THESIS FOR THE DEGREE OF DOCTOR OF PHILOSOPHY

Interacting particle systems in varying environment, stochastic domination in statistical mechanics and optimal pairs trading in finance

MARCUS WARFHEIMER

DEPARTMENT OF MATHEMATICAL SCIENCES
CHALMERS UNIVERSITY OF TECHNOLOGY AND UNIVERSITY OF GOTHENBURG
GÖTEBORG, SWEDEN 2010

Interacting particle systems in varying environment, stochastic domination in statistical mechanics and optimal pairs trading in finance
Marcus Warfheimer
ISBN 978-91-7385-393-4

Doktorsavhandlingar vid
Chalmers tekniska högskola
Ny serie Nr 3074
ISSN 0346-718X

Department of Mathematical Sciences
Chalmers University of Technology and University of Gothenburg
412 96 GÖTEBORG, Sweden
Phone: +46 (0)31-772 10 00

E-mail: `warfheim@chalmers.se`

Printed in Göteborg, Sweden, 2010

Interacting particle systems in varying environment, stochastic domination in statistical mechanics and optimal pairs trading in finance

Marcus Warfheimer

Department of Mathematical Sciences

Chalmers University of Technology and University of Gothenburg

ABSTRACT

In this thesis we first consider the contact process in a randomly evolving environment, introduced by Erik Broman. This process is a generalization of the contact process where the recovery rate can vary between two values. The rate which it chooses is determined by a background process, which evolves independently at different sites. We prove that survival of the process is independent of how we start the background process, that finite and infinite survival are equivalent and finally that the process dies out at criticality.

Second, we consider spin systems on $\mathbb{Z}$ whose rates are again determined by a background process, which is more general than that considered above. We prove that, if the background process has a unique stationary distribution and if the rates satisfy a certain positivity condition, then there are at most two extremal stationary distributions.

Third, we discuss various aspects concerning stochastic domination for the Ising and fuzzy Potts models. We begin by considering the Ising model on the homogeneous tree of degree d, $\mathbb{T}^d$. For given interaction parameters J_1, $J_2 > 0$ and external field $h_1 \in \mathbb{R}$, we compute the smallest external field $\tilde{h}$ such that the plus measure with parameters J_2 and h dominates the plus measure with parameters J_1 and h_1 for all $h \geq \tilde{h}$. Moreover, we discuss continuity of $\tilde{h}$ with respect to the parameters J_1, J_2, h_1 and also how the plus measures are stochastically ordered in the interaction parameter for a fixed external field. Next, we consider the fuzzy Potts model and prove that on $\mathbb{Z}^d$ the fuzzy Potts measures dominate the same set of product measures while on $\mathbb{T}^d$, for certain parameter values, the free and minus fuzzy Potts measures dominate different product measures.

Finally, we study the problem of optimally closing a pair trading strategy when the difference of the underlying assets is assumed to be an Ornstein-Uhlenbeck type process driven by a jump-diffusion process. We prove a verification theorem and analyze a numerical method for the associated free boundary problem. We prove rigorous error estimates, which are used to draw some conclusions from numerical simulations.

Keywords: Interacting particle systems, contact process, randomly evolving environment, spin systems, Ising model, fuzzy Potts model, pairs trading, optimal stopping.

Preface

This thesis consists of the following papers:

- Jeffrey E. Steif and Marcus Warfheimer "The critical contact process in a randomly evolving environment dies out", in *Latin American Journal of Probability and Mathematical Statistics - Alea*, **4** (2008), 337-357.
- Marcus Warfheimer, "Attractive nearest-neighbor spin systems on the integers in a randomly evolving environment".
- Marcus Warfheimer, "Stochastic domination for the Ising and fuzzy Potts models". (Submitted).
- Stig Larsson, Carl Lindberg and Marcus Warfheimer, "Optimal closing of a pair trade with a model containing jumps". (Submitted).

Acknowledgments

First and most I would like to thank my advisor Professor Jeffrey Steif. He has been a great source of inspiration, always there, ready to assist me in any way possible. Moreover, he has presented me to most of the problems in this thesis. Special thanks also goes to my co-advisors Erik Broman, Carl Lindberg and Bernt Wennberg

Next, I would like to thank Stig Larsson for nice cooperation during the work of the last paper.

Life at the Mathematical Sciences would not be the same without the presence of the following colleagues (in fact some of them are former): Anastassia, Daniel, Erik Brodin, Erik J, Jan, Johan T, Marcus, Martin, Mattias, Ottmar, Patrik, Stefan L and Viktor. Thank you for being very friendly and always ready to help me whatever I ask you about. Also, thanks for creating such a good working environment.

Thanks must also go to my friends outside Chalmers, in particular to the people I grew up with and to the members of Fish Quartet. You have inspired me a lot during the years.

Finally, my gratitude goes to my family, especially to my wife Debora and our daughter Moa. I cannot describe by words how much you mean to me. This thesis would not be without your never ending care and support.

Marcus Warfheimer
Göteborg, April 17, 2010

Contents

List of Figures

Part I

INTRODUCTION

1
Introduction

This thesis consists of three parts. Part one (the first two papers) concerns interacting particle systems in a randomly evolving environment, part two (the third paper) concerns stochastic domination in the Ising and fuzzy Potts models and part three (the fourth paper) concerns how to optimally close a pair trading strategy in finance. In this introductory chapter we briefly give some background material to all of these topics and at the end we give a summary of the papers in the thesis. We prefer to present most of the material in a rather informal way; for a mathematically precise description see the relevant references or the papers in the thesis.

