots = \tilde{\alpha}_n 2^{-n} &= \alpha_1 (2^n)^{-1}, \\
\tilde{\alpha}_{n+1} 2^{-(n+1)} = \tilde{\alpha}_{n+2} 2^{-(n+2)} = \cdots = \tilde{\alpha}_{2n} 2^{-2n} &= \alpha_2 (2^n)^{-2},
\end{aligned} \tag{4.26}$$

etcetera. Then it is easy to see that $\tilde{X}$ can be identified with X. Since

$$\tilde{\alpha}_{kn-m} 2^{-(kn-m)} = \alpha_k (2^n)^{-k} \quad\Leftrightarrow\quad \tilde{\alpha}_{kn-m} = \alpha_k 2^{-m} \qquad (k \geq 1,\ 0 \leq m \leq n-1), \tag{4.27}$$

we have, by (1.8),

$$\sum_{k=1}^{\infty} 2^{-k}\log(\tilde\alpha_k) = \sum_{k=1}^{\infty}\sum_{m=0}^{n-1} 2^{-(kn-m)}\log(\tilde\alpha_{kn-m}) = \sum_{k=1}^{\infty}(2^n)^{-k}\sum_{m=0}^{n-1}2^m\big(\log(\alpha_k)-m\log(2)\big)$$
$$=\Big(\sum_{m=0}^{n-1}2^m\Big)\sum_{k=1}^{\infty}(2^n)^{-k}\log(\alpha_k)-\log(2)\Big(\sum_{m=0}^{n-1}m2^m\Big)\sum_{k=1}^{\infty}(2^n)^{-k}>-\infty, \tag{4.28}$$

hence $\delta_{\rm c}>0$ by the statement of the theorem for $N=2$.

In general, if $2^n=M\leq N$, we let $\tilde X$ be the contact process on Ω_M with the same recovery rate and with infection rates given by

$$a(i,j)=\alpha_{|i-j|}N^{-|i-j|}=:\tilde\alpha_{|i-j|}M^{-|i-j|}\qquad (i,j\in\Omega_M,\ i\neq j), \tag{4.29}$$

In an obvious way, we may regard Ω_M as a subset of Ω_N, and by surpressing infections that go outside Ω_M, we may estimate X from below by $\tilde X$. Thus, by what we have just proved, the critical recovery rate for X will be positive provided that

$$\sum_{k=1}^{\infty}M^{-k}\log(\tilde\alpha_k)>-\infty, \tag{4.30}$$

which is satisfied since our assumption (1.8) implies that

$$\sum_{k=1}^{\infty}M^{-k}\log(\tilde\alpha_k)=\sum_{k=1}^{\infty}M^{-k}\log((\tfrac{2}{N})^k\alpha_k)=\sum_{k=1}^{\infty}M^{-k}(k\log(\tfrac{2}{N})+\log(\alpha_k))>-\infty. \tag{4.31}$$

∎

A Coordinate reduction

In this appendix we prove that Lemmas 6 and 7 imply formulas (3.36) (i) and (ii), respectively. The main problem is to invent good notation. Recall that $S_n=\{0,1\}^{\Omega^n}$. For any $x\in S_n$ and $\Delta\subset\Omega^n$, we let

$$x\big|_\Delta:=(x(i))_{i\in\Delta} \tag{A.1}$$

denote the restriction of x to Δ. If Δ,Δ' are disjoint sets, $x\in\{0,1\}^\Delta$ and $x'\in\{0,1\}^{\Delta'}$, then we define $x\,\&\,x'\in\{0,1\}^{\Delta\cup\Delta'}$ by

$$(x\,\&\,x')(i):=\begin{cases}x(i) & \text{if } i\in\Delta,\\ x'(i) & \text{if } i\in\Delta'\end{cases} \tag{A.2}$$

For each $i\in\Omega^{n-1}$, we define $B_i\subset\Omega^n$ by (recall (2.4))

$$B(i):=B_1(i)=\{i'\circ i: i'\in\Omega^n\}. \tag{A.3}$$

Let R,R' be as in Lemma 6. Then we can write

$$\begin{aligned}Rf(x)&=\sum_{x'\in S_1}R(x,x')f(x'),\\ R'f(y)&=\sum_{y'\in S_0}R'(y,y')f(y'),\end{aligned} \tag{A.4}$$

where $R(x,x')$ and $R'(y,y')$ are the matrices of R and R', respectively. We observe that

$$\begin{aligned} R_i f(x) &= \sum_{z\in\{0,1\}^{B(i)}} R(x|_{B(i)}, z) f(x|_{\Omega^n\setminus B(i)} \,\&\, z), \\ R'_i f(y) &= \sum_{z\in\{0,1\}^{\{i\}}} R'(y(i), z) f(y|_{\Omega^n\setminus\{i\}} \,\&\, z), \end{aligned} \tag{A.5}$$

where we identify $\{0,1\}^{B(i)} \cong \{0,1\}^{\Omega^1} = S_1$ and $\{0,1\}^{\{i\}} \cong \{0,1\}^{\Omega^0} = S_0$. Moreover,

$$Pf(x) = \sum_{y\in\{0,1\}^{\Omega^{n-1}}} \Big(\prod_{j\in\Omega^{n-1}} p(x_j, y(j)) \Big) f(y). \tag{A.6}$$

