THESIS FOR THE DEGREE OF DOCTOR OF PHILOSOPHY

Interacting Particle Systems: Percolation, Stochastic Domination and Randomly Evolving Environments

Erik I. Broman

CHALMERS | GÖTEBORG UNIVERSITY

Department of Mathematical Sciences
Chalmers University of Technology and Göteborg University
Göteborg, Sweden 2006

Interacting Particle Systems: Percolation, Stochastic Domination and Randomly Evolving Environments
Erik Broman
ISBN 91-7291-762-8

Doktorsavhandlingar vid Chalmers tekniska högskola. Ny serie nr 2444.
ISSN 0346-718X

Department of Mathematical Sciences
Division of Mathematics
Chalmers University of Technology and Göteborg University
SE-412 96 Göteborg
Sweden
Telephone +46 (0)31 772 1000

Printed in Göteborg, Sweden 2006

Interacting Particle Systems: Percolation, Stochastic Domination and Randomly Evolving Environments

Erik Broman

Department of Mathematical Sciences
Division of Mathematics
Chalmers University of Technology
Göteborg University

Abstract

In this thesis we first analyze the class of one-dependent trigonometric determinantal processes and show that they are all two-block-factors. We do this by constructing the two-block-factors explicitly.

Second we investigate the dynamic stability of percolation for the stochastic Ising model and the contact process. This is a natural extension of what previously has been done for non-interacting particle systems. The main question we ask is: If we have percolation at a fixed time in a time-dependent but time-invariant system, do we have percolation at all times? A key tool in the analysis is the concept of $\epsilon-$movability which we introduce here. We then proceed by developing and relating this concept to others previously studied.

Finally, we introduce a new model which we refer to as the contact process in a randomly evolving environment. By using stochastic domination techniques we will investigate matters of extinction and that of weak and strong survival for this system. We do this by establishing stochastic relations between our new model and the ordinary contact process. In the process, we develop some sharp stochastic domination results for a hidden Markov chain and a continuous time analogue of this.

Keywords: Determinantal processes, k-dependence, k-block-factors, percolation, stochastic Ising models, contact process, $\epsilon-$movability, hidden Markov chain, stochastic domination

Acknowledgments

First and most I would like to thank my adviser professor Jeffrey E. Steif for always taking the time to help me, for many interesting discussions and ideas, and for presenting me almost all of the problems that I have been working on in this thesis.

Thanks also to my co-author professor Olle Häggström for the cooperation during the work on one of the papers of this thesis, and to my co-adviser associate professor Johan Jonasson for reading and commenting on this introduction as well as for many interesting discussions over the years.

I would also like to thank all of my colleagues at the lunch table and especially my lunch-friend Erik Brodin for many non work-related but nonetheless interesting discussions.

Special gratitude must also go to Axel Målqvist and Per Sundell for much inspiration and good friendship during my undergraduate years and also later. They have considerably influenced my choice of career.

I would like to thank my parents for always supporting me no matter how odd my decisions seemed. Finally I would like to thank my sister Malin for all the encouragement and inspiration.

This thesis consists of the following papers:

Paper I: Broman E.I. (2005), One-dependent trigonometric determinantal processes are two-block-factors, *Annals of Probability*, **33** 601-609.

Paper II: Broman E.I. and Steif J.E. (2006), Dynamical stability of percolation for some interacting particle systems and $\epsilon-$movability *Annals of Probability*, **34**.

Paper III: Broman E.I., Häggström O. and Steif J.E. (2006), Refinements of stochastic domination, *Probability Theory and Related Fields*, to appear.

Paper IV: Broman E.I. (2006), Stochastic Domination for a Hidden Markov Chain with Applications to the Contact Process in a Randomly Evolving Environment, *submitted.*

Contents

1 Background

In this section, we will give a short historical overview of the research areas and models studied in this thesis. We will also discuss some of the basic properties of these models. We prefer to give here an informal description rather than make precise mathematical statements. For exact definitions, see the relevant references or the papers of this thesis. We will start with determinantal processes and then proceed with a general description of interacting particle systems. Finally we will discuss the stochastic Ising model and the contact process.

1.1 Determinantal Processes

Since only a relatively small part of this thesis deals with determinantal processes, we will not say much about it.

Determinantal processes arise in numerous contexts such as mathematical physics, random matrix theory and representation theory to name a few. For a general survey see [24] and for the discrete case which we deal with in this thesis see [18]. See also [20] for a detailed analysis of the discrete stationary case.

In the one-dimensional discrete case, let $f : [0,1] \to [0,1]$. It is possible to define a translation invariant probability measure $\mathbb{P}^f$ on $\{0,1\}^{\mathbb{Z}}$ by letting the probability of having 1's at locations $s_1, \ldots, s_n$ be given by the $n \times n$ determinant with entry $\hat{f}(s_j - s_i)$ at position i, j. In a more general setting, one can take $f : \mathbb{T}^d \to [0,1]$ where $\mathbb{T}^d := \mathbb{R}^d/\mathbb{Z}^d$. In this case, the resulting process is indexed by $\mathbb{Z}^d$. Apart from their usefulness in applications, determinantal processes also have the interesting and unusual property of being negatively associated; see [12] for an early reference and [23] for many interesting examples, conjectures and open questions concerning negative association.

1.2 Interacting Particle Systems

The theory of interacting particle systems is a branch of probability theory. It emerged sometime around the late 1960's pioneered by R.L. Dobrushin and F. Spitzer. Within a decade, this field of research had grown remarkably and today it is one of the major branches of probability theory.

The inspiration comes from a number of different sciences such as Physics (the stochastic Ising model), Biology (the contact process) and Sociology (the voter model). Another important example of an interacting particle system is the exclusion process. This model is mostly used to model particle motion. For instance, it can be a lattice gas or traffic flow, where in the latter case the “particles” are cars. Of these four models, this thesis is concerned only with the first two. There has been a number of textbooks written on the subject, some of the most standard being [15] and [16] by T. M Liggett and [2] by R. Durrett.

We will now give an informal description of a general interacting particle system. All interacting particle systems are defined on a graph $G = (S, E)$ where a graph is a set of sites (S, finite or infinite) and a set of edges (E) connecting the sites. Furthermore, the assumption of bounded degree is standard. This means that there exists some number Δ such that the number of edges connected to any site is bounded by Δ. One then considers a random or non-random initial configuration on the sites (or sometimes on the edges) which evolves according to some probabilistic rule. Most common is the case where every site is allowed to take one of two values, for example $\{0, 1\}$ or $\{-1, 1\}$. These values are referred to as the possible states of the site. In the first case the full state space becomes $\{0, 1\}^S$ and in the second it becomes $\{-1, 1\}^S$. The immediate interaction of the system occurs between the sites which are connected by an edge. That is, for every $s, s' \in S$ such that there is an edge between s and s' (often denoted by $(s, s') \in E$ or $s \sim s'$ and referred to as s, s' being neighbors) the rate at which the current state of the site s changes can depend on the current state of the site s' as well as on the current state of s itself. However, it does not depend on the state of any $s'' \in S$ not being a neighbor of s. An important aspect is that the rates are not dependent on the history. Thus, the system is Markovian. The rates at which the changes at a site $s \in S$ occur are usually referred to as flip rates and denoted by $C(s, \sigma)$ where σ is the configuration of the system, i.e. $\sigma \in \{0, 1\}^S$ or $\sigma \in \{-1, 1\}^S$. Informally, we say that the site waits an exponentially distributed time with parameter $C(s, \sigma)$ before changing its state.

We will now proceed by giving a brief description of the two interacting particle systems that we are mainly concerned with in this thesis.

1.2.1 The Stochastic Ising Model

For simplicity, let $G = \mathbb{Z}^d$, i.e. the graph consisting of the sites $\mathbb{Z}^d$ and all nearest neighbor pairs as the edges. The stochastic Ising model was introduced by Glauber in 1963 [5] and is a dynamic version of the classical Ising model. This latter model was introduced by W. Lenz [14] and E. Ising [11] around 1920 as a model in the study of magnetism. To understand the stochastic Ising model one must understand the Ising model, and therefore we start there. The Ising model is a random configuration on $\{-1, 1\}^{\mathbb{Z}^d}$ where each site represents an atom and the state represents the spin of that atom. In the most studied ferromagnetic case, the sites “try” to align with its connected neighbors. The relevant parameters are β, the inverse temperature, and h, the external magnetic field. As β increases, the strength of the neighbor alignment forces increases. The value of h represents a preference of the state $+1$ if $h > 0$ and -1 if $h < 0$, the magnitude of the preference increasing with $|h|$.

The perhaps most interesting behavior occurs when $h = 0$. It is interesting because the system undergoes a phenomenon called phase transition. We will return to this point later. Take $S_n := \Lambda_{n+1} = \{-n-1, \dots, n+1\}^d$ and E_n to be the set of all nearest neighbor pairs of S_n. One can define two Hamiltonians of the system $H_n^{+,\beta}$ and $H_n^{-,\beta}$ by letting, for $\sigma \in \{-1,1\}^{\Lambda_n}$,

$$H_n^{+,\beta}(\sigma) = -\beta \sum_{\substack{(t,t')\in E_n \\ t,t'\in\Lambda_n}} \sigma(t)\sigma(t') - \beta \sum_{\substack{(t,t')\in E_n \\ t\in\Lambda_n \\ t'\in\Lambda_{n+1}\setminus\Lambda_n}} \sigma(t) \tag{1.1}$$

and

$$H_n^{-,\beta}(\sigma) = -\beta \sum_{\substack{(t,t')\in E_n \\ t,t'\in\Lambda_n}} \sigma(t)\sigma(t') + \beta \sum_{\substack{(t,t')\in E_n \\ t\in\Lambda_n \\ t'\in\Lambda_{n+1}\setminus\Lambda_n}} \sigma(t), \tag{1.2}$$

respectively. The physical interpretation of the Hamiltonian is that it measures the energy of the configuration σ. If the configuration has many neighboring sites with opposite alignments, then the first term of the sums of (1.1) and (1.2) becomes large. The second term of these sums gives respectively higher energy to configurations with many -1's and $+1$'s directly inside of the boundary. This is called a boundary condition. Define the probability measures $\mu_n^{+,\beta}$ and $\mu_n^{-,\beta}$ by letting

$$\mu_n^{+,\beta}(\sigma) = \frac{e^{-H_n^{+,\beta}(\sigma)}}{Z} \tag{1.3}$$

for any configuration $\sigma \in \{-1,1\}^{\Lambda_n}$ where Z is a normalization constant. We see that a configuration with a large Hamiltonian (high energy) is given less probability than configurations with a Hamiltonian closer to the minimum value (low energy). Analogously, define $\mu_n^{-,\beta}$ using (1.2) instead of (1.1) in (1.3). It is known (see [15], page 189) that the sequences $\{\mu_n^{+,\beta}\}_n$ and $\{\mu_n^{-,\beta}\}_n$ converge weakly as n tends to infinity; these limits are denoted by $\mu^{+,\beta}$ and $\mu^{-,\beta}$. It is well known ([3], [4]) that there exists $\beta_c \in [0,\infty]$ such that for $0 \le \beta < \beta_c$, we have that $\mu^{-,\beta} = \mu^{+,\beta}$ (and it can then be shown that there is a unique so called Gibbs state) and for $\beta > \beta_c$, $\mu^{-,\beta} \neq \mu^{+,\beta}$. For $\mathbb{Z}^d$ with $d \ge 2$, $\beta_c \in (0,\infty)$. This phenomenon is the phase transition mentioned earlier, and β_c is sometimes referred to as the critical inverse temperature for phase transition in the Ising model. It is quite remarkable that the effect of the boundary conditions on finite boxes survive for $\beta > \beta_c$ as $n \to \infty$. One might have thought that this effect would have disappeared as the size of the boxes grew to infinity. In one dimension this is indeed what happens, i.e. $\beta_c = \infty$. The reason behind this difference in behavior is that in one dimension, the boundary of the "box" always consists of two elements, while in higher dimensions the number of elements of the boundary goes to infinity as n goes to infinity.

For general h, it is possible to define measures $\mu^{+,\beta,h}$ and $\mu^{-,\beta,h}$ in a similar way. However, on $\mathbb{Z}^d$ for any $h \neq 0$, $\mu^{+,\beta,h} = \mu^{-,\beta,h}$ for every $\beta \geq 0$ and so in this case there is no phase transition for any $h \neq 0$. This is not true for all graphs. It is possible to define the corresponding measures $\mu^{-,\beta,h}$, $\mu^{+,\beta,h}$ on any connected graph of bounded degree. In [13] they showed that for any such graph which is also non-amenable (see [13] for a definition) there is phase transition for every $h \neq 0$.

Returning to the stochastic Ising model and $\mathbb{Z}^d$, the flip rates can be taken to be, for general h,

$$C(s,\sigma) = \exp(-\beta \sum_{\substack{t \in \mathbb{Z}^d: \\ (t,s) \in E}} \sigma(t)\sigma(s) - h\sigma(s)). \tag{1.4}$$

Using these flip rates, one can define Markov processes $\Psi^{-,\beta,h}$ and $\Psi^{+,\beta,h}$ with $\mu^{-,\beta,h}$ and $\mu^{+,\beta,h}$ as initial and stationary distributions. We mention that these rates are just one example, although perhaps the most natural one, of possible choices of flip rates such that $\mu^{-,\beta,h}$ and $\mu^{+,\beta,h}$ become stationary distributions for the corresponding Markov process. See [15] pg. 190 for a discussion about this. Furthermore, under the mild assumption of attractiveness, (see [15]) which is satisfied by the flip rates of equation (1.4), it is known that the limit is unique iff $\mu^{-,\beta,h} = \mu^{+,\beta,h}$. This is Theorem 2.16 pg. 195 [15].

1.2.2 The Contact Process

This model was introduced by T. Harris in [10]. For results up to 1985, see [15] and for results between 1985 and 1999 see [16]. It originated as a model for the spread of an infectious disease.

Consider a graph $G = (S, E)$ of bounded degree. In the contact process the state space is $\{0,1\}^S$. We will let 1 represent an infected individual, while a 0 will be used to represent a healthy individual. Let $\lambda > 0$, and define the flip rate intensities to be

$$C(s,\sigma) = \begin{cases} 1 & \text{if} \quad \sigma(s) = 1 \\ \lambda \sum_{(s',s) \in E} \sigma(s') & \text{if} \quad \sigma(s) = 0. \end{cases} \tag{1.5}$$

In words, an individual that is infected becomes healthy at rate one, while a healthy individual becomes infected at rate λ times the number of infected neighbors. Let δ_0, δ_1 denote the measures that put mass one on the configuration of all 0's and all 1's respectively. Observe that δ_0 is a stationary distribution for this model. If everyone is healthy then everyone stays healthy. If we let the initial distribution be $\sigma \equiv 1$, the distribution of this process at time t, which we will denote by $\delta_1 T_\lambda(t)$, is known to converge as t tends to infinity. This is simply because it is a so-called "attractive" process and $\sigma \equiv 1$ is the maximal state; see [15] page 265. This limiting distribution will be referred to as the upper invariant measure for the contact process with parameter λ and will be denoted by ν_λ. We then let Ψ^λ denote the stationary Markov process on $\{0,1\}^S$ with initial (and invariant) distribution ν_λ. For any $s \in S$, one can also define the process $\Psi^{\lambda,\{s\}}$. Here we start with a single infected individual at the site s, and use the same flip rates as before. We say that the process dies out if for any $s \in S$

$$\Psi^{\lambda,\{s\}}(\sigma_t \not\equiv 0 \ \forall \ t \geq 0) = 0,$$

and otherwise it survives. We also say that the process survives strongly if

$$\Psi^{\lambda,\{s\}}(\sigma_t(s) = 1 \text{ i.o.}) > 0.$$

We say that the process survives weakly if it survives but does not survive strongly. It turns out, see [16], that this definition is independent of the choice of s. It is well known that for any graph (see [16] pg. 42) there exists two critical parameter values $0 \leq \lambda_{c1} \leq \lambda_{c2} \leq \infty$ such that

- $\Psi^{\lambda,\{s\}}$ dies out if $\lambda < \lambda_{c1}$
- $\Psi^{\lambda,\{s\}}$ survives weakly if $\lambda_{c1} < \lambda < \lambda_{c2}$
- $\Psi^{\lambda,\{s\}}$ survives strongly if $\lambda > \lambda_{c2}$.

We mention that on $\mathbb{Z}^d$ it is known, see [16], that $\lambda_{c1} = \lambda_{c2}$ and also that the process dies out at this common critical value.

2 Summary of papers

In this section we will summarize the results of the four papers of this thesis as well as putting these results into context.

2.1 Paper I

The first paper, "One-dependent trigonometric determinantal processes are two-block-factors" deals with the following question concerning determinantal processes:

Let the function $f : [0,1] \to [0,1]$ mentioned in Section 1.1, be of the form

$$f(x) = \sum_{k=-m}^{m} a_k e^{-i2\pi kx}.$$

If $\{X_i\}_{i\in\mathbb{Z}}$ has distribution $\mathbb{P}^f$, it is very easy to show that the resulting probability measure will have the property that $\{X_i\}_{i<k}$ is independent of $\{X_i\}_{i\geq k+m}$ for any integer k. This is called m-dependence.

A process $\{X_i\}_{i\in\mathbb{Z}}$ is called a k-block-factor if there exists a function h of k variables and an i.i.d. process $\{Y_i\}_{i\in\mathbb{Z}}$ such that $\{h(Y_i,\ldots,Y_{i+k-1})\}_{i\in\mathbb{Z}}$ and $\{X_i\}_{i\in\mathbb{Z}}$ have the same distribution. Trivially, a k-block-factor is $(k-1)$-dependent but the converse is false in general.

In the first paper it was proved that, for $m = 1$, $\mathbb{P}^f$ is a two-block-factor and the factor map was explicitly constructed. This construction gives an alternative description of the process and one might hope that this will facilitate further analysis. It would be desirable to obtain a similar description of the processes corresponding to $m > 1$. However, at this point there has been no progress in that direction.

2.2 Paper II

Let Ψ be a stationary Markov process with state space $\mathcal{S}$. For an event $\mathcal{A} \subset \mathcal{S}$, let $\mathcal{A}_t$ be the event that $\mathcal{A}$ occurs at time t. A general, classical question in Markov process theory is to ask whether $\Psi(\mathcal{A}_t$ for some $t \geq 0) > 0$? If the answer is no, one says that $\mathcal{A}$ has *zero capacity.* Related is the concept of *exceptional times*: If $\Psi(\mathcal{A}_t) = 1$ for every $t \geq 0$, but $\Psi(\neg\mathcal{A}_t$ for some $t \geq 0) > 0$, we say that the event $\mathcal{A}$ has *exceptional times* with positive probability. Here $\neg\mathcal{A}_t$ is the event that $\mathcal{A}$ does not occur at time t. The relation between these concepts is that under the assumption $\Psi(\mathcal{A}_t) = 1$ for every $t \geq 0$, the statement "$\mathcal{A}$ does not have exceptional times" is equivalent to the statement "$\neg\mathcal{A}$ has zero capacity".

In the second paper of this thesis "Dynamical Stability of Percolation for Some Interacting Particle Systems and ϵ-movability" we studied the question of exceptional times for a particular, interesting event for the contact process and the stochastic Ising model.

To be more precise, let the state space be $\{0,1\}^{\mathbb{Z}^d}$ and ν_λ be the upper invariant measure for the contact process with parameter λ. Let $\mathcal{C}$ be the event that there exists an infinite connected component of infected (i.e. in state 1) sites, referred to as "percolation". For $d \geq 2$, it is known that there exists a critical parameter value $\lambda_p < \infty$ (see [17]) for percolation, i.e. $\nu_\lambda(\mathcal{C}) = 0$ if $\lambda < \lambda_p$ and $\nu_\lambda(\mathcal{C}) = 1$ if $\lambda > \lambda_p$. As in Section 1.2.2 let Ψ^λ be the contact process with stationary distribution ν_λ. We proved that for any $\lambda > \lambda_p$, $\Psi^\lambda(\mathcal{C}_t$ for every $t \geq 0) = 1$. Hence, there are no exceptional times at which we do not percolate. In fact, we proved this result for any graph G of bounded degree satisfying $\lambda_p(G) < \infty$.

Let S be any countable set. For $\sigma, \sigma' \in \{0,1\}^S$ we write $\sigma \preceq \sigma'$ if $\sigma(s) \leq \sigma'(s)$ for every $s \in S$. A function $f : \{0,1\}^S \to \mathbb{R}$ is increasing if $f(\sigma) \leq f(\sigma')$ whenever $\sigma \preceq \sigma'$. For two probability measures μ, ν on $\{0,1\}^S$, we say that μ is stochastically dominated by ν, and write $\mu \preceq \nu$, if for every continuous increasing function f we have that $\mu(f) \leq \nu(f)$.

Given $\mu \preceq \nu$, for $\epsilon > 0$, take a configuration according to ν, and then let each site that has the value 1, independently of everything else, change to a 0 with probability ϵ. Let $\nu^{(-,\epsilon)}$ denote the distribution of the resulting configuration. A natural question to ask is whether for $\epsilon > 0$ $\mu \preceq \nu^{(-,\epsilon)}$. If the answer is yes we say that (μ, ν) is downwards ϵ-movable. If the pair (μ, ν) is downwards ϵ-movable for some $\epsilon > 0$, we say it is downwards movable.

The key to the result that $\Psi^\lambda(\mathcal{C}_t$ for every $t \geq 0) = 1$ for every $\lambda > \lambda_p$ was to prove that for any $\lambda_1 < \lambda_2$ $(\nu_{\lambda_1}, \nu_{\lambda_2})$ is downwards movable.

For the stochastic Ising model the situation is slightly different. Here we have two parameters β and h. For fixed β it is not hard to prove that for $h_1 < h_2$, $\mu^{+,\beta,h_1} \preceq \mu^{+,\beta,h_2}$. This leads to the result that for fixed β, there exists a critical parameter value h_c such that percolation of the state $+1$ occurs (denoted $\mathcal{C}^+$) for $h > h_c$ but not for $h < h_c$. It is also relatively easy to show that the pair $(\mu^{+,\beta,h_1}, \mu^{+,\beta,h_2})$ is downwards movable. Again, this was the key that lead to results such as that for any $h > h_c$, $\Psi^{+,\beta,h}(\mathcal{C}_t^+$ occurs for every $t) = 1$.

The other case, when h is fixed and β varies is very different. It is known that on $\mathbb{Z}^d$, for any h, there are no values of β_1 and β_2 such that $\mu^{+,\beta_1,h} \preceq \mu^{+,\beta_2,h}$. Therefore we had to use other techniques than the downwards $\epsilon-$movability. We used the connection between the Ising model and the random cluster model and also an argument known as a Peierl's argument.

2.3 Paper III

In the third paper, "Refinements of Stochastic Domination", we further investigated the concepts of downwards and upwards ϵ-movability. Upwards ϵ-movability is the symmetric analogue of downwards ϵ-movability. We will start with some relevant definitions and then explain the main result of the paper.

Let S be any countable set and let μ be a probability measure on $\{0,1\}^S$. Let π_p denote the product measure with density $p \in [0,1]$ and define $p_{\max,\mu}$ to be the maximum number p such that $\pi_p \preceq \mu$. If the pair of probability measures $(\pi_{p_{\max,\mu}}, \mu)$ is downwards movable we say that it is non-rigid and otherwise we say that it is rigid. We say that μ is uniformly upwards extractable if there exists a probability measure ν and an $\epsilon > 0$ such that $\mu = \nu^{(+,\epsilon)}$. Finally we say that μ is uniformly insertion tolerant if there exists an $\epsilon > 0$ such that for any $s \in S$, $\mu(\sigma(s) = 1 | \sigma(S \setminus s)) \geq \epsilon$. Defining uniform downwards extractable in the obvious way, we say that μ is uniformly extractable if it is both uniformly upwards and downwards extractable. Similarly, having defined uniform deletion tolerant as the symmetric analogue of uniform insertion tolerant, we say that μ has uniform finite energy if it has both of these properties.

Uniform finite energy and extractability have been studied before, often in the context of information and communication theory and especially in connection to the so called "rate distortion functions". This is of practical importance for instance in the development of audio and image compression techniques. Uniform extractability has also been studied for the Ising model as well as other Markov random fields in [1, 9, 21]. Earlier, in [6, 7, 8], a similar question was studied for Markov chains and autoregressive processes. Of related interest is the result in [9] that for Markov random fields, uniform finite energy implies uniform extractability. Uniform finite energy is also studied in [22] and insertion and deletion tolerance in [19].

The goal of the paper was to provide a connection between these concepts. We consider the following properties:

(I) μ is uniformly upwards extractable.

(II) μ is uniformly insertion tolerant.

(III) μ is rigid.

(IV) There exists a $p > 0$ such that $\pi_p \preceq \mu$.

Our main result concerning these properties is the following. We have that (I) $\Rightarrow$ (II) $\Rightarrow$ (IV) and that (I) $\Rightarrow$ (III) $\Rightarrow$ (IV) while none of the four corresponding reverse implications hold. Also, (III) does not imply (II). Moreover, with $S = \mathbb{Z}$, there exist translation invariant examples for all of the asserted nonimplications. The implications are all fairly easy, while the hard work is in finding counterexamples of the non-implications. Also worth mentioning is that we showed that (II) implies (III) under the extra assumption of μ satisfying an FKG-lattice type condition. We do not know whether or not the result is true without this extra assumption.

2.4 Paper IV

The fourth paper of this thesis is called "Stochastic Domination for a Hidden Markov Chain with Applications to the Contact Process in a Randomly Evolving Environment" (the last part abbreviated CPREE).

In the usual contact process every infected site waits an exponentially distributed time with parameter 1 before becoming healthy. In this paper we instead associate to every site an independent two-state, $\{0,1\}$, background process. Given $\delta_1 < \delta_2$, if the background process is in state 0 we let a site $s \in S$, if infected, become healthy at rate δ_1, while if the background process is in state 1, it becomes healthy at rate δ_2. Furthermore, we normalize the infection rate to be 1. For the CPREE, we investigate matters of survival (both strong and weak) depending on the properties of the background process and the values of δ_1, δ_2.

Much of the analysis comes down to questions concerning stochastic domination of a Poisson-like process with two possible intensities $\alpha_1 < \alpha_2$. We construct this process by starting with a two-state, $\{0,1\}$, background process as above. We then let our Poisson-like process have intensity α_1 if the background process is in state 0 and intensity α_2 if the background process is in state 1. We then try to couple this Poisson-like process with an ordinary Poisson process on $[0,\infty)$ such that if the ordinary Poisson process has an arrival then so does our Poisson-like process. The maximum intensity λ of a Poisson process for which this coupling is possible is called $\lambda_{\max,\mu}$. Analogously we can define $\lambda_{\min,\mu}$. Starting in a discrete setting and then taking limits to get to continuous time, we find the exact values of $\lambda_{\max,\mu}$ and $\lambda_{\min,\mu}$. When we put more and more weight on state 1 in the background process, $\lambda_{\max,\mu}$ approaches α_2 if the rate at which the background process jumps between its two possible states is high enough. In contrast, it turns out that $\lambda_{\min,\mu}$ is always equal to α_2 except for the degenerate case where the background process is in state 0 with probability 1. In this paper we prove these results as well as some other asymptotics. We also obtain results of the same kind for finite time, although in this case we can only get upper and lower bounds rather than an exact closed form expression.

Returning to the CPREE, the distribution at which the recoveries occur is exactly the Poisson-like process just described. By using the stochastic domination results mentioned above, we can compare the CPREE to an ordinary contact process. However, because of the difference in the results of $\lambda_{\max,\mu}$ and $\lambda_{\min,\mu}$, this comparison is straight forward in only one direction. That is, we can make sure that the ordinary contact process is "larger" than the CPREE under suitable, non-trivial conditions. The other direction, that is, trying to couple the CPREE so that it is "smaller" than the ordinary contact process under non-trivial conditions is much harder. However, with some adjustments it can be done. This fact somewhat complicates the analysis of the CPREE as well as (in our opinion) makes it more interesting.

References

[1] Bassalygo, L.A. and Dobrushin, R.L. (1987) Epsilon-entropy of a Gibbs field, (Russian) *Problemy Peredachi Informatsii* **23**, 3–15.

[2] Durrett R. (1988) *Lecture notes on Particle Systems and Percolation*, Wadsworth.

[3] H.-O. Georgii (1988) *Gibbs Measures and Phase Transitions* de Gruyter Studies in Mathematics.

[4] H.-O. Georgii, O. Häggström and C. Maes (2001) The random geometry of equilibrium phases *Phase Transitions and Critical Phenomena*, **18** 1–142.

[5] Glauber R. J. (1963) Time-dependent statistics of the Ising model, *J. Math. Phys.* **4,** 294–307.

[6] Gray, R.M. (1970) Informations rates of autoregressive processes, *IEEE Trans. Information Theory* **IT-16**, 412–421.

[7] Gray, R.M. (1971) Rate distortion functions for finite-state finite-alphabet Markov sources, *IEEE Trans. Information Theory* **IT-17**, 127–134.

[8] Gray, R.M. (1973) Correction to "Information rates of stationary-ergodic finite-alphabet sources", *IEEE Trans. Information Theory* **IT-19**, 573.

[9] Hajek, B. and Berger, T. (1987) A decomposition theorem for binary Markov random fields, *Ann. Probab.* **15**, 1112–1125.

[10] T. Harris (1974) Contact interactions on a lattice. *Ann. Probab.* **2** 969–988.

[11] Ising E. (1925) Beitrag zur Theorie der Ferromagnetismus, *Z. Phys.* **31,** 253–258.

[12] Joag-Dev K. and Proschan, F. (1983) Negative association of random variables with applications *Ann.Statist.* **11**, 286–295.

[13] Jonasson, J. and Steif, J. E. (1999) Amenability and Phase Transition in the Ising model *J. Theor. Probab.*, **12**, 549–559.

[14] Lenz, W. (1920) Beitrag zum verständnis der magnetischen Erscheinungen in festen Körepn, *Phys Zeitschr.* **21**, 613–615.

[15] T. M. Liggett (1985) *Interacting Particle Systems* Springer-Verlag.

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

[17] Liggett, T.M. and Steif, J.E. (2006) Stochastic domination: the contact process, Ising models and FKG measures, *Annales Institut Henri Poincare, Probabilites et Statistiques*, **42**, 223–243.

[18] Lyons R. (2003), Determinantal probability measures, *Publ. Math. Inst. Hautes tudes Sci.* **98**, 167–212.

[19] Lyons, R. and Schramm, O. (1999) Indistinguishability of percolation clusters, *Ann. Probab.* **27**, 1809–1836.

[20] Lyons, R. and Steif, J.E. (2003) Stationary determinantal processes: phase multiplicity, Bernoullicity, entropy, and domination, *Duke Math. Journal*, **120**, 515–575.

[21] Newman, C.M. (1987) Decomposition of binary random fields and zeros of partition functions, *Ann. Probab.* **15**, 1126–1130.

[22] Newman, C.M. and Schulman, L.S. (1981) Infinite clusters in percolation models, *J. Stat. Phys.* **26**, 613–628.

[23] Pemantle, R. (2000) Towards a theory of negative dependence. *J. Math. Phys.* **41**, 1371–1390.

[24] Soshnikov, A. (2000) Determinantal raondom point fields. *Uspekhi Mat. Nauk* **55**, 107–160.

Paper I

One-dependent trigonometric determinantal processes are two-block-factors

Erik I. Broman*

September 29, 2003

Abstract

Given a trigonometric polynomial $f : [0,1] \to [0,1]$ of degree m, one can define a corresponding stationary process $\{X_i\}_{i\in\mathbb{Z}}$ via determinants of the Toeplitz matrix for f. We show that for $m = 1$ this process, which is trivially one-dependent, is a two-block-factor.

AMS subject classification: 60G10

Keywords and phrases: Determinantal processes, k-dependence,k-block-factors

Short title: Determinants and two-block-factors

1 Introduction

We will start by defining a family of probability measures $\mathbf{P}^f$ on the Borel sets of $\{0,1\}^{\mathbb{Z}}$ where $f : [0,1] \to [0,1]$ is a Lebesgue-measurable function (see [9]). For such an f, define the probability of the cylinder sets by

$$\begin{aligned}\mathbf{P}^f[\eta(e_1) = \cdots = \eta(e_k) = 1] &:= \mathbf{P}^f[\{\eta \in \{0,1\}^{\mathbb{Z}} : \eta(e_1) = \cdots = \eta(e_k) = 1\}] \\ &:= \det[\hat{f}(e_j - e_i)]_{1\leq,i,j\leq k},\end{aligned}$$

where $e_1, \ldots, e_k$ are distinct elements in $\mathbb{Z}$ and $k \geq 1$. Here $\hat{f}$ denotes the Fourier coefficients of f, defined by

$$\hat{f}(k) := \int_0^1 f(x)e^{-i2\pi kx}dx.$$

In [9] it is proven that $\mathbf{P}^f$ is indeed a probability measure. In fact they showed this for the more general case of $f : \mathbb{T}^d \to [0,1]$ where $\mathbb{T}^d := \mathbb{R}^d/\mathbb{Z}^d$;

*Research partially supported by the Swedish Natural Science Research Council

in this case the resulting process is indexed by $\mathbb{Z}^d$. This result rests very strongly on the results in [8]. Except for the two definitions below, $\{X_i\}_{i\in\mathbb{Z}}$ will always denote a process distributed according to some measure $\mathbf{P}^f$. Throughout this paper, equality in distribution will be denoted by $=_{\mathcal{D}}$. Let the function $f : [0,1] \to [0,1]$ be of the form

$$f(x) = \sum_{k=-m}^{m} a_k e^{-i2\pi kx}.$$

It is then easily checked that the process $\{X_i\}_{i\in\mathbb{Z}}$ corresponding to the probability measure $\mathbf{P}^f$ is m-dependent according to the definition below.

Definition 1.1 *A process $\{X_i\}_{i\in\mathbb{Z}}$ is called m-dependent if $\{X_i\}_{i<k}$ is independent of $\{X_i\}_{i\geq k+m}$ for all integers k.*

We will also need the definition of an m-block-factor.

Definition 1.2 *The process $\{X_i\}_{i\in\mathbb{Z}}$ is called an m-block-factor if there exists a function h of m variables and an i.i.d. process $\{Y_i\}_{i\in\mathbb{Z}}$ such that $\{X_i\}_{i\in\mathbb{Z}} =_{\mathcal{D}} \{h(Y_i,\dots,Y_{i+m-1})\}_{i\in\mathbb{Z}}$.*

We will as usual not distinguish between the process $\{X_i\}_{i\in\mathbb{Z}}$ and the corresponding probability measure $\mathbf{P}^f$.

Observe that an (m+1)-block-factor is trivially m-dependent. For some time, it was an open question whether all m-dependent processes were in fact (m+1)-block-factors (see [4],[5],[6],[7]). However, in [2] the authors constructed a family of one-dependent processes which are not two-block-factors, and in [3] the authors constructed a one-dependent process which is not a k-block factor for any k. In [1] the authors construct a one-dependent stationary Markov process with five states which is not a two-block-factor, they also prove that this result is sharp in the sense that every one-dependent stationary Markov process with not more that four states is a two-block-factor. In view of the above it is a natural question to ask whether a certain m-dependent process is an (m+1)-block-factor or not.

$\mathbf{P}^f$ as defined above is an m-dependent "trigonometric determinantal probability measure". These probability measures are special cases of general determinantal probability measures, see [10] or [8] for definitions and results. Determinantal processes arise in numerous contexts e.g. mathematical physics, random matrix theory and representation theory to name a few. For a survey see [10], for further results see [8] and for results concerning the discrete stationary case, see [9]. In [9], they ask whether $\mathbf{P}^f$ above is an (m+1)-block-factor. In that paper they say that if one can find sufficiently explicit block factors for all trigonometric polynomials, then one can find explicit factors of i.i.d. processes giving $\mathbf{P}^f$, where f is any function such that $f : \mathbb{T} \to [0,1]$. This in turn would enable one to use more standard probabilistic techniques when studying such a $\mathbf{P}^f$. We answer their question positively for m=1 in Theorem (1.3), constructing an explicit two-block-factor.

Theorem 1.3 *If* $f : [0,1] \to [0,1]$ *is given by*

$$f(x) = b + ae^{-i2\pi x} + ce^{i2\pi x},$$

then the corresponding process $\{X_i\}_{i\in\mathbb{Z}}$ *is a two-block-factor.*

2 Proof of theorem 1.3

Proof of theorem 1.3.

With f as in the statement of the theorem, it follows that $\bar{a} = c$, $b \geq 0$ and hence if $a = a_1 + ia_2$

$$f(x) = b + 2a_1 \cos(2\pi x) + 2a_2 \sin(2\pi x) = b + 2|a| \cos(2\pi x - \phi), \quad (1)$$

for some suitable choice of ϕ. Let, as usual, $\mathbf{P}^f$ be the corresponding probability measure, and write

$$D_k := \det\, [\hat{f}(j-i)]_{1\leq,i,j\leq k+1}$$

where $k \geq 0$.

Note that the process $\{X_i\}_{i\in\mathbb{Z}}$ distributed according to $\mathbf{P}^f$ is obviously stationary. Since $\mathbf{P}^f$ is one-dependent, it is easily seen that it is uniquely determined among the one-dependent processes by the values of

$$\mathbf{P}^f[\eta(i) = \cdots = \eta(i+k) = 1] = \mathbf{P}^f[\eta(1) = \cdots = \eta(1+k) = 1]$$

as k varies over the nonnegative integers.

We have that for $k \geq 2$

$$D_k = \det\, [\hat{f}(j-i)]_{1\leq,i,j\leq k+1} = \begin{vmatrix} b & a & 0 & 0 & 0 & \cdots \\ \bar{a} & b & a & 0 & 0 & \cdots \\ 0 & \bar{a} & b & a & 0 & \cdots \\ 0 & 0 & \bar{a} & b & a & \cdots \\ \vdots & \vdots & \vdots & \ddots & \ddots & \ddots \end{vmatrix} \quad (2)$$

$$= b \begin{vmatrix} b & a & 0 & 0 & 0 & \cdots \\ \bar{a} & b & a & 0 & 0 & \cdots \\ 0 & \bar{a} & b & a & 0 & \cdots \\ 0 & 0 & \bar{a} & b & a & \ldots \\ \vdots & \vdots & \vdots & \ddots & \ddots & \ddots \end{vmatrix} - a \begin{vmatrix} \bar{a} & a & 0 & 0 & 0 & \cdots \\ 0 & b & a & 0 & 0 & \cdots \\ 0 & \bar{a} & b & a & 0 & \cdots \\ 0 & 0 & \bar{a} & b & a & \ldots \\ \vdots & \vdots & \vdots & \ddots & \ddots & \ddots \end{vmatrix}$$

$$= bD_{k-1} - |a|^2 D_{k-2},$$

where the determinant on the left-hand side of the third equality has size $(k+1) \times (k+1)$, and the two on the right-hand side have size $k \times k$. Furthermore

$$D_0 = |b| = b \tag{3}$$

$$D_1 = \begin{vmatrix} b & a \\ \bar{a} & b \end{vmatrix} = b^2 - |a|^2. \tag{4}$$

The characteristic equation corresponding to equation (2) is

$$r^2 - br + |a|^2 = 0, \tag{5}$$

which has two roots

$$r_1 = \frac{b}{2} + \sqrt{\frac{b^2}{4} - |a|^2}, \tag{6}$$

and

$$r_2 = \frac{b}{2} - \sqrt{\frac{b^2}{4} - |a|^2}. \tag{7}$$

Case 1: Assume that $r_1 = r_2 = r$ so that $r = \frac{b}{2}$ and

$$\frac{b^2}{4} = |a|^2$$

and so (since $b, |a| \geq 0$)

$$b = 2|a|.$$

We have by equation (1) that

$$\max_{x \in [0,1]} f(x) = \max_{x \in [0,1]} (b + 2|a| \cos(2\pi x - \phi)) = b + 2|a| = 2b$$

and since $f : [0,1] \to [0,1]$ we get $b \leq 1/2$ and so $|a| \leq 1/4$.