1.1 Interacting particle systems

The field of interacting particle systems is a branch of probability theory. However, the motivation often comes from physical or biological systems. In loose terms, one tries to formulate a mathematical model for objects (particles, people, cars, etc) which interact with each other in a certain way. One way to construct such a model is to place each object at a site in a graph structure and declare that each one of them can be in one of a finite number of different states. (A graph is just a finite or countable set of vertices equipped with a relation that defines which vertices are neighbors.) One then assigns some initial configuration (or distribution) and lets the system evolve according to some probabilistic rules. It is at this point where the interactions come into play. Each object is changing its state at a rate depending on the states of the other (usually neighboring) objects as well as itself.

From a more mathematical point of view, interacting particle systems are a special class of so called Markov processes. Markov processes have the property that given the present state, the future is independent of the past. Denote the set of sites and possible states by S and A respectively. The state space, or configuration space, for our Markov process is then A^S. The most common situation is when A consists of only two elements and that only one coordinate of the process is allowed to change at a time. Such processes are called spin systems. In this situation, the evolution is described by a rate function, $c(x, \eta)$, $x \in S$ and $\eta \in A^S$, which gives the rate at which the coordinate at x flips when the system is in state η. Having something occur "at a rate $c(x, \eta)$" means informally that the time for this to occur has an exponential distribution with mean $1/c(x, \eta)$.

In this generality not much can be said. Therefore one concentrates upon specific types of models of which I will name a few.

The contact process on the d-dimensional lattice $\mathbb{Z}^d$. This process was introduced by Harris [18] and is a model for spread of an infection. The model is such that infected people recover at rate 1 and healthy people are infected with a rate proportional to the number of infected neighbors. The state of the system is described by a configuration $\eta \in \{0, 1\}^{\mathbb{Z}^d}$, where $\eta(x) = 0$ represents that the individual at x is healthy and $\eta(x) = 1$ represents it is infected. Also, the dynamics are specified by the following rate function

$$c(x, \eta) = \begin{cases} 1 & \text{if } \eta(x) = 1 \\ \lambda \sum_{y \sim x} \eta(y) & \text{if } \eta(x) = 0, \end{cases}$$

where $y \sim x$ means that x and y are neighbors and λ is a positive parameter called the infection rate. To simplify notation, we will identify $\{0, 1\}^{\mathbb{Z}^d}$ with subsets of $\mathbb{Z}^d$ by letting $\eta \in \{0, 1\}^{\mathbb{Z}^d}$ correspond to

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

Let η_t be the state of the process at time $t \geq 0$ and denote the distribution of the process with parameter $\lambda > 0$ and initial configuration $A \subseteq \mathbb{Z}^d$ by $\mathbf{P}_\lambda^A$. We say that the process *survives at* λ if

$$\mathbf{P}_\lambda^{\{0\}}[\, \eta_t \neq \emptyset \text{ for all } t \geq 0 \,] > 0;$$

otherwise it is said to *die out at* λ. One can show that

$$\mathbf{P}_\lambda^{\{0\}}[\, \eta_t \neq \emptyset \text{ for all } t \geq 0 \,] = 0$$

for small values of λ and

$$\mathbf{P}_\lambda^{\{0\}}[\, \eta_t \neq \emptyset \text{ for all } t \geq 0 \,] > 0$$

for large values of λ. The first claim follows easily by a comparison with a branssching process and the second, which is somewhat more difficult, follows from a percolation type argument. In words, when we start the process with one site infected, the infection will almost surely eventually disappear for small values of λ and will last forever with positive probability for large values of λ. From this, it is natural to define the critical value:

$$\lambda_c := \inf\{\, \lambda : \mathbf{P}_\lambda^{\{0\}}[\, \eta_t \neq \emptyset \text{ for all } t \geq 0\,] > 0\,\}$$

and the previous statement just means that $0 < \lambda_c < \infty$. A much harder question, and one which had been open for approximately 15 years, is whether the contact process survives or dies out *at* the critical value λ_c. A celebrated theorem by Bezuidenhout and Grimmett gives us the answer.

Theorem 1.1.1 (Bezuidenhout and Grimmett). *The critical contact process dies out.*

For a proof of this, see [1] or [33].

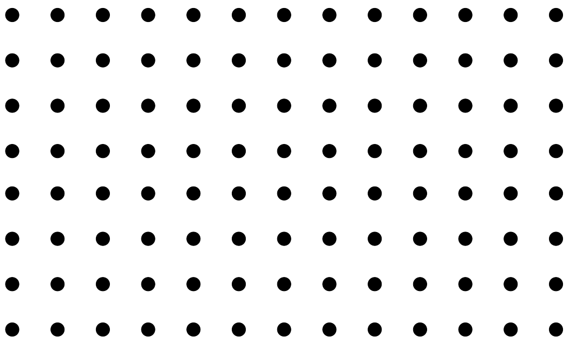

Figure 1.1: *A small portion of the lattice* $\mathbb{Z}^2$.

Remark: We can parameterize the contact process in an equivalent way as follows: Let the recovery rate be $\delta > 0$ and the infection rate be equal to the number of infected neighbors. In other words, we change λ to 1 and let δ be the recovery rate, which of course just corresponds to a time scaling. We will denote the corresponding critical value by δ_c.