Using the identification $x|_{B(i)} \cong x_i$, we calculate

$$\begin{aligned} R_i P f(x) &= \sum_{z\in\{0,1\}^{B(i)}} R(x_i, z) P f(x|_{\Omega^n\setminus B(i)} \,\&\, z) \\ &= \sum_{z\in\{0,1\}^{B(i)}} R(x_i, z) \sum_{y\in\{0,1\}^{\Omega^{n-1}}} \Big(\prod_{j\in\Omega^{n-1}} p((x|_{\Omega^n\setminus B(i)} \,\&\, z)_j, y(j)) \Big) f(y) \\ &= \sum_{y\in\{0,1\}^{\Omega^{n-1}}} \sum_{z\in\{0,1\}^{B(i)}} R(x_i, z) p(z, y(i)) \Big(\prod_{j\in\Omega^{n-1}\setminus\{i\}} p(x_j, y(j)) \Big) f(y) \\ &= \sum_{y\in\{0,1\}^{\Omega^{n-1}}} \sum_{z\in\{0,1\}^{\{i\}}} p(x_i, z) R'(z, y(i)) \Big(\prod_{j\in\Omega^{n-1}\setminus\{i\}} p(x_j, y(j)) \Big) f(y) \\ &= \sum_{y\in\{0,1\}^{\Omega^{n-1}}} \sum_{z\in\{0,1\}^{\{i\}}} p(x_i, y(i)) R'(y(i), z) \Big(\prod_{j\in\Omega^{n-1}\setminus\{i\}} p(x_j, y(j)) \Big) f(y|_{\Omega^{n-1}\setminus\{i\}} \,\&\, z) \\ &= \sum_{y\in\{0,1\}^{\Omega^{n-1}}} \Big(\prod_{j\in\Omega^{n-1}} p(x_j, y(j)) \Big) \sum_{z\in\{0,1\}^{\{i\}}} R'(y(i), z) f(y|_{\Omega^{n-1}\setminus\{i\}} \,\&\, z) = P R'_i f. \end{aligned} \tag{A.7}$$

Here we have used Lemma 6 in the fourth equality. In the fifth equality, we have reordered our sums by relabelling $y(i)$ and z.

The formal proof of formula (3.36) (ii) is similar, but even more cumbersome. Letting I and I'_x be as in Lemma 7, we can write, in matrix notation,

$$\begin{aligned} If(x) &= \sum_{x'\in S_2} I(x;x') f(x') = \sum_{z\in S_1} \sum_{z'\in S_1} I(x_0, x_1; z, z') f(z, z'), \\ I'_x f(y) &= \sum_{y'\in S_1} I'_x(y;y') f(y') = \sum_{z\in\{0,1\}} \sum_{z'\in\{0,1\}} I_{x_0,x_1}(y(0), y(1); z, z') f(z, z'). \end{aligned} \tag{A.8}$$

Then

$$\begin{aligned} I_{ij} f(x) &= \sum_{z\in\{0,1\}^{B(i)}} \sum_{z'\in\{0,1\}^{B(j)}} I(x|_{B(i)}, x|_{B(j)}; z, z') f(x|_{\Omega^n\setminus(B(i)\cup B(j))} \,\&\, z \,\&\, z'), \\ \overline{I}_{ij} f(x,y) &= \sum_{z\in\{0,1\}^{\{i\}}} \sum_{z'\in\{0,1\}^{\{j\}}} I'_{x|_{B(i)}, x|_{B(j)}}(y(i), y(j); z, z') f(y|_{\Omega^{n-1}\setminus\{i,j\}} \,\&\, z \,\&\, z'), \end{aligned} \tag{A.9}$$

and

$$\overline{P} f(x) = \sum_{y\in\{0,1\}^{\Omega^{n-1}}} \Big(\prod_{k\in\Omega^{n-1}} p(x_k, y(k)) \Big) f(x,y). \tag{A.10}$$