With $r_1 = r_2 = r$, it follows from the basic theory of difference equations that the solution to equation (2) is

$$D_k = (C_1 k + C_2) r^k \ \forall k \geq 0,$$

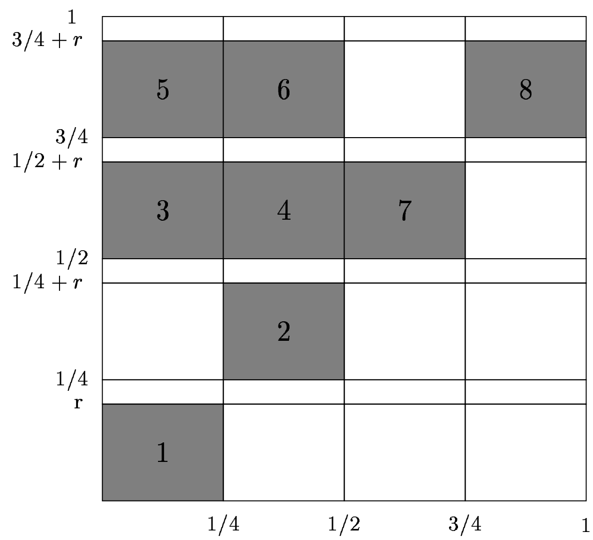

Figure 1: This figure shows A (the shaded area).

for some constants C_1, C_2 yet to be determined. Using (3) and (4), we get that $C_2 = D_0 = b = 2r$ and using this we get $(C_1 + 2r)r = D_1 = b^2 - |a|^2 = b^2 - b^2/4 = 3r^2$. Hence $C_1 = r$ and so

$$D_k = (kr + 2r)r^k \ \forall k \geq 0. \tag{8}$$

We will now construct a two-block-factor which we will show to be distributed according to $\mathbf{P}^f$. Let $\{Y_i\}_{i\in\mathbb{Z}}$ be i.i.d. uniform on $[0,1]$. Define $h : [0,1] \times [0,1] \to [0,1]$ by $h = I_A$ where

$$\begin{aligned} A &= [0, \frac{1}{4}] \times [0, r] \cup [0, \frac{1}{4}] \times [\frac{1}{2}, \frac{1}{2} + r] \cup [0, \frac{1}{4}] \times [\frac{3}{4}, \frac{3}{4} + r] \\ &\cup [\frac{1}{4}, \frac{1}{2}] \times [\frac{1}{4}, \frac{1}{4} + r] \cup [\frac{1}{4}, \frac{1}{2}] \times [\frac{1}{2}, \frac{1}{2} + r] \cup [\frac{1}{4}, \frac{1}{2}] \times [\frac{3}{4}, \frac{3}{4} + r] \\ &\cup [\frac{1}{2}, \frac{3}{4}] \times [\frac{1}{2}, \frac{1}{2} + r] \cup [\frac{3}{4}, 1] \times [\frac{3}{4}, \frac{3}{4} + r]. \end{aligned}$$

A is depicted as the grey area of figure (1). Observe that $r = |a| \leq 1/4$.

We will show that

$$\mathbf{P}[h(Y_i, Y_{i+1}) = \cdots = h(Y_{i+k}, Y_{i+k+1}) = 1] = D_k \ \forall k \geq 0.$$

Since $\{h(Y_i, Y_{i+1})\}_{i\in\mathbb{Z}}$ is one-dependent, this gives us $\{h(Y_i, Y_{i+1})\}_{i\in\mathbb{Z}} =_{\mathcal{D}} \mathbf{P}^f$ as desired. We first observe that the size of the shaded area of figure (1) is $8\frac{1}{4}r = 2r = b$, so that $\mathbf{P}[h(Y_i, Y_{i+1}) = 1] = D_0$.

If $h(Y_i, Y_{i+1}) = \cdots = h(Y_{i+k}, Y_{i+k+1}) = 1$, then (Y_{i+l}, Y_{i+l+1}) must be in one of the boxes marked 1 through 8 of figure (1) $\forall l \in \{0, \ldots, k\}$. If (Y_i, Y_{i+1}) is in the box marked 1, then $Y_{i+1} \in [0, r]$ and so (Y_{i+1}, Y_{i+2}) must be in one of the boxes marked 1, 3 or 5 because otherwise $(Y_{i+1}, Y_{i+2}) \notin A$. Similar "rules" apply if (Y_i, Y_{i+1}) is in one of the other seven boxes. We see that for any ω such that $h(Y_i(\omega), Y_{i+1}(\omega)) = \cdots = h(Y_{i+k}(\omega), Y_{i+k+1}(\omega)) = 1$ there is a natural sequence $j_0 j_1 \cdots j_k(\omega) \in \{1, \ldots, 8\}^{k+1}$ associated to it, where the value of j_l indicates that $(Y_{i+l}(\omega), Y_{i+l+1}(\omega))$ is in the box marked with that value. In any such sequence the number 1 can only be followed by either 1, 3 or 5, as described above, while the number 2 can only be followed by either 2, 4 or 6. Additionally any one of the numbers 3, 4 or 7 must be followed by a 7, while any one of 5, 6 or 8 must be followed by an 8.

We claim that the number of sequences $j_0 j_1 \cdots j_k$ described above is $(4k + 8)$. To see this, observe that every such sequence with $j_k \notin \{1, 2\}$ can be extended into a sequence $j_0 j_1 \cdots j_{k+1}$ in only one way, while if $j_k \in \{1, 2\}$ it can be extended in three ways. Observe also that there are only two sequences $j_0 j_1 \cdots j_k$ ending in 1 or 2.

The set of ω giving a specific sequence $j_0 j_1 \cdots j_k \in \{1, \ldots, 8\}^{k+1}$ has probability $(1/4)r^{k+1}$ since Y_i must be in an interval of length 1/4, while $Y_{i+1}, \ldots, Y_{i+k+1}$ all must be within intervals of length r. Hence the total probability of having $h(Y_i, Y_{i+1}) = \cdots = h(Y_{i+k}, Y_{i+k+1}) = 1$ is $(4k + 8)(1/4)r^{k+1} = (kr + 2r)r^k$. Comparing with equation (8) we see that

$$\mathbf{P}[h(Y_i, Y_{i+1}) = \cdots = h(Y_{i+k}, Y_{i+k+1}) = 1] = D_k$$

$\forall k \geq 0$ and we conclude that $\{h(Y_i, Y_{i+1})\}_{i \in \mathbb{Z}} =_{\mathcal{D}} \mathbf{P}^f$ and so this case is proved.

Case 2: It remains to consider $r_1 \neq r_2$. According to equations (6) and (7)

$$r_1 + r_2 = b,$$

and

$$r_1 r_2 = |a|^2.$$

In this case the solution to equation (2) is, again, from basic difference equation theory,

$$D_k = C_1 r_1^k + C_2 r_2^k \ \forall k \geq 0,$$

for some constants C_1, C_2 yet to be determined. Using this with equation (3) we get

$$C_1 + C_2 = D_0 = r_1 + r_2,$$

and using equation (4) we get

$$C_1 r_1 + C_2 r_2 = D_1 = b^2 - |a|^2 = (r_1 + r_2)^2 - r_1 r_2 = r_1^2 + r_1 r_2 + r_2^2.$$

A straightforward calculation yields

$$C_1 = \frac{r_1^2}{r_1 - r_2}$$

and

$$C_2 = -\frac{r_2^2}{r_1 - r_2}$$

and therefore for $k \geq 1$,

$$D_k = \frac{r_1^{k+2} - r_2^{k+2}}{r_1 - r_2} = \frac{r_1^{k+2} - r_1^{k+1} r_2 + r_2(r_1^{k+1} - r_2^{k+1})}{r_1 - r_2} = r_1^{k+1} + r_2 D_{k-1}. \tag{9}$$

Assume that $b \leq \frac{1}{2}$ so that $2(r_1+r_2) \leq 1$. We will now construct a two-block-factor which we will show to be distributed according to $\mathbf{P}^f$. Let $\{Y_i\}_{i\in\mathbb{Z}}$ be i.i.d. uniform on $[0,1]$ and again take $h : [0,1] \times [0,1] \to [0,1]$ to be the function $h = I_A$ where A is now

$$\begin{aligned} A &= [0, Cr_1] \times [0, r_1] \cup [0, Cr_1] \times [2Cr_1, 2Cr_1 + r_2] \\ &\cup [0, Cr_1] \times [2Cr_1 + Cr_2, 2Cr_1 + Cr_2 + r_2] \\ &\cup [Cr_1, 2Cr_1] \times [Cr_1, Cr_1 + r_1] \\ &\cup [Cr_1, 2Cr_1] \times [2Cr_1, 2Cr_1 + r_2] \\ &\cup [Cr_1, 2Cr_1] \times [2Cr_1 + Cr_2, 2Cr_1 + Cr_2 + r_2] \\ &\cup [2Cr_1, 2Cr_1 + Cr_2] \times [2Cr_1, 2Cr_1 + r_2] \\ &\cup [2Cr_1 + Cr_2, 1] \times [2Cr_1 + Cr_2, 2Cr_1 + Cr_2 + r_2], \end{aligned}$$

and $C = \frac{1}{2(r_1+r_2)} \geq 1$. A is the shaded area of figure (2).

Again we will show that

$$\mathbf{P}[h(Y_i, Y_{i+1}) = \cdots = h(Y_{i+k}, Y_{i+k+1}) = 1] = D_k \ \forall k \geq 0.$$

Since again $\{h(Y_i, Y_{i+1})\}_{i\in\mathbb{Z}}$ is one-dependent this gives us $\{h(Y_i, Y_{i+1})\}_{i\in\mathbb{Z}} =_{\mathcal{D}} \mathbf{P}^f$. We observe that the size of the shaded area of figure (2) equals

$$2Cr_1r_1 + 4Cr_1r_2 + 2Cr_2r_2 = 2C(r_1 + r_2)^2 = r_1 + r_2$$

by our choice of C, and so $\mathbf{P}[h(Y_i, Y_{i+1}) = 1] = D_0$.

For any ω such that $h(Y_i(\omega), Y_{i+1}(\omega)) = \cdots = h(Y_{i+k}(\omega), Y_{i+k+1}(\omega)) = 1$ there is a natural sequence $j_0 j_1 \cdots j_k(\omega) \in \{1, \ldots, 8\}^{k+1}$ associated to it as before. Let $\{\omega : j_0 j_1 \cdots j_k(\omega)\}$ denote the set of ω giving a specific sequence $j_0 j_1 \cdots j_k$, and for convenience we will write $\mathbf{P}[j_0 j_1 \cdots j_k]$ instead of $\mathbf{P}[\{\omega : j_0 j_1 \cdots j_k(\omega)\}]$. Assume that $j_{k-1} \in \{3, 4, 5, 6, 7, 8\}$, we get

$$\mathbf{P}[j_0 j_1 \cdots j_k] = r_2 \mathbf{P}[j_0 j_1 \cdots j_{k-1}]$$

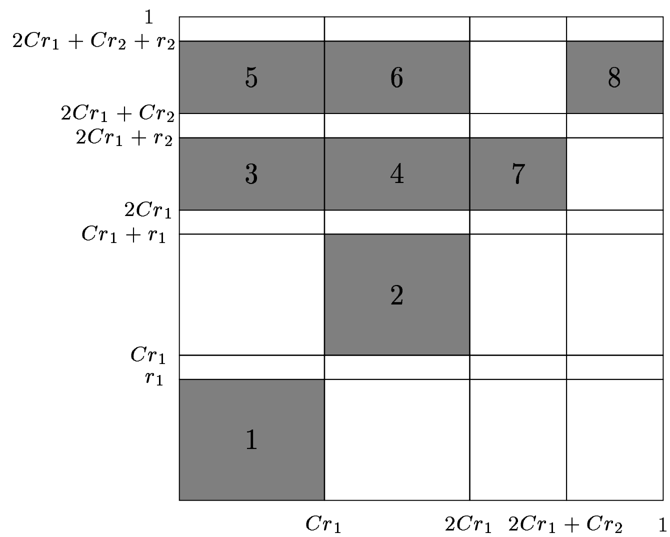

Figure 2: This figure shows A (the shaded area).

since j_k is either 7 or 8 (depending on the value of j_{k-1}). If instead $j_{k-1} = 1$ then j_k must be either $1, 3$ or 5 and of course $j_l = 1$ for all $l \leq (k-1)$. Hence in this case

$$\mathbf{P}[j_0 j_1 \cdots j_k] = r_2 \mathbf{P}[j_0 j_1 \cdots j_{k-1}] = r_2 \mathbf{P}[\underbrace{11 \cdots 1}_{k}] = r_2 C r_1^{k+1}$$

if j_k is equal to 3 or 5 and

$$\mathbf{P}[j_0 j_1 \cdots j_k] = \mathbf{P}[\underbrace{11 \cdots 1}_{k+1}] = C r_1^{k+2}$$

if $j_k = 1$. Similarly if $j_{k-1} = 2$ then j_k must be either $2, 4$ or 5 and of course $j_l = 2$ for all $l \leq (k-1)$. Hence

$$\mathbf{P}[j_0 j_1 \cdots j_k] = r_2 \mathbf{P}[j_0 j_1 \cdots j_{k-1}] = r_2 \mathbf{P}[\underbrace{22 \cdots 2}_{k}] = r_2 C r_1^{k+1}$$

if j_k is equal to 4 or 6 and

$$\mathbf{P}[j_0 j_1 \cdots j_k] = \mathbf{P}[\underbrace{22 \cdots 2}_{k+1}] = C r_1^{k+2}$$

if $j_k = 2$.

Let $\mathcal{A}_k$ be the set of all sequences $j_0 j_1 \cdots j_k$ corresponding to the event $h(Y_i, Y_{i+1}) = \cdots = h(Y_{i+k}, Y_{i+k+1}) = 1$. We have that

$$\begin{aligned}
&\mathbf{P}[h(Y_i, Y_{i+1}) = \cdots = h(Y_{i+k}, Y_{i+k+1}) = 1] \\
&\quad = \sum_{\mathcal{A}_k} \mathbf{P}[j_0 j_1 \cdots j_k] \\
&\quad = \sum_{\substack{\mathcal{A}_k \\ j_{k-1} \notin \{1,2\}}} \mathbf{P}[j_0 j_1 \cdots j_k] + \sum_{\substack{\mathcal{A}_k \\ j_{k-1} \in \{1,2\}}} \mathbf{P}[j_0 j_1 \cdots j_k] \\
&\quad = r_2 \sum_{\substack{\mathcal{A}_{k-1} \\ j_{k-1} \notin \{1,2\}}} \mathbf{P}[j_0 j_1 \cdots j_{k-1}] + 4 r_2 C r_1^{k+1} + 2 C r_1^{k+2} \\
&\quad = r_2 \left(\sum_{\substack{\mathcal{A}_{k-1} \\ j_{k-1} \notin \{1,2\}}} \mathbf{P}[j_0 j_1 \cdots j_{k-1}] + \mathbf{P}[\underbrace{11 \cdots 1}_{k}] + \mathbf{P}[\underbrace{22 \cdots 2}_{k}] \right) \\
&\qquad + 2 r_2 C r_1^{k+1} + 2 C r_1^{k+2} \\
&\quad = r_2 \sum_{\mathcal{A}_{k-1}} \mathbf{P}[j_0 j_1 \cdots j_{k-1}] + 2 C r_1^{k+1} (r_1 + r_2) \\
&\quad = r_2 \mathbf{P}[h(Y_i, Y_{i+1}) = \cdots = h(Y_{i+k-1}, Y_{i+k}) = 1] + r_1^{k+1}.
\end{aligned}$$

Comparing this to equation (9), and using $\mathbf{P}[h(Y_i, Y_{i+1}) = 1] = D_0$ we see that

$$\mathbf{P}[h(Y_i, Y_{i+1}) = \cdots = h(Y_{i+k}, Y_{i+k+1}) = 1] = D_k$$

for all $k \geq 0$, and so this case is also proved.

Finally the case $b > \frac{1}{2}$ remains. Take

$$g(x) = 1 - f(x) = 1 - b - 2|a| \cos(2\pi x - \phi) = 1 - b + 2|a| \cos(2\pi x - \phi'),$$

for some suitable choice of ϕ'. Since $1 - b \leq \frac{1}{2}$, it follows from above that we can construct a two-block-factor $\{h(Y_i, Y_{i+1})\}_{i \in \mathbb{Z}}$ such that

$$\{h(Y_i, Y_{i+1})\}_{i \in \mathbb{Z}} =_{\mathcal{D}} \mathbf{P}^g.$$

With $\tilde{h} = 1 - h$, we get a new two-block-factor $\{\tilde{h}(Y_i, Y_{i+1})\}_{i \in \mathbb{Z}}$ with ones and zeros flipped. Lemma 2.4 in [9] then shows that $\{\tilde{h}(Y_i, Y_{i+1})\}_{i \in \mathbb{Z}}$ has distribution $\mathbf{P}^{1-g}$, which in turn is $\mathbf{P}^f$.

QED

When trying to generalise theorem 1.3 to the case where f is a trigonometric polynomial of degree m, one must consider not only the values of

$$\mathbf{P}^f[\eta(1) = \cdots = \eta(1+k) = 1],$$

but also the values of

$$\mathbf{P}^f[\eta(e_1) = 1 \cdots = \eta(e_k) = 1]$$

where $e_i \in \mathbb{Z}\, \forall i \in \{1, \ldots, k\}$ but where e_i is not necessarily equal to $e_{i-1}+1$. Analysing these new cylinder events adds to the complexity of the problem and therefore, in our opinion, the generalisation of theorem 1.3 (if indeed the generalisation is true) does not seem to be trivial.

Acknowledgements

I would like to thank my supervisor Jeffrey Steif for presenting me the problem and for all the help I received during the writing of this paper.

References

[1] Aaronson J., Gilat D., and Keane M. (1992) On the structure of 1-dependent Markov chains *J. Theoret. Probab.* **5** no.3 545-561

[2] Aaronson J., Gilat D., Keane M., and de Valk V. (1989) An algebraic construction of a class of one-dependent processes. *Ann. Probab.* **17** 128-143

[3] Burton R.M. Goulet M. and Meester R. (1993) On 1-dependent processes and k-block factors. *Ann. Probab.* **21** 2157-2168

[4] Götze F. and Hipp C. (1989) Asymptotic expansions for potential functions of i.i.d. random fields *Probab Theory Related Fields* **82** 349-370

[5] Ibragimov I.A. and Linnik Y.V. (1971) *Independent and Stationary Sequences of Random Variables* Wolters-Noordhoff Publishing, Groningen

[6] Janson S. (1983) Renewal theory for m-dependent variables *Ann. Probab.* **11** 558-568

[7] Janson S. (1984) Runs in m-dependent sequences *Ann. Probab.* **12** 805-818

[8] Lyons R. Determinantal probability measures *Publ. Math. Inst. Hautes Études Sci.,* to appear

[9] Lyons R. and Steif J. Stationary Determinantal Processes: Phase Transitions, Bernoullicity, Entropy, and Domination *Duke Math. Journal,* to appear

[10] Soshnikov A. (2000) Determinantal random point fields *Uspekhi Mat. Nauk* **55**, 107-160

Erik Broman
Dept. of Mathematics
Chalmers University of Technology
412 96 Gothenburg
Sweden
broman@math.chalmers.se
http://www.math.chalmers.se/~broman

Paper II

Dynamical Stability of Percolation for Some Interacting Particle Systems and ϵ-Movability

Erik I. Broman* and Jeffrey E. Steif†
Chalmers University of Technology

January 31, 2005

Abstract

In this paper we will investigate dynamic stability of percolation for the stochastic Ising model and the contact process. We also introduce the notion of downwards and upwards ϵ-movability which will be a key tool for our analysis.

AMS subject classification: 82C43, 82B43, 60K35

Keywords and phrases: percolation, stochastic Ising models, contact process,

Short title: Dynamical Stability for IPS

1 Introduction

Consider bond percolation on an infinite connected locally finite graph G, where for some $p \in [0,1]$, each edge (bond) of G is, independently of all others, open with probability p and closed with probability $1-p$. Write π_p for this product measure. The main questions in percolation theory (see [10]) deal with the possible existence of infinite connected components (clusters) in the random subgraph of G consisting of all sites and all open edges. Write $\mathcal{C}$ for the event that there exists such an infinite cluster. By Kolmogorov's 0-1 law, the probability of $\mathcal{C}$ is, for fixed G and p, either 0 or 1. Since $\pi_p(\mathcal{C})$ is nondecreasing in p, there exists a critical probability $p_c = p_c(G) \in [0,1]$ such that

$$\pi_p(\mathcal{C}) = \begin{cases} 0 & \text{for } p < p_c \\ 1 & \text{for } p > p_c. \end{cases}$$

At $p = p_c$ we can have either $\pi_p(\mathcal{C}) = 0$ or $\pi_p(\mathcal{C}) = 1$, depending on G.

*Research supported by the Swedish Natural Science Research Council

†Research partially supported by the Swedish Natural Science Research Council, NSF grant DMS-010384 and the Göran Gustafsson Foundation (KVA).

In [15], the authors initiated the study of dynamical percolation. In this model, with p fixed, the edges of G switch back and forth according to independent 2 state Markov chains where 0 switches to 1 at rate p and 1 switches to 0 at rate $1-p$. In this way, if we start with distribution π_p, the distribution of the system is at all times π_p. The general question studied in [15] was whether there could exist atypical times at which the percolation structure looks different than at a fixed time.

We record here some of the results from [15]; (i) for any graph G and for any $p < p_c(G)$, there are no times at which percolation occurs, (ii) for any graph G and for any $p > p_c(G)$, there are no times at which percolation does not occur, (iii) there exist graphs which do not percolate for $p = p_c(G)$, but nonetheless, for $p = p_c(G)$, there are exceptional times at which percolation occurs, (iv) there exist graphs which percolate for $p = p_c(G)$, but nonetheless, for $p = p_c(G)$, there are exceptional times at which percolation does not occur, and (v) for $\mathbb{Z}^d$ with $d \geq 19$ with $p = p_c(\mathbb{Z}^d)$, there are no times at which percolation occurs. In addition, it has recently be shown in [23] that for site percolation on the triangular lattice, for $p = p_c = 1/2$, there are exceptional times at which percolation occurs. Given this, a similar result would be expected for $\mathbb{Z}^2$.

The point of the present paper is to initiate a study of dynamical percolation for *interacting* systems where the edges or sites flip at rates which depend on the neighbors. We point out that in a different direction such questions in continuous space, but without interactions, related to continuum percolation have been studied in [2].

Ising model results. Precise definitions of the following Ising model measures and the stochastic Ising model will be given in Section 2. Fix an infinite graph $G = (S, E)$. Let $\mu^{+,\beta,h}$ be the plus state for the Ising model with inverse temperature β and external field h on G (this is a probability measure on $\{-1,1\}^S$). Let $\Psi^{+,\beta,h}$ denote the corresponding stochastic Ising model; (this is a stationary continuous time Markov chain on $\{-1,1\}^S$ with marginal distribution $\mu^{+,\beta,h}$). Let $\mathcal{C}^+$ ($\mathcal{C}^-$) denote the event that there exists an infinite cluster of sites with spin 1 (-1) and let $\mathcal{C}_t^+$ ($\mathcal{C}_t^-$) denote the event that there exists an infinite cluster of sites with spin 1 (-1) at time t. It is known that the family $\mu^{+,\beta,h}$, is, for fixed β, stochastically increasing (to be defined later) in h.

Theorem 1.1 *Consider a graph $G = (S, E)$ of bounded degree. Fix $\beta \geq 0$ and let $h_c = h_c(\beta)$ be defined by*

$$h_c := \inf\{h : \mu^{+,\beta,h}(\mathcal{C}^+) = 1\}.$$

Then for all $h > h_c$,

$$\Psi^{+,\beta,h}(\mathcal{C}_t^+ \text{ occurs for every } t) = 1$$

and for all $h < h_c$

$$\Psi^{+,\beta,h}(\exists t \geq 0 : \mathcal{C}_t^+ \text{ occurs }) = 0.$$

If we modify h_c *to be instead*

$$h_c' := \sup\{h : \mu^{+,\beta,h}(\mathcal{C}^-) = 1\},$$

the same two claims hold with $\mathcal{C}_t^+$ *replaced by* $\mathcal{C}_t^-$ *and with* $h < h_c'$ *and* $h > h_c'$ *reversed.*

This result tells us what happens in the subcritical and supercritical cases (with respect to h with β held fixed). It is the analogue of the easier Proposition 1.1 in [15] where it is proved that if $p < p_c$ ($p > p_c$), then, with probability 1, there is percolation at no time (at all times).

The following easy lemma gives us information about when h_c is non-trivial.

Lemma 1.2 *Assume the graph* G *has bounded degree and let* β *be arbitrary. Then* $h_c > -\infty$*. If* $p_c(site) < 1$*, then* $h_c < \infty$*. Similar results hold if* h_c *is replaced by* h_c'*.*

The following theorems, where we restrict to $\mathbb{Z}^d$, will only discuss the case $h = 0$. However, this will in many cases give us information about the "critical" case $(\beta, h_c(\beta))$ since in a number of situations, $h_c(\beta) = 0$. For example, this is true on all $\mathbb{Z}^d$ with $d \geq 2$ and β sufficiently large. We also mention that while the relationship between h_c and h_c' in Theorem 1.1 might in general be complicated, for $\mathbb{Z}^d$, one easily has that $h_c = -h_c'$; this follows from the known fact that the plus and minus states are the same when $h \neq 0$. When $h = 0$, we will abbreviate $\mu^{+,\beta,0}$ by $\mu^{+,\beta}$ and $\Psi^{+,\beta,0}$ by $\Psi^{+,\beta}$. We point out that while $\mu^{+,\beta,h}$ is stochastically increasing in h for fixed β, there is no such monotonicity in β for fixed h, not even for $h = 0$. Therefore we must use a different approach in the latter case.

We first study percolation of -1's and then percolation of 1's. Let

$$\beta_p(2) := \inf\{\beta : \sum_{l=1}^{\infty} l 3^{l-1} e^{-2\beta l} < \infty\} = \frac{\log 3}{2}.$$

We will refer to $\beta_p(2)$ as the critical inverse temperature of the Peierls regime for $\mathbb{Z}^2$. The choice of $\beta_p(2)$ might at first look quite arbitrary, but it is exactly what is needed to carry out a contour argument (known as Peierls argument) for $\mathbb{Z}^2$. For $d \geq 3$, there is a $\beta_p(d)$, such that for β larger than $\beta_p(d)$, a similar (although topologically more complicated) argument works for $\mathbb{Z}^d$. As a result of this "contour argument", it is well known and easy to show that for $\beta > \beta_p(d)$, we have that

$$\mu^{+,\beta}(\mathcal{C}^-) = 0. \tag{1}$$

Our next result is a dynamical version of (1) and we emphasize that this corresponds to the critical case as it is easy to check that for these β's, $h_c(\beta) = 0$.

Theorem 1.3 *For $\mathbb{Z}^d$ with $d \geq 2$ and $\beta > \beta_p(d)$*

$$\Psi^{+,\beta}(\exists t \geq 0 : \mathcal{C}_t^- \; occurs) = 0.$$

It is well known that $\beta_p(d) \geq \beta_c(d)$, the latter being the critical inverse temperature for the Ising model on $\mathbb{Z}^d$. For $d = 2$, Theorem 1.3 can be extended down to the critical inverse temperature $\beta_c(2)$. First, it is known (see [5]) that on $\mathbb{Z}^2$, for *all* β

$$\mu^{+,\beta}(\mathcal{C}^-) = 0. \tag{2}$$

Our dynamical analogue for $\beta > \beta_c$ is the following where we again point out that this is also a critical case as it is easy to check that for these β's, we also have $h_c(\beta) = 0$.

Theorem 1.4 *For the stochastic Ising model $\Psi^{+,\beta}$ on $\mathbb{Z}^2$ with parameter $\beta > \beta_c$,*

$$\Psi^{+,\beta}(\exists t \geq 0 : \mathcal{C}_t^- \; occurs) = 0.$$

Interestingly, (1) is not always true for $\beta > \beta_c(d)$ although, as stated, it is true for $\mathbb{Z}^2$ or β sufficiently large. In [1], it is shown that for $\mathbb{Z}^d$ with large d, there exists $\beta^+ > \beta_c(d)$ such that the probability in (1) is in fact 1 for all $\beta < \beta^+$. Moreover, they show that for these β, there exists $h > 0$ with

$$\mu^{+,\beta,h}(\mathcal{C}^-) = 1.$$

For such β's, this means that $h_c' > 0$ and hence it immediately follows from Theorem 1.1 that

$$\Psi^{+,\beta}(\mathcal{C}_t^- \text{ occurs for every } t) = 1.$$

Note that for these values of β, the case $h = 0$ is a *non*-critical case.

We next look at percolation of 1's under $\mu^{+,\beta}$. In the above results, we have not discussed the case of percolation of -1's when $\beta \leq \beta_c$. However, by symmetry, this is the same as studying percolation of 1's in this case and so we can now move over to the study of $\mathcal{C}^+$.

First, it is well known that for any graph of bounded degree, $\mu^{+,\beta,h} \neq \mu^{-,\beta,h} \Rightarrow \mu^{+,\beta,h}(\mathcal{C}^+) = 1$. (This is proved in [3] for $\mathbb{Z}^d$; this argument extends to any graph of bounded degree.) In particular, for any graph G of bounded degree and for $\beta > \beta_c(G)$,

$$\mu^{+,\beta}(\mathcal{C}^+) = 1. \tag{3}$$

Our next result is a dynamical version of (3) for $\mathbb{Z}^d$. We mention that this result sometimes corresponds to a critical case and sometimes not. For $\beta > \beta_p(d)$ in $\mathbb{Z}^d$ or $\beta > \beta_c(2)$ in $\mathbb{Z}^2$, we have seen that $h_c = 0$ and so, in these cases, this next result covers the critical case. However, as pointed out, for d large and β just a little higher than β_c, the result in [1] gives us that $h_c < 0$ and hence in this case, this next theorem already follows from Theorem 1.1.

Theorem 1.5 *For the stochastic Ising model $\Psi^{+,\beta}$ on $\mathbb{Z}^d$ with parameter $\beta > \beta_c(d)$,*

$$\Psi^{+,\beta}(\mathcal{C}_t^+ \text{ occurs for every } t) = 1.$$

(The proof we give actually works for any graph of bounded degree). We mention that while $\beta > \beta_c$ is a sufficient condition for (3) to hold, it is certainly not necessary. For example, on $\mathbb{Z}^3$ we have that $\mu^{+,0}(\mathcal{C}^+) = 1$ since $\mu^{+,0} = \pi_{1/2}$ and the critical value for site percolation on $\mathbb{Z}^3$ is less than 1/2. The reason β_c appears is the connection between the Ising model and the random cluster model; β_c corresponds to the critical value for percolation in the corresponding random cluster model (see [13]).

We are now left with the case $\beta \leq \beta_c$. We will not be able to say too much since it is not known in all cases whether one has percolation at a fixed time. We first however have the following easy result for $d \geq 3$. We do not prove this result since it follows easily from the fact that the critical value for site percolation on $\mathbb{Z}^d$ is less than 1/2 for $d \geq 3$ as this gives easily that $h_c(\beta) < 0$ for β sufficiently small and hence Theorem 1.1 is applicable.

Note that the case $\beta = 0$ follows from the result in [15] mentioned above.

Proposition 1.6 *For $d \geq 3$, there exists $\beta_1(d) > 0$ such that for all $\beta < \beta_1(d)$, we have that*

$$\Psi^{+,\beta}(\mathcal{C}_t^+ \text{ occurs for every } t) = 1.$$

Finally, due to work of Higuchi, we can determine what happens with $\beta < \beta_c$ for $\mathbb{Z}^2$. It is shown in [16] that for $\mathbb{Z}^2$, for all $\beta < \beta_c$, we have that $h_c(\beta) > 0$. The following result follows from this fact and Theorem 1.1.

Theorem 1.7 *For $d = 2$, for all $\beta < \beta_c$, we have that*

$$\Psi^{+,\beta}(\exists t \geq 0 : \mathcal{C}_t^+ \text{ occurs }) = 0.$$

We note that even though it is known that for $\mathbb{Z}^2$, $\mu^{+,\beta_c}(\mathcal{C}^+) = 0$, we cannot conclude that

$$\Psi^{+,\beta_c}(\exists t \geq 0 : \mathcal{C}_t^+ \text{ occurs }) = 0$$

since it is known (see [17]) that $h_c(\beta_c) = 0$. In contrast, based on the results in [23], it is interesting to ask

Question 1.8 *For the graph $\mathbb{Z}^2$, is it the case that*

$$\Psi^{+,\beta_c}(\exists t \geq 0 : \mathcal{C}_t^+ \text{ occurs }) = 1?$$

We finally mention that interestingly it is also known (see again [17]) that for $\beta < \beta_c$, $\mu^{+,\beta,h_c(\beta)}(\mathcal{C}^+) = 0$.

Contact process results. Precise definitions of the following items will be given in Section 2. Fix an infinite graph $G = (S, E)$. Consider the contact process on a graph $G = (S, E)$ with parameter λ. Denote by μ_λ the stochastically largest invariant measure, the so-called "upper invariant measure" (this is a probability measure on $\{0,1\}^S$). Let Ψ^λ denote the corresponding stationary contact process (this is a stationary continuous time Markov chain on $\{0,1\}^S$ with marginal distribution μ_λ). If $0 < \lambda_1 < \lambda_2$, it is well known that μ_{λ_1} is stochastically smaller than μ_{λ_2}, denoted by

$$\mu_{\lambda_1} \preceq \mu_{\lambda_2}$$

(see Section 2 for this precise definition).

Theorem 1.9 *Consider the contact process Ψ^λ on a graph $G = (S, E)$, with initial and stationary distribution μ_λ. Let λ_p be defined by*

$$\lambda_p := \inf\{\lambda : \mu_\lambda(\mathcal{C}^+) = 1\}.$$

We have that for all $\lambda > \lambda_p$,

$$\Psi^\lambda(\mathcal{C}_t^+ \text{ occurs for every } t) = 1.$$

In order for this theorem to be nonvacuous, we need to know that $\lambda_p < \infty$ for at least some graph. First, the fact that there exists λ such that $\mu_\lambda(\mathcal{C}^+) > 0$ for $\mathbb{T}^d$ with $d \geq 2$ follows from [12]. Here $\mathbb{T}^d$ is the unique infinite connected graph without circuits and in which each site has exactly $d+1$ neighbours; $\mathbb{T}^d$ is commonly known as the homogenous tree of order d. Combined with a 0-1 law which we develop, Proposition 4.2, we obtain that $\lambda_p < \infty$ in this case. For $\mathbb{Z}^d$ with $d \geq 2$ (as well as for $\mathbb{T}^d$), it is proved in [22] that for large λ, μ_λ stochastically dominates high density product measures which immediately implies that $\lambda_p < \infty$ in these cases.

When we prove Theorem 1.1, we will in fact, prove a more general theorem which holds for a large class of systems. However, this proof will only work for models satisfying the so-called FKG lattice condition (which we call "monotone" in this paper.) We now point out the important fact that for $\lambda < 2$, in 1 dimension, the upper invariant measure for the contact process, while having positive correlations, is *not* monotone (see [20]). These terms are defined in Section 2. (One would also believe it is never monotone whenever the measure is not δ_0.) Hence Theorem 1.9 does not follow from the generalization of Theorem 1.1 which will come later.

ϵ-movability. We now introduce the concepts of upwards and downwards ϵ-movability. While we mainly introduce these as a technical tool to be used in our main results, it turns out that they are of interest in their own right. In [4], the concept of upwards movability is studied for its own sake and related to other well studied concepts such as uniform insertion tolerance.

Let S be a countable set. Take any probability measure μ on $\{-1,1\}^S$ and let X be a $\{-1,1\}^S$ valued random variable with distribution μ. Let Z be a $\{-1,1\}^S$ valued random variable with distribution $\pi_{1-\epsilon}$ and be independent of X. Define $X^{(-,\epsilon)}$ by letting $X^{(-,\epsilon)}(s) = \min(X(s), Z(s))$ for every $s \in S$, and let $\mu^{(-,\epsilon)}$ denote the distribution of $X^{(-,\epsilon)}$. In a similar way, define $X^{(+,\epsilon)}$ by letting $X^{(+,\epsilon)}(s) = \max(X(s), Z(s))$ for every $s \in S$, where Z has distribution π_ϵ and is independent of X. Denote the distribution of $X^{(+,\epsilon)}$ by $\mu^{(+,\epsilon)}$.

Definition 1.10 *Let (μ_1, μ_2) be a pair of probability measures on $\{-1,1\}^S$, where S is a countable set. Assume that*

$$\mu_1 \preceq \mu_2.$$

If

$$\mu_1 \preceq \mu_2^{(-,\epsilon)},$$

then we say that this pair of probability measures is downwards ϵ-movable. If the pair is downwards ϵ-movable for some $\epsilon > 0$, we say that the pair is downwards movable. Analogously, if

$$\mu_1^{(+,\epsilon)} \preceq \mu_2,$$

then we say that the pair (μ_1, μ_2) is upwards ϵ-movable and that it is upwards movable if the pair is upwards ϵ-movable for some $\epsilon > 0$.

For probability measures on $\{0,1\}^S$, we have identical definitions.

The relevance of downward (or upward) ϵ-movability to our dynamical percolation analysis will be explained in Section 5. In Section 3, we will prove ϵ-movability for general monotone systems which will eventually lead to a proof of Theorem 1.1 (and its generalization). We now state a similar and key result for the contact process.

Theorem 1.11 *Let G be a graph of bounded degree, $0 < \lambda_1 < \lambda_2$ and μ_{λ_1}, μ_{λ_2} be the upper invariant measures for the contact process on $\{0,1\}^S$ with parameters λ_1 and λ_2 respectively. Then $(\mu_{\lambda_1}, \mu_{\lambda_2})$ is downwards movable.*

We finally mention how the above questions that we study fall into the context of classical Markov process theory. Let $(\Omega, \mathcal{F}, \mathbb{P})$ be the probability space where a stationary Markov process $\{X_t\}_{t\geq 0}$ taking values in some state space $\mathcal{S}$ is defined. Letting μ denote the distribution of X_t (for any t), consider an event $\mathcal{A} \subseteq \mathcal{S}$ with $\mu(\mathcal{A}) = 1$. Let $\mathcal{A}_t$ be the event that $\mathcal{A}$ occurs at time t. We say that $\mathcal{A}$ is a *dynamicaly stable* event if $\mathbb{P}(\mathcal{A}_t\ \forall t \geq 0) = 1$. In Markov process terminology, this is equivalent to saying that $\mathcal{A}^c$ has *capacity zero*. All the questions in this paper deal with showing, for various models and parameters, that the event that there exists/there does not exist an infinite connected component of sites which are all open is dynamically stable.

The rest of this paper is divided into 9 sections. In Section 2, we will give all necessary preliminaries and precise definitions of our models. Sections 3 and 4 will deal with the concept of ϵ-movability. In Section 3, we develop what will be needed to prove Theorem 1.1 and its generalization. In Section 4, we will prove Theorem 1.11 (which is the key to Theorem 1.9) as well as give a proof that $\lambda_p < \infty$ for trees. In Section 5, we prove 2 elementary lemmas which relate the notion of ϵ-movability to dynamical questions. In the remaining sections, proofs of the remaining results are given. We note that the proof of Theorem 1.4 will use the proof of Theorem 1.5 and hence will come afterwards.

We end with one bit of notation. If μ is a probability measure on some set U, we write $X \sim \mu$ to mean that X is a random variable taking values in U with distribution μ.

2 Models and definitions

Before presenting the interacting particle systems discussed in this paper we will present some definitions and results related to stochastic domination. Let S be any countable set. For $\sigma, \sigma' \in \{-1,1\}^S$ we write $\sigma \preceq \sigma'$ if $\sigma(s) \leq \sigma'(s)$ for every $s \in S$. An increasing function f is a function $f : \{-1,1\}^S \to \mathbb{R}$ such that $f(\sigma) \leq f(\sigma')$ for all $\sigma \preceq \sigma'$. For two probability measures μ, μ' on $\{-1,1\}^S$ we write $\mu \preceq \mu'$ if for every continuous increasing function f we have that $\mu(f) \leq \mu'(f)$. ($\mu(f)$ is shorthand for $\int f(x)d\mu(x)$.) When $\{-1,1\}^S$ is replaced by $\{0,1\}^S$, we have identical definitions. Strassens Theorem (see [19], page 72) states that if $\mu \preceq \mu'$, then there exist random variables X, X' with distribution μ, μ' respectively such that $X \preceq X'$ a.s.

A very useful result is the so called Holley's inequality, which appeared first in [18]. We will present a variant of the theorem by Holley; it is not the most general but is sufficient for our purposes.

Theorem 2.1 *Take S to be a finite set. Let μ, μ' be probability measures on $\{-1,1\}^S$ which assign positive probability to all configurations $\sigma \in \{-1,1\}^S$.*

Assume that

$$\mu(\sigma(s) = 1|\sigma(S \setminus s) = \xi) \leq \mu'(\sigma(s) = 1|\sigma(S \setminus s) = \xi')$$

for every $s \in S$ *and* $\xi \preceq \xi'$ *where* $\xi, \xi' \in \{-1,1\}^{S\setminus s}$. *Then* $\mu \preceq \mu'$.

Proof. See [9] or [13] for a proof.

QED

Two properties of probability measures which are often encountered within the field of interacting particle systems are the monotonicity property and the property of positive correlations presented below.

Definition 2.2 *Take* S *to be a finite set. A probability measure* μ *on* $\{-1,1\}^S$ *which assigns positive probability to every* $\sigma \in \{-1,1\}^S$ *is called monotone if for every* $s \in S$ *and* $\xi \preceq \xi'$ *where* $\xi, \xi' \in \{-1,1\}^{S\setminus s}$,

$$\mu(\sigma(s) = 1|\sigma(S \setminus s) = \xi) \leq \mu(\sigma(s) = 1|\sigma(S \setminus s) = \xi').$$

We point out immediately, that it is known that this is equivalent to the so-called FKG lattice condition.

Definition 2.3 *A probability measure* μ *on* $\{-1,1\}^S$ *is said to have positive correlations if for all bounded increasing functions* $f, g : \{-1,1\}^S \to \mathbb{R}$, *we have*

$$\mu(fg) \geq \mu(f)\mu(g).$$

The following important result is sometimes known as the FKG inequality (see [7]).

Theorem 2.4 *Take* S *to be a finite set. Let* μ *be a monotone probability measure on* $\{-1,1\}^S$ *which assigns positive probability to every configuration. Then* μ *has positive correlations.*

Proof. This was originally proved in [7], see also [9] for a proof.

QED

In this section and also later in this paper we will talk about convergence of probability measures. Convergence will always mean weak convergence, where $\{0,1\}^S$ is given the product topology.