The voter model on the d-dimensional lattice $\mathbb{Z}^d$. This process was introduced independently by Clifford and Sudbury [4] and by Holley and Liggett [19]. Here, the state of the system is described by a configuration $\eta \in \{0,1\}^{\mathbb{Z}^d}$ where $d \geq 1$ and the transition mechanism is described by saying that $\eta(x)$ flips to $1 - \eta(x)$ at rate

$$\frac{1}{2d} \sum_{y \sim x} 1_{\{\eta(y) \neq \eta(x)\}}.$$

One interpretation, made by Holley and Liggett, is to think of the sites in $\mathbb{Z}^d$ as representing voters who can hold either of two political positions, which are denoted by 0 and 1. In this representation the dynamics of the model can be described as follows: A voter waits an exponentially distributed time with mean 1 and then takes the opinion of a neighbor chosen at random. Clearly, if we start the process with all voters in state 0 or all voters in state 1, nothing happens. In mathematical terms the point masses on these two configurations are so called *invariant distributions*, meaning that if we start the process with such a distribution μ, the distribution of the process at any time later is still μ. (Of course, in this case they are also absorbing states.) At this point, one may ask if there are any other invariant distributions? To answer that question it turns out that the dimension d plays a prominent role. Namely, when $d \leq 2$ there are no other than those two above (plus their convex combinations), but when $d \geq 3$ there are in fact other ones. (To people with a little background in probability theory, this result is intimately related to the fact that simple random walk is recurrent when $d \leq 2$ and transient when $d \geq 3$.)

The stochastic Ising model on $\mathbb{Z}^d$. This is a model for magnetism introduced by Glauber [17]. The state space of this model is $\{-1, 1\}^{\mathbb{Z}^d}$. Imagine that atoms are laid out on all of $\mathbb{Z}^d$ and that each of them either can have a spin (state) of $+1$ or -1. The resulting configuration describing the system is an element of $\{-1, 1\}^{\mathbb{Z}^d}$. The dynamics of the evolution is described by declaring a spin $\eta(x)$ at a site x to flip to $-\eta(x)$ at rate

$$\exp\left(-\beta \sum_{y \sim x} \eta(x)\eta(y)\right)$$

where β is a nonnegative parameter called the inverse temperature. Note that the flip rate is higher when the spin at x differs from most of its neighbors than it is when it agrees with most of them. In other words, the system prefers configurations in which the spins tend to be aligned with one another. When $\beta = 0$ it is easy to see that there is only one invariant distribution, namely the product measure μ on $\{-1, 1\}^{\mathbb{Z}^d}$ with density $\frac{1}{2}$ and in addition, for any initial distribution, the distribution at time t converges weakly as $t \to \infty$ to μ. Such a process is called *ergodic*. In the non-ergodic case, i.e. when we have more than one invariant distribution, one says that a phase transition has occurred and each of the invariant distributions corresponds to a different "phase" of the system. The problem is to determine for which choices of β and d the process is ergodic. If $d = 1$, then it turns out that the process is ergodic for all β but when $d \geq 2$ there is a critical value $0 < \beta_d < \infty$ such that the process is ergodic if $\beta < \beta_d$ and non-ergodic if $\beta > \beta_d$.

For further reading about interacting particle systems, there are three standard reference books, namely Liggett [31, 33] and Durrett [10]. The first one deals with the construction of interacting particle systems from given rates as well as more or less all results in the field until 1985. The second one concentrates upon three models; the contact process, the voter model and the so called exclusion process, a model of particle motion, and covers many of the results concerning these models until 1999. The third book considers, among other things, the contact process, the voter model and some variants thereof.

In all the examples above, the dynamics are translation invariant in the sense that the rates only depend on x through $\{\eta(y) : y \sim x\}$. A possible first extension is to allow the rates to depend on x itself. One such example is the so called inhomogeneous contact process where we are given to us a family of rates $\{\delta_x\}_{x\in\mathbb{Z}^d}$ and from them the dynamics of the process is as follows:

$$c(x,\eta) = \begin{cases} \delta_x & \text{if } \eta(x) = 1 \\ \sum_{y\sim x} \eta(y) & \text{if } \eta(x) = 0. \end{cases}$$

A further extension is to allow for more randomness in the model. For example, one possibility is to study the above model with $\{\delta_x\}_{x\in\mathbb{Z}^d}$ taken to be i.i.d. random variables. That suggestion was made by Bramson, Durrett and Schonmann [2] and they called the resulting model *the contact process in a random environment*. For further results concerning that model see for example [26, 32, 36]. You could also extend the last model even more by letting the recovery rates follow some update rule. We then arrive at *the contact process in a randomly evolving environment* which was introduced by Broman [3] and which we will discuss in more detail in Section 1.4.1.

1.2 The ferromagnetic Ising and Potts models

The theory of Gibbs measures goes back to Dobrushin [5–9] and Lanford and Ruelle [28]. It started as a branch of classical statistical physics but can now also be viewed as a part of probability theory. From the physical point of view these measures were proposed as a mathematical description of an equilibrium state of a physical system which consists of a large number of interacting components. In probabilistic terms, a Gibbs measure is just the distribution of a countably infinite family of random variables taking values in some (usually finite) set which admit some prescribed conditional probabilities. To describe these conditional probabilities one has to specify the interaction between the components and that is usually done by a so called Hamiltonian. For an extensive presentation of the theory of Gibbs measures we refer to [15] and for a less extensive one, see [16]. Instead of discussing Gibbs measures in a general context we will now focus on two specific choices of Hamiltonians which for different reasons have attracted a large amount of interest during the last decades.