Using the fact that $x|_{B(i)} \cong x_i$, we calculate

$$
\begin{aligned}
I_{ij}Pf(x) = &\sum_{z\in\{0,1\}^{B(i)}} \sum_{z'\in\{0,1\}^{B(j)}} I(x_i, x_j; z, z')Pf(x|_{\Omega^n\setminus(B(i)\cup B(j))} \,\&\, z \,\&\, z') \\
= &\sum_{z\in\{0,1\}^{B(i)}} \sum_{z'\in\{0,1\}^{B(j)}} I(x_i, x_j; z, z') \\
&\cdot \sum_{y\in\{0,1\}^{\Omega^{n-1}}} \Big(\prod_{k\in\Omega^{n-1}} p((x|_{\Omega^n\setminus(B(i)\cup B(j))} \,\&\, z \,\&\, z')_k, y(k)) \Big) f(y) \\
= &\sum_{y\in\{0,1\}^{\Omega^{n-1}}} \sum_{z\in\{0,1\}^{B(i)}} \sum_{z'\in\{0,1\}^{B(j)}} I(x_i, x_j; z, z')p(z_i, y(i))p(z_j, y(j)) \\
&\cdot \Big(\prod_{k\in\Omega^{n-1}\setminus\{i,j\}} p(x_k, y(k)) \Big) f(y) \\
= &\sum_{y\in\{0,1\}^{\Omega^{n-1}}} \sum_{z\in\{0,1\}^{\{i\}}} \sum_{z'\in\{0,1\}^{\{j\}}} p(x_i, z)p(x_j, z')I'_{x_i,x_j}(z, z'; y(i), y(j)) \\
&\cdot \Big(\prod_{k\in\Omega^{n-1}\setminus\{i,j\}} p(x_k, y(k)) \Big) f(y) \\
= &\sum_{y\in\{0,1\}^{\Omega^{n-1}}} \sum_{z\in\{0,1\}^{\{i\}}} \sum_{z'\in\{0,1\}^{\{j\}}} p(x_i, y(i))p(x_j, y(j))I'_{x_i,x_j}(y(i), y(j); z, z') \\
&\cdot \Big(\prod_{k\in\Omega^{n-1}\setminus\{i,j\}} p(x_k, y(k)) \Big) f(y|_{\Omega^{n-1}\setminus\{i,j\}} \,\&\, z \,\&\, z') \\
= &\sum_{y\in\{0,1\}^{\Omega^{n-1}}} \Big(\prod_{k\in\Omega^{n-1}} p(x_k, y(k)) \Big) \\
&\cdot \sum_{z\in\{0,1\}^{\{i\}}} \sum_{z'\in\{0,1\}^{\{j\}}} I'_{x_i,x_j}(y(i), y(j); z, z')f(y|_{\Omega^{n-1}\setminus\{i,j\}} \,\&\, z \,\&\, z') \\
= &\overline{P}\,\overline{I}_{ij}f(x).
\end{aligned}
\tag{A.11}
$$

Here we have used Lemma 7 in the fourth equality, and in the fifth equality, we have reordered our sums by relabelling $y(i), y(j), z$, and z'.

References

[Daw00] D.A. Dawson. Stochastic models of evolving information systems, (Ottawa, Canada, 1998), 1–14, *CMS Conf. Proc.* 26, AMS, 2000.

[DG93] D.A. Dawson and A. Greven. Hierarchical models of interacting diffusions: multiple time scale phenomena, phase transition and pattern of cluster-formation. *Probab. Theory Relat. Fields* 96, 435–473, 1993.

[DK96] P. Donnelly and T.G. Kurtz. A countable representation of the Fleming-Viot measure-valued diffusion. *Ann. Probab.* 24(2), 698–742, 1996.

[DK99] P. Donnelly and T.G. Kurtz. Genealogical processes for Fleming-Viot models with selection and recombination. *Ann. Appl. Probab.* 9(4), 1091–1148, 1999.

[Dur88] R. Durrett. Lecture notes on particle systems and percolation. Wadsworth & Brooks/Cole, Pacific Grove, 1988.

[FS04] K. Fleischmann and J.M. Swart. Trimmed trees and embedded particle systems. *Ann. Probab.* 32(3A), 2179–2221, 2004.

[HL78] R.A. Holley and T.M. Liggett. The survival of the contact process. Ann. Probab. 6, 198–206, 1978.

[Kur98] T.G. Kurtz. Martingale problems for conditional distributions of Markov processes. Electronic J. Probab. 3, Paper no. 9, 1–29, 1998.

[Lig85] T.M. Liggett. *Interacting Particle Systems.* Springer-Verlag, New York, 1985.

[Lig95] T.M. Liggett. Improved upper bounds for the contact process critical value. *Ann. Probab.* 23, 697–723, 1995.

[Lig99] T.M. Liggett. *Stochastic Interacting Systems: Contact, Voter and Exclusion Process.* Springer-Verlag, Berlin, 1999.

[RP81] L.C.G. Rogers and J.W. Pitman. Markov functions. *Ann. Probab.* 9(4), 573–582, 1981.

[SF83] S. Sawyer and J. Felsenstein. Isolation by distance in a hierarchically clustered population. *J. Appl. Probab.* 20 1–10, 1983.

[Swa07] J.M. Swart. *Extinction versus unbounded growth.* Habilitation Thesis of the University Erlangen-Nürnberg, 2007. ArXiv:math/0702095v1.

[Swa08] J.M. Swart. The contact process seen from a typical infected site. Accepted in *J. Theor. Probab.* ArXiv:math.PR/0507578v4.