2.1 The Ising model

Take $G = (S, E)$, where $|S| < \infty$. The Ising measure $\mu^{\beta,h}$ on $\{-1,1\}^S$ at inverse temperature $\beta \geq 0$, external field h and with free boundary conditions is defined as follows. For any configuration $\sigma \in \{-1,1\}^S$, let

$$H^{\beta,h}(\sigma) = -\beta \sum_{\substack{\{t,t'\}\in E \\ t,t'\in S}} \sigma(t)\sigma(t') - h\sum_{t\in S}\sigma(t). \tag{4}$$

$H^{\beta,h}$ is called the Hamiltonian. Define $\mu^{\beta,h}$ by assigning the probability

$$\mu^{\beta,h}(\sigma) = \frac{e^{-H^{\beta,h}(\sigma)}}{Z} \tag{5}$$

to any configuration $\sigma \in \{-1,1\}^S$ where Z is a normalization constant. Of course Z depends on the graph and the values β and h, but this will not be important for us and therefore not reflected in the notation.

Take $S_n := \Lambda_{n+1} = \{-n-1,\dots,n+1\}^d$ and E_n to be the set of all nearest neighbor pairs of S_n. Given a configuration ξ on $\{-1,1\}^{\mathbb{Z}^d\setminus\Lambda_n}$, let, for $\sigma \in \{-1,1\}^{\Lambda_n}$,

$$H_n^{\xi,\beta,h}(\sigma) = -\beta \sum_{\substack{\{t,t'\}\in E_n \\ t,t'\in\Lambda_n}} \sigma(t)\sigma(t') - h\sum_{t\in\Lambda_n}\sigma(t) - \beta \sum_{\substack{\{t,t'\}\in E_n \\ t\in\Lambda_n \\ t'\in\Lambda_{n+1}\setminus\Lambda_n}} \sigma(t)\xi(t') \tag{6}$$

be our Hamiltonian. Here ξ is called a boundary condition. Again we define a probability measure using (5) but using the Hamiltonian of (6) instead. This Ising measure will be denoted by $\mu_n^{\xi,\beta,h}$. The cases $\xi \equiv 1$ and $\xi \equiv -1$ are especially important and the corresponding Ising measures are denoted by $\mu_n^{+,\beta,h}$ and $\mu_n^{-,\beta,h}$ respectively. We view $\mu_n^{+,\beta,h}$ ($\mu_n^{-,\beta,h}$) as a probability measure on $\{-1,1\}^{\mathbb{Z}^d}$ by letting, with probability 1, the configuration be identically 1 (-1) outside Λ_n. It is known (see [19], page 189) that the sequences $\{\mu_n^{+,\beta,h}\}$ and $\{\mu_n^{-,\beta,h}\}$ converge as n tends to infinity; these limits are denoted by $\mu^{+,\beta,h}$ and $\mu^{-,\beta,h}$.

The same kind of construction can be carried out on any infinite connected locally finite graph $G = (S, E)$. One defines a Hamiltonian analogous to the one in (6) but with Λ_n replaced by any $\Lambda \subseteq S$ where $|\Lambda| < \infty$. With $\xi \equiv 1$ or $\xi \equiv -1$, one then considers the corresponding limits of Ising measures as $\Lambda \uparrow S$, the limit turning out to be independent of the particular choice of sequence. See for instance [9] for how this is carried out in detail. Fix $h = 0$ and abbreviate $\mu^{+,\beta,0}$ and $\mu^{-,\beta,0}$ by $\mu^{+,\beta}$ and $\mu^{-,\beta}$. It is well known ([8], [9]) that for any graph, there exists $\beta_c \in [0,\infty]$ such that for $0 \leq \beta < \beta_c$, we have that $\mu^{-,\beta} = \mu^{+,\beta}$ (and there is then a unique so called Gibbs state) and for $\beta > \beta_c$, $\mu^{-,\beta} \neq \mu^{+,\beta}$. For $\mathbb{Z}^d$ with $d \geq 2$, and many other graphs, $\beta_c \in (0,\infty)$. β_c is sometimes referred to as the critical inverse temperature for phase transition in the Ising model. Furthermore in [14], the author shows that if G is of bounded degree, the condition $\beta_c < \infty$ is equivalent to the condition $p_c < 1$, where p_c is the critical parameter value for site percolation on G. It is easy to see that for any graph of bounded degree $p_c > 0$ (see the proof of Theorem 1.10 of [10]). This in turn implies via the connection between the random cluster model and the Ising model, described below, that $\beta_c > 0$ for any graph of bounded degree.

2.2 Spin Systems.

A configuration $\sigma \in \{-1,1\}^S$ can be seen as particles on a discrete set S having one of two different "spins" represented by -1 and 1. To this we will add a stochastic dynamics, and assume that the system is described by "flip rate intensities" which we will denote by $\{C(s,\sigma)\}_{s\in S,\ \sigma\in\{-1,1\}^S}$. $C(s,\sigma)$ represents the rate at which site s changes its state when the present configuration is σ. Of course $C(s,\sigma) \geq 0\ \forall s \in S, \sigma \in \{-1,1\}^S$, and we assume that the interaction is nearest neighbour in the sense that the flip rate of a site $s \in S$ only depends on the configuration σ at s and at sites t with $\{s,t\} \in E$. We will limit ourselves to only allow one site flip in every transition and we will only consider flip rate intensities such that

$$\sup_{s,\sigma} C(s,\sigma) < \infty.$$

In many cases we will consider translation invariant systems and then this last condition will hold trivially. Furthermore we will always assume the trivial condition that for every $s \in S$

$$\sup_{\sigma:\sigma(s)=0} C(s,\sigma(s)) > 0, \quad \sup_{\sigma:\sigma(s)=1} C(s,\sigma(s)) > 0.$$

We will call such an object a spin system (see [19] or [6] for results concerning general spin systems). Given such rates, one can obtain a Markov process Ψ on $\{-1,1\}^S$ governed by these flip rates; see [19]. Such a Markov process with a specified initial distribution μ on $\{-1,1\}^S$ will be denoted by Ψ^μ. Given a Markov process, μ will be called an invariant distribution for the process if the projections of Ψ^μ onto $\{-1,1\}^S$ at any fixed time $t \geq 0$ is μ. In this case, Ψ^μ will be a stationary Markov process on $\{-1,1\}^S$ all of whose marginal distributions are μ. Of course the state space $\{-1,1\}^S$ can be exchanged for either $\{0,1\}^S$ or $\{0,1\}^E$.

Sometimes we will work with two different sets of flip rates $\{C_1(s,\sigma)\}_{s\in S,\ \sigma\in\{-1,1\}^S}$ and $\{C_2(s,\sigma)\}_{s\in S,\ \sigma\in\{-1,1\}^S}$, governing two Markov processes Ψ_1 and Ψ_2 respectively. We will write $C_1 \preceq C_2$ if the following conditions are satisfied;

$$C_2(s,\sigma_2) \geq C_1(s,\sigma_1)\ \forall s \in S,\ \forall \sigma_1 \preceq \sigma_2 \text{ s.t. } \sigma_1(s) = \sigma_2(s) = 0, \qquad (7)$$

and

$$C_1(s,\sigma_1) \geq C_2(s,\sigma_2)\ \forall s \in S,\ \forall \sigma_1 \preceq \sigma_2 \text{ s.t. } \sigma_1(s) = \sigma_2(s) = 1. \qquad (8)$$

The point of $C_1 \preceq C_2$ is that a coupling of Ψ_1 and Ψ_2 will then exist for which $\{(\eta,\delta) : \eta(s) \leq \delta(s) \forall s \in S\}$ is invariant for the process; see [19].

2.3 Stochastic Ising models

We will now briefly discuss stochastic Ising models. We will omit most details; for an extensive discussion and analysis see again [19]. Consider $G_n = (S_n, E_n)$ defined in the subsection 2.1. Given β and h, it is possible to construct flip rates C_n^+ on $\{-1,1\}^{S_n}$ for which $\mu_n^{+,\beta,h}$ is reversible and invariant. We denote by $\Psi_n^{+,\beta,h}$ the corresponding stationary Markov process with initial distribution $\mu_n^{+,\beta,h}$. One possible choice of flip rate intensities are that for every $s \in \Lambda_n$ and $\sigma \in \{-1,1\}^S$,

$$C_n^+(s,\sigma) = \exp[-\beta(\sum_{\substack{t\in\Lambda_n: \\ \{t,s\}\in E_n}} \sigma(t)\sigma(s) + \sum_{\substack{t\in\Lambda_{n+1}\setminus\Lambda_n: \\ \{t,s\}\in E_n}} \sigma(s)) - h\sigma(s)].$$

Sites in $\Lambda_{n+1} \setminus \Lambda_n$ are kept fixed at 1. Observe that if $s \in \Lambda_{n-1}$, the second sum is over an empty set. A straightforward calculation gives

$$C_n^+(s,\sigma)\mu_n^{+,\beta,h}(\sigma) = C_n^+(s,\sigma_s)\mu_n^{+,\beta,h}(\sigma_s), \qquad (9)$$

where

$$\sigma_s(t) = \begin{cases} \sigma(t) & \text{if } t \neq s \\ -\sigma(t) & \text{if } t = s. \end{cases}$$

This shows that indeed $\mu_n^{+,\beta,h}$ is reversible and invariant for C_n^+. Any family of spin rates satisfying (9) is called a stochastic Ising model (on our finite set). One can show that there exists a limiting distribution $\Psi^{+,\beta,h}$ of $\Psi_n^{+,\beta,h}$ when n tends to infinity; see [19], Theorem 2.2, page 17 and Theorem 2.7, page 139. Furthermore $\Psi^{+,\beta,h}$ is a stationary Markov process on $\{-1,1\}^{\mathbb{Z}^d}$ with marginal distribution $\mu^{+,\beta,h}$ governed by flip rate intensities

$$C(s,\sigma) = \exp(-\beta \sum_{\substack{t\in\mathbb{Z}^d: \\ \{t,s\}\in E}} \sigma(t)\sigma(s) - h\sigma(s)); \qquad (10)$$

see [19] Theorem 2.7 page 139. It is also possible to construct $\Psi^{+,\beta,h}$ directly on $\{-1,1\}^{\mathbb{Z}^d}$ without going through the limiting procedure. Furthermore there are several possible choices of flip rate intensities that can be used to construct a stationary and reversible Markov process on $\{-1,1\}^{\mathbb{Z}^d}$ with marginal distribution $\mu^{+,\beta,h}$. In [19], a stochastic Ising model is defined to be any spin system with flip rate intensities $\{C(s,\sigma)\}_{s\in\mathbb{Z}^d,\sigma\in\{-1,1\}^{\mathbb{Z}^d}}$ satisfying that for each $s \in \mathbb{Z}^d$

$$C(s,\sigma)\exp(\beta \sum_{\substack{\{t,s\}\in E\\ t\in\mathbb{Z}^d}} \sigma(t)\sigma(s) + h\sigma(s)) \tag{11}$$

is independent of $\sigma(s)$. Therefore, when we refer to a stochastic Ising model $\Psi^{+,\beta,h}$ with marginal distribution $\mu^{+,\beta,h}$, we will have this definition in mind. It is particularly easy to see that (11) (or the condition of detailed balance as it is often referred to) is satisfied for the flip rate intensities of (10) but there are many other rates satisfying this. It is known that the set of so called Gibbs states are exactly the same as the class of reversible measures with respect to the flip rates satisfying (11); see [19] page 190-196. Note also that the condition specified in (11) with $\mathbb{Z}^d$ replaced by Λ_n is equivalent to that of (9) (modified with the boundary condition removed).

While we defined above stochastic Ising models on $\{-1,1\}^{\mathbb{Z}^d}$, this construction can be done on more general graphs (see [19]).

2.4 The random cluster model

Unlike all other models in this paper, the random cluster model deals with configurations on the edges E of a graph $G=(S,E)$. We will review the definition of the regular random cluster measure on general finite graphs and the "wired" random cluster measure on $\Lambda_n \subseteq \mathbb{Z}^d$. We will also recall the limiting measures and in the next subsection the connection between the random cluster model and the Ising model. In doing so we will follow the outlines of [9] and [13] closely.

Take a finite graph $G=(S,E)$. Define the random cluster measure $\nu_G^{p,q}$ on $\{0,1\}^E$ with parameters $p\in[0,1]$ and $q>0$ as the probability measure which assigns to the configuration $\eta\in\{0,1\}^E$ the probability

$$\nu_G^{p,q}(\eta) = \frac{q^{k(\eta)}}{Z}\prod_{e\in E} p^{\eta(e)}(1-p)^{1-\eta(e)}. \tag{12}$$

Here Z is again a normalization constant and $k(\eta)$ is the number of connected components of η. From now on we will always take $q=2$ and therefore we will suppress q in the notation.

Take $G_n=(S_n,E_n)$, where $S_n=\Lambda_{n+1}\subseteq\mathbb{Z}^d$ and E_n is the set of all nearest neighbour pairs of Λ_{n+1}. Write ν_n^p for $\nu_{G_n}^p$, and define

$$\tilde{\nu}_n^p(\cdot) = \nu_n^p(\cdot|\text{ all edges of } E_n \text{ with both end sites in } \Lambda_{n+1}\setminus\Lambda_n \text{ are present}). \tag{13}$$

This is the so called "wired" random cluster measure. It is called "wired" since all edges of the boundary are present. It is immediate from the defining equations (12) and (13) that for $e \in E_n$ and any $\xi \in \{0,1\}^{E_n \setminus e}$

$$\tilde{\nu}_n^p(\eta(e) = 1 | \eta(E_n \setminus e) = \xi) = \begin{cases} p, & \text{if the endpoints of } e \text{ are} \\ & \text{connected in } \xi, \\ \frac{p}{2-p} & \text{otherwise.} \end{cases} \tag{14}$$

One can show (see [9] or [13]) that when n tends to infinity, the probability measures $\{\tilde{\nu}_n^p\}_{n \in \mathbb{N}^+}$ converge to a probability measure $\tilde{\nu}^p$. Furthermore, the construction of $\tilde{\nu}_n^p$ on $\{0,1\}^{E_n}$ can be done on any finite subgraph by connecting all sites of the boundary of the graph with each other. As a consequence, we can also define random cluster measures on more general graphs than $\mathbb{Z}^d$, see for example [11].

2.5 The random cluster model and the Ising model

Take $G_n = (S_n, E_n)$ as in Section 2.4. As in [13], let $\mathbf{P}_n^p$ be the probability measure on $\{-1,1,\}^{S_n} \times \{0,1\}^{E_n}$ defined in the following way.

1. Assign each site of $\Lambda_{n+1} \setminus \Lambda_n$ and every edge with both endpoints in $\Lambda_{n+1} \setminus \Lambda_n$ the value 1.

2. Assign each site of Λ_n the value 1 or -1 with equal probability, assign each edge with not more than one endpoint in $\Lambda_{n+1} \setminus \Lambda_n$ the value 0 or 1 with probabilities $1-p$ and p respectively. Do this independently for all sites and edges.

3. Condition on the event that no two sites with different spins have an open edge connecting them.

One can then check that $\mathbf{P}_n^p(\sigma, \{0,1\}^{E_n}) = \mu_n^{+,\beta}(\sigma)$ with $\beta = -\log(1-p)/2$, and that $\mathbf{P}_n^p(\{-1,1\}^{S_n}, \eta) = \tilde{\nu}_n^p(\eta)$. Here, $\mathbf{P}_n^p(\sigma, \{0,1\}^{E_n})$ is just the marginal in the first coordinate of $\mathbf{P}_n^p$. The same kind of construction can be carried out on any finite graph $G = (S, E)$.

2.6 The contact process

Consider a graph $G = (S, E)$ of bounded degree. In the contact process the state space is $\{0,1\}^S$. Let $\lambda > 0$, and define the flip rate intensities to be

$$C(s, \sigma) = \begin{cases} 1 & \text{if} \quad \sigma(s) = 1 \\ \lambda \displaystyle\sum_{(s',s) \in E} \sigma(s') & \text{if} \quad \sigma(s) = 0. \end{cases}$$

If we let the initial distribution be $\sigma \equiv 1$, the distribution of this process at time t which we will denote by $\delta_1 T_\lambda(t)$ is known to converge as t tends to infinity. This is simply because it is a so called "attractive" process and $\sigma \equiv 1$ is the maximal state and $\{\delta_1 T_\lambda(t)\}$ is stochastically decreasing; see

[19] page 265. This limiting distribution will be referred to as the upper invariant measure for the contact process with parameter λ and will be denoted by μ_λ. We then let Ψ^λ denote the stationary Markov process on $\{0,1\}^S$ with initial (and invariant) distribution μ_λ.

3 ϵ-movability for monotone measures

In this section, we prove movability results for classes of monotone measures. The finite case is covered by Lemma 3.3, while the countable case is discussed in Proposition 3.4. In this section, we will always assume that our measures have full support.

For any $|S| < \infty$, $s \in S$, $\xi \in \{0,1\}^{S\setminus s}$ and probability measure μ on $\{0,1\}^S$ write $\mu^{(*,\epsilon)}(i|\xi)$ for $\mu^{(*,\epsilon)}(\sigma(s) = i|\sigma(S \setminus s) = \xi)$, $\mu^{(*,\epsilon)}(i \cap \xi)$ for $\mu^{(*,\epsilon)}(\{\sigma(s) = i\} \cap \{\sigma(S \setminus s) = \xi\})$ and $\mu^{(*,\epsilon)}(\xi)$ for $\mu^{(*,\epsilon)}(\sigma(S \setminus s) = \xi)$. Here, $*$ can represent either $+$ or $-$ and $i \in \{0,1\}$. Note that s is suppressed in the notation and so should be understood from context.

We begin with an easy lemma whose proof is left to the reader. The idea is that if the configuration outside of s is ξ under $\mu^{(-,\epsilon)}$, it must have been at least as large under μ "before flipping some 1's to 0's"; then use monotonicity.

Lemma 3.1 *Assume that μ is a monotone probability measure on $\{0,1\}^S$ where $|S| < \infty$. Take $s \in S$ and let $\xi \in \{0,1\}^{S\setminus s}$. Then, for any $\epsilon > 0$, we have that*

$$\mu^{(-,\epsilon)}(1|\xi) \geq (1-\epsilon)\mu(1|\xi)$$

and that

$$\mu^{(+,\epsilon)}(0|\xi) \geq (1-\epsilon)\mu(0|\xi).$$

The next lemma will be used to prove lemma 3.3.

Lemma 3.2 *Assume that μ is a monotone probability measure on $\{0,1\}^S$ where $|S| < \infty$. For any $\epsilon > 0$, $\mu^{(-,\epsilon)}$ is also monotone.*

Proof. Let $s \in S$ be arbitrary, $X \sim \mu$ and $X^{(-,\epsilon)} \sim \mu^{(-,\epsilon)}$. For any $\delta, \eta \in \{0,1\}^{S\setminus s}$ define the probability measures μ_δ and μ_η on $\{0,1\}^{S\setminus s}$ by letting $\mu_\delta(\mathcal{A}) = \mathbb{P}(X \in \mathcal{A}|X^{(-,\epsilon)}(S \setminus s) \equiv \delta)$ and $\mu_\eta(\mathcal{A}) = \mathbb{P}(X \in \mathcal{A}|X^{(-,\epsilon)}(S \setminus s) \equiv \eta)$ for every event $\mathcal{A}$ in $\{0,1\}^{S\setminus s}$, respectively. We will prove that

$$\mu_\delta \preceq \mu_\eta \quad \forall \delta \preceq \eta. \tag{15}$$

This will give us (since $\mathbb{P}(X(s) = 1|X(S \setminus s) \equiv \eta)$ is an increasing function of η) that

$$\mathbb{P}(X^{(-,\epsilon)}(s) = 1|X^{(-,\epsilon)}(S \setminus s) \equiv \eta)$$

$$= (1-\epsilon)\int_{\tilde\eta\in\{0,1\}^{S\setminus s}} \mathbb{P}(X(s)=1|X(S\setminus s)\equiv\tilde\eta)d\mu_\eta(\tilde\eta)$$
$$\geq (1-\epsilon)\int_{\tilde\eta\in\{0,1\}^{S\setminus s}} \mathbb{P}(X(s)=1|X(S\setminus s)\equiv\tilde\eta)d\mu_\delta(\tilde\eta)$$
$$= \mathbb{P}(X^{(-,\epsilon)}(s)=1|X^{(-,\epsilon)}(S\setminus s)\equiv\delta).$$

Since s was choosen arbitrarily this would prove the statement.

We now prove (15). Define for $\eta \preceq \tilde\eta$ $d(\tilde\eta,\eta) := |\{t\in S\setminus s : \tilde\eta(t)=1\}| - |\{t\in S\setminus s : \eta(t)=1\}|$ and $d(\tilde\eta,0) = |\{t\in S\setminus s : \tilde\eta(t)=1\}|$. Here $|\cdot|$ denotes cardinality. Let $\mu_{S\setminus s}(\eta) = \mathbb{P}(X(S\setminus s)\equiv\eta)$ and define $\mu^{(-,\epsilon)}_{S\setminus s}$ similarly. We have that for $\eta\preceq\tilde\eta$:

$$\mu_\eta(\tilde\eta) \tag{16}$$
$$= \mathbb{P}(X^{(-,\epsilon)}(S\setminus s)\equiv\eta|X(S\setminus s)\equiv\tilde\eta)\frac{\mu_{S\setminus s}(\tilde\eta)}{\mu^{(-,\epsilon)}_{S\setminus s}(\eta)}$$
$$= \epsilon^{d(\tilde\eta,\eta)}(1-\epsilon)^{d(\eta,0)}\frac{\mu_{S\setminus s}(\tilde\eta)}{\mu^{(-,\epsilon)}_{S\setminus s}(\eta)}. \tag{17}$$

It is well known that μ being monotone implies that for every $\tilde\delta,\ \tilde\eta$

$$\mu_{S\setminus s}(\tilde\eta\vee\tilde\delta)\mu_{S\setminus s}(\tilde\eta\wedge\tilde\delta) \geq \mu_{S\setminus s}(\tilde\eta)\mu_{S\setminus s}(\tilde\delta). \tag{18}$$

By a simple modification of Theorem 2.9 pg 75 of [19], it is enough for us to show that

$$\mu_\eta(\tilde\eta\vee\tilde\delta)\mu_\delta(\tilde\eta\wedge\tilde\delta) \geq \mu_\eta(\tilde\eta)\mu_\delta(\tilde\delta) \tag{19}$$

for all $\tilde\eta\succeq\eta$ and $\tilde\delta\succeq\delta$ to show (15). An elementary calculation shows that

$$d(\tilde\eta\vee\tilde\delta,\eta) + d(\tilde\eta\wedge\tilde\delta,\delta) = d(\tilde\eta,\eta) + d(\tilde\delta,\delta). \tag{20}$$

We therefore get

$$\mu_\eta(\tilde\eta\vee\tilde\delta)\mu_\delta(\tilde\eta\wedge\tilde\delta)$$
$$= \epsilon^{d(\tilde\eta\vee\tilde\delta,\eta)+d(\tilde\eta\wedge\tilde\delta,\delta)}(1-\epsilon)^{d(\eta,0)+d(\delta,0)}\frac{\mu_{S\setminus s}(\tilde\eta\vee\tilde\delta)}{\mu^{(-,\epsilon)}_{S\setminus s}(\eta)}\frac{\mu_{S\setminus s}(\tilde\eta\wedge\tilde\delta)}{\mu^{(-,\epsilon)}_{S\setminus s}(\delta)}$$
$$\geq \epsilon^{d(\tilde\eta,\eta)+d(\tilde\delta,\delta)}(1-\epsilon)^{d(\eta,0)+d(\delta,0)}\frac{\mu_{S\setminus s}(\tilde\eta)}{\mu^{(-,\epsilon)}_{S\setminus s}(\eta)}\frac{\mu_{S\setminus s}(\tilde\delta)}{\mu^{(-,\epsilon)}_{S\setminus s}(\delta)} = \mu_\eta(\tilde\eta)\mu_\delta(\tilde\delta),$$

where (16) is used in the first and last equality and equations (18) and (20) are used in the inequality.

QED

Lemma 3.3 *Let μ_1, μ_2 be probability measures on $\{0,1\}^S$ where $|S| < \infty$. Assume that μ_2 is monotone and that*

$$A := \inf_{\substack{s \in S \\ \xi \in \{0,1\}^{S \setminus s}}} [\mu_2(\sigma(s) = 1 | \sigma(S \setminus s) \equiv \xi) - \mu_1(\sigma(s) = 1 | \sigma(S \setminus s) \equiv \xi)] > 0.$$

Then for any choice of $\epsilon > 0$, such that

$$A > \frac{1}{1-\epsilon} - 1,$$

we have

$$\mu_1 \preceq \mu_2^{(-,\epsilon)}.$$

Hence (μ_1, μ_2) is downwards movable.

Proof. Monotonicity of μ_2, Lemma 3.1, the definition of A and our choice of ϵ give us that for any $s \in S$ and $\xi \in \{0,1\}^{S \setminus s}$

$$\begin{aligned} &\mu_2^{(-,\epsilon)}(1|\xi) \\ &\quad \geq (1-\epsilon)\mu_2(1|\xi) \geq (1-\epsilon)(A + \mu_1(1|\xi)) \\ &\quad \geq (1-\epsilon)\frac{\mu_1(1|\xi)}{1-\epsilon} = \mu_1(1|\xi). \end{aligned}$$

By Lemma 3.2, $\mu_2^{(-,\epsilon)}$ is monotone and so $\forall \tilde{\xi} \preceq \xi$,

$$\mu_1(1|\tilde{\xi}) \leq \mu_2^{(-,\epsilon)}(1|\tilde{\xi}) \leq \mu_2^{(-,\epsilon)}(1|\xi).$$

The proof is completed by the use of Holley's inequality, Theorem 2.1.

QED

Proposition 3.4 *Let S be any finite or countable set and consider $(S_n)_{n \in \mathbb{N}^+}$, a collection of sets such that $|S_n| < \infty$ $\forall n \in \mathbb{N}^+$ and $S_n \uparrow S$. Let $(\mu_{1,n})_{n \in \mathbb{N}^+}$, $(\mu_{2,n})_{n \in \mathbb{N}^+}$, be two collections of probability measures, where $\mu_{1,n}, \mu_{2,n}$ are probability measures on $\{0,1\}^{S_n}$ for every $n \in \mathbb{N}^+$. Furthermore, assume that all of the probability measures $(\mu_{1,n})_{n \in \mathbb{N}^+}$ $((\mu_{2,n})_{n \in \mathbb{N}^+})$ are monotone, that $\mu_{1,n} \to \mu_1$ and that $\mu_{2,n} \to \mu_2$. Set*

$$A_n := \inf_{\substack{s \in S_n \\ \xi \in \{0,1\}^{S_n \setminus s}}} [\mu_{2,n}(\sigma(s) = 1 | \sigma(S \setminus s) \equiv \xi) - \mu_{1,n}(\sigma(s) = 1 | \sigma(S \setminus s) \equiv \xi)].$$

If

$$\inf_{n \in \mathbb{N}^+} A_n > 0,$$

then (μ_1, μ_2) is upwards (downwards) movable.

Proof. Take $\epsilon > 0$ such that

$$\inf_{n\in\mathbb{N}^+} A_n > \frac{1}{1-\epsilon} - 1.$$

With this choice of ϵ, Lemma 3.3 says that $(\mu_{1,n}, \mu_{2,n})$ is upwards (downwards) ϵ-movable. Since $\mu_{1,n} \to \mu_1$ and $\mu_{2,n} \to \mu_2$ we easily get that $\mu_{2,n}^{(-,\epsilon)} \to \mu_2^{(-,\epsilon)}$ and $\mu_{1,n}^{(+,\epsilon)} \to \mu_1^{(+,\epsilon)}$. Furthermore since the relations

$$\mu_{1,n} \preceq \mu_{2,n}^{(-,\epsilon)}$$

and

$$\mu_{1,n}^{(+,\epsilon)} \preceq \mu_{2,n}$$

are easily seen to be preserved under weak limits, we get that

$$\mu_1 \preceq \mu_2^{(-,\epsilon)} \text{ and } \mu_1^{(+,\epsilon)} \preceq \mu_2.$$

QED

4 ϵ-movability for the contact process and a 0-1 Law

The conditions in our next proposition might seem overly technical; however, these represent the essential features of the contact process (after a small suitable time rescaling) and therefore we feel it is instructive to highlight these features. In Proposition 4.1 and Lemmas 5.1, 5.2 and 8.1 we will use the so-called graphical representation to define our processes; see for instance [19] page 172.

Proposition 4.1 *Let μ_1 and μ_2 be two probability measures defined on $\{0,1\}^S$, where S is a countable set. Assume that $\mu_1 \preceq \mu_2$ and that there exists two stationary Markov processes Ψ_1 and Ψ_2, governed by flip rate intensities $\{C_1(s,\sigma_1)\}_{s\in S,\sigma_1\in\{0,1\}^S}$ and $\{C_2(s,\sigma_2)\}_{s\in S,\sigma_2\in\{0,1\}^S}$ respectively, and with marginal distributions μ_1 and μ_2. Assume that $C_1 \preceq C_2$ (conditions (7) and (8) of the introduction). Consider the following conditions;*

1. *There exists an $\epsilon_1 > 0$ such that*

$$C_2(s,\sigma_2) - C_1(s,\sigma_1) \geq \epsilon_1 \quad (21)$$
$$\forall s \in S,\ \forall \sigma_2 \succeq \sigma_1 \text{ s.t. } \sigma_2(s) = 0 \text{ and } C_1(s,\sigma_1) \neq 0.$$

2. *There exists an $\epsilon_2 > 0$ such that*

$$C_1(s,\sigma_1) - C_2(s,\sigma_2) \geq \epsilon_2 \quad (22)$$
$$\forall s \in S,\ \forall \sigma_2 \succeq \sigma_1 \text{ s.t. } \sigma_1(s) = 1 \text{ and } C_2(s,\sigma_2) \neq 0.$$

3. There exists an $\epsilon_3 > 0$ such that

$$C_1(s, \sigma_1) \geq \epsilon_3 \ \forall s \in S, \ \forall \sigma_1 \ s.t. \ \sigma_1(s) = 1, \tag{23}$$

4. There exists an $\epsilon_4 > 0$ such that

$$C_2(s, \sigma_2) \geq \epsilon_4 \ \forall s \in S, \ \forall \sigma_2 \ s.t. \ \sigma_2(s) = 0. \tag{24}$$

If conditions 1 2 and 3 are satisfied, then (μ_1, μ_2) is downwards movable. If conditions 1 2 and 4 are satisfied, then (μ_1, μ_2) is upwards movable.

Proof. We will prove the first statement, the second follows by symmetry. Define

$$\lambda := \sup_{s, \sigma_2 : \sigma_2(s) = 0} C_2(s, \sigma_2) + \sup_{s, \sigma_1 : \sigma_1(s) = 1} C_1(s, \sigma_1).$$

Our aim is to construct a coupling of the processes $\{X_{1,t}\}_{t \geq 0} \sim \Psi_1$ and $\{X_{2,t}\}_{t \geq 0} \sim \Psi_2$ such that $X_{1,t} \preceq X_{2,t} \ \forall t \geq 0$ in such a way that we prove the proposition. Before presenting the actual coupling we will discuss the idea behind it. For every site $s \in S$ associate an independent Poisson process with parameter λ. Next, let $\{U_{s,k}\}_{s \in S, k \geq 1}$ and $\{U'_{s,k}\}_{s \in S, k \geq 1}$ be independent uniform $[0,1]$ random variables also independent of the Poisson processes. If τ is an arrival time for the Poisson process at site s, we write $U_{s,\tau}$ for $U_{s,k}$ where k is such that τ is the kth arrival of the Poisson process at site s. Now, let τ be an arrival time for the Poisson process associated to a site s. For $i \in \{1, 2\}$, let X_{i,τ^-} and X_{i,τ^+} denote the configurations before and after the arrival. We will let the outcome of $U_{s,\tau}$ decide what happens with the $\{X_{2,t}\}_{t \geq 0}$ process at time $t = \tau$, and then we will let $U'_{s,\tau}$ together with $U_{s,\tau}$ decide what happens with the $\{X_{1,t}\}_{t \geq 0}$ process at time $t = \tau$. As we will see, we will do this so that $X_{1,t} \preceq X_{2,t}$ for all $t \geq 0$. Furthermore, we will do this in such a way that there exists an $\epsilon \in (0, 1)$ such that if $U'_{s,\tau} \geq 1 - \epsilon$, then $X_{1,\tau^+}(s) = 0$ regardless of the outcome of $U_{s,\tau}$. Consider now the process $\{X_t^\epsilon\}_{t \geq 0}$ we get by taking $X_0^\epsilon(s) = 1$ for every $s \in S$ and letting $\{X_t^\epsilon(s)\}_{t \geq 0}$ be updated at every arrival time τ for the Poisson process associated to s, and updated in such a way that $X_{\tau^+}^\epsilon(s) = 0$ if $U'_{s,\tau} \geq 1 - \epsilon$, and $X_{\tau^+}^\epsilon(s) = 1$ if $U'_{s,\tau} < 1 - \epsilon$. Of course the distribution of X_t^ϵ will converge to $\pi_{1-\epsilon}$. Observe that whenever $X_t^\epsilon(s) = 0$ we have that $X_{1,t}(s) = 0$. Therefore we can conclude that

$$X_{1,t} \preceq \min(X_{2,t}, X_t^\epsilon) \ \forall t \geq 0. \tag{25}$$

Furthermore since the process $\{X_t^\epsilon\}_{t \geq 0}$ does not depend on any $U_{s,\tau}$ we have that $X_t^\epsilon(s)$ is conditionally independent of $X_{2,t}$ if there has been an arrival for the Poisson process associated to s before time t. Let s_i, $i \in \{1, \ldots, n\}$ be distinct sites in S and let $\mathcal{A}_t$ be the event that all Poisson processes associated to s_1 through s_n have had an arrival by time t. Of course $\mathbb{P}(\mathcal{A}_t) = (1-e^{-\lambda t})^n$ and so we get that

$$\begin{aligned}
&\mathbb{P}(X_{2,t}X_t^\epsilon(s_1) = \cdots = X_{2,t}X_t^\epsilon(s_n) = 1) \\
&\quad = \mathbb{P}(X_{2,t}X_t^\epsilon(s_1) = \cdots = X_{2,t}X_t^\epsilon(s_n) = 1|\mathcal{A}_t)\mathbb{P}(\mathcal{A}_t) \\
&\qquad + \mathbb{P}(X_{2,t}X_t^\epsilon(s_1) = \cdots = X_{2,t}X_t^\epsilon(s_n) = 1|\mathcal{A}_t^c)\mathbb{P}(\mathcal{A}_t^c) \\
&\quad = \mathbb{P}(X_{2,t}(s_1) = \cdots = X_{2,t}(s_n) = 1|\mathcal{A}_t) \\
&\qquad \times \mathbb{P}(X_t^\epsilon(s_1) = \cdots = X_t^\epsilon(s_n) = 1|\mathcal{A}_t)\mathbb{P}(\mathcal{A}_t) \\
&\qquad + \mathbb{P}(X_{2,t}X_t^\epsilon(s_1) = \cdots = X_{2,t}X_t^\epsilon(s_n) = 1|\mathcal{A}_t^c)\mathbb{P}(\mathcal{A}_t^c) \\
&\quad = \mathbb{P}(X_{2,t}(s_1) = \cdots = X_{2,t}(s_n) = 1|\mathcal{A}_t)\mathbb{P}(\mathcal{A}_t)(1-\epsilon)^n \\
&\qquad + \mathbb{P}(X_{2,t}X_t^\epsilon(s_1) = \cdots = X_{2,t}X_t^\epsilon(s_n) = 1|\mathcal{A}_t^c)\mathbb{P}(\mathcal{A}_t^c) \\
&\quad = \mathbb{P}(\{X_{2,t}(s_1) = \cdots = X_{2,t}(s_n) = 1\} \cap \mathcal{A}_t)(1-\epsilon)^n \\
&\qquad + \mathbb{P}(X_{2,t}X_t^\epsilon(s_1) = \cdots = X_{2,t}X_t^\epsilon(s_n) = 1|\mathcal{A}_t^c)\mathbb{P}(\mathcal{A}_t^c) \\
&\quad \geq (\mathbb{P}(X_{2,t}(s_1) = \cdots = X_{2,t}(s_n) = 1) - \mathbb{P}(\mathcal{A}_t^c))(1-\epsilon)^n \\
&\qquad + \mathbb{P}(X_{2,t}X_t^\epsilon(s_1) = \cdots = X_{2,t}X_t^\epsilon(s_n) = 1|\mathcal{A}_t^c)\mathbb{P}(\mathcal{A}_t^c) \\
&\quad = \mathbb{P}(X_{2,t}(s_1) = \cdots = X_{2,t}(s_n) = 1)(1-\epsilon)^n \\
&\qquad + \mathbb{P}(\mathcal{A}_t^c)(\mathbb{P}(X_{2,t}X_t^\epsilon(s_1) = \cdots = X_{2,t}X_t^\epsilon(s_n) = 1|\mathcal{A}_t^c) - (1-\epsilon)^n) \\
&\quad = \mu_2^{(-,\epsilon)}(\sigma(s_1) = \cdots = \sigma(s_n) = 1) \\
&\qquad + \mathbb{P}(\mathcal{A}_t^c)(\mathbb{P}(X_{2,t}X_t^\epsilon(s_1) = \cdots = X_{2,t}X_t^\epsilon(s_n) = 1|\mathcal{A}_t^c) - (1-\epsilon)^n) \\
&\quad \xrightarrow{t\to\infty} \mu_2^{(-,\epsilon)}(\sigma(s_1) = \cdots = \sigma(s_n) = 1).
\end{aligned}$$

In addition

$$\begin{aligned}
&\mathbb{P}(X_{2,t}(s_1) = \cdots = X_{2,t}(s_n) = 1 \cap \mathcal{A}_t)(1-\epsilon)^n \\
&\quad \leq \mathbb{P}(X_{2,t}(s_1) = \cdots = X_{2,t}(s_n) = 1)(1-\epsilon)^n \\
&\quad = \mu_2^{(-,\epsilon)}(\sigma(s_1) = \cdots = \sigma(s_n) = 1).
\end{aligned}$$

Hence, by inclusion exclusion, we have that the distribution of $\min(X_{2,t}, X_t^\epsilon)$ approaches $\mu_2^{(-,\epsilon)}$ as t tends to infinity. So by first taking the limit in (25), we get that $\mu_1 \preceq \mu_2^{(-,\epsilon)}$, as desired.

Now to the construction. Take $X_{1,0} \sim \mu_1$, $X_{2,0} \sim \mu_2$, such that $X_{1,0} \preceq X_{2,0}$. Let τ be an arrival time for the Poisson process associated to s. Take $U_{s,\tau}$ and $U'_{s,\tau}$. The following transition rules apply:

$X_{2,\tau-}$	$X_{2,\tau+}$	if
0	1	$U_{s,\tau} \leq \frac{C_2(s,X_{2,\tau-})}{\lambda}$
1	0	$U_{s,\tau} \geq \frac{\lambda - C_2(s,X_{2,\tau-})}{\lambda}$.

It is easy to check that the process $\{X_{2,t}\}_{t\geq 0}$ thus constructed will have the right flip-rate intensities. The construction of $\{X_{1,t}\}_{t\geq 0}$ is slightly more complicated. If $C_2(s, X_{2,\tau^-}) = 0$ and $X_{2,\tau^-}(s) = 0$ then it follows from (7) that $C_1(s, X_{1,\tau^-}) = 0$, and in that case we interpret $\frac{C_1(s,X_{1,\tau^-})}{C_2(s,X_{2,\tau^-})}$ as 0. Observe that $C_2(s, X_{2,\tau^-})$ can be 0 when $X_{2,\tau^-}(s) = 1$ but it will not cause any problems. With these observations in mind, these are the transition rules we apply:

$(X_{1,\tau^-}, X_{2,\tau^-})$	$(X_{1,\tau^+}, X_{2,\tau^+})$	if
$(0,0)$	$(1,1)$	$U_{s,\tau} \le \frac{C_2(s,X_{2,\tau^-})}{\lambda}$ and $U'_{s,\tau} \le \frac{C_1(s,X_{1,\tau^-})}{C_2(s,X_{2,\tau^-})}$
$(0,0)$	$(0,1)$	$U_{s,\tau} \le \frac{C_2(s,X_{2,\tau^-})}{\lambda}$ and $U'_{s,\tau} > \frac{C_1(s,X_{1,\tau^-})}{C_2(s,X_{2,\tau^-})}$
$(0,0)$	$(0,0)$	otherwise
$(0,1)$	$(0,0)$	$U_{s,\tau} \ge \frac{\lambda - C_2(s,X_{2,\tau^-})}{\lambda}$
$(0,1)$	$(1,1)$	$U_{s,\tau} < \frac{\sup_{s,\sigma_2:\sigma_2(s)=0} C_2(s,\sigma_2)}{\lambda}$ and $U'_{s,\tau} \le \frac{C_1(s,X_{1,\tau^-})}{\sup_{s,\sigma_2:\sigma_2(s)=0} C_2(s,\sigma_2)}$
$(0,1)$	$(0,1)$	otherwise
$(1,1)$	$(0,0)$	$U_{s,\tau} \ge \frac{\lambda - C_2(s,X_{2,\tau^-})}{\lambda}$
$(1,1)$	$(0,1)$	$U_{s,\tau} < \frac{\lambda - C_2(s,X_{2,\tau^-})}{\lambda}$ and $U'_{s,\tau} \ge \frac{\lambda - C_1(s,X_{1,\tau^-})}{\lambda - C_2(s,X_{2,\tau^-})}$
$(1,1)$	$(1,1)$	otherwise

It is not difficult to check that all flip rate intensities are correct and that $X_{1,t} \preceq X_{2,t}$ for all $t \ge 0$. Observe that by the definition of λ, the events $\left\{U_{s,\tau} \ge \frac{\lambda - C_2(s,X_{2,\tau^-})}{\lambda}\right\}$ and $\left\{U_{s,\tau} < \frac{\sup_{s,\sigma_2:\sigma_2(s)=0} C_2(s,\sigma_2)}{\lambda}\right\}$ are disjoint when $(X_{1,\tau^-}, X_{2,\tau^-}) = (0,1)$.