The ferromagnetic Ising model is a simplified mathematical description of a ferromagnetic substance such as iron, cobalt or nickel. It was introduced by Wilhelm Lenz in the 1920's [29] and first investigated by Ernst Ising [24]. In the same way as for the stochastic Ising model we think of the atoms as laid out on the d-dimensional lattice and that the spin of each of them is allowed to take two possible orientations, $+1$ (up) and -1 (down). Moreover, there are two parameters $J \geq 0$, $h \in \mathbb{R}$ in this model. The first one describes the strength of the interaction between neighboring spins and the second the affect of an external field. For given J, h and configuration $\sigma \in \{-1,1\}^{\mathbb{Z}^d}$

the Hamiltonian for this model is given by the so called Ising potential:

$$\Phi_A^{J,h}(\sigma) = \begin{cases} -J\sigma(x)\sigma(y) & \text{if } A = \{x,y\}, \text{ where } \langle x,y\rangle \in \mathbb{E}^d, \\ -h\sigma(x) & \text{if } A = \{x\}, \\ 0 & \text{otherwise.} \end{cases}$$

Here $A \subseteq \mathbb{Z}^d$, $\mathbb{E}^d$ denotes the set of edges in the d-dimensional lattice and $\langle x,y\rangle$ is the edge connecting x and y. From this potential the Gibbs measures are defined as follows: A probability measure μ on $\{-1,1\}^{\mathbb{Z}^d}$ is said to be a Gibbs measure (or sometimes Gibbs state) for the ferromagnetic Ising model with parameters $h \in \mathbb{R}$ and $J \geq 0$ if it admits conditional probabilities such that for all finite $U \subseteq \mathbb{Z}^d$, all $\sigma \in \{-1,1\}^U$ and all $\eta \in \{-1,1\}^{\mathbb{Z}^d \setminus U}$

$$\begin{aligned} &\mu(X(U) = \sigma \mid X(\mathbb{Z}^d \setminus U) = \eta) \\ &= \frac{1}{Z_{J,h}^{U,\eta}} \exp\Bigg[J\Bigg(\sum_{\langle x,y\rangle \in \mathbb{E}^d, x,y \in U} \sigma(x)\sigma(y) + \sum_{\langle x,y\rangle \in \mathbb{E}^d, x \in U, y \in \partial U} \sigma(x)\eta(y) \Bigg) \\ &\qquad + h \sum_{x \in U} \sigma(x) \Bigg]. \end{aligned}$$

Here $Z_{J,h}^{U,\eta}$ is a normalizing constant and ∂U is the outer boundary of U defined formally as

$$\partial U = \{\, x \in \mathbb{Z}^d \setminus U : \text{ There exists } y \in U \text{ such that } \langle x,y\rangle \in \mathbb{E}^d \,\}.$$

In words, μ is a Gibbs measure for the Ising model if it has prescribed conditional distributions inside any finite region given that the configuration is held fixed outside and these conditional distributions are given by the right hand side of the above expression. A natural question from both a physical and mathematical point of view is if this definition uniquely determines the Gibbs measure, or stated otherwise, is it possible to have two different measures with the same prescribed conditional distributions? It turns out that if $d = 1$ or $h \neq 0$ there is only one Gibbs measure but interestingly, when $d \geq 2$ and J is large enough there exists more than one Gibbs measure. When such a phenomena of multiple Gibbs measures occurs one says that the system undergoes a *phase transition*. For a proof of the above statement as well as a survey in the study of phase transitions for the Ising model we refer to [20].

Although we have chosen to discuss the Ising model on the d-dimensional lattice only the above definitions make perfect sense for other types of graphs too. It turns out that the question of phase transition is highly dependent on the graph structure. As an example if the underlying graph is the homogeneous tree of degree d the system can in fact undergo a phase transition even when $h \neq 0$, see [15].

A natural generalization of the ferromagnetic Ising model is the (ferromagnetic) Potts model in which the spins are allowed to take $q \geq 2$ (rather than just two) different states. We confine ourselves to the case with no external field and for simplicity we let $\mathbb{Z}^d$ be the underlying graph. For $\sigma \in \{1, \ldots, q\}^{\mathbb{Z}^d}$, the interaction potential for the Potts model is given by

$$\Psi_A^J(\sigma) = \begin{cases} -2JI_{\{\sigma(x)=\sigma(y)\}} & \text{if } A = \{x, y\}, \text{ where } \langle x, y\rangle \in \mathbb{E}^d, \\ 0 & \text{otherwise.} \end{cases}$$

In words, this interaction favors configurations where many neighboring pairs of spins agree. In a similar way as for the Ising model we can define the notion of Gibbs measures and study phase transitions etc, see [16] and the references therein.

1.3 Pairs trading

Since this part is relatively small we will not say much about it. Pairs trading was developed at Morgan Stanley in the late 1980's, and today it is one of the most common investment strategies in the financial industry. The idea behind pairs trading is quite intuitive: the investor finds two assets, for which the prices have moved together historically. When the price spread widens, the investor takes a short position in the outperforming asset, and a long position in the underperforming one with the hope that the spread will converge again, generating a profit. However, the trader should be aware of the risk of drifting. This happens when the two correlated stock prices start to drift apart. Therefore, in practice the investor typically chooses in advance a stop-loss level, which corresponds to the level of loss above which the investor will close the pair trade and take the loss.

For a historical evaluation of pairs trading see [14] and for books that treat the applied aspects of pairs trading we refer to [11, 38, 39].

1.4 Summary of papers

1.4.1 Paper I

In this paper we consider the so called *contact process in a randomly evolving environment* (CPREE), introduced by Broman [3]. This process is a generalization of the contact process, where the recovery rate is allowed to vary between two values, δ_0 and δ_1. (Recall the equivalent parameterization of the contact process.) The rate which is chosen is determined by a background process, which evolves independently at different sites. To be precise, we consider the Markov process $\{(B_t, C_t)\}_{t\geq 0}$ with state space $\{0,1\}^{\mathbb{Z}^d} \times \{0,1\}^{\mathbb{Z}^d}$ which performs transitions according to the following rates at a site $x \in \mathbb{Z}^d$:

transition	rate
$(0,0) \to (0,1)$	$\sum_{y \sim x} C(y)$
$(1,0) \to (1,1)$	$\sum_{y \sim x} C(y)$
$(0,1) \to (0,0)$	δ_0
$(1,1) \to (1,0)$	δ_1
$(0,0) \to (1,0)$	γp
$(0,1) \to (1,1)$	γp
$(1,0) \to (0,0)$	$\gamma(1-p)$
$(1,1) \to (0,1)$	$\gamma(1-p)$

where $d \geq 1$, $\gamma, \delta_0, \delta_1 > 0$ with $\delta_1 \leq \delta_0$ and $p \in [0,1]$. In other words, at each site x independently, $\{B_t(x)\}_{t\geq 0}$ is a 2-state Markov chain with infinitesimal matrix

$$\begin{pmatrix} -\gamma p & \gamma p \\ \gamma(1-p) & -\gamma(1-p) \end{pmatrix}$$

which in turn determines the recovery rate of $\{C_t(x)\}_{t\geq 0}$ in the following way. For each x and t, the recovery rate at time t and site x is δ_0 or δ_1 depending on whether $B_t(x) = 0$ or $B_t(x) = 1$. Also, the infection rate is always the number of infected neighbors. (Actually Broman did this on a more general graph, but here we will only consider $\mathbb{Z}^d$.) Broman referred to $\{B_t\}_{t\geq 0}$ as *the background process* and the whole process $\{(B_t, C_t)\}_{t\geq 0}$ as *the contact process in a randomly evolving environment* (CPREE). Let $\{C_t^{\mu,\nu}\}_{t\geq 0}$ denote the right marginal when the initial distribution of the whole process is $\mu \times \nu$. Furthermore, let $\mathbf{P}_p$ denote the measure governing the process for the parameters p, γ, δ_0 and δ_1, where γ, δ_0 and δ_1 are considered fixed. Also, denote the product measure with density $q \in [0,1]$ by π_q. Broman defined the critical value

$$p_c := \inf \left\{ p : \mathbf{P}_p[\, C_t^{\pi_p,\{0\}} \neq \emptyset \, \forall t > 0 \,] > 0 \right\}$$

(p_c is taken to be 1 if no p satisfies this) and proved that if $\delta_1 < \delta_c < \delta_0$ and $\gamma > \max(2d, \delta_c - \delta_1)$, then $p_c \in (0,1)$. (Recall the definition of δ_c from the remark after Theorem 1.1.1.) At the end of his paper he asked whether the critical value is affected if we vary the initial distribution of the background process. Our first result answers this question. Given γ, $\delta_0, \delta_1 > 0$ with $\delta_1 \leq \delta_0$, $q \in [0,1]$ and $A \subseteq \mathbb{Z}^d$ with $|A| < \infty$, define

$$p_c(q, A) := \inf \left\{ p : \mathbf{P}_p[\, C_t^{\pi_q,A} \neq \emptyset \, \forall t > 0 \,] > 0 \right\}.$$

Theorem 1.4.1. *Given $A, A' \subseteq \mathbb{Z}^d$ with $|A|$, $|A'| < \infty$ and p, q, $q' \in [0,1]$,*

$$\mathbf{P}_p[\, C_t^{\pi_q,A} \neq \emptyset \, \forall t > 0 \,] > 0 \iff \mathbf{P}_p[\, C_t^{\pi_{q'},A'} \neq \emptyset \, \forall t > 0 \,] > 0.$$

In particular, $p_c(q, A)$ is independent of both q and A.

We will let p_c denote this common value. (Recall, p_c of course depends on γ, δ_0 and δ_1.) Also, if $\mathbf{P}_p[\,C_t^{\pi_q,A} \neq \emptyset\,\forall t > 0\,] > 0$ holds (which we now know is independent of q and A), we say that $\{C_t\}$ *survives at* p; otherwise it is said to *die out at* p.

Standard arguments yield that the limiting distribution starting from all 1's exists and we will denote the limit by $\bar{\nu}_p$. This measure gives us another natural way to define a critical value:

$$p_c' := \inf\{\,p : \bar{\nu}_p \neq \pi_p \times \delta_\emptyset\,\}.$$

For general attractive interacting particle systems it might or might not be the case that these two critical values coincide. However, for the ordinary contact process this is the case (due to its self-duality) and our next result shows that this is also true in our situation.

Theorem 1.4.2. *$\{C_t\}$ survives at p if and only if $\bar{\nu}_p \neq \pi_p \times \delta_\emptyset$. In particular $p_c = p_c'$.*

Our final result is a generalization of Theorem 1.1.1.