We now want to show that there exists an $\epsilon > 0$ so that $U'_{s,\tau} \ge 1 - \epsilon$, implies that $X_{1,\tau^+}(s) = 0$. Note that if $(X_{1,\tau^-}, X_{2,\tau^-}) = (0,0)$ and $C_1(s, X_{1,\tau^-}) > 0$ ($\Rightarrow C_2(s, X_{2,\tau^-}) > 0$) then

$$\frac{C_1(s, X_{1,\tau^-})}{C_2(s, X_{2,\tau^-})} \le \frac{C_2(s, X_{2,\tau^-}) - \epsilon_1}{C_2(s, X_{2,\tau^-})} \le 1 - \frac{\epsilon_1}{\sup_{s,\sigma_2:\sigma_2(s)=0} C_2(s,\sigma_2)} < 1$$

and if $(X_{1,\tau^-}, X_{2,\tau^-}) = (0,0)$ and $C_1(s, X_{1,\tau^-}) = 0$ then

$$\frac{C_1(s, X_{1,\tau^-})}{C_2(s, X_{2,\tau^-})} = 0.$$

Furthermore if $(X_{1,\tau^-}, X_{2,\tau^-}) = (0,1)$ and $C_1(s, X_{1,\tau^-}) > 0$, then

$$\frac{C_1(s, X_{1,\tau^-})}{\sup_{s,\sigma_2:\sigma_2(s)=0} C_2(s,\sigma_2)} \leq 1 - \frac{\epsilon_1}{\sup_{s,\sigma_2:\sigma_2(s)=0} C_2(s,\sigma_2)} < 1$$

while again if $(X_{1,\tau^-}, X_{2,\tau^-}) = (0,1)$ and $C_1(s, X_{1,\tau^-}) = 0$, then the 0 never changes to a 1. Finally if $(X_{1,\tau^-}, X_{2,\tau^-}) = (1,1)$ and $C_2(s, X_{2,\tau^-}) > 0$ ($\Rightarrow C_1(s, X_{1,\tau^-}) > 0$), then

$$\frac{\lambda - C_1(s, X_{1,\tau^-})}{\lambda - C_2(s, X_{2,\tau^-})} \leq \frac{\lambda - C_2(s, X_{2,\tau^-}) - \epsilon_2}{\lambda - C_2(s, X_{2,\tau^-})} \leq 1 - \frac{\epsilon_2}{\lambda - C_2(s, X_{2,\tau^-})} \leq 1 - \frac{\epsilon_2}{\lambda},$$

and if $(X_{1,\tau^-}, X_{2,\tau^-}) = (1,1)$ and $C_2(s, X_{2,\tau^-}) = 0$,

$$\frac{\lambda - C_1(s, X_{1,\tau^-})}{\lambda - C_2(s, X_{2,\tau^-})} \leq \frac{\lambda - \epsilon_3}{\lambda} = 1 - \frac{\epsilon_3}{\lambda} < 1.$$

Therefore, whenever

$$U'_{s,\tau} \geq \max\left(1 - \frac{\epsilon_1}{\sup_{s,\sigma_2:\sigma_2(s)=0} C_2(s,\sigma_2)}, 1 - \frac{\epsilon_2}{\lambda}, 1 - \frac{\epsilon_3}{\lambda}\right),$$

we have that $X_{1,\tau^+}(s) = 0$ regardless of the outcome of $U_{s,\tau}$. Therefore (μ_1, μ_2) is downwards ϵ-movable where

$$\begin{aligned} \epsilon &:= 1 - \max\left(1 - \frac{\epsilon_1}{\sup_{s,\sigma_2:\sigma_2(s)=0} C_2(s,\sigma_2)}, 1 - \frac{\epsilon_2}{\lambda}, 1 - \frac{\epsilon_3}{\lambda}\right) \\ &= \min\left(\frac{\epsilon_1}{\sup_{s,\sigma_2:\sigma_2(s)=0} C_2(s,\sigma_2)}, \frac{\epsilon_2}{\lambda}, \frac{\epsilon_3}{\lambda}\right). \end{aligned}$$

QED

Proof of Theorem 1.11. Take $\delta > 0$ such that $\lambda_1(1+\delta) < \lambda_2$ and consider the process $\{X_t\}_{t\geq 0}$ constructed in the following way. Take $X_0 \equiv 1$ and let the process evolve with flip rate intensities

$$C_1(s,\sigma) = \begin{cases} 1+\delta & \text{if} \quad \sigma(s) = 1 \\ \lambda_1(1+\delta)\sum_{s'\sim s} \sigma(s') & \text{if} \quad \sigma(s) = 0. \end{cases} \tag{26}$$

Denote the limiting distribution of X_t as t tends to infinity by $\mu_{1+\delta,\lambda_1(1+\delta)}$. It is easy to see that this process is just a time-scaling of the contact process constructed in Section 2.6 with parameter λ_1. Recall that that process had limiting distribution μ_{λ_1}, the upper invariant measure for the contact process. Thus we have $\mu_{\lambda_1} = \mu_{1+\delta,\lambda_1(1+\delta)}$. By Proposition 4.1 with C_1 as above and C_2 as in Section 2.6 with parameter λ_2, there exists an $\epsilon > 0$ such that

$$\mu_{1+\delta,\lambda_1(1+\delta)} \preceq \mu_{\lambda_2}^{(-,\epsilon)}.$$

Hence $(\mu_{\lambda_1}, \mu_{\lambda_2})$ is downwards movable.

QED

For the rest of this section we will only consider the graph $\mathbb{T}^d$ for $d \geq 2$. The following is a 0-1 law for the upper invariant measure for the contact process.

Proposition 4.2 *Let $\mathcal{A} \subseteq \{0,1\}^{\mathbb{T}^d}$ where $d \geq 2$ be a set which is invariant under all graph automorphisms on $\mathbb{T}^d$. Then, for $\lambda > 0$, we have that*

$$\mu_\lambda(\mathcal{A}) \in \{0,1\}.$$

Proof. Let $\epsilon > 0$. By elementary measure theory, there exists a cylinder event $\mathcal{B}$ depending on finitely many coordinates such that

$$\mu_\lambda(\mathcal{A}\Delta\mathcal{B}) \leq \epsilon. \tag{27}$$

Let supp$\mathcal{B}$ denote the finite number of coordinates with respect to which $\mathcal{B}$ is measurable. Letting $\{T_\lambda(t)\}_{t\geq 0}$ denote the Markov semigroup for the contact process with parameter λ, we have that $\delta_1 T_\lambda(t) \to \mu_\lambda$ and also that $\mu_\lambda \preceq \delta_1 T_\lambda(t)$ for every $t \geq 0$. Choose t so that for all (equivalently some) sites s

$$\delta_1 T_\lambda(t)(\eta(s) = 1) \leq \mu_\lambda(\eta(s) = 1) + \frac{\epsilon}{2|\text{supp}\mathcal{B}|}.$$

It follows easily that if m is any coupling of $\delta_1 T_\lambda(t)$ and μ_λ which is concentrated on $\{(\eta,\delta) : \eta \preceq \delta\}$, then for any finite set S of sites

$$m(\{(\eta,\delta) : \eta(s) \neq \delta(s) \text{ occurs for some } s \in S) \leq \frac{|S|\epsilon}{2|\text{supp}\mathcal{B}|}.$$

In particular, if E is any event depending on at most $2|\text{supp}\mathcal{B}|$ sites, then

$$|\delta_1 T_\lambda(t)(E) - \mu_\lambda(E)| \leq \epsilon. \tag{28}$$

For this fixed t, Theorem 4.6 page 35 of [19] shows that there exists an automorphism $\gamma \in AUT(\mathbb{T}^d)$ such that

$$|\delta_1 T_\lambda(t)(\mathcal{B} \cap \gamma\mathcal{B}) - \delta_1 T_\lambda(t)(\mathcal{B})\delta_1 T_\lambda(t)(\gamma\mathcal{B})| \leq \epsilon. \tag{29}$$

Furthermore, since μ_λ is invariant under automorphisms (27) implies that

$$\mu_\lambda(\gamma\mathcal{A}\Delta\gamma\mathcal{B}) \leq \epsilon,$$

and since $\mathcal{A} = \gamma\mathcal{A}$, we have

$$\mu_\lambda(\mathcal{A}\Delta\gamma\mathcal{B}) \leq \epsilon.$$

It follows that

$$\mu_\lambda(\mathcal{B}\Delta\gamma\mathcal{B}) \leq \mu_\lambda(\mathcal{A}\Delta\gamma\mathcal{B}) + \mu_\lambda(\mathcal{A}\Delta\mathcal{B}) \leq 2\epsilon.$$

Next, (28) implies that

$$|\delta_1 T_\lambda(t)(\mathcal{B}\Delta\gamma\mathcal{B}) - \mu_\lambda(\mathcal{B}\Delta\gamma\mathcal{B})| \leq \epsilon,$$

and so

$$\delta_1 T_\lambda(t)(\mathcal{B}\Delta\gamma\mathcal{B}) \leq 3\epsilon. \tag{30}$$

We get that

$$\begin{aligned}|\mu_\lambda(\mathcal{A}) - \mu_\lambda(\mathcal{A})^2| &= |\mu_\lambda(\mathcal{A}) - \mu_\lambda(\mathcal{A})\mu_\lambda(\gamma\mathcal{A})| \\ &\leq |\mu_\lambda(\mathcal{B}) - \mu_\lambda(\mathcal{B})\mu_\lambda(\gamma\mathcal{B})| + 3\epsilon \\ &\leq |\delta_1 T_\lambda(t)(\mathcal{B}) - \delta_1 T_\lambda(t)(\mathcal{B})\delta_1 T_\lambda(t)(\gamma\mathcal{B})| + 6\epsilon \\ &\leq |\delta_1 T_\lambda(t)(\mathcal{B}) - \delta_1 T_\lambda(t)(\mathcal{B}\cap\gamma\mathcal{B})| + 7\epsilon \\ &\leq \delta_1 T_\lambda(t)(\mathcal{B}\Delta\gamma\mathcal{B}) + 7\epsilon \leq 10\epsilon.\end{aligned}$$

Where we used (27), (28) and (29) for the three first inequalities and (30) in the last. Since $\epsilon > 0$, was choosen arbitrarily we get that

$$\mu_\lambda(\mathcal{A}) = \mu_\lambda(\mathcal{A})^2$$

and so $\mu_\lambda(\mathcal{A}) \in \{0, 1\}$.

QED

Remarks: The above proof works for any transitive and even quasi-transitive graph. For the case of $\mathbb{Z}^d$, this was proved in Proposition 2.16 page 143 of [19]. It is mentioned there that while $\delta_1 T_\lambda(t)$ is ergodic for each t, one cannot conclude immediately the ergodicity of μ_λ because the class of ergodic processes is not weakly closed. We point out however that there is another important notion of convergence given by the $\bar{d}$-metric (see [24] page 89 for definition) on stationary processes. Convergence in this metric is stronger than weak convergence and weaker than convergence in the total variation norm. It is also known that the ergodic processes are $\bar{d}$-closed and that weak convergence together with stochastic ordering implies

$\bar{d}$-convergence. In this way, one can conclude ergodicity of μ_λ using the $\bar{d}$-metric giving an alternative proof of Proposition 2.16 of [19]. In fact, the proof of Proposition 4.2 is essentially based on this idea. However, because of the open question listed below, it is not so easy to formulate the $\bar{d}$-metric for tree indexed processes and so we choose a more hands on approach. Observe that the crucial property of $\bar{d}$-convergence which is essentially used in the above proof is that for each fixed k, one has uniform convergence of the probability measures (in say the total variation norm) over all sets which depend on at most k points. (The point is that the k points can lie anywhere and hence this is much stronger than weak convergence).

Open Question related to defining the $\bar{d}$-metric for tree indexed processes: Assume that μ and ν are two automorphism invariant probability measures on $\{0,1\}^{\mathbb{T}^d}$ such that $\mu \preceq \nu$. Does there exist a $\mathbb{T}^d$-invariant coupling (X,Y) with $X \sim \mu$, $Y \sim \nu$ and $X \preceq Y$?

Proposition 4.3 *On $\mathbb{T}^d$, $d \geq 2$ there exists a λ_p such that for all $\lambda > \lambda_p$*

$$\mu_\lambda(\mathcal{C}^+) = 1.$$

Proof. By Theorem 1.33(c), page 275 in [19], for sufficiently large λ, $\mu_\lambda(\eta(s) = 1) \geq 2/3$. By [12] we have that if $\mu_\lambda(\eta(s) = 1) \geq 2/3$, then

$$\mu_\lambda(\mathcal{C}^+) > 0.$$

Finally, Proposition 4.2 then implies that

$$\mu_\lambda(\mathcal{C}^+) = 1.$$

QED

5 Relationship between ϵ-movability and dynamics

In the general setup we have a family of stationary Markov processes parametrised by one or two parameters, e.g. the contact processes Ψ^λ (λ is here the only parameter) or a stochastic Ising model $\Psi^{+,\beta,h}$ (β and h being the parameters). Many of the proofs in this paper will involve comparing the marginal distributions of these Markov processes for two different values of one of the involved parameters. Let p be the parameter and let $p_1 < p_2$. Assume that the marginal distributions are μ_{p_1} and μ_{p_2} respectively and that $\mu_{p_1} \preceq \mu_{p_2}$. Lemmas 5.1 and 5.2 shows that there is a close connection between showing that (μ_{p_1}, μ_{p_2}) is downwards ϵ-movable and that the infimum of the second process over a short time interval is stochastically larger than the first process.

Let Ψ^μ be a stationary Markov process on $\{0,1\}^S$ with marginal distribution μ and let $\{X_t\}_{t\geq 0} \sim \Psi^\mu$. For $\delta > 0$ and $s \in S$ define

$$X_{\inf,\delta}(s) := \inf_{t\in[0,\delta]} X_t(s),$$

and denote the distribution of $X_{\inf,\delta}$ by $\mu_{\inf,\delta}$. Similarly define

$$X_{\sup,\delta}(s) := \sup_{t\in[0,\delta]} X_t(s),$$

and denote the distribution of $X_{\sup,\delta}$ by $\mu_{\sup,\delta}$.

Lemma 5.1 *Take S to be the sites of a bounded degree graph. Let $\{C(s,\sigma)\}_{s\in S,\ \sigma\in\{-1,1\}^S}$ be the flip rate intensities for a stationary Markov process Ψ^μ on $\{-1,1\}^S$ with marginal distribution μ. Let*

$$\lambda := \sup_{(s,\sigma)} C(s,\sigma).$$

For any $\tau > 0$, if we set $\epsilon := 1 - e^{-\lambda\tau}$, we have that

$$\mu^{(-,\epsilon)} \preceq \mu_{\inf,\tau}.$$

Similarly, we get that

$$\mu_{\sup,\tau} \preceq \mu^{(+,\epsilon)}.$$

Proof. We will prove the first statement, the second statement follows by symmetry. Take $\tau > 0$. For every $s \in S$ associate an independent Poisson process with parameter λ. Define $\{(X_t^1, X_t^2)\}_{t\geq 0}$ in the following way. Let $X_0^1 \equiv X_0^2 \sim \mu$, and take t' to be an arrival time for the Poisson process of a site s. For $i \in \{1,2\}$, let $X_{t',-}^i$ and $X_{t',+}^i$ denote the configurations before and after the arrival. We let $X_{t',+}^1(s) \neq X_{t',-}^1(s)$ with probability $C(s, X_{t',-}^1)/\lambda$ and we let $X_{t',+}^2(s) = 0$ and finally we let $X_{t',+}^1(S \setminus s) \equiv X_{t',-}^1(S \setminus s)$, $X_{t',+}^2(S \setminus s) \equiv X_{t',-}^2(S \setminus s)$. Do this independently for all arrival times for all Poisson processes of all sites. Observe that once $X_t^2(s)$ is 0, it remains so. Note also that $X_\tau^1 \sim \mu$, $X_\tau^2 \sim \mu^{(-,\epsilon)}$. Furthermore if $X_t^1(s) = 0$ for some $t \in [0,\tau]$ the construction guarantees that $X_\tau^2(s) = 0$ and therefore $X_\tau^2 \preceq X_{\inf,\tau}^1 \sim \mu_{\inf,\tau}$.

QED

Lemma 5.2 *Take S to be the sites of any bounded degree graph. Let $\{C(s,\sigma)\}_{s\in S,\ \sigma\in\{-1,1\}^S}$ be the flip rate intensities of a stationary Markov process Ψ^μ on $\{-1,1\}^S$ with marginal distribution μ. Define*

$$\lambda_1 := \inf_{s,\sigma:\sigma(s)=1} C(s,\sigma).$$

If $\lambda_1 > 0$ *then for any* $0 < \epsilon < 1$*, if we set* $\tau := -\frac{\log(1-\epsilon)}{\lambda_1}$*, we have that*

$$\mu_{\inf,\tau} \preceq \mu^{(-,\epsilon)}.$$

Similarly, defining $\lambda_2 := \inf_{s,\sigma:\sigma(s)=0} C(s,\sigma)$*, if* $\lambda_2 > 0$*, then for any* $0 < \epsilon < 1$*, if we set* $\tau := -\frac{\log(1-\epsilon)}{\lambda_2}$*, we have that*

$$\mu^{(+,\epsilon)} \preceq \mu_{\sup,\tau}.$$

Proof. We will prove the first statement, the second statement follows by symmetry. For every $s \in S$ associate an independent Poisson process with parameter $\lambda := \sup_{(s,\sigma)} C(s,\sigma)$. Next, let $\{U_{s,k}\}_{s\in S,k\geq 1}$ be independent uniform $[0,1]$ random variables also independent of the Poisson processes. If t' is an arrival time for the Poisson process at site s, we write $U_{s,t'}$ for $U_{s,k}$ where k is such that t' is the kth arrival of the Poisson process at site s. Define $\{(X_t^1, X_t^2)\}_{t\geq 0}$ in the following way. Let $X_0^1 \equiv X_0^2 \sim \mu$, and take t' to be an arrival time for the Poisson process of a site s. We let $X_{t',+}^1(s) \neq X_{t',-}^1(s)$ if $U_{s,t'} \leq C(s, X_{t',-}^1)/\lambda$. Furthermore we let $X_{t',+}^2(s) = 0$ if $U_{s,t'} \leq \lambda_1/\lambda$ or $X_{t',-}^2(s) = 0$, and finally we let $X_{t',+}^1(S \setminus s) \equiv X_{t',-}^1(S \setminus s)$, $X_{t',+}^2(S \setminus s) \equiv X_{t',-}^2(S \setminus s)$. Do this independently for all arrival times for all Poisson processes of all sites. Clearly $X_\tau^1 \sim \mu$ and $X_\tau^2 \sim \mu^{(-,\epsilon)}$. Furthermore, if $X_\tau^2(s) = 0$, then either $X_0^1(s) = X_0^2(s) = 0$ or there exists a $t \in [0,\tau]$ such that t is an arrival time for the Poisson process associated to s and $U_{s,t} \leq \lambda_1/\lambda$. Since $\lambda_1 \leq C(s, X_{t-}^1)$ if $X_{t-}^1(s) = 1$, we get that either $X_{t+}^1(s)$ or $X_{t-}^1(s)$ is 0 and therefore $X_{\inf,\tau}^1 \preceq X_\tau^2$.

QED

To illustrate why the condition $\lambda_1 > 0$ of Lemma 5.2 is needed, consider the case $\mu = \pi_p$ for some $p > 0$. With $\epsilon > 0$, if we assume the trivial dynamics $C(s,\sigma) = 0$ for all s, σ, we will of course not have that $\mu_{\inf,\tau} \preceq \mu^{(-,\epsilon)}$ for any $\tau > 0$.

6 Proof of Theorem 1.9

Proof of Theorem 1.9. Take $\lambda > \lambda_p$ and let $\lambda' = (\lambda + \lambda_p)/2$. By Theorem 1.11 there exists an $\epsilon > 0$ such that $(\mu_{\lambda'}, \mu_\lambda)$ is downwards ϵ-movable. Lemma 5.1 gives us that there exists a $\tau > 0$ such that $\mu_\lambda^{(-,\epsilon)} \preceq \mu_{\lambda,\inf,\tau}$ and hence that $\mu_{\lambda'} \preceq \mu_{\lambda,\inf,\tau}$. Therefore, since $\mathcal{C}^+$ is an increasing event and $\lambda' > \lambda_p$, we have that

$$1 = \mu_{\lambda'}(\mathcal{C}^+) \leq \mu_{\lambda,\inf,\tau}(\mathcal{C}^+)$$

and so

$$\Psi^\lambda(\mathcal{C}_t^+ \ \forall t \in [0,\tau]) = 1.$$

The theorem now follows from countable additivity.

QED

7 Proof of Theorem 1.1

In this section we will deal with stationary distributions for interacting particle systems which are monotone in the sense of Definition 2.2.

Let $G = (S, E)$ be a countable connected locally finite graph and let $\Lambda \subseteq S$ be connected and $|\Lambda| < \infty$. Let $\{\mu_\Lambda^p\}_{p \in I}$, where $I \subseteq \mathbb{R}$ be a family of probability measures on $\{-1,1\}^\Lambda$ such that

$$\mu_\Lambda^{p_1} \preceq \mu_\Lambda^{p_2} \; \forall p_1 \leq p_2.$$

Assume that there exist stationary Markov processes Ψ_Λ^p governed by flip rate intensities $\{C_{p,\Lambda}(s,\sigma)\}_{s \in \Lambda, \sigma \in \{-1,1\}^\Lambda}$ and with marginal distributions μ_Λ^p. Furthermore assume that there exists limiting distributions Ψ^p of Ψ_Λ^p and μ^p of μ_Λ^p as $\Lambda \uparrow S$. Assume that μ_Λ^p are monotone for every p and Λ. For $p_1 < p_2$, let

$$A_{\Lambda,p_1,p_2} := \inf_{\substack{s \in \Lambda \\ \xi \in \{-1,1\}^{\Lambda \setminus s}}} [\mu_\Lambda^{p_2}(\sigma(s) = 1 | \sigma(\Lambda \setminus s) \equiv \xi) - \mu_\Lambda^{p_1}(\sigma(s) = 1 | \sigma(\Lambda \setminus s) \equiv \xi)]$$

and assume that for all $p_1 < p_2$

$$\inf_{\Lambda \subseteq S} A_{\Lambda,p_1,p_2} > 0.$$

For fixed $p_1 < p_2$ there exists by Proposition 3.4 an $\epsilon > 0$ such that (μ^{p_1}, μ^{p_2}) is both upwards and downwards ϵ-movable. Next, by Lemma 5.1 there exists a $\tau > 0$ such that

$$\mu^{p_2,(-,\epsilon)} \preceq \mu_{\inf,\tau}^{p_2},$$

and therefore

$$\mu^{p_1} \preceq \mu_{\inf,\tau}^{p_2}. \tag{31}$$

Theorem 7.1 *Consider the setup just described. Let $\mathcal{A}$ be an increasing event on $\{-1,1\}^S$ and let $\mathcal{A}_t$ be the event that $\mathcal{A}$ occurs at time t.*
(1) Let $a \in \mathbb{R}$. If

$$\mu^p(\mathcal{A}) = 1$$

for all $p \in I$ with $p > a$, then

$$\Psi^p(\mathcal{A}_t \textit{ occurs for every } t) = 1$$

for all $p \in I$ with $p > a$.
(2) Let $a \in \mathbb{R}$. If

$$\mu^p(\mathcal{A}) = 0$$

for all $p \in I$ with $p < a$, then

$$\Psi^p(\mathcal{A}_t \textit{ occurs for some } t) = 0$$

for all $p \in I$ with $p < a$.

Proof. We prove only (1) as (2) is proved in an identical way. Take $p > a$ and let $p_2 = (p+a)/2$. By the argument leading towards (31), there exists $\tau > 0$ such that

$$\mu^{p_2}(\mathcal{A}) \leq \mu^{p}_{\inf,\tau}(\mathcal{A}).$$

By using $\mu^{p_2}(\mathcal{A}) = 1$ and

$$\mu^{p}_{\inf,\tau}(\mathcal{A}) \leq \Psi^p(\mathcal{A}_t \text{ occurs for every } t \in [0,\tau]),$$

we get by countable additivity that

$$\Psi^p(\mathcal{A}_t \text{ occurs for every } t) = 1.$$

QED

We will now be able to prove Theorem 1.1 easily.

Proof of Theorem 1.1. We prove only the very first statement; all the other statements are proved in a similar manner. We fix $\beta \geq 0$ and then h will correspond to our parameter p in the above set up. For any $\Lambda \subseteq S$, any $s \in \Lambda$ and any $\xi \in \{-1,1\}^{\Lambda \setminus s}$, we have that

$$\mu_{\Lambda}^{+,\beta,h}(\sigma(s) = 1 | \sigma(\Lambda \setminus s) = \xi) = \frac{1}{1 + e^{-2\beta(\sum_{t:t\sim s}\xi(t)) - 2h}}, \tag{32}$$

where we let $\xi(t) = 1$ if $t \in \Lambda^c$ in order to take the boundary condition into account. It is obvious from (32) and the definition of monotonicity that $\mu_{\Lambda}^{+,\beta,h}$ is monotone for any h and Λ. Letting $h_1 < h_2$, it is immediate that

$$A_{\Lambda,h_1,h_2} = \inf_{\substack{s\in\Lambda \\ \xi\in\{-1,1\}^{\Lambda\setminus s}}} \left[\frac{1}{1 + e^{-2\beta(\sum_{t:t\sim s}\xi(t)) - 2h_2}} - \frac{1}{1 + e^{-2\beta(\sum_{t:t\sim s}\xi(t)) - 2h_1}}\right] > 0,$$

where again $\xi(t) = 1$ for all $t \in \Lambda^c$. It is not hard to see that this strict inequality must hold uniformly in Λ; i.e.,

$$\inf_{\Lambda \subseteq S} A_{\Lambda,h_1,h_2} > 0.$$

It follows that all of the assumptions of Theorem 7.1 hold and part (1) of that result gives us what we want.

QED

Proof of Lemma 1.2. Fix $\beta \geq 0$. Given any $p \in (0,1)$, it is easy to see that there exists a real number h_2 such that for all $h \geq h_2$, for $s \in S$ and for all $\xi \in \{-1,1\}^{S \setminus s}$

$$\mu^{+,\beta,h}(\sigma(s) = 1 | \sigma(S \setminus s) = \xi) \geq p$$

and hence $\pi_p \preceq \mu^{+,\beta,h}$. It is also easy to see that there exists a real number h_1 such that for all $h < h_1$, for $s \in S$ and for all $\xi \in \{-1,1\}^{S\setminus s}$

$$\mu^{+,\beta,h}(\sigma(s) = 1 | \sigma(S \setminus s) = \xi) \leq p$$

and hence $\mu^{+,\beta,h} \preceq \pi_p$. The statements of the lemma easily follow from these facts.

QED

8 Proof of Theorem 1.3

In this section we will use a variant of the so called Peierls argument to prove Theorem 1.3. We prove this only for $\mathbb{Z}^2$; the proof (with more complicated topological details) can be carried out for $\mathbb{Z}^d$ with $d \geq 3$.

We will write $0 \overset{-,t}{\longleftrightarrow} \partial\Lambda_L$ for the event that there exists a path of sites in state -1 connecting the origin to $\partial\Lambda_L := \Lambda_{L+1} \setminus \Lambda_L$ at time t and we will write $0 \overset{-,t}{\longleftrightarrow} \infty$ for the event that there exists an infinite path of sites in state -1 containing the origin at time t. We will also write $0 \overset{+,t}{\longleftrightarrow} \partial\Lambda_L$ and $0 \overset{+,t}{\longleftrightarrow} \infty$ for the obvious analogous events. We will first need Lemma 8.1 and the concept of a dual graph. The dual graph $G_n^{dual} = (S_n^{dual}, E_n^{dual})$ of $G_n = (S_n, E_n)$ consists of the set of sites $S_n^{dual} := \{-n - \frac{1}{2}, \ldots, n + \frac{1}{2}\}^2$ and E_n^{dual} which is the set of nearest neighbor pairs of S_n^{dual}. In this paper we will only work with the edges of the dual graph. An edge $e \in E_n^{dual}$ crosses one (and only one) edge $f \in E_n$ and the end sites of this edge f will be called the sites (of G_n) associated to e. For a random spin configuration X on $\{-1,1\}^{S_n}$ define a random edge configuration Y on $\{0,1\}^{E_n^{dual}}$ in the following way:

$$Y(e) = \begin{cases} 0 & \text{if} \quad X(t) = X(s) \\ 1 & \text{if} \quad X(t) \neq X(s), \end{cases} \tag{33}$$

where s, t are the sites associated to edge $e \in E_n^{dual}$. In figure (1) we have drawn a configuration $\sigma \in \{-1,1\}^{S_1}$ and the induced edge configuration on $\{0,1\}^{E_1^{dual}}$.

Assume that the sites evolve according to the flip rate intensities $\{C_n(s,\sigma)\}_{s \in S_n,\ \sigma \in \{-1,1\}^{S_n}}$. Consider γ, a (finite) path of edges in the dual graph. Take γ' to be a subset of γ. Assume that all edges of γ' are absent and all edges of $\gamma \setminus \gamma'$ are present at $t = 0$. We want to estimate the probability of the event that all edges of γ' are present at some point (not necessarily all at the same time) during some time interval $[0, \tau]$. In other words we want to estimate the probability of the event $\{Y_{\sup,\tau}(\gamma') \equiv 1 | Y_0(\gamma') \equiv 0, Y_0(\gamma \setminus \gamma') \equiv 1\}$.

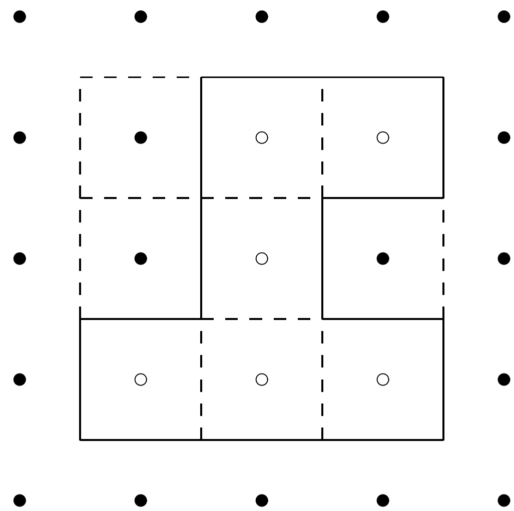

Figure 1: S_1 and the edges of it's dual graph. A solid circle marks a site with spin 1, while an empty circle has spin -1. A solid line is a present edge of the dual graph, and a dashed line is an absent edge of the dual graph.

Lemma 8.1 *Let $\{C_n(s,\sigma)\}_{s\in S_n,\ \sigma\in\{-1,1\}^{S_n}}$ be the flip rate intensities for a stationary Markov process on $\{-1,1\}^{S_n}$ and let Y_t be defined as above. Let*

$$\lambda := \sup_{(s,\sigma)} C_n(s,\sigma)\ (<\infty).$$

For any $\tau>0$ and any $\gamma'\subseteq E_n^{dual}$,

$$\mathbf{P}(Y_{\sup,\tau}(\gamma')\equiv 1|Y_0(\gamma')\equiv 0, Y_0(E_n^{dual}\setminus\gamma'))\leq (4(1-e^{-\lambda\tau})^{1/4})^{|\gamma'|}.$$

Proof.

Take $\tau>0$. For every $s\in S_n$ associate an independent Poisson process with parameter λ. Define $\{X_t\}_{t\geq 0}$ in the following way. Let $X_0\sim\mu$ and take t' to be an arrival time for the Poisson process of a site s. We let $X_{t',+}(s)\neq X_{t',-}(s)$ with probability $C(s,X_{t',-})/\lambda$. Do this independently for all arrival times for all Poisson processes associated to the different sites. It is immediate that $X_\tau\sim\mu$. Let s_i, $i\in\{1,\dots,l\}$ be distinct sites of S_n. The event $\{X_{\inf,\tau}(s_i)\neq X_{\sup,\tau}(s_i)\ \forall i\in\{1,\dots,l\}\}$ is contained in the event that every Poisson process associated to the sites $s_i, i\in\{1,\dots,l\}$ have had at least one arrival by time τ. The probability that a particular site has had an arrival by time τ is $1-e^{-\lambda\tau}$. Furthermore this event is independent of the Poisson processes for all other sites. Therefore

$$\mathbf{P}(X_{\inf,\tau}(s_i)\neq X_{\sup,\tau}(s_i)\ \forall i\in\{1,\dots,l\})\leq (1-e^{-\lambda\tau})^l. \tag{34}$$

Given γ', consider the set of all sites associated to some edge of γ' and let $n_{\gamma'}$ be the cardinality of that set. Observe that $n_{\gamma'} \le 2|\gamma'|$ and that in order for the event $\{Y_{\sup,\tau}(\gamma') \equiv 1 | Y_0(\gamma') \equiv 0, Y_0(E_n^{dual} \setminus \gamma')\}$ to occur, at least $|\gamma'|/4$ of the sites associated to γ' must flip during $[0,\tau]$. This is because one site is associated to at most 4 edges. Denote the event that at least $|\gamma'|/4$ of the sites associated to γ' flips during $[0,\tau]$ by $\mathcal{A}_{\tau,\gamma'}$. Take $\tilde{S}$ to be a subset of the sites associated to γ' such that $|\tilde{S}| \ge |\gamma'|/4$. By (34), the probability that all of these sites flips during $[0,\tau]$ is less than $(1-e^{-\lambda\tau})^{|\tilde{S}|} \le (1-e^{-\lambda\tau})^{|\gamma'|/4}$. To conclude, observe that the number of subsets of the sites associated to γ' is bounded by $2^{2|\gamma'|}$. Hence, the probability of the event $\mathcal{A}_{\tau,\gamma'}$ must be less than $(1-e^{-\lambda\tau})^{|\gamma'|/4}2^{2|\gamma'|}$, and so

$$\begin{aligned}&\mathbf{P}(Y_{\sup,\tau}(\gamma') \equiv 1 | Y_0(\gamma') \equiv 0, Y_0(E_n^{dual} \setminus \gamma')) \\ &\quad \le \mathbf{P}(\mathcal{A}_{\tau,\gamma'}) \le ((1-e^{-\lambda\tau})^{1/4}4)^{|\gamma'|}.\end{aligned}$$

QED

Proof of Theorem 1.3. We will prove the theorem for $d=2$. For $\beta > \beta_p$, choose $\delta_1 > 0$ so that $\beta' := \beta\frac{2-\delta_1}{2} > \beta_p$ and hence

$$\sum_{l=1}^{\infty} l 3^{l-1} e^{-2\beta' l} < \infty.$$

Next, choose N and $\epsilon < 1/2$ such that $\frac{4}{N} \le \delta_1$, and $\epsilon^{\frac{1}{N}} \le e^{-\beta(2-\delta_1)}$ and let τ be such that $\epsilon = 4(1-e^{-\lambda\tau})^{1/4}$. Let $\delta > 0$ be arbitrary and choose L so that

$$3\sum_{l=L}^{\infty} l 3^{l-1} e^{-2\beta' l} < \delta.$$

Let $\mathcal{E}_{L,\tau}$ be the event that $0 \overset{-,t}{\longleftrightarrow} \partial\Lambda_L$, for some $t \in [0,\tau]$. Let $\Psi_n^{+,\beta}$ be defined as in Section 2.3. We will show that

$$\Psi_n^{+,\beta}(\mathcal{E}_{L,\tau}) < \delta \ \forall n > L.$$

Since $\Psi_n^{+,\beta}(\mathcal{E}_{L,\tau}) \to \Psi^{+,\beta}(\mathcal{E}_{L,\tau})$, (see Section 2.3) we get that $\Psi^{+,\beta}(\mathcal{E}_{L,\tau}) \le \delta$. Letting $L \to \infty$ and $\delta \to 0$, we get that

$$\Psi^{+,\beta}(\exists t \in [0,\tau] : 0 \overset{-,t}{\longleftrightarrow} \infty) = 0,$$

and then by countable additivity

$$\Psi^{+,\beta}(\exists t \ge 0 : 0 \overset{-,t}{\longleftrightarrow} \infty) = 0.$$

It is well known (see [8]) that if all sites in $\Lambda_{n+1} \setminus \Lambda_n$ takes the value $+1$,

$$\begin{aligned}&\mathcal{E}_{L,\tau} &(35)\\ &\quad \subseteq \{\exists \gamma \subseteq E_n^{dual}, t \in [0,\tau] : |\gamma| \ge L,\ \gamma \text{ surrounds the origin}, Y_t(\gamma) \equiv 1\} \\ &\quad \subseteq \{\exists \gamma \subseteq E_n^{dual} : |\gamma| \ge L,\ \gamma \text{ surrounds the origin}, Y_{\sup,\tau}(\gamma) \equiv 1\}.\end{aligned}$$

To prove $\Psi_n^{+,\beta}(\mathcal{E}_{L,\tau}) < \delta$, consider γ with $|\gamma| = l$ a contour in E_n^{dual} surrounding the origin. By Lemma 8.1, $\mathbf{P}(Y_{\sup,\tau}(\gamma') \equiv 1 | Y_0(\gamma') \equiv 0, Y_0(\gamma \setminus \gamma') \equiv 1) \leq \epsilon^{|\gamma'|}$ whenever $\gamma' \subseteq \gamma$. We get

$$
\begin{aligned}
&\mathbf{P}(Y_{\sup,\tau}(\gamma) \equiv 1) \qquad (36)\\
&= \sum_{k=0}^{l} \sum_{\substack{\gamma' \subseteq \gamma \\ |\gamma'|=k}} \mathbf{P}(Y_0(\gamma') \equiv 0, Y_0(\gamma \setminus \gamma') \equiv 1) \\
&\qquad \times \mathbf{P}(Y_{\sup,\tau}(\gamma') \equiv 1 | Y_0(\gamma') \equiv 0, Y_0(\gamma \setminus \gamma') \equiv 1) \\
&\leq \sum_{k=0}^{l} \sum_{\substack{\gamma' \subseteq \gamma \\ |\gamma'|=k}} \mathbf{P}(Y_0(\gamma') \equiv 0, Y_0(\gamma \setminus \gamma') \equiv 1) \epsilon^k \\
&= \sum_{k=0}^{l} \mathbf{P}(\{\text{all edges except k of } \gamma \text{ are present at } t = 0\}) \epsilon^k \\
&= \sum_{k=0}^{l/N} \mathbf{P}(\{\text{all edges except k of } \gamma \text{ are present at } t = 0\}) \epsilon^k \\
&+ \sum_{k=l/N+1}^{l} \mathbf{P}(\{\text{all edges except k of } \gamma \text{ are present at } t = 0\}) \epsilon^k.
\end{aligned}
$$

Obviously, l/N need not be an integer, but correcting for this is trivial and is left for the reader.

We need to estimate $\mathbf{P}(\{\text{all edges except k of } \gamma \text{ are present at } t = 0\})$. For this purpose, define T: $\{-1,1\}^{S_n} \to \{-1,1\}^{S_n}$, by

$$
(T\sigma)(s) = \begin{cases} \sigma(s) & \text{if } \text{ s is not in the domain bounded by } \gamma \\ -\sigma(s) & \text{if } \text{ s is in the domain bounded by } \gamma \end{cases}
$$

for all $\sigma \in \{-1,1\}^{S_n}$. Let $E_k = \{\sigma : \text{ all edges except k of } \gamma \text{ are present}\}$. Since $H_n^{+,\beta}$ of (6) gives a contribution of $-\beta$ for adjacent pairs of equal spin and $+\beta$ for adjacent pairs of unequal spin, we have that for $\sigma \in E_k$, $H_n^{+,\beta}(T\sigma) = H_n^{+,\beta}(\sigma) - 2\beta(|\gamma| - k) + 2\beta k = H_n^{+,\beta}(\sigma) - 2\beta|\gamma| + 4\beta k$.