Theorem 1.4.3. *If $\{C_t\}$ survives at $p > 0$, then there exists $\delta > 0$ so that it survives at $p - \delta$. In particular, if $p_c \in (0, 1]$, then the critical contact process in a randomly evolving environment dies out.*

1.4.2 Paper II

Recall that spin systems are interacting particle systems where each coordinate has two possible states and only one coordinate changes in each transition. In this paper we consider spin systems on $\mathbb{Z}$ in a randomly evolving environment, where the environment is more general than in the previous paper. To describe the process we are dealing with in mathematical terms, let $c_0(x,\eta)$, $c_1(x,\eta)$ and $b(x,\eta)$ be given rate functions and define a Markov process $\{(\beta_t,\eta_t)\}_{t\geq 0}$ on $\{0,1\}^{\mathbb{Z}} \times \{0,1\}^{\mathbb{Z}}$ with the dynamics at a site x specified in the following way:

transition	rate
$(\beta,\eta) \to (\beta,\eta_x)$	$c_0(x,\eta)$ if $\beta(x) = 0$
$(\beta,\eta) \to (\beta,\eta_x)$	$c_1(x,\eta)$ if $\beta(x) = 1$
$(\beta,\eta) \to (\beta_x,\eta)$	$b(x,\beta)$

Here, for given $\eta \in \{0,1\}^{\mathbb{Z}^d}$ and $x \in \mathbb{Z}$, η_x is the element in $\{0,1\}^{\mathbb{Z}}$ defined by

$$\eta_x(y) = \begin{cases} \eta(y) & \text{if } y \neq x \\ 1 - \eta(x) & \text{if } y = x. \end{cases}$$

As before, the left marginal $\{\beta_t\}_{t\geq 0}$ will be referred to as the background process. Furthermore, we will assume that the rate functions are translation invariant, that

$$\begin{aligned} c_0(x,\eta) &\leq c_1(x,\eta) \quad \text{if} \quad \eta(x) = 0, \\ c_1(x,\eta) &\leq c_0(x,\eta) \quad \text{if} \quad \eta(x) = 1, \end{aligned} \tag{1}$$

that $c_0(x,\eta)$ and $c_1(x,\eta)$ only depend on η through $\eta(x-1)$, $\eta(x)$ and $\eta(x+1)$ and that c_0, c_1 and b satisfy the following attractivity condition:

Definition 1.4.1. A spin system on $\mathbb{Z}$ with rate function $c(x,\eta)$ is said to be attractive if whenever $\eta \le \eta'$,

$$\begin{aligned} c(x,\eta) &\le c(x,\eta') \quad \text{if} \quad \eta(x) = \eta'(x) = 0, \\ c(x,\eta) &\ge c(x,\eta') \quad \text{if} \quad \eta(x) = \eta'(x) = 1. \end{aligned} \tag{2}$$

(Here, $\le$ refers to the usual partial ordering on $\{0,1\}^{\mathbb{Z}}$, i.e., $\eta \le \eta'$ if and only if $\eta(x) \le \eta'(x)$ for all $x \in \mathbb{Z}$.) This condition is exactly what is needed to be able to couple two copies, with initial configurations stochastically ordered, so that the two copies continue to be ordered for all times. Furthermore, note that we can equivalently view our process on $\{0,1\}^{\mathbb{Z}\times\{0,1\}}$ and that the conditions (1) and (2) just means that the whole process is attractive on that space. (Definition 1.4.1 can of course be generalizad to $\{0,1\}^S$ where S is countable.) The attractivity can be used to show (via monotonicity) the existence of two extremal stationary distributions ν_0 and ν_1 defined by

$$\nu_0 = \lim_{t\to\infty} \delta_0 S(t) \qquad \nu_1 = \lim_{t\to\infty} \delta_1 S(t),$$

where δ_0 and δ_1 denote the point masses corresponding to the elements $\eta \equiv 0$ and $\eta \equiv 1$ in $\{0,1\}^{\mathbb{Z}\times\{0,1\}}$ and $\{S(t)\}_{t\ge 0}$ denotes the semigroup associated to $\{(\beta_t,\eta_t)\}_{t\ge 0}$. The main result in this paper is that, if the background process has a unique stationary distribution and the rates c_0, c_1 satisfy a certain positivity condition, then ν_0 and ν_1 are the only extremal stationary distributions for the process.

To state our result we need a bit more notation. The assumptions on c_0 and c_1 imply that they together can be described by at most 16 different parameters. To describe the values we will use the following notation: $c_i(001) = c_i(x,\eta)$ when $\eta(x-1) = 0$, $\eta(x) = 0$ and $\eta(x+1) = 1$ etc. Define

$$\begin{aligned} C_1 = \{\, & c_i(100) + c_j(110),\ c_i(001) + c_j(011), \\ & c_i(011) + c_j(110),\ c_i(100) + c_j(001),\ i = 0,1,\ j = 0,1 \,\} \end{aligned}$$

and let

$$C = \min(C_1).$$

Moreover, let $\mathcal{I}$ denote the set of invariant distributions for $\{(\beta_t,\eta_t)\}_{t\ge 0}$ and let $\mathcal{I}_e$ denote its extreme points.

Before we state our main result, we want to emphasize that the case with no background process has been studied before by Liggett. The proof of our main result follows closely the ideas of his proof. To state his result let $c(x,\eta)$ be a rate function for an attractive, translation invariant, nearest-neighbor spin system $\{\eta_t\}_{t\ge 0}$ on $\{0,1\}^{\mathbb{Z}}$ and define $\mu_i = \lim_{t\to\infty} \delta_i T(t)$, $i = 0,1$, where δ_i is the point mass corresponding to the element $\eta \equiv i$ in $\{0,1\}^{\mathbb{Z}}$ and $\{T(t)\}_{t\ge 0}$ denotes the semigroup associated to $\{\eta_t\}_{t\ge 0}$. Moreover, let $\mathcal{J}_e$ denote the extreme points of the set of stationary distributions for $\{\eta_t\}_{t\ge 0}$.