Hence, for $\sigma \in E_k$

$$
\mu_n^{+,\beta}(\sigma) = \frac{e^{-H_n^{+,\beta}(\sigma)}}{Z} = \frac{e^{-H_n^{+,\beta}(T\sigma) - 2\beta|\gamma| + 4\beta k}}{Z},
$$

and so

$$
\begin{aligned}
&\mu_n^{+,\beta}(E_k) \\
&\quad = \sum_{\sigma \in E_k} \mu_n^{+,\beta}(\sigma) = e^{-2\beta l + 4\beta k} \sum_{\sigma \in E_k} \frac{e^{-H_n^{+,\beta}(T\sigma)}}{Z}
\end{aligned}
$$

$$\leq e^{-2\beta l+4\beta k} \sum_{\sigma \in \{-1,1\}^{S_n}} \frac{e^{-H_n^{+,\beta}(T\sigma)}}{Z} = e^{-2\beta l+4\beta k},$$

where the last equality follows from T being bijective. We then get that

$$\sum_{k=0}^{l/N} \mathbf{P}(\{\text{all edges except k of } \gamma \text{ are present at } t=0\})\epsilon^k \tag{37}$$
$$\leq \sum_{k=0}^{l/N} e^{-2\beta l+4\beta k}\epsilon^k \leq e^{-2\beta l+\frac{4\beta l}{N}} \sum_{k=0}^{l/N} \epsilon^k \leq 2e^{-2\beta l+\frac{4\beta l}{N}}$$
$$\leq 2e^{-\beta(2-\delta_1)l} = 2e^{-2\beta' l}.$$

Furthermore

$$\sum_{k=l/N+1}^{l} \mathbf{P}(\{\text{all edges except k of } \gamma \text{ are present at } t=0\})\epsilon^k \tag{38}$$
$$\leq \epsilon^{l/N} \sum_{k=l/N+1}^{l} \mathbf{P}(\{\text{all edges except k of } \gamma \text{ are present at } t=0\})$$
$$\leq \epsilon^{l/N} \leq e^{-\beta(2-\delta_1)l} = e^{-2\beta' l},$$

where we use that $\{\text{all edges except k of } \gamma \text{ are present at } t=0\}$ are disjoint events for different k. Hence (36), (37) and (38) combined gives us

$$\mathbf{P}(Y_{\sup,\tau}(\gamma) \equiv 1) \leq 3e^{-2\beta' l}$$

and so by (35), for all $n > L$,

$$\Psi_n^{+,\beta}(\mathcal{E}_{L,\tau})$$
$$\leq \Psi_n^{+,\beta}(\exists \gamma \subseteq E_n^{dual} : |\gamma| \geq L,\ \gamma \text{ surrounds the origin}, Y_{\sup,\tau}(\gamma) \equiv 1)$$
$$\leq \sum_{l=L}^{\infty} l3^{l-1} 3e^{-2\beta' l} < \delta,$$

where the second to last inequality follows from the fact that the number of contours around the origin of length l is at most $l3^{l-1}$, (see [8]).

QED

Remark: For $\mathbb{Z}^d$, the proof is generalized by noting that the number of connected surfaces of size l surrounding the origin is at most $C(d)^l$, for some constant $C(d)$. The arguments are the same but the "topological details" are messier.

9 Proof of Theorem 1.5

We will start this subsection by presenting a theorem by T.M. Liggett, R.H. Schonmann and A.M. Stacey ([21]).

Theorem 9.1 *Let G=(S,E) be a graph with a countable set of sites in which every site has degree at most* $\Delta \geq 1$, *and in which every finite connected component of* G *contains a site of degree strictly less than* Δ. *Let* $p, \alpha, r \in [0,1], q = 1-p$, *and suppose that*

$$\begin{aligned} (1-\alpha)(1-r)^{\Delta-1} &\geq q, \\ (1-\alpha)\alpha^{\Delta-1} &\geq q. \end{aligned}$$

If $\mu \in G(p)$, *then* $\pi_{\alpha r} \preceq \mu$. *In particular, if* $q \leq (\Delta-1)^{\Delta-1}/\Delta^{\Delta}$, *then* $\pi_\rho \preceq \mu$, *where*

$$\rho = \left(1 - \frac{q^{1/\Delta}}{(\Delta-1)^{(\Delta-1)/\Delta}}\right)(1-(q(\Delta-1))^{1/\Delta}).$$

Here $G(p)$ denotes the set of probability measures on $\{-1,1\}^S$ such that if $\mu \in G(p)$, $X \sim \mu$ then for any site $s \in S$

$$\mathbb{P}[X(s) = 1 | \sigma(\{X(t) : \{s,t\} \notin E\})] \geq p \text{ a.s.}$$

Observe that when $p \to 1 \Rightarrow q \to 0$ and so $\rho \to 1$. The above theorem is stated as the original in [21]. However, by considering the line-graph of $G = (S,E)$, it can be restated in the following way.

Corollary 9.2 *Let* $\tilde{G} = (\tilde{S}, \tilde{E})$ *be any countable graph of degree at most* Δ. *For each* $0 < \rho < 1$ *there exists a* $0 < p < 1$ *where* $p = p(\Delta, \rho)$ *such that if* $Y \sim \nu$ *where* ν *is a probability measure on the edges of* $\tilde{G}$ *such that for every edge* $e \in \tilde{E}$

$$\mathbb{P}[Y(e) = 1 | \sigma(\{Y(f) : e \not\sim f\})] \geq p \text{ a.s.}$$

we have that $\pi_\rho^{\tilde{E}} \preceq \nu$.

By $e \not\sim f$ we of course mean that the edges e and f does not have any endpoints in common. Here, $\pi_\rho^{\tilde{E}}$ is the product measure with density ρ on the edges of $\tilde{G}$.

Consider a graph $G = (S,E)$ and a subgraph $G' = (S', E')$ where $S' = S$ and $E' \subset E$. Let $X \sim \pi_p$ on S. We declare an edge $e \in E'$ to be closed if any of the endpoints takes the value 0 under X. Corollary 9.2 gives us that for any $\rho < 1$ there is a $p < 1$ such that this method of closing edges dominates independent bond percolation with density ρ on E'. Observe that we can choose p independent of E' since the maximal degree of E' is bounded above by the maximal degree of E.

Let $(X, Y) \sim \mathbf{P}_n^p$, defined in Section 2.5. Close every $e \in E_n$ such that $Y(e) = 1$ independently with probability ϵ thus creating $(X, Y^{(-,\epsilon)})$. Compare this to closing every site in S_n independently with parameter ϵ' (creating $X^{(-,\epsilon')}$) and defining

$$Y^{\epsilon'}(e) = \begin{cases} 1 & \text{if } Y(e) = 1 \text{ and neither one of the endpoints of } e \text{ flips} \\ 0 & \text{otherwise.} \end{cases}$$

By the arguments of the last paragraph we see that for a fixed ϵ there exists an ϵ' (that we can choose independent of (X, Y) and n) such that the first way (i.e. independent bond percolation) of removing edges is stochastically dominated by the latter. Hence

$$\begin{aligned} &\mathbf{P}_n^p((X, Y^{(-,\epsilon)}) \in (\{-1,1\}^{S_n}, \cdot) | (X,Y)) \\ &\quad \preceq \mathbf{P}_n^p((X^{(-,\epsilon')}, Y^{\epsilon'}) \in (\{-1,1\}^{S_n}, \cdot) | (X,Y)). \end{aligned}$$

By averaging over all possible (X, Y), the next lemma follows.

Lemma 9.3 *With notation as above, for any $\epsilon > 0$ there exists $\epsilon' > 0$ independent of n such that*

$$\mathbf{P}_n^p((X, Y^{(-,\epsilon)}) \in (\{-1,1\}^{S_n}, \cdot)) \preceq \mathbf{P}_n^p((X^{(-,\epsilon')}, Y^{\epsilon'}) \in (\{-1,1\}^{S_n}, \cdot)).$$

Observe that

$$\mathbf{P}_n^p((X, Y^{(-,\epsilon)}) \in (\{-1,1\}^{S_n}, \cdot)) =_{\mathcal{D}} \tilde{\nu}_n^{p,(-,\epsilon)}(\cdot) \tag{39}$$

and that

$$\mathbf{P}_n^p((X^{(-,\epsilon')}, Y^{\epsilon'}) \in (\cdot, \{-1,1\}^{E_n})) =_{\mathcal{D}} \mu_n^{+,\beta,(-,\epsilon')}(\cdot). \tag{40}$$

We are now ready to prove Theorem 1.5.

Proof of Theorem 1.5. For any choice of $\beta > \beta_c$ take $p = 1 - e^{-2\beta}$ and let $\delta \in (0, p - p_c)$. Now, (14) and Holley's inequality implies that

$$\tilde{\nu}_n^{p-\delta} \preceq \tilde{\nu}_n^p \ \forall n \in \mathbb{N}^+.$$

Since by (14) both $\tilde{\nu}_n^{p-\delta}$ and $\tilde{\nu}_n^p$ are monotone, there exists by Lemma 3.3 (it is easy to check that all other conditions of that lemma are satisfied) an $\epsilon > 0$ such that

$$\tilde{\nu}_n^{p-\delta} \preceq \tilde{\nu}_n^{p,(-,\epsilon)} \ \forall n \in \mathbb{N}^+. \tag{41}$$

In [13] they show that the limit $\lim_n \tilde{\nu}_n^{p-\delta}(0 \longleftrightarrow \partial\Lambda_n)$ exists and that

$$\lim_n \tilde{\nu}_n^{p-\delta}(0 \longleftrightarrow \partial\Lambda_n) > 0. \tag{42}$$

Here $\{0 \longleftrightarrow \partial\Lambda_n\}$ denotes the event that there exists a path of present edges connecting the origin to $\partial\Lambda_n := \Lambda_{n+1} \setminus \Lambda_n$. Since $\{0 \longleftrightarrow \partial\Lambda_n\}$ is an increasing event on the edges, Lemma 9.3 guarantees the existence of an $\epsilon' > 0$ such that

$$\begin{aligned}
&\tilde{\nu}_n^{p,(-,\epsilon)}(0 \longleftrightarrow \partial\Lambda_n) \\
&\quad = \mathbf{P}_n^p((X, Y^{(-,\epsilon)}) \in (\{-1,1\}^{S_n}, 0, \longleftrightarrow \partial\Lambda_n)) \\
&\quad \le \mathbf{P}_n^p((X^{(-,\epsilon')}, Y^{\epsilon'}) \in (\{-1,1\}^{S_n}, 0 \longleftrightarrow \partial\Lambda_n)) \ \forall n \in \mathbb{N}^+.
\end{aligned}$$

If there exists a path of present edges connecting the origin to the boundary $\partial\Lambda_n$ under Y, all the sites of this path must have the value 1 under X. Similarly for $(X^{(-,\epsilon')}, Y^{\epsilon'})$, if there exists a path of present edges connecting the origin to the boundary $\partial\Lambda_n$ under $Y^{\epsilon'}$, all the sites of this path must have the value 1 under $X^{(-,\epsilon')}$. Hence

$$\begin{aligned}
&\mathbf{P}_n^p((X^{(-,\epsilon')}, Y^{\epsilon'}) \in (\{-1,1\}^{S_n}, 0 \longleftrightarrow \partial\Lambda_n)) \\
&\quad = \mathbf{P}_n^p((X^{(-,\epsilon')}, Y^{\epsilon'}) \in (0 \overset{+}{\longleftrightarrow} \partial\Lambda_n, 0 \longleftrightarrow \partial\Lambda_n)) \\
&\quad \le \mathbf{P}_n^p((X^{(-,\epsilon')}, Y^{\epsilon'}) \in (0 \overset{+}{\longleftrightarrow} \partial\Lambda_n, \{0,1\}^{E_n})) \\
&\quad = \mu_n^{+,\beta,(-,\epsilon')}(0 \overset{+}{\longleftrightarrow} \partial\Lambda_n).
\end{aligned}$$

Of course

$$\mu_n^{+,\beta,(-,\epsilon')}(0 \overset{+}{\longleftrightarrow} \partial\Lambda_n) \le \mu_n^{+,\beta,(-,\epsilon')}(0 \overset{+}{\longleftrightarrow} \partial\Lambda_L) \ \forall L < n.$$

Therefore, for any L we have that

$$\begin{aligned}
0 &< \lim_n \tilde{\nu}_n^{p-\delta}(0 \longleftrightarrow \partial\Lambda_n) \\
&\le \lim_n \mu_n^{+,\beta,(-,\epsilon')}(0 \overset{+}{\longleftrightarrow} \partial\Lambda_L) = \mu^{+,\beta,(-,\epsilon')}(0 \overset{+}{\longleftrightarrow} \partial\Lambda_L),
\end{aligned}$$

and so

$$0 < \lim_L \mu^{+,\beta,(-,\epsilon')}(0 \overset{+}{\longleftrightarrow} \partial\Lambda_L) = \mu^{+,\beta,(-,\epsilon')}(0 \overset{+}{\longleftrightarrow} \infty).$$

The limit in L exists since $\{0 \overset{+}{\longleftrightarrow} \partial\Lambda_{L_2}\} \subseteq \{0 \overset{+}{\longleftrightarrow} \partial\Lambda_{L_1}\}$ for $L_1 \le L_2$. Since $\mu^{+,\beta}$ is ergodic (see [19] page 143 and 195) it follows that $\mu^{+,\beta,(-,\epsilon')}$ must also be ergodic. This is because $\mu^{+,\beta,(-,\epsilon')}$ can be expressed as a function of two independent processes, one being $\mu^{+,\beta}$ and the other a product measure. We conclude that

$$\mu^{+,\beta,(-,\epsilon')}(\mathcal{C}^+) = 1. \tag{43}$$

By Lemma 5.1, there exists a $\tau > 0$ such that

$$\mu^{+,\beta,(-,\epsilon')} \preceq \mu^{+,\beta}_{\text{inf},\tau}$$

and therefore

$$\mu^{+,\beta}_{\inf,\tau}(\mathcal{C}^+) = 1.$$

Therefore

$$\Psi^{+,\beta}(\mathcal{C}_t^+ \text{ occurs for every } t \in [0,\tau]) = 1.$$

Finally using countable additivity

$$\Psi^{+,\beta}(\mathcal{C}_t^+ \text{ occurs for every t}) = 1.$$

QED

10 Proof of Theorem 1.4

The aim of this section is to prove Theorem 1.4. For that we will use Theorem 1.5 and Lemma 10.1. We will not prove Lemma 10.1 since it follows immediately from the proof of Lemma 11.12 in [10] due to Y. Zhang.

A probability measure μ on $\{-1,1\}^S$ is said to have the finite energy property if all conditional probabilities on finite sets are strictly positive.

Lemma 10.1 *Take μ to be any probability measure on $\{-1,1\}^{\mathbb{Z}^2}$ which has positive correlations and the finite energy property. Assume further that μ is invariant under translations, rotations and reflections in the coordinate axes. If $\mu(\mathcal{C}^+) = 1$, then $\mu(\mathcal{C}^-) = 0$.*

Proof of Theorem 1.4. Fix $\beta > \beta_c$. By (43), there exists $\epsilon > 0$ such that

$$\mu^{+,\beta,(-,\epsilon)}(\mathcal{C}^+) = 1.$$

Since $\mu^{+,\beta}$ and $\pi_{1-\epsilon}$ both have positive correlations, it follows that $\mu^{+,\beta,(-,\epsilon)}$ has positive correlations. This is because (see [19], page 78) the product of two probability measures which have positive correlations also has positive correlations. Furthermore, a collection of increasing functions of random variables which have positive correlations also has positive correlations. In addition, the finite energy property is easily seen to hold for $\mu^{+,\beta,(-,\epsilon)}$. Using this we can by Lemma 10.1 conclude that

$$\mu^{+,\beta,(-,\epsilon)}(\mathcal{C}^-) = 0.$$

By Lemma 5.1 there exists a $\tau > 0$ such that $\mu^{+,\beta,(-,\epsilon)} \preceq \mu^{+,\beta}_{\inf,\tau}$ and hence

$$\mu^{+,\beta}_{\inf,\tau}(\mathcal{C}^-) = 0.$$

It follows that

$$\Psi^{+,\beta}(\exists t \in [0,\tau] : \mathcal{C}_t^- \text{ occurs}) = 0,$$

and by countable additivity, we conclude

$$\Psi^{+,\beta}(\exists t \geq 0 : \mathcal{C}_t^- \text{ occurs}) = 0.$$

QED

Acknowledgement: We thank the referee for a very careful reading and for providing a number of suggestions.

References

[1] M. Aizenman, J. Bricmont and J. L. Lebowitz (1987) Percolation of the Minority Spins in High-Dimensional Ising Models. *J. Stat. Phys.* **49**, 859.

[2] J. van den Berg, R. Meester, D. G. White (1997) Dynamic Boolean models. *Stochastic Process. Appl.* **69** no. 2, 247–257.

[3] J. Bricmont, J. L. Lebowitz and C. Maes (1987) Percolation in strongly correlated systems: the massless Gaussian field. *J. Statist. Phys.* **48** no. 5-6, 1249–1268.

[4] E. I. Broman, O. Häggström and J. E. Steif, Refinements of stochastic domination, (2005), preprint.

[5] A. Coniglio, C. R. Nappi and F. Peruggi, L. Russo (1976) Percolation and phase transitions in the Ising model. *Comm. Math. Phys.* **51** no. 3, 315–323.

[6] R. Durrett (1988) Lecture notes on particle systems and percolation *The Wadsworth & Brooks/Cole Statistics/Probability Series. Wadsworth & Brooks/Cole Advanced Books & Software, Pacific Grove, CA, viii+335 pp.*

[7] C. M. Fortuin, P. W. Kasteleyn and J. Ginibre (1971) Correlation inequalities on some partially ordered sets *Commun. Math. Phys.* **22** 89-103.

[8] H.-O. Georgii (1988) Gibbs Measures and Phase Transitions *de Gruyter Studies in Mathematics* **9**.

[9] H.-O. Georgii, O. Häggström and C. Maes (2001) The random geometry of equilibrium phases *Phase Transitions and Critical Phenomena*, **18** (C. Domb and J.L. Lebowitz, eds), pp 1-142, Academic Press, London.

[10] G. Grimmett (1999) Percolation, *Second Edition Springer-Verlag.*

[11] O. Häggström. (1996) The random-cluster model on a homogeneous tree *Prob. Theory Relat. Fields* **104** 231-253.

[12] O. Häggström. (1997) Infinite clusters in dependent automorphism invariant percolation on trees *Annals of Probability* **25** No. 3, 1423-1436.

[13] O. Häggström. (1998) Random-Cluster Representations in the Study of Phase Transitions *Markov Processes and Related Fields* **4** 275-321.

[14] O. Häggström. (2000) Markov random fields and percolation on general graphs *Adv. Appl. Prob.* **32** 39-66.

[15] O. Häggström, Y. Peres and J. E. Steif (1997) Dynamical percolation *Annals Inst. Henri Poincare*, 497-528.

[16] Y. Higuchi (1993) Coexistence of infinite (∗)-clusters. II. Ising percolation in two dimensions. *Probab. Theory Related Fields* **97** 1–33.

[17] Y. Higuchi (1993) A sharp transition for the two-dimensional Ising percolation. *Probab. Theory Related Fields* **97** 489–514.

[18] R. Holley (1974) Remarks on the FKG inequalities *Commun. Math. Phys.* **36** 227-231.

[19] T. M. Liggett (1985) Interacting Particle Systems *Springer-Verlag.*

[20] T. M. Liggett (1994) Survival and Coexistence in Interacting Particle Systems *Probability and phase transition, Kluwer* 209-226.

[21] T. M. Liggett, R. H. Schonmann and A. M. Stacey (1997) Domination by product measures, *Annals of Probability*, **25** No. 1, 71-95.

[22] T. M. Liggett and J. E. Steif, Stochastic Domination: The Contact Process, Ising Models, FKG Measures and Exchangeable Processes, (2004), preprint.

[23] O. Schramm and J. E. Steif, Quantitative noise sensitivity and exceptional times for percolation, in preparation.

[24] P. C. Shields (1996) The Ergodic Theory of Discrete Sample Paths *Graduate Studies in Mathematics* **13**.

Erik I. Broman
Dept. of Mathematics
Chalmers University of Technology
412 96 Gothenburg
Sweden
`broman@math.chalmers.se`
`www.math.chalmers.se/~broman`

Jeffrey E. Steif
Dept. of Mathematics
Chalmers University of Technology

412 96 Gothenburg
Sweden
steif@math.chalmers.se
www.math.chalmers.se/~steif

Paper III

Refinements of stochastic domination

Erik I. Broman*, Olle Häggström†, and Jeffrey E. Steif‡

October 3, 2005

Abstract

In a recent paper by two of the authors, the concepts of upwards and downwards ϵ-movability were introduced, mainly as a technical tool for studying dynamical percolation of interacting particle systems. In this paper, we further explore these concepts which can be seen as refinements or quantifications of stochastic domination, and we relate them to previously studied concepts such as uniform insertion tolerance and extractability.

AMS subject classification: 60G99.

Keywords and phrases: finite energy, stochastic domination, extractability, rigidity

1 Introduction

In [3], certain refinements of stochastic domination were introduced, which we call upwards and downwards ϵ-movability; see Definition 1.1 below. These concepts were introduced mainly as a technical tool in the analysis of dynamical percolation for interacting particle systems, but they turn out to be interesting in their own right.

The purpose of the present paper is to relate them to other concepts that have arisen in a number of problems and that we feel belong to the same circle of ideas. These include finite energy [16] and insertion and deletion tolerance [13]; see Definition 1.5. Later, we also define the term *extractability*; see Definition 1.6. Although the term is our own, this concept does have a

*Dept. of Mathematics, Chalmers University of Technology, S-412 96 Göteborg, Sweden, `http://www.math.chalmers.se/~broman/`, research supported by the Swedish Research Council.

†Dept. of Mathematics, Chalmers University of Technology, S-412 96 Göteborg, Sweden, `http://www.math.chalmers.se/~olleh/`, research supported by the Swedish Research Council.

‡Dept. of Mathematics, Chalmers University of Technology, S-412 96 Göteborg, Sweden, `http://www.math.chalmers.se/~steif/`, research supported by the Swedish Research Council and the Göran Gustafsson Foundation (KVA).

history; in particular, there has been interest in finding lower bounds on ϵ for which ϵ-extractability holds. The question of uniform extractability has been studied for the Ising model as well as other Markov random fields in [1, 8, 15]. Earlier, in [5, 6, 7], a similar question was studied for Markov chains and autoregressive processes. Of related interest is the result in [8] that for Markov random fields, uniform finite energy implies uniform extractability.

Let S be a countable set. For $p \in [0,1]$, let $\pi_p = \prod_{s\in S} p\delta_1 + (1-p)\delta_0$ be the standard product measure with density p. When talking about product measures on $\{0,1\}^S$, we will always mean these uniform ones (with the same p for every $s \in S$).

Let μ be an arbitrary probability measure on $\{0,1\}^S$. For $\epsilon \in (0,1)$, we will let $\mu^{(+,\epsilon)}$ denote the distribution of the process obtained by first choosing an element of $\{0,1\}^S$ according to μ and then independently changing each 0 to a 1 with probability ϵ. Similarly, we will let $\mu^{(-,\epsilon)}$ denote the distribution of the process obtained by first choosing an element of $\{0,1\}^S$ according to μ and then independently changing each 1 to a 0 with probability ϵ. Finally, for $\delta \in (0,1)$, we let $\mu^{(-,\epsilon,+,\delta)}$ denote the distribution of the process obtained by first choosing an element of $\{0,1\}^S$ according to μ and then independently changing each 0 to a 1 with probability δ and each 1 to a 0 with probability ϵ.

It is elementary to check that for any $\epsilon \in [0,1)$, $\mu_1^{(+,\epsilon)} = \mu_2^{(+,\epsilon)}$ or $\mu_1^{(-,\epsilon)} = \mu_2^{(-,\epsilon)}$ implies that $\mu_1 = \mu_2$.

For $\sigma, \sigma' \in \{0,1\}^S$ we write $\sigma \preceq \sigma'$ if $\sigma(s) \le \sigma'(s)$ for every $s \in S$. A function $f : \{0,1\}^S \to \mathbb{R}$ is increasing if $f(\sigma) \le f(\sigma')$ whenever $\sigma \preceq \sigma'$. For two probability measures μ, μ' on $\{0,1\}^S$, we say that μ **is stochastically dominated by** μ', and write $\mu \preceq \mu'$, if for every continuous increasing function f we have that $\mu(f) \le \mu'(f)$. ($\mu(f)$ is shorthand for $\int f d\mu$.) By Strassens theorem (see [9, p. 72]), this is equivalent to the existence of random variables $X, X' \in \{0,1\}^S$ such that $X \sim \mu$, $X' \sim \mu'$, and $X \preceq X'$ a.s.; here and throughout, "$\sim$" means "has distribution".

For a probability measure μ on $\{0,1\}^S$, define $p_{\max,\mu}$ by

$$p_{\max,\mu} := \sup\{p \in [0,1] : \pi_p \preceq \mu\};$$

the supremum is easily seen to be achieved.

Definition 1.1 *Let (μ_1, μ_2) be a pair of probability measures on $\{0,1\}^S$, where S is a countable set. Assume that $\mu_1 \preceq \mu_2$. If, given $\epsilon > 0$, we have*

$$\mu_1 \preceq \mu_2^{(-,\epsilon)},$$

then we say that the pair (μ_1, μ_2) is **downwards ϵ-movable**. *(μ_1, μ_2) is said to be* **downwards movable** *if it is downwards ϵ-movable for some $\epsilon > 0$. Analogously, if, given $\epsilon > 0$, we have*

$$\mu_1^{(+,\epsilon)} \preceq \mu_2,$$

then we say that the pair (μ_1, μ_2) is **upwards ϵ-movable**, *and we say that (μ_1, μ_2) is* **upwards movable** *if the pair is upwards ϵ-movable for some $\epsilon > 0$.*

Note that if we restrict to the case where both μ_1 and μ_2 are product measures, then these concepts become trivial.

Following is a natural example where a stochastically ordered pair of probability measures is neither downwards nor upwards movable. We assume that the reader is familiar with the Ising model; for a definition and survey, see, e.g., [4] or [9].

Example 1.2 Let $\mu^{+,\beta}$ and $\mu^{-,\beta}$ be the plus and minus states for the Ising model on $\mathbb{Z}^d$ with zero external field at inverse temperature $\beta > 0$. It is well known that $\mu^{-,\beta} \preceq \mu^{+,\beta}$, and it is known (see [12]) that $p_{\max,\mu^{+,\beta}} = p_{\max,\mu^{-,\beta}}$. Assume that the pair $(\mu^{-,\beta}, \mu^{+,\beta})$ is upwards ϵ-movable for some $\epsilon > 0$. It then follows that

$$(\pi_{p_{\max,\mu^{+,\beta}}})^{(+,\epsilon)} = (\pi_{p_{\max,\mu^{-,\beta}}})^{(+,\epsilon)} \preceq (\mu^{-,\beta})^{(+,\epsilon)} \preceq \mu^{+,\beta},$$

which of course contradicts the definition of $p_{\max,\mu^{+,\beta}}$. Therefore the pair is not upwards movable. By symmetry of the model, it is not downwards movable either. $\square$

Next, we provide an easy example of a pair of measures which is downwards but not upwards movable.

Example 1.3 Let $\nu = \frac{1}{2}\pi_q + \frac{1}{2}\delta_0$ and $\mu = \frac{1}{2}\pi_p + \frac{1}{2}\delta_0$ where $q < p$ and where δ_0 is the measure which puts probability 1 on the configuration of all zeros. It is trivial to check that when $|S| = \infty$, (ν, μ) is downwards but not upwards movable. $\square$

In [3] a considerable amount of effort was spent on trying to show downwards movability when the pair considered was two stationary distributions, corresponding to two different parameter values, for some specific interacting particle system. In particular, the so called contact process (see [10] for definitions and a survey) was investigated. Considering (μ_1, μ_2), where μ_i is the upper invariant measure for the contact process with infection rate λ_i, it was shown in [3] that if $\lambda_1 < \lambda_2$, then the pair is downwards movable.

Another result from [3] is that if $\mu_1 \preceq \mu_2$, μ_2 satisfies the FKG lattice condition (see [9, p. 78]) and

$$\inf_{\substack{\tilde{S} \subset S \\ |\tilde{S}| < \infty}} \inf_{\substack{s \in \tilde{S} \\ \xi \in \{0,1\}^{\tilde{S} \setminus s}}} [\mu_2(\sigma(s) = 1 | \sigma(\tilde{S} \setminus s) \equiv \xi) - \mu_1(\sigma(s) = 1 | \sigma(\tilde{S} \setminus s) \equiv \xi)] > 0$$

then (μ_1, μ_2) is downwards movable. This however is not sufficient to get the result for the contact process mentioned above since by [11], the upper invariant measure for the contact process on $\mathbb{Z}$ does not satisfy the FKG lattice condition when $\lambda < 2$.

In the present paper, we will concentrate on the case where μ_1 is a product measure but μ_2 is not. We now proceed with some further explanations and definitions needed to state our main results, Theorems 1.7 and 1.10 below.

If $p_{\max,\mu} = 0$, then trivially $(\pi_{p_{\max,\mu}}, \mu)$ is downwards movable but not upwards movable. Assume next that μ is a probability measure with $p_{\max,\mu} > 0$. If $p \in [0, p_{\max,\mu})$, then the pair (π_p, μ) is trivially upwards movable. It is also easy to see that it is downwards movable by arguing as follows. By Strassen's theorem, we may choose $X \sim \mu$ and $Y \sim \pi_{p_{\max,\mu}}$ such that $X \geq Y$ a.s. Then choose $\epsilon > 0$ such that $(1-\epsilon)p_{\max,\mu} > p$, and let $Z \sim \pi_{1-\epsilon}$ be independent of both X and Y. We obtain $\min(X, Z) \geq \min(Y, Z)$ a.s., and since $\min(Y, Z) \sim \pi_{p_{\max,\mu}(1-\epsilon)}$ we conclude that $\pi_p \preceq \mu^{(-,\epsilon)}$, as desired.

The final case we are left with (when one of the measures is a uniform product measure) is $(\pi_{p_{\max,\mu}}, \mu)$ with $p_{\max,\mu} > 0$. This pair is by definition not upwards movable, but we believe it is an interesting question to ask if it is downwards movable and this question motivates the following definition.

Definition 1.4 *We say that μ is* **nonrigid** *if the pair $(\pi_{p_{\max,\mu}}, \mu)$ is downwards movable and otherwise we will say that μ is* **rigid**.

All uniform product measures other than δ_0 are trivially rigid while all μ such that $p_{\max,\mu} = 0$ are trivially nonrigid. Heuristically, it is natural to expect that as long as $p_{\max,\mu} > 0$, then typically μ should be rigid. This issue turns out to be quite intricate, however; see Proposition 3.1 and Theorem 3.2 below.

Definition 1.5 *We say that μ is* **ϵ-insertion tolerant** *if for any $s \in S$, we have that*

$$\mu(\sigma(s) = 1 | \sigma(S \setminus s)) \geq \epsilon \text{ a.s.} \tag{1}$$

We say that μ is **uniformly insertion tolerant** *if it is ϵ-insertion tolerant for some $\epsilon > 0$. The analogous notions of* **ϵ-deletion tolerant** *and* **uniformly deletion tolerant** *are defined similarly (the "1" is replaced by "0"). Finally, we say μ has* **finite ϵ-energy** *if it is both ϵ-insertion tolerant and ϵ-deletion tolerant, and that it has* **uniform finite energy** *if it has finite ϵ-energy for some $\epsilon > 0$.*

Closely related are the following notions of extractability.

Definition 1.6 *We call μ* **ϵ-upwards extractable** *if there exists a probability measure ν such that $\mu = \nu^{(+,\epsilon)}$. We call μ* **uniformly upwards extractable** *if it is ϵ-upwards extractable for some $\epsilon > 0$. The notions of* **ϵ-downwards extractable** *and* **uniformly downwards extractable** *are defined analogously (the "+" is replaced by "−"). Finally, μ is called* **ϵ-extractable** *if there exists a probability measure ν such that $\mu = \nu^{(-,\epsilon,+,\epsilon)}$, and it is called* **uniformly extractable** *if it is ϵ-extractable for some $\epsilon > 0$.*

We are now equipped with all the definitions needed to state our main theorem. We refer to Figure 1 for a comprehensive diagram over the implications and non-implications that the theorem asserts.

Theorem 1.7 *Let S be a countable set and consider the following properties of a probability measure μ on $\{0,1\}^S$:*

(I) *μ is uniformly upwards extractable.*

(II) *μ is uniformly insertion tolerant.*

(III) *μ is rigid.*

(IV) *There exists a $p > 0$ such that $\pi_p \preceq \mu$.*

We then have that (I) $\Rightarrow$ (II) $\Rightarrow$ (IV) *and that* (I) $\Rightarrow$ (III) $\Rightarrow$ (IV) *while none of the four corresponding reverse implications hold. Also,* (III) *does not imply* (II). *Moreover, with $S = \mathbb{Z}$, there exist translation invariant examples for all of the asserted nonimplications.*

In addition, it turns out that (IV) does not even imply "(II) or (III)"; see Remark 3.4. Note that we have not managed to work out whether or not (II) implies (III).

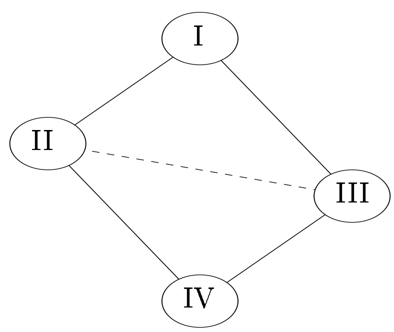

Figure 1. Hasse diagram of the implications between properties (I), (II), (III) and (IV) in Theorem 1.7: we have proved that one property implies another iff there is a downwards path in the diagram from the former to the latter. We do not know whether the dashed line between (II) and (III) should be there or not, i.e., whether or not uniform insertion tolerance implies rigidity. As will be seen in Theorem 1.9, the desired implication (II) $\Rightarrow$ (III) holds under an additional FKG-like assumption. If we restrict to finite S, then some of the implications will turn into equivalences; see Theorem 1.10.

Some of the asserted implications are easy: (I) trivially implies (II). The implication (III) $\Rightarrow$ (IV) is also trivial as we saw. It is a direct application of Holley's inequality (see, e.g., [4, Theorem 4.8]) to see that ϵ-insertion tolerance implies that $\pi_\epsilon \preceq \mu$, whence (II) implies (IV). Thus, apart from the implication (I) $\Rightarrow$ (III) (which is in fact not so hard either), we see all the implications claimed in the theorem. Therefore our interest in Theorem 1.7 is more in the counterexamples showing the distinction between some of these properties rather than in the implications.

As mentioned above, we do not know in general whether (II) implies (III). However Theorem 1.9 provides us with a partial answer, telling us that this is true under the extra assumption of μ being downwards FKG, a property weaker than satisfying the FKG lattice condition and defined as follows.

Definition 1.8 *A measure μ on $\{0,1\}^S$ is downwards FKG if for any finite $S' \subset S$ and any increasing subsets A, B*

$$\mu(A \cap B | \sigma(S') \equiv 0) \geq \mu(A | \sigma(S') \equiv 0) \mu(B | \sigma(S') \equiv 0).$$

The concept of downwards FKG was made explicit in [12], and was further studied in [2], where it was proved that the upper invariant measure for the contact process is downwards FKG.

Theorem 1.9 *Let μ be a translation invariant downwards FKG measure on $\{0,1\}^{\mathbb{Z}^d}$. Then* (II) *implies* (III).

Of course, some of the nonimplications in Theorem 1.7 can become implications under appropriate auxiliary assumptions. For instance, for probability measures μ on $\{0,1\}^{\mathbb{Z}}$ satisfying translation invariance and conditional negative association (see [14] for a definition of the latter), (IV) implies (III). This follows readily from results in [14]; we omit the proof.

Another situation in which further implications between the various properties arise, is when S is taken to be finite. By the support of a measure μ on $\{0,1\}^S$, denoted $\mathrm{supp}(\mu)$, we mean $\{\xi \in \{0,1\}^S : \mu(\sigma(S) \equiv \xi) > 0\}$.

Theorem 1.10 *Let S be finite, and consider properties* (I)–(IV) *of probability measures on $\{0,1\}^S$. We then have*

$$\text{(I)} \Leftrightarrow \text{(II)} \Leftrightarrow \mathrm{supp}(\mu) \textit{ is an up-set,} \tag{2}$$

and

$$\text{(III)} \Leftrightarrow \text{(IV)} \Leftrightarrow \mu(\sigma(S) \equiv 1) > 0. \tag{3}$$

Consequently, the properties in (2) *imply those in* (3) *but not vice versa. Note in particular that if we are in the full support case, then* (I)–(IV) *all hold.*

The rest of this paper is organized as follows. In Sections 2–3, we will establish a number of auxiliary results, in Section 4, we prove Theorem 1.9 and in Section 5, we tie things together giving proofs of Theorems 1.7 and 1.10. Finally, in Section 6, we list some open problems.

2 Uniform insertion tolerance and upwards extractability

In this section we focus on uniform upwards extractability (property (I)) and uniform insertion tolerance (property (II)). Proposition 2.1 provides an equivalence between these properties when S is finite, while Theorem 2.2 exhibits a contrasting scenario for S countable.

Proposition 2.1 *If S is finite and μ is a probability measure on $\{0,1\}^S$, then the following are equivalent:*

(i) *uniform insertion tolerance,*

(ii) *uniform upwards extractability, and*

(iii) $\mathrm{supp}(\mu)$ *is an up-set.*

Theorem 2.2 *For S countably infinite, there exists a probability measure μ on $\{0,1\}^S$ that is uniformly insertion tolerant but not uniformly upwards extractable. Moreover, we can take μ to be a translation invariant measure on $\{0,1\}^{\mathbb{Z}}$.*

Proof of Proposition 2.1. (i) $\Rightarrow$ (iii) and (ii) $\Rightarrow$ (i) are immediate, and so it only remains to show (iii) $\Rightarrow$ (ii).

In what follows, given a configuration $\sigma \in \{0,1\}^S$, $|\sigma|$ will be the number of 1's in σ. If there is to exist a ν such that $\mu = \nu^{(+,\epsilon)}$ with $\epsilon \in [0,1)$, it is not hard to see that we must have

$$\nu(\sigma) = \sum_{\tilde{\sigma} \preceq \sigma} (-\epsilon)^{|\sigma|-|\tilde{\sigma}|}(1-\epsilon)^{|\tilde{\sigma}|-|S|}\mu(\tilde{\sigma}) \ \forall \sigma \in \{0,1\}^S. \tag{4}$$

This can be verified through a direct calculation, but it is easier to calculate $\nu^{(+,\epsilon)}(\sigma)$ and check that it is indeed equal to $\mu(\sigma)$, as follows.

$$\begin{aligned}
\nu^{(+,\epsilon)}(\sigma) &= \sum_{\sigma_1 \preceq \sigma} \epsilon^{|\sigma|-|\sigma_1|}(1-\epsilon)^{|S|-|\sigma|}\nu(\sigma_1) \\
&= \sum_{\sigma_1 \preceq \sigma} \epsilon^{|\sigma|-|\sigma_1|}(1-\epsilon)^{|S|-|\sigma|} \sum_{\sigma_2 \preceq \sigma_1} (-\epsilon)^{|\sigma_1|-|\sigma_2|}(1-\epsilon)^{|\sigma_2|-|S|}\mu(\sigma_2) \\
&= \sum_{\sigma_1 \preceq \sigma} \sum_{\sigma_2 \preceq \sigma_1} \epsilon^{|\sigma|-|\sigma_1|}(-\epsilon)^{|\sigma_1|-|\sigma_2|}(1-\epsilon)^{|\sigma_2|-|\sigma|}\mu(\sigma_2) \\
&= \sum_{\sigma_2} (1-\epsilon)^{|\sigma_2|-|\sigma|}\mu(\sigma_2) \sum_{\sigma_1 : \sigma_2 \preceq \sigma_1 \preceq \sigma} \epsilon^{|\sigma|-|\sigma_1|}(-\epsilon)^{|\sigma_1|-|\sigma_2|}.
\end{aligned}$$

If we fix σ_2, then the binomial theorem gives that the last summation is equal to 0 unless $\sigma_2 = \sigma$ in which case it is equal to 1. We therefore easily obtain that $\nu^{(+,\epsilon)}(\sigma) = \mu(\sigma)$ for every σ.

What remains is to check that $\nu(\sigma) \geq 0$ for all σ. From (4) it is immediate that $\nu(\sigma) = 0$ for every $\sigma \notin \mathrm{supp}(\mu)$ since $\mathrm{supp}(\mu)$ is an up-set. For $\sigma \in \mathrm{supp}(\mu)$ on the other hand, it is easy to see that if we do this construction for different ϵ's, then we get

$$\lim_{\epsilon \to 0} \nu(\sigma) = \mu(\sigma).$$

Since $\mu(\sigma) > 0$ for all $\sigma \in \mathrm{supp}(\mu)$ and $|S| < \infty$, for $\epsilon > 0$ small enough, we get that $\nu(\sigma) > 0$ for all $\sigma \in \mathrm{supp}(\mu)$. This shows that μ is ϵ-upwards extractable for all such ϵ. □

Proof of Theorem 2.2. Let $S = \cup_{k=2}^{\infty} S_k$, where

$$S_k = ((k,1),(k,2),\ldots,(k,k)).$$

We will take the probability measure μ on $\{0,1\}^S$ to be the product measure

$$\mu = \mu_2 \times \mu_3 \times \cdots \tag{5}$$

where each μ_k is a probability measure on $\{0,1\}^{S_k}$. The μ_k's are constructed as follows, drawing heavily on an example in [8]. For $\sigma \in \{0,1\}^{S_k}$, let

$$\mu_k(\sigma) = \begin{cases} \frac{4}{3}2^{-k} & \text{if the number of 1's in } \sigma \text{ is even} \\ \frac{2}{3}2^{-k} & \text{if the number of 1's in } \sigma \text{ is odd.} \end{cases} \tag{6}$$

We may think of μ_k as the distribution of a $\{0,1\}^{S_k}$-valued random variable X_k obtained by first tossing a biased coin with heads-probability $\frac{2}{3}$, and if heads pick the components of X_k i.i.d. $(\frac{1}{2},\frac{1}{2})$ conditioned on an even number of 1's, while if tails pick the components i.i.d. $(\frac{1}{2},\frac{1}{2})$ conditioned on an odd number of 1's. One can also check that this distribution is the same as choosing all but (an arbitrary) one of the variables according to $\pi_{1/2}$ and then taking the last variable to be 1 with probability $1/3$ $(2/3)$ if there are an even (odd) number of 1's in the other bits. This last description immediately implies that μ_k is $\frac{1}{3}$-insertion tolerant. Because of the product structure in (5), this property is inherited by μ, which therefore is uniformly insertion tolerant.