Theorem 1.4.4 (Liggett). *Suppose*

$$c(x,\eta) + c(x,\eta_x) > 0 \quad \textit{whenever} \quad \eta(x-1) \neq \eta(x+1). \tag{3}$$

Then $\mathcal{J}_e = \{\mu_0, \mu_1\}$.

For a proof, see [30] or [31, p. 145-152]. In fact, he also proved that if condition (3) fails, then $\mathcal{J}_e$ contains infinitely many points, see [31, p. 145].

Theorem 1.4.5. *Suppose that the background process is ergodic and* $C > 0$*. Then* $\mathcal{I}_e = \{\nu_0, \nu_1\}$.

Remarks:

(i) From [31, p. 152] we get that Theorem 1.4.4 is equivalent to the statement that (3) and

$$\begin{aligned} c(011) + c(110) &> 0 \\ c(100) + c(001) &> 0 \end{aligned}$$

implies $\mathcal{J}_e = \{\mu_0, \mu_1\}$. By letting $c = c_0 = c_1$, it is now clear that Theorem 1.4.5 covers Theorem 1.4.4.

(ii) The hypotheses in Theorem 1.4.5 are true for the CPREE studied in the first paper. Indeed, if c_1 and c_2 satisfy (1) and are symmetric under reflections, i.e.

$$\begin{aligned} c_i(100) &= c_i(001) \\ c_i(110) &= c_i(011), \quad i = 0, 1 \end{aligned}$$

then $C > 0$ if and only if $c_0(001) > 0$ and $c_1(011) > 0$.

(iii) Note that we are not assuming independence or even nearest-neighbor interaction between coordinates in the background process.

1.4.3 Paper III

In [34], various results were proved concerning stochastic domination (defined below) for the Ising model with no external field on $\mathbb{Z}^d$ and on the homogeneous binary tree $\mathbb{T}^2$ (i.e. the unique infinite tree where each site has 3 neighbors). As an example, the following distinction between $\mathbb{Z}^d$ and $\mathbb{T}^2$ was shown: On $\mathbb{Z}^d$, the plus and minus states (to be defined later) dominate the same set of product measures, while on $\mathbb{T}^2$ that statement fails completely except in the case when we have a unique phase. In this paper we study stochastic domination for the Ising model in the case of nonzero external field and also for the so called fuzzy Potts model.

Definition 1.4.2 (Stochastic domination). Given a finite or countable set V and probability measures μ_1, μ_2 on $\{-1,1\}^V$, we say that μ_2 dominates μ_1 (written $\mu_1 \leq \mu_2$ or $\mu_2 \geq \mu_1$) if

$$\int f\, d\mu_1 \leq \int f\, d\mu_2$$

for all real-valued, continuous and increasing functions f on $\{-1,1\}^V$.

Here, increasing for a function $f : \{-1,1\}^V \to \mathbb{R}$ means that $f(\eta) \leq f(\eta')$ whenever $\eta \leq \eta'$. It can be shown that a necessary and sufficient condition for two probability measures μ_1, μ_2 to satisfy $\mu_1 \leq \mu_2$ is that there exists a coupling measure ν on $\{-1,1\}^V \times \{-1,1\}^V$ with first and second marginals equal to μ_1 and μ_2 respectively and

$$\nu(\,(\eta,\xi) : \eta \leq \xi\,) = 1.$$

(For a proof, see for example [31, p 72-74].) Given any set $S \subseteq \mathbb{R}$ and a family of probability measures $\{\mu_s\}_{s\in S}$ indexed by S, we will say that the map $S \ni s \mapsto \mu_s$ is increasing if $\mu_{s_1} \leq \mu_{s_2}$ whenever $s_1 < s_2$.

Results for the Ising model

For the Ising model with parameters $J > 0$, $h \in \mathbb{R}$ on a general graph of bounded degree standard monotonicity arguments based on Holley's theorem (see [16]) can be used to show that there exist two particular Gibbs states $\mu_h^{J,+}$ and $\mu_h^{J,-}$, called the plus and the minus state, which are extreme with respect to the stochastic ordering in the sense that

$$\mu_h^{J,-} \leq \mu \leq \mu_h^{J,+} \quad \text{for any other } \mu \in \mathcal{G}(J,h).$$

(Here, $\mathcal{G}(J,h)$ denotes the set of Gibbs state for the Ising model with parameters $J > 0$ and $h \in \mathbb{R}$.) To simplify the notation, we will write $\mu^{J,+}$ for $\mu_0^{J,+}$ and $\mu^{J,-}$ for $\mu_0^{J,-}$. (Of course, the plus and minus state are also highly dependent on the graph G, but we suppress that in the notation.) In [34] the authors studied, among other things, stochastic domination between the plus measures $\{\mu^{J,+}\}_{J>0}$ in the case when $G = \mathbb{T}^2$. For example, they showed that the map $(0,\infty) \ni J \mapsto \mu^{J,+}$ is increasing when $J > J_c$ and proved the existence of and computed the smallest $J > J_c$ such that $\mu^{J,+}$ dominates $\mu^{J',+}$ for all $0 < J' \leq J_c$. (On $\mathbb{Z}^d$, the fact that $\mu^{J_1,+}$ and $\mu^{J_2,+}$ are not stochastically ordered when $J_1 \neq J_2$ gives that such a J does not even exist in that case.) Our first result deals with the following question: Given J_1, $J_2 > 0$, $h_1 \in \mathbb{R}$, can we find the smallest external field $\tilde{h} = \tilde{h}(J_1, J_2, h_1)$ with the property that $\mu_h^{J_2,+}$ dominates $\mu_{h_1}^{J_1,+}$ for all $h \geq \tilde{h}$? To clarify the question a bit more, note that an easy application of Holley's theorem tells us that for fixed $J > 0$, the map $\mathbb{R} \ni h \mapsto \mu_h^{J,+}$ is increasing. Hence, for given J_1, J_2 and h_1 as above the set