It remains to show that μ is not uniformly upwards extractable. To this end, let X be a $\{0,1\}^S$-valued random variable with distribution μ, and for $k = 2,3,\ldots$ let Y_k denote the number of 1's in $X(S_k)$. It is immediate from (6) that

$$\mathbb{P}(Y_k \text{ is even}) = \frac{2}{3} \tag{7}$$

for each k. Using our last description of μ_k, the weak law of large numbers implies that

$$\frac{Y_k}{k} \to \frac{1}{2} \text{ in probability as } k \to \infty.$$

Hence, in particular,

$$\lim_{k\to\infty} \mathbb{P}(Y_k \le k-m) \;=\; 1 \qquad (8)$$

for any fixed m.

Now assume (for contradiction) that $\mu = \nu^{(+,\epsilon)}$ for some fixed $\epsilon > 0$; since μ being ϵ_2-upwards extractable implies it is ϵ_1-upwards extractable for $\epsilon_1 < \epsilon_2$, we may without loss of generality assume that $\epsilon \le 1/3$. Pick X' according to ν; we may then suppose that X has been obtained from X' by randomly switching 0's to 1's independently with probability ϵ. The intuition behind the argument leading up to a contradiction is that the process of independently flipping 0's to 1's will cancel all preferences of ending up with an even number of 1's.

If $X'(S_k)$ contains precisely l 0's, then the conditional probability (given X') that an even number of these switch to 1's when going from X' to X is easily seen to equal

$$\tfrac{1}{2} + \tfrac{1}{2}(1-2\epsilon)^l\,.$$

The easiest way to see this is using an equivalent random mechanism where each 0 independently "updates" with probability 2ϵ and then all the sites which have updated then independently actually switch to a 1 with probability 1/2. It follows that the conditional probability (again given X') that Y_k is odd is at least

$$\min\{\tfrac{1}{2} + \tfrac{1}{2}(1-2\epsilon)^l, \tfrac{1}{2} - \tfrac{1}{2}(1-2\epsilon)^l\} \;=\; \tfrac{1}{2} - \tfrac{1}{2}(1-2\epsilon)^l.$$

Now pick m large enough so that $\frac{1}{2} - \frac{1}{2}(1-2\epsilon)^m > \frac{5}{12}$. Since $X' \preceq X$ a.s., we get from (8) that

$$\lim_{k\to\infty} \mathbb{P}(A_k) \;=\; 1$$

where A_k is the event that there are at least m 0's in $X'(S_k)$. This gives

$$\begin{aligned}\lim_{k\to\infty} \mathbb{P}(Y_k \text{ is odd}) &\ge \lim_{k\to\infty} \mathbb{P}(Y_k \text{ is odd} \mid A_k)\mathbb{P}(A_k)\\ &\ge (\tfrac{1}{2} - \tfrac{1}{2}(1-2\epsilon)^m) \lim_{k\to\infty} \mathbb{P}(A_k) > \tfrac{5}{12}\,.\end{aligned}$$

This clearly contradicts (7).

We now translate this example into the setting of translation invariant distributions on $\{0,1\}^{\mathbb{Z}}$.

Begin with randomly designating either all even integers or all odd integers (each with probability $\frac{1}{2}$) in the index set $\mathbb{Z}$ to represent copies of S_2. Assume that we happened to choose the even integers (the other case is handled analogously). Then we toss another fair coin to decide whether to put i.i.d. copies of $X(S_2)$ on the pairs $\{\ldots,(-4,-2),(0,2),(4,6),\ldots\}$ in $\mathbb{Z}$, or on $\{\ldots(-2,0),(2,4),(6,8),\ldots\}$. Then use one more fair coin to decide whether $\{\ldots-3,1,5,9,\ldots\}$ or $\{\ldots,-1,3,7,11,\ldots\}$ should be designated for i.i.d. copies of $X(S_3)$, and once this is decided toss a fair three-sided coin to choose one of the three possible placements of the length-3 blocks in this subsequence to put these copies. And so on.

This makes the resulting process X^* translation invariant. Also, since the property of ϵ-insertion tolerance is obviously closed under convex combinations, we easily obtain that X^* is $\frac{1}{3}$-insertion tolerant and therefore uniformly insertion tolerant.

Furthermore, for any $k \geq 2$, we may apply (7) to the i.i.d. copies of $X(S_k)$ to deduce that with probability 1 there will exist $i \in \{0, 1, \ldots, k2^{k-1} - 1\}$ such that

$$\lim_{n\to\infty} \frac{1}{n} \sum_{j=1}^{n} \mathbf{1}_{B_{i,j,k}} = \frac{2}{3} \tag{9}$$

where $B_{i,j,k}$ denotes the event that the number of 1's in

$$\{i+jk2^{k-1}, i+jk2^{k-1}+2^{k-1}, i+jk2^{k-1}+2\cdot 2^{k-1}, \ldots, i+jk2^{k-1}+(k-1)2^{k-1}\}$$

is even. The right way to think of i is that it is the first place to the right of the origin where a copy of $X(S_k)$ starts. The summation variable j on the other hand, makes us jump to the starting points of all the other copies of $X(S_k)$ to the right of the origin. Furthermore, by arguing as in for the non-translation invariant construction, we have that if X^* is uniformly upwards extractable, then for large k the limit in (9) will be less than $1 - \frac{5}{12} = \frac{7}{12}$ for all $i \in \{0, 1, \ldots, k2^{k-1} - 1\}$. But this contradicts (9), so we can conclude that X^* is not uniformly upwards extractable. □

Note, finally, that the examples in the above proof also show that uniform finite energy does not imply uniform extractability.

3 Rigidity

We now proceed to discuss the issue of when a measure is rigid. As mentioned in the introduction, any measure which does not dominate a nontrivial product measure is trivially nonrigid and so it would be more interesting to have a nonrigid measure which dominates a nontrivial product measure; such a measure is provided in Theorem 3.2 below.

Proposition 3.1 *If S is finite and μ is a probability measure on $\{0,1\}^S$, then the following are equivalent.*

(i) *μ dominates π_p for some $p > 0$,*

(ii) *μ is rigid, and*

(iii) *$\mu(\sigma(S) \equiv 1) > 0$.*

This does not extend to infinite S, as shown in the following result.

Theorem 3.2 *For S countably infinite, there exists a μ which dominates a nontrivial product measure π_p but is nevertheless nonrigid. Moreover, we can take μ to be a translation invariant measure on $\{0,1\}^{\mathbb{Z}}$.*

Proof of Proposition 3.1. It is easy to see that the condition that μ dominates π_p for some $p > 0$ is equivalent to the condition that $\mu(\sigma(S) \equiv 1) > 0$. Also, recall that if μ is rigid it must dominate a non-trivial product measure.

To make the proof complete, it only remains to show that (i) and (iii) of the statement imply that μ is rigid. We have $\pi_{p_{\max,\mu}} \preceq \mu$, so that

$$\pi_{p_{\max,\mu}}(A) \;\leq\; \mu(A) \tag{10}$$

for all increasing events $A \subseteq \{0,1\}^S$. We next claim that

$$\exists A \neq \emptyset, \{0,1\}^S \text{ such that } A \text{ is increasing and } \pi_{p_{\max,\mu}}(A) = \mu(A). \tag{11}$$

To see this, note that if we had strict inequality in (10) for all such nontrivial increasing events A, then we could find a sufficiently small $\delta > 0$ so that

$$\pi_{p_{\max,\mu}+\delta}(A) \;<\; \mu(A)$$

for all such A (this uses the finiteness of S), contradicting the definition of $p_{\max,\mu}$. Now, for such an A we have that $\mu(A) \geq \mu(\sigma(S) \equiv 1) > 0$ and hence for any $\epsilon > 0$

$$\mu^{(-,\epsilon)}(A) \;<\; \mu(A)$$

(again because S is finite), which in combination with (11) yields

$$\pi_{p_{\max,\mu}} \;\not\preceq\; \mu^{(-,\epsilon)}\,.$$

Since $\epsilon > 0$ was arbitrary, μ is rigid. □

It will be convenient for the proof of Theorem 3.2 to have the following lemma, whose elementary proof we omit.

Lemma 3.3 *For $k \geq 1$, $p \in (0,1)$ and $m \in \{0,1,\ldots,k\}$, write $\rho_{k,p,m}$ for the distribution of a Binomial(k,p) random variable conditioned on taking value at least m. For $p_1 \leq p_2$, we have*

$$\rho_{k,p_1,m} \;\preceq\; \rho_{k,p_2,m}\,.$$

Proof of Theorem 3.2. As in the proof of Theorem 2.2, we take $S = \cup_{k=2}^{\infty} S_k$ where $S_k = ((k,1),(k,2),\ldots,(k,k))$, and the probability measure μ on $\{0,1\}^S$ to be the product measure

$$\mu \;=\; \mu_2 \times \mu_3 \times \cdots$$

where each μ_k is a probability measure on $\{0,1\}^{S_k}$. This time, we take the μ_k's to be as follows. For $\sigma \in \{0,1\}^{S_k}$, set

$$\mu_k(\sigma) = \begin{cases} k^{-1}2^{-k} & \text{if the number of 1's in } \sigma \text{ is exactly 1,} \\ 1-2^{-k} & \text{if } \sigma = (1,1,1,\ldots,1), \\ 0 & \text{otherwise.} \end{cases} \qquad (12)$$

We now make three claims about the μ_k measures:

CLAIM 1. $p_{\max,\mu_k} \geq \frac{1}{2}$ for all k.

CLAIM 2. $\lim_{k\to\infty} p_{\max,\mu_k} = \frac{1}{2}$.

CLAIM 3. For any fixed $\epsilon < \frac{1}{2}$, we have for all k sufficiently large that

$$\mu_k^{(-,\epsilon)} \succeq \pi_{\frac{1}{2}}$$

where $\pi_{\frac{1}{2}}$ is product measure with $p = \frac{1}{2}$ on $\{0,1\}^{S_k}$.

We slightly postpone proving the claims, and first show how they imply the existence of a nonrigid measure that dominates $\pi_{\frac{1}{2}}$.

Let us modify S and μ slightly by setting, for $m \geq 2$,

$$\tilde{S}_m = \cup_{k=m}^{\infty} S_k$$

and

$$\tilde{\mu}_m = \mu_m \times \mu_{m+1} \times \cdots \qquad (13)$$

so that in other words $\tilde{\mu}_m$ is the probability measure on $\{0,1\}^{\tilde{S}_m}$ which arises by projecting μ on $\{0,1\}^{\tilde{S}_m}$.

Using the product structure (13), we get from CLAIM 1 that $p_{\max,\tilde{\mu}_m} \geq \frac{1}{2}$ (for any m), and from CLAIM 2 that $p_{\max,\tilde{\mu}_m} \leq \frac{1}{2}$ (for any m). Hence

$$p_{\max,\tilde{\mu}_m} = \frac{1}{2}$$

for any m. Fixing $\epsilon \in (0,1/2)$, we can also deduce from (13) and CLAIM 3 that

$$\tilde{\mu}_m^{(-,\epsilon)} \succeq \pi_{\frac{1}{2}} = \pi_{p_{\max,\tilde{\mu}_m}} \qquad (14)$$

for m sufficiently large. For such m we thus have that $\tilde{\mu}_m$ is nonrigid.

It remains to prove CLAIM 1, CLAIM 2 and CLAIM 3.

CLAIM 1 is the same as saying that $\mu_k \succeq \pi_{\frac{1}{2}}$. This is immediate to verify, but the best way to think about it is as follows. Suppose that we pick $X_k \in \{0,1\}^{S_k}$ according to $\pi_{\frac{1}{2}}$, and if $X_k = (0,0,\ldots,0)$ then we switch one of the 0's (chosen uniformly at random) to a 1, while otherwise we switch *all* 0's to 1's. The resulting random element of $\{0,1\}^{S_k}$ then has distribution μ_k.

To prove CLAIM 2, it suffices (in view of CLAIM 1) to prove that

$$\limsup_{k\to\infty} p_{\max,\mu_k} \le \frac{1}{2}$$

and to this end it is enough to show for any $\delta > 0$ that

$$\mu_k \not\preceq \pi_{\frac{1}{2}+\delta} \tag{15}$$

for all sufficiently large k. Let A_k denote the event of seeing at most one 1 in $\{0,1\}^{S_k}$; then A_k is a decreasing event and its complement $\neg A_k$ is increasing. Now simply note that

$$\frac{\mu_k(A_k)}{\pi_{\frac{1}{2}+\delta}(A_k)} = \frac{(\frac{1}{2})^k}{(\frac{1}{2}-\delta)^k + k(\frac{1}{2}+\delta)(\frac{1}{2}-\delta)^{k-1}} \tag{16}$$

which tends to ∞ as $k \to \infty$. Hence, taking k large enough gives $\mu_k(A_k) > \pi_{\frac{1}{2}+\delta}(A_k)$, so that $\mu_k(\neg A_k) < \pi_{\frac{1}{2}+\delta}(\neg A_k)$ and (15) is established, proving CLAIM 2.

To prove CLAIM 3, note first that both $\pi_{\frac{1}{2}}$ and $\mu_k^{(-,\epsilon)}$ are invariant under permutations of S_k, so that it suffices to show for k large that

$$\mu_k^{(-,\epsilon)}(B_n) \le \pi_{\frac{1}{2}}(B_n) \tag{17}$$

for $n = 0, 1, \ldots, k-1$, where B_n is the event of seeing at most n 1's in S_k. For $n = 0$ we get

$$\frac{\mu_k^{(-,\epsilon)}(B_0)}{\pi_{\frac{1}{2}}(B_0)} = \frac{(\frac{1}{2})^k\epsilon + (1-(\frac{1}{2})^k)\epsilon^k}{(\frac{1}{2})^k} \tag{18}$$

while for $n = 1$

$$\frac{\mu_k^{(-,\epsilon)}(B_1)}{\pi_{\frac{1}{2}}(B_1)} = \frac{(\frac{1}{2})^k + (1-(\frac{1}{2})^k)(\epsilon^k + k\epsilon^{k-1}(1-\epsilon))}{(k+1)(\frac{1}{2})^k}. \tag{19}$$

The right-hand sides of (18) and (19) tend to ϵ and 0, respectively, as $k \to \infty$, so (17) is verified for $n = 0$ and 1 (and k large enough). To verify (17) for $n \ge 2$ (and all such k), define two random variables Y and Y' as the number of 1's in two random elements of $\{0,1\}^{S_k}$ with respective distributions $\mu_k^{(-,\epsilon)}$ and $\pi_{\frac{1}{2}}$. Note that Y conditioned on taking value at least 2 has the same distribution as a Bin $(k, 1-\epsilon)$ random variable conditional on taking value at least 2, while the conditional distribution of Y' given that it is at least 2, is that of a Bin $(k, \frac{1}{2})$ variable conditioned on being at least 2. Defining $\rho_{k,(1-\epsilon),2}$ and $\rho_{k,\frac{1}{2},2}$ as in Lemma 3.3, we thus have for $n \in \{2, \ldots, k-1\}$ that

$$\mu_k^{(-,\epsilon)}(B_n) = 1 - (1-\mu_k^{(-,\epsilon)}(B_1))(1-\rho_{k,(1-\epsilon),2}(B_n)) \tag{20}$$

and

$$\pi_{\frac{1}{2}}(B_n) \;=\; 1-(1-\pi_{\frac{1}{2}}(B_1))(1-\rho_{k,\frac{1}{2},2}(B_n))\,. \tag{21}$$

But we have already seen that $\mu_k^{(-,\epsilon)}(B_1) \leq \pi_{\frac{1}{2}}(B_1)$, and Lemma 3.3 tells us that $\rho_{k,(1-\epsilon),2}(B_n) \leq \rho_{k,\frac{1}{2},2}(B_n)$, so (20) and (21) yield

$$\mu_k^{(-,\epsilon)}(B_n) \;\leq\; \pi_{\frac{1}{2}}(B_n)\,,$$

and CLAIM 3 is established.

Finally, we translate this example into the setting of translation invariant distributions on $\{0,1\}^{\mathbb{Z}}$. The measure $\tilde{\mu}_m$ can be turned into a translation invariant measure $\tilde{\mu}_m^*$ on $\{0,1\}^{\mathbb{Z}}$ by the same independent-copy-and-paste procedure as in Theorem 2.2. The property

$$\pi_{\frac{1}{2}} \;\preceq\; (\tilde{\mu}_m^*)^{(-,\epsilon)}$$

is obviously inherited from (14). Thus, in order to show that $\tilde{\mu}_m^*$ is nonrigid, it only remains to show that it does not stochastically dominate $\pi_{\frac{1}{2}+\delta}$ for any $\delta > 0$. This follows using (16) by an argument analogous to (9) in Theorem 2.2: If we pick k depending on δ as in the justification of CLAIM 2, then, under $\tilde{\mu}_m^*$, certain infinite arithmetic progressions will have subsequences of length k which contain at most one 1 often enough (under spatial averaging) that the corresponding event has $\pi_{\frac{1}{2}+\delta}$-measure 0. We omit the details. □

Remark 3.4 The measure $\tilde{\mu}_m$ is obviously not uniformly insertion tolerant, and we have thus demonstrated the existence of a measure for which property (IV) holds while neither (II) nor (III) does. □

Remark 3.5 For any $p \in (0,1)$, the construction above can be modified by replacing 2^{-k} by p^k in (12). Proceeding as in the rest of the proof yields the result that for any $p, \epsilon \in (0,1)$ such that $p+\epsilon < 1$, there exists a measure μ on $\{0,1\}^S$ where S is countably infinite, with the property that $p_{\max,\mu} = p$ and

$$\pi_{p_{\max,\mu}} \preceq \mu^{(-,\epsilon)}\,.$$

This is obviously sharp. □

4 Further results on rigidity

In this section, we continue the study of rigidity, and prove Theorem 1.9.

The proof of Theorem 1.9 will make use of the following technical lemma.

Lemma 4.1 *Let μ be a measure on $\{0,1\}^{\mathbb{Z}^d}$. Assume that it is δ-insertion tolerant for some $\delta > 0$. If for some $p \in (0,1)$ and $\epsilon > 0$*

$$\mu^{(-,\epsilon)}(\sigma(\{1,\ldots,n\}^d) \equiv 0) \le (1-p)^{n^d} \text{ for all } n \ge 0, \tag{22}$$

then there exists $p' > p$ such that

$$\mu(\sigma(\{1,\ldots,n\}^d) \equiv 0) \le (1-p')^{n^d} \text{ for all } n \ge 0.$$

Proof. Let $X \sim \mu$ and $Z \sim \pi_{1-\epsilon}$ be independent and let $X^{(-,\epsilon)} = \min(X,Z)$. It is easy to see using the δ-insertion tolerance that for any $s \in \{1,\ldots,n\}^d$, and any $\zeta \in \{0,1\}^{\{1,\ldots,n\}^d \setminus s}$

$$\begin{aligned}&\mathbb{P}(X(s) = 1 \cap X(\{1,\ldots,n\}^d \setminus s) \equiv \zeta)\\ &\quad\ge \frac{\delta}{1-\delta}\mathbb{P}(X(s) = 0 \cap X(\{1,\ldots,n\}^d \setminus s) \equiv \zeta).\end{aligned}$$

Iterating this, we get that for any $\xi \in \{0,1\}^{\{1,\ldots,n\}^d}$

$$\mathbb{P}(X(\{1,\ldots,n\}^d) \equiv \xi) \ge \left(\frac{\delta}{1-\delta}\right)^{|\xi|}\mathbb{P}(X(\{1,\ldots,n\}^d) \equiv 0).$$

Here $|\xi|$ denotes the cardinality of the set $\{s \in \{1,\ldots,n\}^d : \xi(s) = 1\}$. An elementary and straightforward calculation will now yield that

$$\mathbb{P}(X^{(-,\epsilon)}(\{1,\ldots,n\}^d) \equiv 0) \ge \left(1 + \frac{\epsilon\delta}{1-\delta}\right)^{n^d}\mathbb{P}(X(\{1,\ldots,n\}^d) \equiv 0).$$

Using this in combination with (22) proves the lemma. □

Proof of Theorem 1.9. The case $p_{\max,\mu} = 1$ is trivial and we therefore assume that $p_{\max,\mu} \in (0,1)$. In [12], it is shown that if μ is downwards FKG and if

$$\mu(\sigma(\{1,\ldots,n\}^d) \equiv 0) \le (1-p)^{n^d} \text{ for all } n \ge 0, \tag{23}$$

then $\pi_p \preceq \mu$. Therefore if $\pi_{p_{\max,\mu}} \preceq \mu^{(-,\epsilon)}$ for some $\epsilon > 0$, then (22) trivially holds (with $p = p_{\max,\mu}$) and so we can conclude from Lemma 4.1 and the above result in [12] that $\pi_{p'} \preceq \mu$ for some $p' > p_{\max,\mu}$, a contradiction. □

5 Proof of main result

Lemma 5.1 *If μ is uniform upwards extractable, then for any $\epsilon > 0$ there exists a $\delta > 0$ such that $(\mu^{(-,\epsilon)})^{(+,\delta)} \preceq \mu$.*

Proof. Let ν and $\alpha > 0$ be such that $\mu = \nu^{(+,\alpha)}$. One can easily compute that for any α, ϵ, and δ, we have that

$$((\mu^{(+,\alpha)})^{(-,\epsilon)})^{(+,\delta)} = \mu^{(-,\epsilon(1-\delta),+,\alpha(1-\epsilon)+\alpha\epsilon\delta+(1-\alpha)\delta)}.$$

Now, given $\epsilon > 0$, choose $\delta > 0$ such that $\alpha(1-\epsilon) + \alpha\epsilon\delta + (1-\alpha)\delta < \alpha$. We therefore get that

$$(\mu^{(-,\epsilon)})^{(+,\delta)} = ((\nu^{(+,\alpha)})^{(-,\epsilon)})^{(+,\delta)} \preceq \nu^{(-,\epsilon(1-\delta),+,\alpha)} \preceq \nu^{(+,\alpha)} = \mu.$$

□

Lemma 5.2 *Given a probability measure μ on $\{0,1\}^S$, assume that for every $\epsilon > 0$, there exists a $\delta > 0$ such that $(\mu^{(-,\epsilon)})^{(+,\delta)} \preceq \mu$. Then μ is rigid.*

Proof. The case $p_{\max,\mu} = 1$ is trivial, and we will therefore assume that $p_{\max,\mu} \in (0,1)$. Assume for contradiction that μ is nonrigid. Then there exists an $\epsilon > 0$ such that $\pi_{p_{\max,\mu}} \preceq \mu^{(-,\epsilon)}$. By assumption there exists a $\delta > 0$ such that $(\mu^{(-,\epsilon)})^{(+,\delta)} \preceq \mu$. Hence $(\pi_{p_{\max,\mu}})^{(+,\delta)} \preceq (\mu^{(-,\epsilon)})^{(+,\delta)} \preceq \mu$. Since $p_{\max,\mu} < 1$, $(\pi_{p_{\max,\mu}})^{(+,\delta)}$ is a product measure with density strictly larger than $p_{\max,\mu}$. This is a contradiction. □

Our next example provides us with an example of a μ which is on one hand rigid but on the other hand not uniformly insertion tolerant. It is a variant of [14, Remark 6.4] and shows that the reverse statement of Lemma 5.1 is false.

Example 5.3 Let $\{X_i\}_{i\in\mathbb{N}}$ be defined in the following way. For every even $i \geq 0$, let independently (X_i, X_{i+1}) be $(1,1)$ or $(0,0)$ with probability 1/2 each. Let μ_e denote the distribution of this process. For $\epsilon, \delta > 0$ let $\{X_i^{(-,\epsilon(1-\delta),+,\delta)}\}_{i\in\mathbb{N}}$ be a sequence of random variables with distribution $\mu_e^{(-,\epsilon(1-\delta),+,\delta)} = (\mu_e^{(-,\epsilon)})^{(+,\delta)}$. By noting that for any $\epsilon > 0$ there exists a $\delta > 0$ such that for even i

$$\mathbb{P}(\max(X_i^{(-,\epsilon(1-\delta),+,\delta)}, X_{i+1}^{(-,\epsilon(1-\delta),+,\delta)}) = 1) < \tfrac{1}{2},$$

we see that for the same choice of ϵ, δ we get that $(\mu_e^{(-,\epsilon)})^{(+,\delta)} \preceq \mu_e$. Lemma 5.2 gives us that μ_e is rigid. However, it is easy to see that μ_e is not uniform insertion tolerant. Furthermore it is possible to make this example translation invariant by some easy manipulations. □

Proof of Theorem 1.7. Lemma 5.1 together with Lemma 5.2 shows that property (I) implies property (III) and all the other implications were indicated in the introduction. As far as all of the reversed implications claimed not to hold, we continue as follows. Example 5.3 together with Lemma 5.2 shows that (III) does not imply (II) (and hence that (III) does not imply (I) and that (IV) does not imply (II)). Theorem 3.2 implies that (IV) does not imply (III). Finally, Theorem 2.2 shows that (II) does not imply (I). Also, all of these examples were translation invariant measures on $\{0,1\}^{\mathbb{Z}}$. □

Proof of Theorem 1.10. This follows immediately from Propositions 2.1 and 3.1. □

Of some interest in this context is the following result, which is an easy consequence of Lemma 5.1.

Corollary 5.4 *Assume that* (μ_1, μ_2) *is downwards movable and that* μ_2 *is uniformly upwards extractable. Then* (μ_1, μ_2) *is also upwards movable.*

Although it would take us too far afield to discuss details, let us mention that the processes studied in [14] provide a nice source of examples illustrating the various concepts in this paper. For example, one can find there examples of processes which are rigid but are not uniformly insertion tolerant.

6 Open problems

We end the paper with a list of open problems.

1. Does property (II) imply (III) in Theorem 1.7?

2. Let S be countable and μ be a uniformly insertion tolerant probability measure on $\{0,1\}^S$. Is it the case that for all $\epsilon > 0$, there exists a $\delta > 0$ such that $(\mu^{(-,\epsilon)})^{(+,\delta)} \preceq \mu$? A positive answer to this question would of course yield a positive answer to question 1.

3. Is the reverse statement of Lemma 5.2 true?

References

[1] Bassalygo, L.A. and Dobrushin, R.L. (1987) Epsilon-entropy of a Gibbs field, (Russian) *Problemy Peredachi Informatsii* **23**, 3–15.

[2] van den Berg, J., Häggström O. and Kahn, J. (2005) Some conditional correlation inequalities for percolation and related processes, *Random Structures Algorithms*, to appear.

[3] Broman, E.I. and Steif, J.E. (2004) Dynamical stability of percolation for some interacting particle systems and ϵ-movability, *Ann. Probab.*, to appear.

[4] Georgii, H.-O., Häggström, O. and Maes, C. (2001) The random geometry of equilibrium phases, *Phase Transitions and Critical Phenomena, Volume 18* (C. Domb and J.L. Lebowitz, eds), pp 1–142, Academic Press, London.

[5] Gray, R.M. (1970) Informations rates of autoregressive processes, *IEEE Trans. Information Theory* **IT-16**, 412–421.

[6] Gray, R.M. (1971) Rate distortion functions for finite-state finite-alphabet Markov sources, *IEEE Trans. Information Theory* **IT-17**, 127–134.

[7] Gray, R.M. (1973) Correction to "Information rates of stationary-ergodic finite-alphabet sources", *IEEE Trans. Information Theory* **IT-19**, 573.

[8] Hajek, B. and Berger, T. (1987) A decomposition theorem for binary Markov random fields, *Ann. Probab.* **15**, 1112–1125.

[9] Liggett, T.M. (1985) *Interacting Particle Systems* Springer, New York.

[10] Liggett, T.M. (1999) *Stochastic Interacting Systems: Contact, Voter and Exclusion Processes*, Springer, New York.

[11] Liggett, T.M. (1994) Survival and coexistence in interacting particle systems, *Probability and Phase Transition* (ed. G. Grimmett), Kluwer, Dodrecht, pp 209–226.

[12] Liggett, T.M. and Steif, J.E. (2004) Stochastic domination: the contact process, Ising models and FKG measures, *Annales Institut Henri Poincare, Probabilites et Statistiques*, to appear.

[13] Lyons, R. and Schramm, O. (1999) Indistinguishability of percolation clusters, *Ann. Probab.* **27**, 1809–1836.

[14] Lyons, R. and Steif, J.E. (2003) Stationary determinantal processes: phase multiplicity, Bernoullicity, entropy, and domination, *Duke Math. Journal*, **120**, 515–575.

[15] Newman, C.M. (1987) Decomposition of binary random fields and zeros of partition functions, *Ann. Probab.* **15**, 1126–1130.

[16] Newman, C.M. and Schulman, L.S. (1981) Infinite clusters in percolation models, *J. Stat. Phys.* **26**, 613–628.

Paper IV

Stochastic Domination for a Hidden Markov Chain with Applications to the Contact Process in a Randomly Evolving Environment

Erik I. Broman*

February 27, 2006

Abstract

In this paper we introduce and study the contact process in a randomly evolving environment. This can be thought of as a contact process depending on a two-state background process, where the rate of recovery at any time $t \geq 0$ is chosen according to the current state of this background process. By using stochastic domination techniques we will investigate matters of extinction and that of weak and strong survival. All of the results will be obtained in three steps. First, we will analyze a certain hidden Markov chain for which we obtain some sharp stochastic domination results. This process is constructed by using two i.i.d. sequences, both with state space $\{0, 1\}$ but with different densities. We then let our process be a mix of these two, depending on the state of a two-state background process. Secondly we will exploit these results to get an analogue in continuous time. This analogue is constructed by using two ordinary Poisson processes with different intensities, where the one we use depends on a continuous time two-state background process. Finally we will use these continuous time results to analyze the contact process in a randomly evolving environment.

AMS subject classification: 82C22, 60K35

Keywords and phrases: contact process, stochastic domination, hidden markov chain

Short title: Stoch. dom. and a randomly evolving c.p.

1 Introduction

The first part of this introduction will discuss the concept of stochastic domination and then move on to state our discrete time results. We will then

*Dept. of Mathematics, Chalmers University of Technology, S-412 96 Göteborg, Sweden, url: `http://www.math.chalmers.se/~broman/`, e-mail: broman@math.chalmers.se
research supported by the Swedish Research Council.

proceed by defining the Contact Process in a Randomly Evolving Environment, from now on referred to as CPREE, that we introduce in this paper. We would like to point out that a model called the Contact Process in a Random Environment (or CPRE) has been studied before. The first papers concerning this latter model were [2] and [7], and then further studies were carried out in for instance [1], [8], [13] and [14]. However the random environments in these papers are static while here they change over time.

In this paper we are concerned with models on connected graphs $G = (S, E)$ of bounded degree, in which every site $s \in S$ can take values 0 or 1. Here σ and ξ will mainly denote configurations on S, i.e. $\sigma, \xi \in \{0,1\}^S$. We say that $\xi \preceq \tilde{\xi}$ if $\xi(s) \leq \tilde{\xi}(s)$ for every $s \in S$. An increasing function f is a function $f : \{0,1\}^S \to \mathbb{R}$ such that $f(\xi) \leq f(\tilde{\xi})$ for all $\xi \preceq \tilde{\xi}$. For two probability measures μ, μ' on $\{0,1\}^S$, we write $\mu \preceq \mu'$ if for every continuous increasing function f we have that $\mu(f) \leq \mu'(f)$. ($\mu(f)$ is shorthand for $\int f(x) d\mu(x)$.) Strassens Theorem (see [9], page 72) states that if $\mu \preceq \mu'$, then there exist random variables X, X' with distributions μ, μ' respectively, defined on the same probability space, such that $X \preceq X'$ a.s.

We will need the following standard definition.

Definition 1.1 *Let S be such that $|S| < \infty$ and let μ be a probability measure on $\{0,1\}^S$ with full support. μ is said to be monotone, if for every $s \in S$ and any $\xi, \tilde{\xi} \in \{0,1\}^{S \setminus s}$ such that $\xi \preceq \tilde{\xi}$, one has that*

$$\mu(\sigma(s) = 1 | \sigma(S \setminus s) \equiv \xi) \leq \mu(\sigma(s) = 1 | \sigma(S \setminus s) \equiv \tilde{\xi}).$$

If $|S| = \infty$, we say that a probability measure μ on $\{0,1\}^S$ is monotone if the restriction of μ to any finite subset of S is monotone.

For $p \in [0,1]$, let each site $s \in S$, independently of all others, take value 1 with probability p and 0 with probability $1-p$. Write π_p for this product measure on $\{0,1\}^S$. For any probability measure μ on $\{0,1\}^S$ define $p_{\max,\mu}$ by

$$p_{\max,\mu} := \sup\{p \in [0,1] : \pi_p \preceq \mu\}.$$

The supremum is easily seen to be obtained, which motivates the notation. Similarly define

$$p_{\min,\mu} := \inf\{p \in [0,1] : \mu \preceq \pi_p\}.$$

Next, informally we will here think of $\{B_n\}_{n=1}^\infty$ as a background process which influences another process $\{X_n\}_{n=1}^\infty$. Formally, fix $0 \leq \alpha_0 \leq \alpha_1 \leq 1$ and let $\{B_n\}_{n=1}^\infty$ be any process with state space $\{0,1\}$. Conditioned on $\{B_n\}_{n=1}^\infty$ let the process $\{X_n\}_{n=1}^\infty$, also with state space $\{0,1\}$, be a sequence of conditionally independent random variables where the (conditional) distribution of X_k is

$$\begin{array}{lll} \text{if} & \text{then} & \text{with prob.} \\ B_k = 0 & X_k = 1 & \alpha_0 \\ B_k = 1 & X_k = 1 & \alpha_1, \end{array} \qquad (1)$$

for every $k \geq 1$.

We will say that μ is translation invariant on $\mathbb{N}$ if for every $l \geq 1$, $k \geq 0$ and any $\xi \in \{0,1\}^{\{1,\dots,l\}}$

$$\mu(\sigma(1,\dots,l) \equiv \xi) = \mu(\sigma(k+1,\dots,k+l) \equiv \xi).$$

In section 2 we will prove the following proposition.

Proposition 1.2 *Assume that the distribution of* $\{B_n\}_{n=1}^{\infty}$ *is monotone and translation invariant. Then the sequence*

$$\{\mathbb{P}(X_n = 1 | X_{n-1} = \cdots = X_1 = 0)\}_{n\geq 1},$$

is decreasing in n*. In addition the limit equals* $p_{\max,\mu}$*, where* μ *is the distribution of* $\{X_n\}_{n=1}^{\infty}$.

The proof is an easy consequence of results from [5] and [11].

We are now ready to define the discrete background process that we will use throughout this paper. For $p, \gamma \in [0,1]$, define the Markov chain $\{B_n\}_{n=1}^{\infty}$ in the following way:

$$B_1 = \begin{cases} 1 & \text{w.p.} \quad p \\ 0 & \text{w.p.} \quad 1-p, \end{cases} \tag{2}$$

and for $k \geq 2$,

$$\begin{array}{lll} \text{if} & \text{then} & \text{w.p.} \\ B_{k-1} = 0 & B_k = 1 & \gamma p \\ B_{k-1} = 1 & B_k = 0 & \gamma(1-p). \end{array} \tag{3}$$

In other words, B_k takes the same value as B_{k-1} unless there is an update which happens independently with probability γ. If an update occurs, $B_k = 1$ with probability p, and $B_k = 0$ with probability $1-p$. The resulting joint process $\{(B_n, X_n)\}_{n=1}^{\infty}$, with the second marginal defined through (1), is an example of a so called hidden Markov chain. The main theorem of Section 3 is the following, here μ refers to the distribution of $\{X_n\}_{n=1}^{\infty}$ with the background process as above.

Theorem 1.3 *We have that*

$$p_{\max,\mu} = \frac{1}{2}\left(1 - C - \sqrt{(1-C)^2 - 4D}\right),$$

where

$$C = (1 - \alpha_0 - \alpha_1) - \gamma(1 - \alpha_0 - (1-p)(\alpha_1 - \alpha_0))$$

and

$$D = \alpha_0\alpha_1 + \gamma(\alpha_1(1-\alpha_0) - (1-p)(\alpha_1 - \alpha_0)).$$

Furthermore

$$p_{\min,\mu} = \frac{1}{2}\left(1 + C' + \sqrt{(1-C')^2 - 4D'}\right),$$

where C' and D' are as C and D but with α_0, α_1, p and γ replaced by $1 - \alpha_1, 1 - \alpha_0, 1 - p$ and γ respectively.

The proof of this theorem unfortunately involves some tedious (but straightforward) calculations; however this result is needed for all the other results of this paper.

From the results of Section 3, we will in Section 4 prove our next result. First define

$$X_n^c := \sum_{i=1}^{n} X_i \quad \forall n \in \mathbb{N},$$

where c indicates that we are counting the number of 1's up to time n. The pair of processes $\{(B_t, X_t)\}_{t\geq 0}$, to be defined below, will be the continuous time analogue of the pair of processes $\{(B_n, X_n^c)\}_{n=1}^{\infty}$. We will let $B_0 = 1$ with probability p and $B_0 = 0$ with probability $1-p$. Thereafter the process waits an exponentially distributed time with parameter $\gamma > 0$ and updates its status. After an update, the process takes value 1 with probability p and 0 with probability $1-p$, and all of this is done independently of everything else. Having defined $\{B_t\}_{t\geq 0}$ we define $\{X_t\}_{t\geq 0}$ by starting with an ordinary Poisson(α_1)-process on $[0, \infty)$ and thin it in the following way. If the Poisson(α_1)-process has an arrival at some time $\tau \in [0, \infty)$ and if $B_\tau = 0$ then, independently of everything else

$$X_\tau = \begin{cases} X_{\tau^-} + 1 & \text{with prob. } \frac{\alpha_0}{\alpha_1} \\ X_{\tau^-} & \text{otherwise.} \end{cases}$$

If instead $B_\tau = 1$, we let $X_\tau = X_{\tau^-} + 1$. In words, $\{X_t\}_{t\geq 0}$ is the counting process for a type of Poisson process where the parameter comes from $\{B_t\}_{t\geq 0}$. Analogous to the definition of $p_{\max,\mu}$, define $\lambda_{\max,\mu}$, where μ here refers to the distribution of $\{X_t\}_{t\geq 0}$, in the following way. $\lambda_{\max,\mu}$ is the maximum real number λ such that a Poisson(λ)-process can be coupled with the process $\{X_t\}_{t\geq 0}$ so that if the Poisson(λ)-process has an arrival at time $\tau \in [0, \infty)$ then so does the $\{X_t\}_{t\geq 0}$ process. In other words, there exists $\{X'_t\}_{t\geq 0}$ with distribution Poisson(λ) coupled with $\{X_t\}_{t\geq 0}$ such that

$$X_t - X'_t \text{ is non-decreasing in } t.$$

Define $\lambda_{\min,\mu}$ to be the minimal real number λ such that a Poisson(λ)-process can be coupled with the process $\{X_t\}_{t\geq 0}$ so that if the $\{X_t\}_{t\geq 0}$ process has an arrival at time $\tau \in [0, \infty)$ then so does the Poisson(λ)-process. Observe that $\lambda_{\max,\mu} = \lambda_{\max,\mu}(\alpha_0, \alpha_1, \gamma, p)$. We will write out the arguments in most equations, but not in more general discussions. Trivially $\alpha_0 \leq \lambda_{\max,\mu} \leq \lambda_{\min,\mu} \leq \alpha_1$. For future convenience let $\mathrm{Poi}^{\gamma,p}_{\alpha_0,\alpha_1}$ denote the distribution of $\{X_t\}_{t\geq 0}$ and Poi_λ denote the distribution of a Poisson process with intensity λ. The coupling described above is a form of stochastic domination and we will write

$$\mathrm{Poi}_{\lambda_{\max,\mu}} \preceq \mathrm{Poi}^{\gamma,p}_{\alpha_0,\alpha_1}, \tag{4}$$

and

$$\mathrm{Poi}^{\gamma,p}_{\alpha_0,\alpha_1} \preceq \mathrm{Poi}_{\lambda_{\min,\mu}}. \tag{5}$$

Define

$$\begin{aligned}\bar{\lambda} &= \bar{\lambda}(\alpha_0,\alpha_1,\gamma,p) \\ &:= \frac{1}{2}(\alpha_0+\alpha_1+\gamma-\sqrt{(\alpha_1-\alpha_0-\gamma)^2+4\gamma(1-p)(\alpha_1-\alpha_0)}).\end{aligned}$$

Theorem 1.4 *Let $\{(B_t,X_t)\}_{t\geq 0}$ be as above. For every choice of $\alpha_0,\alpha_1,\gamma > 0$ with $\alpha_0 \leq \alpha_1$ and $p \in [0,1]$ we have that*

$$\lambda_{\max,\mu}(\alpha_0,\alpha_1,\gamma,p) = \bar{\lambda}, \tag{6}$$

and for $p > 0$

$$\lambda_{\min,\mu}(\alpha_0,\alpha_1,\gamma,p) = \alpha_1.$$

Remarks:

- Note the apparent lack of symmetry between $\lambda_{\max,\mu}$ and $\lambda_{\min,\mu}$. Informally, consider for a moment the model to be a point process, where the process is 0 unless there is an arrival, in which case it takes the value 1. We can then see that the true symmetric statement of the $\lambda_{\max,\mu}$ result would concern a model which is 1 unless there is an arrival in which case it takes the value 0. This however does not correspond to the result concerning $\lambda_{\min,\mu}$.

- We will show in Section 4 that $\lambda_{\max,\mu}(\alpha_0,\alpha_1,\gamma,p) \to \min(\alpha_1,\alpha_0+\gamma)$ as $p \to 1$. Hence if $\gamma > \alpha_1 - \alpha_0$, then $\lambda_{\max,\mu}(\alpha_0,\alpha_1,\gamma,p) \to \alpha_1$ as $p \to 1$ which one would expect. In contrast, for every $p > 0$ $\lambda_{\min,\mu}(\alpha_0,\alpha_1,\gamma,p) = \alpha_1$ and so $\lambda_{\min,\mu}(\alpha_0,\alpha_1,\gamma,p) \not\to \alpha_0$ as $p \to 0$ as one might have suspected; this gives a discontinuity at $p = 0$. Also, it is trivial to show that $\lambda_{\max,\mu}(\alpha_0,\alpha_1,\gamma,p) \to \alpha_0$ as $p \to 0$.

- Furthermore, for fixed $0 < p < 1$, it follows from the proof of Proposition 1.9, where we take the limit $\gamma \to \infty$ in equation (6) above, that $\lim_{\gamma\to\infty} \lambda_{\max,\mu}(\alpha_0,\alpha_1,\gamma,p) = p\alpha_1 + (1-p)\alpha_0$. This is exactly what one should expect, since as γ grows larger, the suppressing of possible arrivals becomes "increasingly independent". Whenever a possible arrival occurs, the background process has with very high probability been updated since the last possible arrival, and if so the new arrival is suppressed independently of everything else. Also, by letting $\gamma \to 0$ in equation (6) we get that $\lambda_{\max,\mu}(\alpha_0, \alpha_1, 0, p) \to \alpha_0$. This last result is also natural. As γ becomes smaller, we will find longer and longer time intervals in which the background process is in the lower state. Therefore the Poisson process we dominate must have lower and lower density.

It is natural to ask for quantitative versions of Theorem 1.4 for finite time, and in fact we will need such results to prove Theorem 1.4. Therefore, for $T > 0$, let $\lambda^T_{\max,\mu}(\alpha_0, \alpha_1, \gamma, p)$ denote the maximum intensity of the Poisson process which the second marginal of the truncated process $\{(B_t, X_t)\}_{t\in[0,T]}$ can dominate. Define $\lambda^T_{\min,\mu}(\alpha_0, \alpha_1, \gamma, p)$ analogously. We feel that this bound is interesting in its own right and we therefore present it in our next theorem together with a lower bound on $\lambda^T_{\max,\mu}(\alpha_0, \alpha_1, \gamma, p)$ and a result for $\lambda^T_{\min,\mu}(\alpha_0, \alpha_1, \gamma, p)$ (this last result will follow from the proof of Theorem 1.4).

Let

$$L = L(\alpha_0, \alpha_1, \gamma, p) := \sqrt{(\alpha_1 - \alpha_0 - \gamma)^2 + 4\gamma(1-p)(\alpha_1 - \alpha_0)}.$$

Theorem 1.5 *For every choice of* $\alpha_0, \alpha_1, \gamma, T > 0$ *with* $\alpha_0 \le \alpha_1$ *and* $p \in (0,1)$ *we have that*

$$\lambda^T_{\max,\mu}(\alpha_0, \alpha_1, \gamma, p) \ge \bar{\lambda} + (p\alpha_1 + (1-p)\alpha_0 - \bar{\lambda})e^{-TL}. \tag{7}$$

Furthermore there exists a constant $E > 0$, *depending on* $\alpha_1, \alpha_0, \gamma$ *and* p, *such that*

$$\lambda^T_{\max,\mu}(\alpha_0, \alpha_1, \gamma, p) \le \bar{\lambda} + \frac{1}{T}(p\alpha_1 + (1-p)\alpha_0 - \bar{\lambda})\frac{1 - e^{-TE}}{E}. \tag{8}$$

Finally

$$\lambda^T_{\min,\mu} = \lambda_{\min,\mu} = \alpha_1.$$

Remark: Observe that the right-hand side of equation (7) tends to $p\alpha_1 + (1-p)\alpha_0$ as T tends to 0, and that it tends to $\bar{\lambda} = \lambda_{\max,\mu}(\alpha_0, \alpha_1, \gamma, p)$ as T tends to infinity. Both results are of course what you would expect. The same is true for the upper bound of equation (8).

1.1 Models

1.1.1 The contact process

In this section we will discuss some basic concepts concerning the contact process, see [9] for results up to 1985 and [10] for results between 1985 and

1999. Consider a graph $G = (S, E)$ of bounded degree. In the contact process the state space is $\{0,1\}^S$. We will let 1 represent an infected individual, while a 0 will be used to represent a healthy individual. Let $\lambda > 0$, and define the flip rate intensities to be

$$C(s,\sigma) = \begin{cases} 1 & \text{if} \quad \sigma(s) = 1 \\ \lambda \sum_{(s',s)\in E} \sigma(s') & \text{if} \quad \sigma(s) = 0. \end{cases} \tag{9}$$

By flip rate intensities, informally, we mean as usual that every site $s \in S$ waits an exponentially distributed time with parameter $C(s,\sigma)$ before changing its status. Here, $\mathbf{1}_0, \mathbf{1}_1$ will denote the measures that put mass one on the configuration of all 0's and all 1's respectively. If we let the initial distribution be $\sigma \equiv 1$, the distribution of this process at time t, which we will denote by $\mathbf{1}_1 T_\lambda(t)$, is known to converge as t tends to infinity. This is simply because it is a so-called "attractive" process and $\sigma \equiv 1$ is the maximal state; see [9] page 265. This limiting distribution will be referred to as the upper invariant measure for the contact process with parameter λ and will be denoted by ν_λ. We then let Ψ^λ denote the stationary Markov process on $\{0,1\}^S$ with initial (and invariant) distribution ν_λ. One can also choose to start the process with any set $A \subset S$ of infected individuals and then use the flip rate intensities above. Denote this latter process by $\Psi^{\lambda,A}$. We say that the process dies out if for any $s \in S$

$$\Psi^{\lambda,\{s\}}(\sigma_t \not\equiv 0 \ \forall \ t \geq 0) = 0,$$

and otherwise it survives. We also say that the process survives strongly if

$$\Psi^{\lambda,\{s\}}(\sigma_t(s) = 1 \text{ i.o.}) > 0.$$

We say that the process survives weakly if it survives but does not survive strongly. These and all other statements like it, made here and later, are independent of the specific choice of the site s; see [10]. We will use the same definition of survival for some closely related processes below. It is well known that for any graph (see [10] pg. 42) there exists two critical parameter values $0 \leq \lambda_{c1} \leq \lambda_{c2} \leq \infty$ such that

- $\Psi^{\lambda,\{s\}}$ dies out if $\lambda < \lambda_{c1}$
- $\Psi^{\lambda,\{s\}}$ survives weakly if $\lambda_{c1} < \lambda < \lambda_{c2}$
- $\Psi^{\lambda,\{s\}}$ survives strongly if $\lambda > \lambda_{c2}$.

The above description of the contact process with flip rate intensities chosen according to (9) is standard. However for our purposes it is more convenient to use the following flip rate intensities. Let $\delta > 0$ and

$$C(s,\sigma) = \begin{cases} \delta & \text{if } \sigma(s) = 1 \\ \sum_{(s',s)\in E} \sigma(s') & \text{if } \sigma(s) = 0. \end{cases} \tag{10}$$

This is just a time scaling of the original model. We will denote the upper invariant measure by ν_δ and the corresponding process starting with distribution ν_δ by Ψ_δ. If we instead choose to start with a specific set $A \subset S$ of infected individuals, we denote the corresponding process by Ψ_δ^A. We will let the distribution of the process at time $t \geq 0$ be denoted by $\nu_{\delta,t}^A$. At some point we need to consider the process $\Psi_\delta^{\lambda,A}$, this is exactly like the model just described except for a λ inserted in front of the sum in equation (10).

As above, it follows that there exists $0 \leq \delta_{c1} \leq \delta_{c2} \leq \infty$ such that

- $\Psi_\delta^{\{s\}}$ dies out if $\delta > \delta_{c2}$
- $\Psi_\delta^{\{s\}}$ survives weakly if $\delta_{c1} < \delta < \delta_{c2}$
- $\Psi_\delta^{\{s\}}$ survives strongly if $\delta < \delta_{c1}$.

We point out that on $G = \mathbb{Z}^d$ it is known (see [3]) that $\delta_{c1} = \delta_{c2}$. It is also well known (see [10]) that on any homogeneous tree of degree larger than or equal to 2, this is not the case.

1.1.2 CPREE

This model is a pair of processes $\{(B_t, Y_t)\}_{t\geq 0}$ with state space $\{\{0,1\} \times \{0,1\}\}^S$. The second coordinate of $\{(B_t, Y_t)\}_{t\geq 0}$ will represent whether an individual is infected or not, while the first coordinate will represent how prone the individual is to recover. With a slight abuse of notation we have chosen the first coordinate to be denoted by $\{B_t\}_{t\geq 0}$ even though a process with this notation was already defined previously in the introduction. However, at every site $s \in S$, the marginal of the $\{B_t\}_{t\geq 0}$ process defined in this subsection (denoted by $\{B_t(s)\}_{t\geq 0}$) will be independent of the rest of the $\{B_t\}_{t\geq 0}$ process defined here, and have distribution according to the process with the same notation defined earlier. It will be clear from context which of these two we are referring to.

For any $A \subset S$, let $Y_0(s) = 1$ iff $s \in A$, and let $B_0 \sim \pi_p$. For $0 \leq \delta_0 < \delta_1 < \infty$, $\gamma > 0$ and $p \in [0,1]$, let the flip rate intensities $C(s, (B_t, Y_t))$ of a site $s \in S$ be

from	to	with intensity
$(0,0)$	$(0,1)$	$\sum_{(s',s)\in E} Y_t(s')$

$(0,0)$	$(1,0)$	γp
$(0,1)$	$(0,0)$	δ_0
$(0,1)$	$(1,1)$	γp
$(1,0)$	$(0,0)$	$\gamma(1-p)$
$(1,0)$	$(1,1)$	$\sum_{(s',s)\in E} Y_t(s')$
$(1,1)$	$(0,1)$	$\gamma(1-p)$
$(1,1)$	$(1,0)$	δ_1.

Denote the distribution of $\{Y_t\}_{t\geq 0}$ by $\Psi^{\gamma,p,A}_{\delta_0,\delta_1}$ and the distribution at a fixed time t by $\nu^{\gamma,p,A}_{\delta_0,\delta_1,t}$. The definition of dying out, surviving weakly and surviving strongly is the same as for the ordinary contact process. At this point a question naturally arises. For fixed $\delta_0, \delta_1, \gamma$ and p, do the initial state of the background process have any effect on this definition? This point is raised in Section 6, where we list some open questions. Note that we are here assuming that $B_0 \sim \pi_p$ which then is included in the definition.

We will write

$$\Psi^{\gamma,p,A}_{\delta_0,\delta_1} \preceq \Psi^A_\delta$$

when we mean that there exists a process $\{Y_t\}_{t\geq 0}$ with distribution as above and a process $\{Y'_t\}_{t\geq 0}$ with distribution Ψ^A_δ coupled such that

$$Y_t(s) \leq Y'_t(s) \quad \forall s \in S \text{ and } \forall t \geq 0,$$

and use the obvious notation for all similar types of situations. This stochastic ordering also implies that

$$\nu^{\gamma,p,A}_{\delta_0,\delta_1,t} \preceq \nu^A_{\delta,t} \quad \forall t \geq 0.$$

It is easy to show that $\Psi^{\gamma,p,A}_{\delta_0,\delta_1}$ is in this sense stochastically decreasing in p. We have already introduced this notation for continuous time processes in (4) and (5). There it was a relation between jump processes indexed by $t \geq 0$, while here it is a relation between processes with state space $\{0,1\}^S$. It will be clear from the context which one we are referring to. It will be useful to observe that with this definition, the recovery process at every site of our CPREE is in fact a $\mathrm{Poi}^{\gamma,p}_{\delta_0,\delta_1}$ process as defined earlier. This explains the relation between Theorem 1.4 and our next result.

We will let Δ_G denote the maximum degree of a graph G of bounded degree. We can now list our main results concerning this model:

Theorem 1.6 *Let $G=(S,E)$ be any graph of bounded degree and $A \subset S$, be such that $|A| < \infty$. For any $\delta < \min(\delta_1, \delta_0+\gamma)$ there exists a $p = p(\delta,\delta_0,\delta_1,\gamma) \in (0,1)$ large enough so that*

$$\Psi^{\gamma,p,A}_{\delta_0,\delta_1} \preceq \Psi^A_\delta.$$

Remark: To prove this theorem we couple the recovery rates of the CPREE with the recovery rates of an ordinary contact process. One would perhaps hope to get this result for any $\delta < \delta_1$, rather then for $\delta < \min(\delta_1, \delta_0 + \gamma)$. However, the specific coupling here cannot in general be done for every $\delta < \delta_1$.

Our next result also uses Theorem 1.4. However, it does so in a different way. The reason for this is that a straightforward approach would need a result for $\lambda_{\min,\mu}(\alpha_0, \alpha_1, \gamma, p)$ analogous to the one we have for $\lambda_{\max,\mu}(\alpha_0, \alpha_1, \gamma, p)$. However this is false since $\lambda_{\min,\mu}(\alpha_0, \alpha_1, \gamma, p)$ is equal to α_0 for any $p > 0$. Here, let $\Psi^{\gamma,p,A}_{\delta_0,\delta_1,B_0(A)\equiv 0}$ denote the distribution of $\Psi^{\gamma,p,A}_{\delta_0,\delta_1}$ conditioned on the event that $B_0(s) = 0$ for every $s \in A$.

Theorem 1.7 *Let $G = (S, E)$ be any graph of bounded degree. Let $A \subset S$, be such that $|A| < \infty$ and $\gamma \geq \Delta_G$. For any choice of $\delta > \delta_0$ and $\lambda < 1$ there exists a $p = p(\delta, \lambda, \delta_0, \delta_1, \gamma) \in (0, 1)$ small enough so that*

$$\Psi^{\lambda,A}_{\delta} \preceq \Psi^{\gamma,p,A}_{\delta_0,\delta_1,B_0(A)\equiv 0}.$$

Remarks: It is unfortunate that we need the assumption that $B_0(s) = 0$ for every $s \in A$. However, this is of no importance when we later apply the theorem to prove Theorem 1.8 stated below. Furthermore, it would be nice if we could remove the assumption $\gamma \geq \Delta_G$. However we do not conjecture whether this is possible or not. Also, as $\lambda \to 1$, the proof requires that $p \to 0$. Finally, for $\lambda = 1$ the proof does not work. Again we do not conjecture whether the result is valid for $\lambda = 1$ or not.

We are now ready to state the main theorem concerning the CPREE model of this paper. Results $1 - 3$ use an easy coupling argument while $4 - 6$ follow from applications of Theorems 1.6 and 1.7. Here, any statements similar to $p_{c1} < p < p_{c2}$ in the case that $p_{c1} = p_{c2}$ should be interpreted as empty statements.

Theorem 1.8 *Let $s \in S$, $0 \leq \delta_0 \leq \delta_1 < \infty$ and consider the process $\Psi^{\gamma,p,\{s\}}_{\delta_0,\delta_1}$. We have the following results:*

1. *Assume that $\delta_{c1} < \delta_0 < \delta_{c2} < \delta_1$. There exists $p_{c2} = p_{c2}(\delta_0, \delta_1, \gamma) \in [0, 1]$ such that $\Psi^{\gamma,p,\{s\}}_{\delta_0,\delta_1}$ dies out if $p > p_{c2}$ and survives weakly if $p < p_{c2}$.*

2. *Assume that $\delta_0 < \delta_{c1} \leq \delta_{c2} < \delta_1$. There exists $p_{c2} = p_{c2}(\delta_0, \delta_1, \gamma) \in [0, 1]$ and $p_{c1} = p_{c1}(\delta_0, \delta_1, \gamma) \in [0, 1]$ such that $p_{c1} \leq p_{c2}$ and $\Psi^{\gamma,p,\{s\}}_{\delta_0,\delta_1}$ dies out if $p > p_{c2}$ survives weakly if $p_{c1} < p < p_{c2}$ and survives strongly if $p < p_{c1}$.*

3. *Assume that* $\delta_0 < \delta_{c1} < \delta_1 < \delta_{c2}$. *There exists* $p_{c1} = p_{c1}(\delta_0, \delta_1, \gamma) \in [0,1]$ *such that* $\Psi^{\gamma,p,\{s\}}_{\delta_0,\delta_1}$ *survives strongly if* $p < p_{c1}$ *and survives weakly if* $p > p_{c1}$.

4. *In case number 1, if* $\gamma > \delta_{c2} - \delta_0$ *then* $p_{c2} < 1$ *and if* $\gamma \geq \Delta_G$, *then* $p_{c2} > 0$.

5. *In case number 2, if* $\gamma > \delta_{c2} - \delta_0$ *then* $p_{c2} < 1$, *if* $\gamma > \delta_{c1} - \delta_0$ *then* $p_{c1} < 1$ *and if* $\gamma \geq \Delta_G$, *then* $p_{c1}, p_{c2} > 0$.

6. *In case number 3, if* $\gamma > \delta_{c1} - \delta_0$ *then* $p_{c1} < 1$ *and if* $\gamma \geq \Delta_G$, *then* $p_{c1} > 0$.

Remarks:

- We do not include trivial cases like $\delta_{c1} < \delta_0 < \delta_1 < \delta_{c2}$ in the statement.

- One might suspect that the condition $\gamma > \delta_{c2} - \delta_0$ should in fact be $\gamma > \delta_1 - \delta_0$, considering the statement of Theorem 1.6. The point is however that we must only be able to choose the δ of Theorem 1.6 to be larger than δ_{c2}, not δ_1.

- We would like to point out that even if we only show that $p_{c1}, p_{c2} < 1$ whenever $\gamma > \delta_{c1} - \delta_0$, $\gamma > \delta_{c2} - \delta_0$ respectively, there is no apparent reason why this should not be true for all $\gamma > 0$. Similarly for $p_{c1}, p_{c2} > 0$.

- By using the renormalization methods of [3] it seems like it is possible to show that on $\mathbb{Z}^d$, for every $\gamma > 0$, $p_{c1}, p_{c2} > 0$. However, since such a proof would need to explain substantial parts of the renormalization argument and is a result only valid for $\mathbb{Z}^d$ we choose not to explore it further.

The rest of the paper is organized as follows. Proposition 1.2 is proved in Section 2. This is then used to prove Theorem 1.3 in Section 3. In Section 4 we use a limiting argument to conclude Theorem 1.4 from Theorem 1.3. We then exploit Theorem 1.4 to prove Theorems 1.6 and 1.7 in Section 5. Finally, these last two theorems, will be used to prove our main result, Theorem 1.8 in the same section.

We exploit the techniques for proving Theorem 1.6 further to conclude the following results concerning p_{c1} and p_{c2}.

Proposition 1.9 *Fix* $i \in \{1,2\}$ *and assume that* $\delta_0 < \delta_{ci}$. *We have for* p_{ci},

$$\limsup_{\gamma \to \infty} p_{ci}(\delta_0, \delta_1, \gamma) \leq \frac{\delta_{ci} - \delta_0}{\delta_1 - \delta_0}.$$

Remark: We conjecture that the limit exists and that

$$\lim_{\gamma\to\infty} p_{ci}(\delta_0,\delta_1,\gamma) = \frac{\delta_{ci}-\delta_0}{\delta_1-\delta_0}.$$

We cannot prove this with the techniques of this paper; this is closely related to the remarks after Theorem 1.4. However, the intuition why it should be true is that as $\gamma \to \infty$, the recovery process should become increasingly similar to an ordinary Poisson($\delta_0 + p(\delta_1 - \delta_0)$)-process (see the remarks of Theorem 1.4). In turn, our CPREE then should become more and more like an ordinary contact process with recovery rate $\delta_0 + p(\delta_1 - \delta_0)$. Therefore we should get that p_c solves the equation $\delta_c = \delta_0 + p(\delta_1 - \delta_0)$.

As γ tends to 0 we can unfortunately not conclude anything about p_{ci}. The reason is yet again connected to the remark after Theorem 1.4. We know that for $\gamma = 0$, $\lambda_{\min,\mu}(\alpha_0, \alpha_1, 0, p) = \alpha_1$ from Theorem 1.4 and that $\lambda_{\max,\mu}(\alpha_0, \alpha_1, 0, p) = \alpha_0$. Therefore the stochastic domination techniques we use in this paper do not yield any nontrivial results. We also point out that the case $\gamma = 0$, corresponds to the CPRE and we therefore refer to the papers mentioned in the first paragraph of the introduction.

We also have the following easy result about p_{c1}, p_{c2}.

Proposition 1.10 *We have that for any $\gamma > 0$ and $\delta_1 > \delta_{ci} > \delta_0$, where $i \in \{1, 2\}$*

$$\lim_{\delta_0\uparrow\delta_{ci}} p_{ci}(\delta_0,\delta_1,\gamma) = 0.$$

Remark: One would of course expect that

$$\lim_{\delta_1\downarrow\delta_{ci}} p_{ci}(\delta_0,\delta_1,\gamma) = 1.$$

However, it is not possible to prove this the same way as we prove Proposition 1.10; again this is a fact that propagates from Theorem 1.4.

2 Proof of Proposition 1.2

The proof of Theorem 1.2 will require the following two results, the first uses Lemma 3.2 of [5] and the second is a restatement (which is more suitable for our purposes) of Theorem 1.2 of [11].

Lemma 2.1 *If $\{B_n\}_{n=1}^{\infty}$ is monotone then $\{X_n\}_{n=1}^{\infty}$ is monotone.*

Proof Let $\{Z_n\}_{n=1}^{\infty} \sim \pi_{1-\frac{1-\alpha_1}{1-\alpha_0}}$ and $\{Z'_n\}_{n=1}^{\infty} \sim \pi_{\alpha_0}$ be independent. Observe that $\{X_n\}_{n=1}^{\infty}$ has the same distribution as $\{\max(\min(B_n, Z_n), Z'_n)\}_{n=1}^{\infty}$. It follows from Lemma 3.2 of [5] that $\{\min(B_n, Z_n)\}_{n=1}^{\infty}$ is monotone. It then follows similarly that $\{\max(\min(B_n, Z_n), Z'_n)\}_{n=1}^{\infty}$ is monotone.

QED

Lemma 2.2 *Let μ be a translation invariant measure on $\{0,1\}^{\mathbb{N}}$ which is monotone. Then the following two statements are equivalent.*

1.

$$\pi_p \preceq \mu$$

2. For any $n \in \mathbb{N}$

$$\mu(\sigma(n) = 1|\sigma(1,\ldots,n-1) \equiv 0) \geq p.$$

Proof of Proposition 1.2. Let

$$A_n := \mathbb{P}(X_n = 1|X_{n-1} = \cdots = X_1 = 0).$$

Since $\{X_n\}_{n=1}^{\infty}$ is monotone (Lemma 2.1) and translation invariant it is easy to see that A_n is decreasing in n, and therefore the limit $A = \lim_{n\to\infty} A_n$ exists. It is now an easy consequence of Lemma 2.2 that this limit is equal to $p_{\max,\mu}$.

QED

The above results show that when the assumptions of the theorem hold then

$$\inf_{n\in\mathbb{N},\xi\in\{0,1\}^{n-1}} \mathbb{P}(X_n = 1|(X_{n-1},\ldots,X_1) \equiv \xi) = p_{\max,\mu}.$$

It is very easy to find examples for which this statement is not true. For instance let $(X,Y) \in \{0,1\} \times \{0,1\}$ and $\mathbb{P}(X = Y = 1) = \mathbb{P}(X = Y = 0) = 1/2$. This dominates a product measure with positive density but in this case $p_{\max,\mu} = 0$.

3 Discrete time domination results

This section is devoted to the proof of Theorem 1.3. $\{(B_n, X_n)\}_{n=1}^{\infty}$ are the processes defined in the introduction. We start with the following lemma; we do not include the elementary proof.

Lemma 3.1 *The process $\{B_n\}_{n=1}^{\infty}$ is monotone.*

We will also need the following lemma which gives us a recursion formula of A_n expressed in terms of A_{n-1}.

Lemma 3.2 *We have that*

$$A_n = \frac{CA_{n-1} + D}{1 - A_{n-1}}, \tag{11}$$

with C, D as in Theorem 1.3

Proof. The proof is straightforward, however it involves some tedious calculations. We have

$$\begin{aligned} A_n &= \frac{\mathbb{P}(X_n = 1, X_{n-1} = 0 | (X_{n-2}, \dots, X_1) \equiv 0)}{\mathbb{P}(X_{n-1} = 0 | (X_{n-2}, \dots, X_1) \equiv 0)} \\ &= \frac{\mathbb{P}(X_n = 1, X_{n-1} = 0 | (X_{n-2}, \dots, X_1) \equiv 0)}{1 - A_{n-1}}. \end{aligned}$$

Observe that

$$\begin{aligned} &\mathbb{P}(X_n = 1 | (X_{n-1}, \dots, X_1) \equiv 0) \\ &\quad = \mathbb{P}(X_n = 1 | B_n = 1, (X_{n-1}, \dots, X_1) \equiv 0) \mathbb{P}(B_n = 1 | (X_{n-1}, \dots, X_1) \equiv 0) \\ &\quad\quad + \mathbb{P}(X_n = 1 | B_n = 0, (X_{n-1}, \dots, X_1) \equiv 0) \mathbb{P}(B_n = 0 | (X_{n-1}, \dots, X_1) \equiv 0) \\ &\quad = \alpha_1 \mathbb{P}(B_n = 1 | (X_{n-1}, \dots, X_1) \equiv 0) \\ &\quad\quad + \alpha_0 (1 - \mathbb{P}(B_n = 1 | (X_{n-1}, \dots, X_1) \equiv 0)) \\ &\quad = \alpha_0 + (\alpha_1 - \alpha_0) \mathbb{P}(B_n = 1 | (X_{n-1}, \dots, X_1) \equiv 0). \end{aligned}$$

Therefore

$$\begin{aligned} &\mathbb{P}(B_n = 1 | (X_{n-1}, \dots, X_1) \equiv 0) \\ &\quad = \frac{\mathbb{P}(X_n = 1 | (X_{n-1}, \dots, X_1) \equiv 0) - \alpha_0}{\alpha_1 - \alpha_0} = \frac{A_n - \alpha_0}{\alpha_1 - \alpha_0}. \end{aligned}$$

Furthermore, using the above we get

$$\begin{aligned} &\mathbb{P}(X_n = 1, X_{n-1} = 0 | (X_{n-2}, \dots, X_1) \equiv 0) \\ &\quad = \mathbb{P}(X_n = 1, X_{n-1} = 0 | B_{n-1} = 1, (X_{n-2}, \dots, X_1) \equiv 0) \\ &\quad\quad \times \mathbb{P}(B_{n-1} = 1 | (X_{n-2}, \dots, X_1) \equiv 0) \\ &\quad\quad + \mathbb{P}(X_n = 1, X_{n-1} = 0 | B_{n-1} = 0, (X_{n-2}, \dots, X_1) \equiv 0) \\ &\quad\quad \times \mathbb{P}(B_{n-1} = 0 | (X_{n-2}, \dots, X_1) \equiv 0) \\ &\quad = \mathbb{P}(X_n = 1 | X_{n-1} = 0, B_{n-1} = 1) \\ &\quad\quad \times \mathbb{P}(X_{n-1} = 0 | B_{n-1} = 1) \frac{A_{n-1} - \alpha_0}{\alpha_1 - \alpha_0} \\ &\quad\quad + \mathbb{P}(X_n = 1 | X_{n-1} = 0, B_{n-1} = 0) \\ &\quad\quad \times \mathbb{P}(X_{n-1} = 0 | B_{n-1} = 0) \left(1 - \frac{A_{n-1} - \alpha_0}{\alpha_1 - \alpha_0}\right) \\ &\quad = [\alpha_1(1 - \gamma(1-p)) + \alpha_0 \gamma (1-p)](1 - \alpha_1) \frac{A_{n-1} - \alpha_0}{\alpha_1 - \alpha_0} \\ &\quad\quad + [\alpha_1 \gamma p + \alpha_0 (1 - \gamma p)](1 - \alpha_0) \left(1 - \frac{A_{n-1} - \alpha_0}{\alpha_1 - \alpha_0}\right) \\ &\quad = \frac{A_{n-1} - \alpha_0}{\alpha_1 - \alpha_0} \Big([\alpha_1(1 - \gamma(1-p)) + \alpha_0 \gamma (1-p)](1 - \alpha_1) \\ &\quad\quad - [\alpha_1 \gamma p + \alpha_0 (1 - \gamma p)](1 - \alpha_0) \Big) + [\alpha_1 \gamma p + \alpha_0 (1 - \gamma p)](1 - \alpha_0). \end{aligned}$$

Finally observing that

$$
\begin{aligned}
&\frac{1}{\alpha_1-\alpha_0}\Big[[\alpha_1(1-\gamma(1-p)+\alpha_0\gamma(1-p)](1-\alpha_1)\\
&\quad-[\alpha_1\gamma p+\alpha_0(1-\gamma p)](1-\alpha_0)\Big]\\
&\quad=\frac{1}{\alpha_1-\alpha_0}\Big[[\alpha_1-(\alpha_1-\alpha_0)\gamma(1-p)]\\
&\quad(1-\alpha_1)-[(\alpha_1-\alpha_0)\gamma p+\alpha_0](1-\alpha_0)\Big]\\
&\quad=\frac{1}{\alpha_1-\alpha_0}\Big[[\alpha_1(1-\alpha_1)-\alpha_0(1-\alpha_0)\Big]-\gamma(1-p)(1-\alpha_1)-\gamma p(1-\alpha_0)\\
&\quad=(1-\alpha_0-\alpha_1)-\gamma(1-p)(1-\alpha_1)-\gamma p(1-\alpha_0)\\
&\quad=(1-\alpha_0-\alpha_1)-\gamma[1-\alpha_0-(1-p)(\alpha_1-\alpha_0)],
\end{aligned}
$$

and that

$$
\begin{aligned}
&-\frac{\alpha_0}{\alpha_1-\alpha_0}\Big([\alpha_1(1-\gamma(1-p))+\alpha_0\gamma(1-p)](1-\alpha_1)\\
&\qquad-[\alpha_1\gamma p+\alpha_0(1-\gamma p)](1-\alpha_0)\Big)+[\alpha_1\gamma p+\alpha_0(1-\gamma p)](1-\alpha_0)\\
&\quad=\frac{\alpha_1}{\alpha_1-\alpha_0}[\alpha_1\gamma p+\alpha_0(1-\gamma p)](1-\alpha_0)\\
&\qquad-\frac{\alpha_0}{\alpha_1-\alpha_0}[\alpha_1(1-\gamma(1-p))+\alpha_0\gamma(1-p)](1-\alpha_1)\\
&\quad=\frac{1}{\alpha_1-\alpha_0}\Big[[\alpha_0+(\alpha_1-\alpha_0)\gamma p](1-\alpha_0)\alpha_1\\
&\qquad-[\alpha_1-(\alpha_1-\alpha_0)\gamma(1-p)](1-\alpha_1)\alpha_0\Big]\\
&\quad=\frac{1}{\alpha_1-\alpha_0}\Big[\alpha_0(1-\alpha_0)\alpha_1-\alpha_1(1-\alpha_1)\alpha_0\Big]\\
&\qquad+\gamma p(1-\alpha_0)\alpha_1+\gamma(1-p)(1-\alpha_1)\alpha_0\\
&\quad=\alpha_0\alpha_1+\gamma p(1-\alpha_0)\alpha_1+\gamma(1-p)(1-\alpha_1)\alpha_0\\
&\quad=\alpha_0\alpha_1+\gamma[(p-\alpha_0+(1-p)\alpha_0)\alpha_1+(1-p)\alpha_0-(1-p)\alpha_0\alpha_1]\\
&\quad=\alpha_0\alpha_1+\gamma[\alpha_1(1-\alpha_0)-(1-p)(\alpha_1-\alpha_0)],
\end{aligned}
$$

completes the proof.

QED

We are now ready to prove Theorem 1.3.
Proof of Theorem 1.3. From Proposition 1.2 we know that the limit $A=\lim_{n\to\infty}A_n$ exists, and therefore we can take the limit of both sides of equation (11) (A_n is easily seen to be uniformly bounded away from 1) to

conclude that

$$A = \lim_{n\to\infty} A_n = \lim_{n\to\infty} \frac{CA_{n-1} + D}{1 - A_{n-1}} = \frac{CA + D}{1 - A}.$$

This gives us that

$$A - A^2 = CA + D,$$

and therefore

$$A^2 + (C-1)A + D = 0,$$

solving this equation gives

$$\begin{aligned} A &= \frac{1}{2}\left(1 - C \pm \sqrt{(1-C)^2 - 4D}\right) \\ &= \frac{1}{2}\left(1 - C \pm \sqrt{(2\alpha_1 - (1-C))^2 - 4(D + \alpha_1^2 - \alpha_1(1-C))}\right). \end{aligned}$$

We will now proceed to rule out one of the solutions.

$$\begin{aligned} & D + \alpha_1^2 - \alpha_1(1-C) \\ &\quad = \alpha_1^2 - \alpha_1 + \alpha_0\alpha_1 + \gamma(\alpha_1(1-\alpha_0) - (1-p)(\alpha_1 - \alpha_0)) \\ &\qquad + \alpha_1((1 - \alpha_0 - \alpha_1) - \gamma(1 - \alpha_0 - (1-p)(\alpha_1 - \alpha_0))) \\ &\quad = \gamma(\alpha_1(1-\alpha_0) - (1-p)(\alpha_1 - \alpha_0)) - \gamma\alpha_1(1 - \alpha_0 - (1-p)(\alpha_1 - \alpha_0)) \\ &\quad = \gamma(-(1-p)(\alpha_1 - \alpha_0) + (1-p)\alpha_1(\alpha_1 - \alpha_0)) \\ &\quad = -\gamma(1-p)(\alpha_1 - \alpha_0)(1 - \alpha_1). \end{aligned}$$

Using that $\gamma(1-p)((\alpha_1 - \alpha_0)(1-\alpha_1)) \geq 0$ we get

$$\begin{aligned} & \frac{1}{2}\left(1 - C + \sqrt{(2\alpha_1 - (1-C))^2 - 4(D + \alpha_1^2 - \alpha_1(1-C))}\right) \\ &\quad \geq \frac{1}{2}\left(1 - C + \sqrt{(2\alpha_1 - (1-C))^2}\right) \geq \frac{1}{2}(1 - C + (2\alpha_1 - (1-C))) = \alpha_1. \end{aligned}$$

Obviously we cannot have

$$A \geq \alpha_1,$$

since already (for $\gamma, p \in (0,1)$)

$$A_2 = \mathbb{P}(X_2 = 1 | X_1 = 0) < \alpha_1$$

and $A \leq A_n$ for every n. We conclude that

$$A = \frac{1}{2}\left(1 - C - \sqrt{(1-C)^2 - 4D}\right).$$

Using Proposition 1.2 we then conclude that $p_{\max,\mu} = A$. Finally, the result for $p_{\min,\mu}$ follows from an easy symmetry argument.

QED

Observe that when $\alpha_0 = \alpha_1 = \alpha$, $\{X_n\}_{n=1}^{\infty}$ is i.i.d. and $p_{\max,\mu} = p_{\min,\mu} = \alpha$. Note that in this case $C = 1 - 2\alpha - \gamma(1-\alpha)$ and $D = \alpha^2 + \gamma\alpha(1-\alpha)$, and so

$$\begin{aligned} & p_{\max,\mu} \\ & \quad = \frac{1}{2}\left(2\alpha + \gamma(1-\alpha) - \sqrt{(2\alpha+\gamma(1-\alpha))^2 - 4(\alpha^2+\gamma\alpha(1-\alpha)}\right) \\ & \quad = \frac{1}{2}\left(2\alpha + \gamma(1-\alpha) - \sqrt{(\gamma(1-\alpha))^2}\right) = \alpha, \end{aligned}$$

as we should get. Similarly one can check that that $p_{\min,\mu} = \alpha$.

Furthermore if we choose $\gamma = 1$, $\{X_n\}_{n=1}^{\infty}$ is again i.i.d. and we would expect to get that $p_{\max,\mu} = p_{\min,\mu} = \alpha_0 + p(\alpha_1 - \alpha_0)$. Again, this is easy to check. Finally, as $\gamma \to 0$ we get that $p_{\max,\mu} \to \alpha_0$ and $p_{\min,\mu} \to \alpha_1$. It is not hard to see why this is what we should expect.

4 Continuous time domination results

In this section we prove Theorem 1.4.

For $T > 0$, let $D_{\mathbb{N}}[0,T]$ be the set of functions from $[0,T]$ to $\mathbb{N}$ that are right-continuous and have left limits. Let $D_{\mathbb{N}}[0,\infty)$ be defined in the same way, but with $[0,T]$ replaced by $[0,\infty)$. Let a function be called a count path if it is a non-decreasing function that takes integer values and has jumps of size 1. Define $D_c \subset D_{\mathbb{N}}[0,1]$ to be the set of count paths. D_c is closed under the Skorokhod topology, see [4] pg. 137.

Let $\alpha_0, \alpha_1, \gamma > 0$ and let m be such that $\alpha_{0,m} := \alpha_0/m$, $\alpha_{1,m} := \alpha_1/m$, $\gamma_m := \gamma/m \in (0,1)$. Consider the model in the last section with α_0, α_1 and γ replaced by $\alpha_{0,m}, \alpha_{1,m}$ and γ_m respectively (p is not changed). Denote the corresponding processes by $\{(B_n^m, X_n^m)\}_{n=1}^{\infty}$ but consider only the truncated part $\{(B_n^m, X_n^m)\}_{n=1}^{m}$. As in the introduction, let

$$X_n^{c,m} = \sum_{i=1}^{n} X_i^m \text{ for } n \in \{1, \cdots, m\}.$$

Define the continuous time version $\{(B_t^m, X_t^m)\}_{t\in[0,1]}$ by letting

$$(B_t^m, X_t^m) = (B_n^m, X_n^{c,m}) \text{ for } t \in [n-1, n)/m \text{ and } n \in \{1, \cdots, m\}, \tag{12}$$

and $(B_{t=1}^m, X_{t=1}^m) = (B_m^m, X_m^{c,m})$. According to Theorem 1.3, we can couple the $\{(B_n^m, X_n^m)\}_{n=1}^{m}$ process with an i.i.d. process $\{Y_n^m\}_{n=1}^{m}$ with density $p_{\max,\mu_m}$ (where μ_m denotes the distribution of $\{X_n^m\}_{n=1}^{\infty}$) such that

$$Y_n^m \leq X_n^m \quad \forall n \in \{1, \cdots, m\}. \tag{13}$$

Here $p_{\max,\mu_m}$ is given by Theorem 1.3. Define $\{Y_n^{c,m}\}_{n=1}^{\infty}$ in the obvious way and the continuous time version $\{Y_t^m\}_{t\in[0,1]}$ by letting

$$Y_t^m = Y_n^{c,m} \text{ for } t \in [n-1,n)/m, \quad n \in \{1,\cdots,m\}$$

and $Y_{t=1}^m = Y_m^{c,m}$. We get from equation (13) that

$$X_t^m - Y_t^m \text{ is non-decreasing in } t \ \ \forall n \in \{1,\cdots,m\}. \tag{14}$$

We state the following lemma; the proof is an elementary exercise in convergence in the Skorokhod topology.

Lemma 4.1 *The set $\{(f,g) \in D_c \times D_c : f-g$ is non-decreasing$\}$ is closed in the product Skorokhod topology.*

Consider now $\{(B_t, X_t)\}_{t\in[0,1]}$ defined in section 1. It is easy to see that the flip rate intensities corresponding to $\{(B_t, X_t)\}_{t\in[0,1]}$ are

from	to	with intensity
$(0,k)$	$(1,k)$	γp
$(0,k)$	$(0,k+1)$	α_0
$(1,k)$	$(0,k)$	$\gamma(1-p)$
$(1,k)$	$(1,k+1)$	α_1,

(15)

for any $k \geq 0$. Observe that for $\{(B_n^m, X_n^{c,m})\}_{n=1}^{\infty}$ we have the transition probabilities

from	to	with probability
$(0,k)$	$(0,k)$	$(1-\alpha_0/m)(1-\gamma p/m)$
$(0,k)$	$(1,k)$	$(\gamma p/m)(1-\alpha_1/m)$
$(0,k)$	$(0,k+1)$	$(1-\gamma p/m)\alpha_0/m$
$(0,k)$	$(1,k+1)$	$\gamma p \alpha_1/m^2$
$(1,k)$	$(1,k)$	$(1-\alpha_1/m)(1-\gamma(1-p)/m)$
$(1,k)$	$(0,k)$	$(\gamma/m)(1-p)(1-\alpha_0/m)$
$(1,k)$	$(1,k+1)$	$(\alpha_1/m)(1-\gamma(1-p)/m)$
$(1,k)$	$(0,k+1)$	$\gamma(1-p)\alpha_0/m^2$.

(16)

Using the flip rate intensities of equations (15) and (16), it is a standard result to show the next lemma. Again we omit the proof. However see for instance [6] for a survey on the convergence of Markov processes in general.

Lemma 4.2 *The sequence of processes $\{(B_t^m, X_t^m)\}_{t\in[0,1]}$ defined above and indexed by m, converges weakly to the Markov process $\{(B_t, X_t)\}_{t\in[0,1]}$.*

QED

We are now ready to prove our main results of this section. We will start by proving the following lemma.

Lemma 4.3 *With the assumptions of Theorem 1.4, we have that*

$$\lambda_{\max,\mu}(\alpha_0,\alpha_1,\gamma,p) \geq \bar{\lambda}.$$

Proof We will start by constructing the coupling on the finite time-interval $[0,1]$ and then argue that we can extend it to infinite time.

Let $\{(B_t^m, X_t^m, Y_t^m)\}_{t\in[0,1]}$ be any sequence of processes indexed by m where, as indicated by the notation, the marginals $\{(B_t^m, X_t^m)\}_{t\in[0,1]}$ and $\{Y_t^m\}_{t\in[0,1]}$ have the distribution of the processes defined at the beginning of this section. Furthermore assume that these marginals are coupled so that $X_t^m - Y_t^m$ is non-decreasing for every m. Obviously the marginal $\{Y_t^m\}_{t\in[0,1]}$ converges weakly to a Poisson process $\{Y_t\}_{t\in[0,1]}$ with intensity

$$\begin{aligned}
&\lim_{m\to\infty} mp_{\max,\mu_m} \\
&\quad = \lim_{m\to\infty}\frac{1}{2}(\alpha_0+\alpha_1+\gamma(1-\frac{1}{m}\alpha_0-(1-p)\frac{1}{m}(\alpha_1-\alpha_0))) \\
&\quad -\frac{1}{2}\Big((\alpha_0+\alpha_1+\gamma(1-\frac{1}{m}\alpha_0-(1-p)\frac{1}{m}(\alpha_1-\alpha_0)))^2 \\
&\quad -4(\alpha_0\alpha_1+\gamma(\alpha_1(1-\frac{1}{m}\alpha_0)-(1-p)(\alpha_1-\alpha_0)))\Big)^{1/2} \\
&\quad = \frac{1}{2}\left(\alpha_0+\alpha_1+\gamma-\sqrt{(\alpha_0+\alpha_1+\gamma)^2-4(\alpha_0\alpha_1+\gamma(\alpha_0+p(\alpha_1-\alpha_0)))}\right) \\
&\quad = \frac{1}{2}\left(\alpha_0+\alpha_1+\gamma-\sqrt{(\alpha_1-\alpha_0-\gamma)^2+4\gamma(1-p)(\alpha_1-\alpha_0)}\right) = \bar{\lambda}.
\end{aligned}$$

Lemma 4.2 shows that also the sequence $\{(B_t^m, X_t^m)\}_{t\in[0,1]}$ converges weakly. It can then be argued that the sequence $\{(B_t^m, X_t^m, Y_t^m)\}_{t\in[0,1]}$ is tight and so there exists a subsequence $\{\{(B_t^{m(k)}, X_t^{m(k)}, Y_t^{m(k)})\}_{t\in[0,1]}\}_{k=1}^{\infty}$ that converges weakly to some process $\{(\tilde{B}_t,\tilde{X}_t,\tilde{Y}_t)\}_{t\in[0,1]}$. Of course, the marginal distribution $\{(\tilde{B}_t,\tilde{X}_t)\}_{t\in[0,1]}$ must be equal to the distribution of $\{(B_t,X_t)\}_{t\in[0,1]}$, and the marginal distribution $\{\tilde{Y}_t\}_{t\in[0,1]}$ must be equal to the distribution of $\{Y_t\}_{t\in[0,1]}$. Furthermore using Lemma 4.1 we conclude that

$$X_t - Y_t \text{ is non-decreasing.} \tag{17}$$

It not hard to see that we can adapt the proof to work for any time-interval $[0,T]$. It is then easy to construct the coupling on $D_{\mathbb{N}}[0,\infty)$. Hence we have established that

$$\lambda_{\max,\mu}(\alpha_0,\alpha_1,\gamma,p) \geq \bar{\lambda}.$$

QED

Considering $\{(B^m_n, X^m_n)\}_{n=1}^{\infty}$, let for every $m, i \geq 1$ $A^m_i := \mathbb{P}(X^m_i = 1 | X^m_{i-1} = \cdots = X^m_1 = 0)$, and let $A^m := p_{\max,\mu_m} = \lim_{i\to\infty} A^m_i$. In our next lemma, we will need that Tm (where $T > 0$) is an integer, which will not always be the case. However, adjusting the proofs for this is trivial and we therefore leave it to the reader. The same comment applies for other results to follow.

Lemma 4.4 *For any $T > 0$,*

$$\lim_{m\to\infty} mA^m_{Tm} = \bar{\lambda} + (p\alpha_1 + (1-p)\alpha_0 - \bar{\lambda})\frac{e^{-(\alpha_0+\alpha_1+\gamma)T}}{e^{-\bar{\lambda}T}\mathbb{P}(X_t = 0\ \forall t \in [0,T])}.$$

Proof Let C^m, D^m denote C, D of Theorem 1.3 with parameters α_0/m, α_1/m, γ/m and p. By Lemma 3.2, for any n,

$$\begin{aligned}
A^m_n &- A^m \\
&= \frac{C^m A^m_{n-1} + D^m}{1 - A^m_{n-1}} - \frac{C^m A^m + D^m}{1 - A^m} \\
&= \frac{(C^m A^m_{n-1} + D^m)(1 - A^m) - (C^m A^m + D^m)(1 - A^m_{n-1})}{(1 - A^m_{n-1})(1 - A^m)} \\
&= \frac{C^m (A^m_{n-1} - A^m) + D^m (A^m_{n-1} - A^m)}{(1 - A^m_{n-1})(1 - A^m)} \\
&= (A^m_{n-1} - A^m)\frac{C^m + D^m}{(1 - A^m_{n-1})(1 - A^m)} \\
&= \cdots = (A^m_1 - A^m)\left(\frac{C^m + D^m}{1 - A^m}\right)^{n-1}\frac{1}{\prod_{k=1}^{n-1}(1 - A^m_k)}.
\end{aligned}$$

Furthermore

$$\begin{aligned}
C^m &+ D^m \\
&= (1 - \alpha_0/m - \alpha_1/m) - \gamma/m(1 - \alpha_0/m - (1-p)(\alpha_1/m - \alpha_0/m)) \\
&\quad + \alpha_0\alpha_1/m^2 + \gamma/m(\alpha_1/m(1 - \alpha_0/m) - (1-p)(\alpha_1/m - \alpha_0/m)) \\
&= 1 - \alpha_0/m - \alpha_1/m + \alpha_0\alpha_1/m^2 - \gamma/m(1 - \alpha_0/m - \alpha_1/m(1 - \alpha_0/m)) \\
&= (1 - \alpha_0/m)(1 - \alpha_1/m)(1 - \gamma/m).
\end{aligned}$$

Recall also that we in Lemma 4.3 proved that $\bar{\lambda} = \lim_{m\to\infty} mA^m$. We get that

1. $\lim_{m\to\infty}(C^m + D^m)^{Tm-1} = e^{-(\alpha_0+\alpha_1+\gamma)T}$
2. $\lim_{m\to\infty}(1 - A^m)^{Tm-1} = \lim_{m\to\infty} e^{(Tm-1)\log(1-A^m)}$
   $= \lim_{m\to\infty} e^{(Tm-1)(-A^m + \mathcal{O}((A^m)^2))} = e^{-\bar{\lambda}T}$
3. $\lim_{m\to\infty}\prod_{k=1}^{Tm-1}(1 - A^m_k) = \lim_{m\to\infty}\mathbb{P}(X^m_t = 0\ \forall t \in [0, T - 1/m])$
   $= \mathbb{P}(X_t = 0\ \forall t \in [0,T])$
4. $mA^m_1 = m(p\alpha_1/m + (1-p)\alpha_0/m) = p\alpha_1 + (1-p)\alpha_0$.

Therefore

$$\begin{aligned}&\lim_{m\to\infty} mA^m_{Tm} &(18)\\ &= \lim_{m\to\infty} mA^m + (mA^m_1 - mA^m)\left(\frac{C^m + D^m}{1 - A^m}\right)^{Tm-1} \frac{1}{\prod_{k=1}^{Tm-1}(1 - A^m_k)}\\ &= \bar\lambda + (p\alpha_1 + (1-p)\alpha_0 - \bar\lambda)\frac{e^{-(\alpha_0+\alpha_1+\gamma)T}}{e^{-\bar\lambda T}\mathbb{P}(X_t = 0\ \forall t \in [0,T])},\end{aligned}$$

as desired.

QED

Next we prove the upper bound in Theorem 1.5 of $\lambda^T_{\max,\mu}(\alpha_0, \alpha_1, \gamma, p)$.

Lemma 4.5 *For every choice of $\alpha_0, \alpha_1, \gamma, T > 0$, with $\alpha_0 \le \alpha_1$ and $p \in (0,1)$ we have that there exists a constant $E > 0$, depending on $\alpha_1, \alpha_0, \gamma$ and p such that*

$$\lambda^T_{\max,\mu}(\alpha_0, \alpha_1, \gamma, p) \le \bar\lambda + \frac{1}{T}(p\alpha_1 + (1-p)\alpha_0 - \bar\lambda)\frac{1 - e^{-TE}}{E}.$$

Proof of Lemma 4.5 We have that

$$\begin{aligned}\mathbb{P}(X^m_t = 0\ \forall t \in [0,T]) &= \prod_{k=1}^{Tm}(1 - A^m_k) &(19)\\ &= e^{\sum_{k=1}^{Tm}\log(1-A^m_k)} = e^{-\sum_{k=1}^{Tm} A^m_k + \mathcal{O}((A^m_k)^2)} = e^{\mathcal{O}(1/m) - \sum_{k=1}^{Tm} A^m_k}.\end{aligned}$$

Using equation (19) it is easy to see that it suffices to get an estimate on $\sum_{k=1}^{Tm} A^m_k$. To that end, let $n > 0$ be an integer and let $T_k := kT/n$ for $k \in \{1, \dots, n\}$. Using that for fixed m, A^m_k is decreasing in k, we get that

$$\begin{aligned}\sum_{k=1}^{Tm} A^m_k &= \sum_{k=1}^{T_1 m} A^m_k + \sum_{k=T_1 m+1}^{T_2 m} A^m_k + \cdots + \sum_{k=T_{n-1}m+1}^{T_n m} A^m_k\\ &\le T_1 m A^m_1 + (T_2 - T_1) m A^m_{T_1 m} + \cdots + (T_n - T_{n-1}) m A^m_{T_{n-1}m}.\end{aligned}$$

Using equation (18), that $mA^m_1 = p\alpha_1 + (1 - p\alpha_0)$ and that $(T_k - T_{k-1}) = T/n$ for every k, we get that

$$\begin{aligned}\lim_{m\to\infty}\sum_{k=1}^{Tm} A^m_k &\le \frac{T}{n}(p\alpha_1 + (1-p)\alpha_0) &(20)\\ &+ \sum_{k=1}^{n-1}\frac{T}{n}\left[\bar\lambda + (p\alpha_1 + (1-p)\alpha_0 - \bar\lambda)\frac{e^{-(\alpha_0+\alpha_1+\gamma)T_k}}{e^{-\bar\lambda T_k}\mathbb{P}(X_t = 0\ \forall t \in [0,T_k])}\right].\end{aligned}$$

Note that the existence of this limit follows from the existence of the limit on the left hand side of equation (19). We observe that trivially $\mathbb{P}(X_t = 0 \ \forall t \in [0,T]) \geq e^{-\alpha_1 T}$ and so we get that

$$\begin{aligned}
&\frac{e^{-(\alpha_0+\alpha_1+\gamma)T_k}}{e^{-\bar{\lambda}T_k}\mathbb{P}(X_t = 0 \ \forall t \in [0,T_k])} \\
&\quad \leq \exp(-(\alpha_0+\alpha_1+\gamma)T_k + \bar{\lambda}T_k + \alpha_1 T_k) \\
&\quad = \exp(\frac{-T_k}{2}\left(\alpha_0+\alpha_1+\gamma+\sqrt{(\alpha_1-\alpha_0-\gamma)^2+4\gamma(1-p)(\alpha_1-\alpha_0)}\right) + \alpha_1 T_k) \\
&\quad = \exp(\frac{-T_k}{2}(\alpha_0+\alpha_1+\gamma+|\alpha_1-\alpha_0-\gamma|+2E) + \alpha_1 T_k) \\
&\quad = \exp(T_k(\alpha_1 - \max(\alpha_1, \alpha_0+\gamma) - E)) \leq e^{-ET_k},
\end{aligned}$$

where E solves the equation

$$|\alpha_1-\alpha_0-\gamma| + 2E = \sqrt{|\alpha_1-\alpha_0-\gamma|^2+4\gamma(1-p)(\alpha_1-\alpha_0)}.$$

We get that

$$\begin{aligned}
&\sum_{k=1}^{n-1}\frac{T}{n}\frac{e^{-(\alpha_0+\alpha_1+\gamma)T_k}}{e^{-\bar{\lambda}T_k}\mathbb{P}(X_t = 0 \ \forall t \in [0,T_k])} \qquad (21)\\
&\quad \leq \frac{T}{n}\sum_{k=1}^{n-1} e^{-ET_k} = \frac{T}{n}\sum_{k=1}^{n-1}(e^{-ET/n})^k \\
&\quad = \frac{T}{n}\left(\frac{1-e^{-TE}}{1-e^{-TE/n}} - 1\right) = \frac{T}{n}\left(\frac{e^{-TE/n}-e^{-TE}}{1-e^{-TE/n}}\right) \\
&\quad = \frac{T}{n}\left(\frac{e^{-TE/n}-e^{-TE}}{TE/n+\mathcal{O}(1/n^2)}\right) = \left(\frac{e^{-TE/n}-e^{-TE}}{E+\mathcal{O}(1/n)}\right).
\end{aligned}$$

Combining equations (20) and (21) and taking the limit as n tends to infinity (after taking the limit as m tends to infinity), we get that

$$\lim_{m\to\infty}\sum_{k=1}^{Tm} A_k^m \leq T\bar{\lambda} + (p\alpha_1 + (1-p)\alpha_0 - \bar{\lambda})\frac{1-e^{-TE/2}}{E}.$$

Combining equation (19) with this yields

$$\begin{aligned}
\mathbb{P}(X_t = 0 \ \forall t \in [0,T]) &= \lim_{m\to\infty}\mathbb{P}(X_t^m = 0 \ \forall t \in [0,T]) \\
&\geq \exp\left(-\left(T\bar{\lambda} + (p\alpha_1 + (1-p)\alpha_0 - \bar{\lambda})\frac{1-e^{-TE/2}}{E}\right)\right).
\end{aligned}$$

Finally we conclude that

$$\lambda_{\max,\mu}^T(\alpha_0,\alpha_1,\gamma,p) \leq \bar{\lambda} + \frac{1}{T}(p\alpha_1 + (1-p)\alpha_0 - \bar{\lambda})\frac{1-e^{-TE/2}}{E}.$$

QED

Remark: It is interesting that in the above proof we "lift ourselves up by the boots" by using a simple estimate for $\mathbb{P}(X_t = 0 \ \forall t \in [0,T])$ to obtain a better one.

Proof of Theorem 1.4. The first statement follows immediately from Lemmas 4.3 and 4.5, by letting T tend to infinity.

We can of course trivially conclude that $\lambda_{\min,\mu}(\alpha_0, \alpha_1, \gamma, p) \leq \alpha_1$. To see why we have equality consider the event

$$\{\text{There are at least } k \text{ arrivals during } [0,1]\}.$$

Let $\alpha < \alpha_1$, we have that

$$\mathrm{Poi}_\alpha(\text{There are at least } k \text{ arrivals during } [0,1]) = \sum_{l=k}^{\infty} e^{-\alpha} \frac{\alpha^l}{l!}.$$

We also see that

$$\mathrm{Poi}^{\gamma,p}_{\alpha_0,\alpha_1}(\text{There are at least } k \text{ arrivals during } [0,1]) \geq pe^{-\gamma} \sum_{l=k}^{\infty} e^{-\alpha_1} \frac{\alpha_1^l}{l!}.$$

Since

$$\frac{\sum_{l=k}^{\infty} e^{-\alpha_1} \frac{\alpha_1^l}{l!}}{\sum_{l=k}^{\infty} e^{-\alpha} \frac{\alpha^l}{l!}} \longrightarrow_{k \to \infty} \infty,$$

we get that for every $\alpha < \alpha_1$, $\gamma > 0$ and $p > 0$ there exists a k such that

$$\begin{aligned} &\mathrm{Poi}_\alpha(\text{There are at least } k \text{ arrivals during } [0,1]) \\ &\quad < \mathrm{Poi}^{\gamma,p}_{\alpha_0,\alpha_1}(\text{There are at least } k \text{ arrivals during } [0,1]). \end{aligned}$$

Obviously this contradicts

$$\mathrm{Poi}^{\gamma,p}_{\alpha_0,\alpha_1} \preceq \mathrm{Poi}_\alpha,$$

and so $\lambda_{\min,\mu}(\alpha_0, \alpha_1, \gamma, p) \geq \alpha_1$.

QED

In the next section we will need the following easy corollary to Theorem 1.4.

Corollary 4.6 *For any $\delta < \min(\delta_1, \delta_0 + \gamma)$ we can find a $0 < p < 1$ close enough to one so that*

$$\mathrm{Poi}_\delta \preceq \mathrm{Poi}^{\gamma,p}_{\delta_0,\delta_1}.$$

Proof. We just need to observe that

$$\begin{aligned}
&\lim_{p\to 1}\lambda_{\max,\mu}(\delta_0,\delta_1,\gamma,p) \qquad (22)\\
&\quad=\lim_{p\to 1}\frac{1}{2}(\delta_0+\delta_1+\gamma-\sqrt{(\delta_1-\delta_0-\gamma)^2+4\gamma(1-p)(\delta_1-\delta_0)}))\\
&\quad=\frac{1}{2}(\delta_1+\delta_0+\gamma-|\delta_1-\delta_0-\gamma|)=\min(\delta_1,\delta_0+\gamma).
\end{aligned}$$

QED

Observe that

$\mathrm{Poi}_{\delta_0,\delta_1}^{\gamma,p}$(There are no arrivals in $[0,t]$) $\geq (1-p)e^{-\gamma t}e^{-\delta_0 t}=(1-p)e^{-(\gamma+\delta_0)t}$

and that

$$\mathrm{Poi}_{\delta}(\text{There are no arrivals in } [0,t])=e^{-\delta t}.$$

Therefore, if $\delta>\gamma+\delta_0$, we have for fixed p and some t that

$$e^{-\delta t}\leq (1-p)e^{-(\gamma+\delta_0)t},$$

and so we cannot have that

$$\mathrm{Poi}_{\delta}\preceq \mathrm{Poi}_{\delta_0,\delta_1}^{\gamma,p},$$

which is an alternative way to see why the limit in equation (22) cannot simply be equal to δ_1.

Proof of Theorem 1.5. The upper bound is just Lemma 4.5.

For the lower bound, we start by observing that using Theorem 1.4 we trivially get that $\mathbb{P}(X_t=0\ \forall t\in[0,T])\leq e^{-\bar{\lambda}T}$. Therefore by equation (18)

$$\begin{aligned}
&\lim_{m\to\infty} mA_{Tm}^m\\
&\quad\geq \bar{\lambda}+(p\alpha_1+(1-p)\alpha_0-\bar{\lambda})\frac{e^{-(\alpha_0+\alpha_1+\gamma)T}}{e^{-2T\bar{\lambda}}}\\
&\quad=\bar{\lambda}+(p\alpha_1+(1-p)\alpha_0-\bar{\lambda})e^{-TL}.
\end{aligned}$$

We therefore need to show that $\lambda_{\max,\mu}^T\geq\lim_{m\to\infty}mA_{Tm}^m$. Observe that the second marginal of the discrete time process $\{(B_n^m,X_n^m)\}_{n=1}^{Tm}$ trivially dominates an i.i.d. sequence of density A_{Tm}^m. Therefore, going through a limiting procedure very similar to the one of Lemma 4.3, we get that the second marginal of $\{(B_t,X_t)\}_{t\in[0,T]}$ dominates a Poisson process with intensity $\lim_{m\to\infty}mA_{Tm}^m$ on the time interval $[0,T]$.

The result for $\lambda_{\min,\mu}^T$ follows as in the proof of Theorem 1.4.

QED

5 CPREE-results

Proof of Theorem 1.6. For every site $s \in S$ the recoveries of the $\Psi^{\gamma,p,A}_{\delta_0,\delta_1}$ process at that site has the same distribution as the arrivals of a $\mathrm{Poi}^{\gamma,p}_{\delta_0,\delta_1}$ process. By Corollary 4.6 we can couple the processes $\Psi^{\gamma,p,A}_{\delta_0,\delta_1}$ and Ψ^{A}_{δ} so that at every site the former has a recovery whenever the latter does. Furthermore, coupling the infection rates are of course trivial. This gives the result.

QED

For $A \subset S$ such that $|A| < \infty$, let $\Psi^{\gamma,p,A}_{\delta_0,\infty,B(A)\equiv 0}$ denote the CPREE where a site $s \in S$ *always* is healthy (i.e. in state 0) as long as the background process of the site s is in state 1. That is, we do not allow the site to become infected if the background process of the site is in state 1. More precisely, for any graph $G = (S, E)$ let $\{(B_t, Y_t)\}_{t\geq 0}$ be a pair of processes with state space $\{\{0,1\} \times \{0,1\}\}^S$ such that $B_0 \sim \pi_p$ conditioned on the event that $B_0(s) = 0$ for every $s \in A$, and let $Y_0(s) = 1$ iff $s \in A$. Observe that the conditioning does not affect the probability of $B_0(s)$ being 0 or 1 for any $s \notin A$. Let the pair evolve according to the following flip rate intensities at any site s.

from	to	with intensity
$(0,0)$	$(1,0)$	γp
$(0,0)$	$(0,1)$	$\sum_{(s',s)\in E} Y_t(s')$
$(0,1)$	$(0,0)$	δ_0
$(0,1)$	$(1,0)$	γp
$(1,0)$	$(0,0)$	$\gamma(1-p)$.

Observe that with this definition the state $(1,1)$ is not allowed. Informally, this can be interpreted as letting the rate of recovery when $B_t(s) = 1$ be infinite, hence the notation.

Proof of Theorem 1.7. We start by observing that it is easy to see from the definitions of $\Psi^{\gamma,p,A}_{\delta_0,\delta_1,B_0(A)\equiv 0}$ and $\Psi^{\gamma,p,A}_{\delta_0,\infty,B(A)\equiv 0}$ that

$$\Psi^{\gamma,p,A}_{\delta_0,\infty,B(A)\equiv 0} \preceq \Psi^{\gamma,p,A}_{\delta_0,\delta_1,B_0(A)\equiv 0}.$$

We will construct $\{(B_t, Y_t)\}_{t\geq 0}$ to have distribution $\Psi^{\gamma,p,A}_{\delta_0,\infty,B(A)\equiv 0}$ for some p close to 0, and couple it with a process $\{Y'_t\}_{t\geq 0}$ such that $\{Y_t\}_{t\geq 0}$ stochastically dominates $\{Y'_t\}_{t\geq 0}$. It will be easy to see that in turn $\{Y'_t\}_{t\geq 0}$ will stochastically dominate $\Psi^{\lambda,A}_{\delta}$.

We now proceed to the actual construction. Let $B_0 \sim \pi_p$, conditioned on the event that $B_0(s) = 0$ for every $s \in A$. For every site $s \in S$, associate an independent process $\{B_t(s), X_t(s)\}_{t\geq 0}$ such that $\{1 - B_t(s), X_t(s)\}_{t\geq 0}$ is the model of Theorem 1.4 with $\alpha_0 = 0$, $\alpha_1 = \Delta_G$ and with p replaced by $1 - p$. We get from Theorem 1.4 that

$$\begin{aligned}
&\lambda_{\max,\mu}(0, \Delta_G, \gamma, 1-p) \\
&\quad = \frac{1}{2}(0 + \Delta_G + \gamma - \sqrt{(\Delta_G - 0 - \gamma)^2 + 4\gamma(1 - (1-p))(\Delta_G - 0)}) \\
&\quad = \frac{1}{2}(\Delta_G + \gamma - \sqrt{(\Delta_G - \gamma)^2 + 4\Delta_G\gamma p})).
\end{aligned}$$

That is, we can couple the pair of processes $\{B_t(s), X_t(s)\}_{t\geq 0}$ with a Poisson process $\{X'_t(s)\}_{t\geq 0}$ with intensity $\lambda_{\max,\mu}(0, \Delta_G, \gamma, 1-p)$ such that if this latter process has an arrival then so does $\{X_t(s)\}_{t\geq 0}$. There is a slight issue with $s \in A$, where we have conditioned that $B_0(s) = 0$. However, this corresponds to conditioning that the background process of Theorem 1.4 starts in state 1, and it is not hard to see that the conclusion of the theorem is still valid in this case. Informally, if we in this theorem start with the background process in state 1, this means that we are starting in the higher intensity state, and so it becomes "easier to dominate". It is easy to make this statement precise.

Let for every $s \in S$, $\{D_t(s)\}_{t\geq 0}$ be a Poisson process with intensity δ_0 and consider some quadruple $\{B_t(s), X_t(s), X'_t(s), D_t(s)\}_{t\geq 0}$ with marginal distributions as indicated by the notation. We now proceed to construct $\{(B_t, Y_t)\}_{t\geq 0}$ (the first marginal is of course already defined) and $\{Y'_t\}_{t\geq 0}$ from these four processes. Let $Y_0(s) = Y'_0(s)$ for every $s \in S$ and let $Y_0(s) = Y'_0(s) = 1$ iff $s \in A$. Let for every $s \in S$ $\{(B_t(s), Y_t(s))\}_{t\geq 0}$ and $\{Y'_t(s)\}_{t\geq 0}$ denote the marginals of the processes $\{(B_t, Y_t)\}_{t\geq 0}$ and $\{Y'_t\}_{t\geq 0}$ at the site s.

Let $N(s, Y_\tau(s)), N(s, Y'_\tau(s))$ denote the number of neighbors of the site s that are infected at time τ under $\{Y_t\}_{t\geq 0}$ and $\{Y'_t\}_{t\geq 0}$ respectively. Recall that by definition, for any $s \in S$, $Y_0(s) = Y'_0(s) = 0$ if $B_0(s) = 1$. We will write $X_\tau(s) \neq X_{\tau^-}(s)$, $X'_\tau(s) \neq X'_{\tau^-}(s)$ and $D_\tau(s) \neq D_{\tau^-}(s)$ to indicate that these processes have an arrival at time τ. Observe that by construction, for every $s \in S$ and $t \geq 0$, if $X'_\tau(s) \neq X'_{\tau^-}(s)$ then $X_\tau(s) \neq X_{\tau^-}(s)$. We will also write $B_{\tau^-}(s) < B_\tau(s)$ when we mean that the B_t process flips from 0 to 1 at time τ.

At time τ, $\{(Y_t(s), Y'_t(s))\}_{t\geq 0}$ will change:

from	to	if	
			(23)
$(1,1)$	$(0,0)$	$D_\tau(s) \neq D_{\tau^-}(s)$ or $B_{\tau^-}(s) < B_\tau(s)$	
$(1,0)$	$(0,0)$	$D_\tau(s) \neq D_{\tau^-}(s)$ or $B_{\tau^-}(s) < B_\tau(s)$	

and also:

$$\begin{array}{llcc}
\text{from} & \text{to} & \text{with probability} & \text{if} \\
(0,0) & (1,1) & N(s,Y'_\tau(s))/\Delta_G & X'_\tau(s) \neq X'_{\tau^-}(s) \\
(0,0) & (1,0) & (N(s,Y_\tau(s)) - N(s,Y'_\tau(s)))/\Delta_G & X'_\tau(s) \neq X'_{\tau^-}(s) \\
(0,0) & (1,0) & N(s,Y_\tau(s))/\Delta_G & X'_\tau(s) = X'_{\tau^-}(s), \\
 & & & X_\tau(s) \neq X_{\tau^-}(s) \\
(1,0) & (1,1) & N(s,Y'_\tau(s))/\Delta_G & X'_\tau(s) \neq X'_{\tau^-}(s).
\end{array} \tag{24}$$

No other transitions are allowed. Note that by construction $\{X_t(s)\}_{t\geq 0}$ and $\{X'_t(s)\}_{t\geq 0}$ only have arrivals when $\{B_t(s)\}_{t\geq 0}$ is in state 0. Therefore, these rates make sure that $\{Y_t\}_{t\geq 0}$ and $\{Y'_t\}_{t\geq 0}$ are in state 0 when $\{B_t(s)\}_{t\geq 0}$ is in state 1. Note also that since $N(s,Y'_0(s)) = N(s,Y_0(s))$ for every $s \in S$, the rates make sure that

$$Y'_t(s) \leq Y_t(s) \quad \forall s \in S, \ t \geq 0,$$

and that $N(s,Y'_t(s)) \leq N(s,Y_t(s))$ for every $s \in S$ and $t \geq 0$.

It remains to check that $\{(B_t,Y_t)\}_{t\geq 0}$ and $\{Y'_t\}_{t\geq 0}$ have the right distribution. As noted above $\{Y_t\}_{t\geq 0}$ is 0 if $\{B_t\}_{t\geq 0}$ is 1. Furthermore it is easy to see that when $\{B_t\}_{t\geq 0}$ is 0, $\{Y_t\}_{t\geq 0}$ flips from 0 to 1 at rate $N(s,Y_\tau(s))$ and from 1 to 0 at rate δ_0. It is also easy to see that $\{Y'_t\}_{t\geq 0}$ flips from 1 to 0 at a rate which is the minimum of two exponentially distributed times with parameters δ_0 and γp, the latter being the rate at which $\{B_t\}_{t\geq 0}$ flips from 0 to 1. Hence $\{Y'_t\}_{t\geq 0}$ flips from 1 to 0 at rate $\delta_0 + \gamma p$ and by choosing p small enough this is less that δ. It also not hard to see that $\{Y'_t\}_{t\geq 0}$ flips from 0 to 1 at a rate $\lambda_{\max,\mu}(0,\Delta_G,\gamma,1-p)N(s,Y'_t(s))/\Delta_G$. Furthermore by choosing p perhaps even smaller, we get that

$$\begin{aligned}
&\lambda_{\max,\mu}(0,\Delta_G,\gamma,1-p)N(s,Y'_t(s))/\Delta_G \\
&\quad = \frac{N(s,Y'_t(s))}{2\Delta_G}(\Delta_G + \gamma - \sqrt{(\Delta_G-\gamma)^2 + 4\Delta_G\gamma p})) \geq \lambda N(s,Y'_t(s)).
\end{aligned}$$

Here we used that $\gamma \geq \Delta_G$. Therefore $\{Y'_t\}_{t\geq 0}$ is a contact process with infection rate larger than λ and with recovery rate less that δ, and so the distribution of $\{Y'_t\}_{t\geq 0}$ dominates $\Psi_\delta^{\lambda,A}$.

QED

We are now ready to prove Theorem 1.8.

Proof of Theorem 1.8. We will start with the existence of p_{c1} and p_{c2}.

Let $0 < p_1 \leq p_2 < \infty$ and let $\{B_t^1\}_{t\geq 0}$, $\{B_t^2\}_{t\geq 0}$ be two background processes with parameters p_1, p_2 respectively. Let B_0^1 have distribution π_{p_1} and B_0^2 have distribution π_{p_2} and couple them so that $B_0^1(s) \leq B_0^2(s)$ for every $s \in S$. It is easy to see that we can then couple the processes so that

$$B_t^1(s) \leq B_t^2(s) \ \forall t \geq 0, \ \forall s \in S.$$

Using these processes to construct $\{(B_t^1, Y_t^1)\}_{t\geq 0}$ and $\{(B_t^2, Y_t^2)\}_{t\geq 0}$ with distributions $\Psi^{\gamma,p_1,A}_{\delta_0,\delta_1}$ and $\Psi^{\gamma,p_2,A}_{\delta_0,\delta_1}$, it is easy to see that we couple the marginals $\{Y_t^1\}_{t\geq 0}$, $\{Y_t^2\}_{t\geq 0}$ so that

$$Y_t^2(s) \leq Y_t^1(s) \ \forall t \geq 0,\ \forall s \in S.$$

This establishes the existence of p_{c1} and p_{c2}.

We now proceed to prove the part of statement 4 concerning $p_{c2} > 0$. We will then argue that the rest of the statements follow in the same or similar ways. For the process $\Psi^{\gamma,p,A}_{\delta_0,\delta_1}$, associate to every site $s \in S$, a Poisson process with intensity $\gamma+\delta_1+\Delta_G$. This Poisson process will be the underlying process that controls *everything* that can happen at the site, possible updates for the background process, possible recoveries and possible infections. The point of using these Poisson processes is that if it at some site s, does not have an arrival during some time interval $[0,T]$, it gives no information about the state of the background process at time T at that site.

Let B be the set of sites s such that either $s \in A$ or s has a neighbor which is in A. Let E_1 be the following event: None of the Poisson$(\gamma + \delta_1 + \Delta_G)$-processes associated to the sites of $B \setminus A$ have any arrivals during the time interval [0,1]. All of the sites of A have exactly one arrival during the time interval [0,1], and this arrival results in that the background process is updated to a 0. It is easy to see that the event E_1 has strictly positive probability, and it is also easy to see that if E_1 occurs this gives us no extra information about the status of $B_1(s)$ for any $s \notin A$. Therefore if E_1 occurs we have the situation at $t = 1$ that B_1 have distribution π_p conditioned on the event that $B_1(s) = 0$ for every $s \in A$, and $Y_1(s) = 1$ iff $s \in A$. By choosing $\delta > \delta_0$ close enough to δ_0 and $\lambda < 1$ close enough to 1 so that the contact process $\Psi^{\gamma,A}_{\delta}$ survives weakly, and then using Theorem 1.7 to couple our process $\Psi^{\gamma,p,A}_{\delta_0,\delta_1}$ above $\Psi^{\gamma,A}_{\delta}$ for $t \geq 1$ the statement follows.

All of the statements about $p_{c1}, p_{c2} > 0$ are proved in exactly the same way. All of the statements about $p_{c1}, p_{c2} < 1$ are proved in a similar way, but follow even easier since we can use Theorem 1.6 directly without worrying about the initial status of the background process at the sites of A.

QED

Proof of Proposition 1.9. We will show the theorem for p_{c2}, the proof for p_{c1} is identical. First, we use Taylor expansion to see that

$$\begin{aligned}
&\lim_{\gamma\to\infty} \lambda_{\max,\mu}(\delta_0, \delta_1, \gamma, p) \\
&\quad = \lim_{\gamma\to\infty} \frac{1}{2}(\delta_0 + \delta_1 + \gamma - \sqrt{(\delta_0 + \delta_1 + \gamma)^2 - 4(\delta_0\delta_1 + \gamma(\delta_0 + p(\delta_1 - \delta_0)))}) \\
&\quad = \lim_{\gamma\to\infty} \frac{1}{2}(\delta_0 + \delta_1 + \gamma)\left(1 - \sqrt{1 - 4\frac{\delta_0\delta_1 + \gamma(\delta_0 + p(\delta_1 - \delta_0))}{(\delta_0 + \delta_1 + \gamma)^2}}\right)
\end{aligned}$$

$$= \lim_{\gamma\to\infty} \frac{1}{2}(\delta_0+\delta_1+\gamma)\left(1-\left(1-\frac{4\frac{\delta_0\delta_1+\gamma(\delta_0+p(\delta_1-\delta_0))}{(\delta_0+\delta_1+\gamma)^2}}{2}\right)+\mathcal{O}(\frac{1}{\gamma^2})\right)$$

$$= \lim_{\gamma\to\infty} \frac{\delta_0\delta_1+\gamma(\delta_0+p(\delta_1-\delta_0))}{\delta_0+\delta_1+\gamma}+\mathcal{O}(\frac{1}{\gamma}) = \delta_0+p(\delta_1-\delta_0).$$

It is now clear from Theorem 1.4, the proof of Theorem 1.6 and the above calculation that given any $\epsilon > 0$, we can find γ' large enough so that with $\delta = \delta_0 + p(\delta_1-\delta_0) - \epsilon$ we have that for all $\gamma \geq \gamma'$

$$\Psi^{\gamma,p,A}_{\delta_0,\delta_1} \preceq \Psi^A_\delta$$

and so the $\Psi^{\gamma,p,A}_{\delta_0,\delta_1}$ dies out if $\delta_0+p(\delta_1-\delta_0)-\epsilon > \delta_{c2}$. This is the same as saying that for any $\epsilon > 0$ there exists γ' large enough so that for all $\gamma \geq \gamma'$, if

$$p > \frac{\delta_{c2}-\delta_0+\epsilon}{\delta_1-\delta_0},$$

the process dies out. Therefore for every $\gamma \geq \gamma'$ we have that

$$p_{c2}(\delta_0,\delta_1,\gamma) \leq \frac{\delta_{c2}-\delta_0+\epsilon}{\delta_1-\delta_0}.$$

We can therefore conclude that

$$\limsup_{\gamma\to\infty} p_{c2}(\delta_0,\delta_1,\gamma) \leq \frac{\delta_{c2}-\delta_0}{\delta_1-\delta_0}.$$

QED

Proof of Proposition 1.10. We show the proposition for p_{c2}, the proof for p_{c1} is identical. Observe (for example by using Taylor-expansion) that for any p,

$$\begin{aligned}
&\lambda_{\max,\mu}(\delta_0,\delta_1,\gamma,p)\\
&\quad= \frac{1}{2}(\delta_0+\delta_1+\gamma-\sqrt{(\delta_1-\delta_0-\gamma)^2+4\gamma(1-p)(\delta_1-\delta_0)})\\
&\quad= \frac{1}{2}(\delta_0+\delta_1+\gamma-\sqrt{(\delta_0-\delta_1-\gamma)^2-4\gamma p(\delta_1-\delta_0)})\\
&\quad\geq \frac{1}{2}(\delta_0+\delta_1+\gamma-|\delta_0-\delta_1-\gamma|+2p\gamma(\delta_1-\delta_0)) = \delta_0+p\gamma(\delta_1-\delta_0).
\end{aligned}$$

Therefore, for every $p > 0$, we can choose $\delta_0 < \delta_{c2}$ sufficiently close to δ_{c2} so that $\lambda_{\max,\mu}(\delta_0,\delta_1,\gamma,p) > \delta_{c2}$. Therefore, as above, the process $\Psi^{\gamma,p,A}_{\delta_0,\delta_1}$ dies out and therefore

$$\lim_{\delta_0\uparrow\delta_{c2}} p_{c2}(\delta_0,\delta_1,\gamma) < p.$$

Since $p > 0$ was arbitrary, $\lim_{\delta_0\uparrow\delta_{c2}} p_{c2}(\delta_0,\delta_1,\gamma) = 0$ and we are done.

QED

6 Open questions

We here list some open questions related to the results of this paper.

1. Do either of the critical values p_{c1} and p_{c2} depend on the initial state of the background process?

2. Instead of studying the CPREE model one could study other interacting particle systems such as a stochastic Ising model in a random evolving environment.

3. Is it possible to generalize the model used for the background process in some way? For instance, can we analyze the situation where we allow more than 2 different states?

Acknowledgment: I thank Professor Jeffrey E. Steif for all the suggestions he had during the work on this paper. I would also like to thank him for patiently reading this manuscript many times before its completion.

References

[1] E.D. Andjel (1992) Survival of multidimensional contact process in random environments, *Bol. Soc. Bras. Mat.* **23**, 109-119.

[2] M. Bramson, R. Durrett and R.H. Schonmann (1991) The contact process in a random environment, *Ann. Probab.*, **19**, 960-983.

[3] C. Bezuidenhout and G. Grimmett (1990) The critical contact process dies out, *Ann. Prob.*, **18**, 1462-1482.

[4] P. Billingsley *Convergence of Probability Measures 2:nd Ed.* John Wiley & Sons, Inc. 1999

[5] Broman, E.I. and Steif, J.E. (2006) Dynamical stability of percolation for some interacting particle systems and ϵ-movability, *Ann. Probab.*, to appear.

[6] S. N. Ethier, T. G. Kurtz *Markov Processes: Characterization and Convergence* John Wiley & Sons, Inc. 1986

[7] T. M. Liggett (1991) Spatially inhomogeneous contact process, *Spatial Stochastic Processes. A Festschrift in honor of the Seventieth Birthday of Ted Harris*, 105-140.

[8] T. M. Liggett (1992) The survival of one dimensional contact process in random environments, *Ann. Probab.* **20**, 696-723.

[9] T. M. Liggett *Interacting Particle Systems* Springer-Verlag 1985.

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

[11] T.M. Liggett and J. E. Steif (2006) Stochastic Domination: The Contact Process, Ising Models and FKG Measures. *Annales Institut Henri Poincare, Probabilites et Statistiques*, to appear.

[12] R. Lyons and J. E. Steif (2003) Stationary Determinantal Processes: Phase Multiplicity, Bernoullicity, Entropy, and Domination *Duke Math. Journal* **Vol 120**, 515-575.

[13] N. Madras, R. Schinazi and R. Schonmann (1994) On the critical behavior of the contact process in deterministic inhomogeneous random environment, *Ann. Probab.* **22**, 1140-1159.

[14] C. M. Newman and S. Volchan (1996) Persistent survival of one-dimensional contact process in random environments, *Ann. Probab.*, **24** 411-421.