$$\{\, h \in \mathbb{R} : \mu_h^{J_2,+} \geq \mu_{h_1}^{J_1,\pm} \,\}$$

is an infinite interval and we want to find the left endpoint of that interval (possibly $-\infty$ or $+\infty$ at this stage). For a general graph of bounded degree not much can be said, but we have the following easy bounds on $\tilde{h}$.

Proposition 1.4.6. *Consider the Ising model on a general graph $G = (V, E)$ of bounded degree. Define*

$$\tilde{h} = \tilde{h}(J_1, J_2, h_1) = \inf\{\, h \in \mathbb{R} : \mu_h^{J_2,+} \geq \mu_{h_1}^{J_1,+} \,\}.$$

Then

$$h_1 - N(J_1 + J_2) \leq \tilde{h} \leq h_1 + N|J_1 - J_2|,$$

where $N = \sup_{x \in V} N_x$ and N_x is the number of neighbors of the site $x \in V$.

We will now consider the case when $G = \mathbb{T}^d$, the homogeneous d-ary tree, defined as the unique infinite tree where each site has exactly $d + 1 \geq 3$ neighbors. The parameter d is fixed in all that we will do and so we suppress that in the notation. To state our results, we need to recall some more facts, all of which can be found in [15, p. 247-255]. For $J > 0$, define

$$\phi_J(t) = \frac{1}{2} \log \frac{\cosh(t + J)}{\cosh(t - J)}, \quad t \in \mathbb{R}.$$

Given $J > 0$ and $h \in \mathbb{R}$, there is a one-to-one correspondence $t \mapsto \mu$ between the real solutions of the equation

$$t = h + d\,\phi_J(t)$$

and the completely homogeneous Markov chains in $\mathcal{G}(J, h)$ (see [15] for a definition). Let $t_\pm(J, h)$ denote the real numbers which correspond to the plus and minus measure respectively. We will write $t_\pm(J)$ instead of $t_\pm(J, 0)$. Furthermore, let

$$h^*(J) = \max_{t \geq 0} \big(d\,\phi_J(t) - t\big)$$

and denote by $t^*(J)$ the $t \geq 0$ where the function $t \mapsto d\,\phi_J(t) - t$ attains its unique maximum.

Theorem 1.4.7. *Consider the Ising model on $\mathbb{T}^d$ and let $J_1,\ J_2 > 0,\ h_1 \in \mathbb{R}$ be given. Define*

$$f_\pm(J_1, J_2, h_1) = \inf\{\, h \in \mathbb{R} : \mu_h^{J_2,+} \geq \mu_{h_1}^{J_1,\pm} \,\}$$
$$g_\pm(J_1, J_2, h_1) = \inf\{\, h \in \mathbb{R} : \mu_h^{J_2,-} \geq \mu_{h_1}^{J_1,\pm} \,\}$$

and denote $\tau_\pm = \tau_\pm(J_1, J_2, h_1) = t_\pm(J_1, h_1) + |J_1 - J_2|$. Then the following holds:

$$(4)\quad f_\pm(J_1, J_2, h_1) = \begin{cases} -h^*(J_2) & \textit{if } t_-(J_2, -h^*(J_2)) \leq \tau_\pm < t^*(J_2) \\ \tau_\pm - d\,\phi_{J_2}(\tau_\pm) & \textit{if } \tau_\pm \geq t^*(J_2) \textit{ or } \tau_\pm < t_-(J_2, -h^*(J_2)) \end{cases}$$

$$(5)\quad g_{\pm}(J_1, J_2, h_1) = \begin{cases} h^*(J_2) & \textit{if } -t^*(J_2) < \tau_{\pm} \leq t_+(J_2, h^*(J_2)) \\ \tau_{\pm} - d\,\phi_{J_2}(\tau_{\pm}) & \textit{if } \tau_{\pm} \leq -t^*(J_2) \textit{ or } \tau_{\pm} > t_+(J_2, h^*(J_2)) \end{cases}$$

Remark: By looking at the formulas (4) and (5), we see that there are functions ψ, $\theta : (0, \infty) \times \mathbb{R} \mapsto \mathbb{R}$ such that

$$f_{\pm}(J_1, J_2, h_1) = \psi(J_2, \tau_{\pm}(J_1, J_2, h_1)) \quad \text{and}$$
$$g_{\pm}(J_1, J_2, h_1) = \theta(J_2, \tau_{\pm}(J_1, J_2, h_1)).$$

(Of course, $\psi(J_2, t)$ and $\theta(J_2, t)$ are just (4) and (5) with t instead of $\tau_{\pm}$.) It is easy to check that for fixed $J_2 > 0$, the maps $t \mapsto \psi(J_2, t)$ and $t \mapsto \theta(J_2, t)$ are continuous. A picture of these functions when $J_2 = 2$, $d = 4$ can be seen in Figure 1.2.

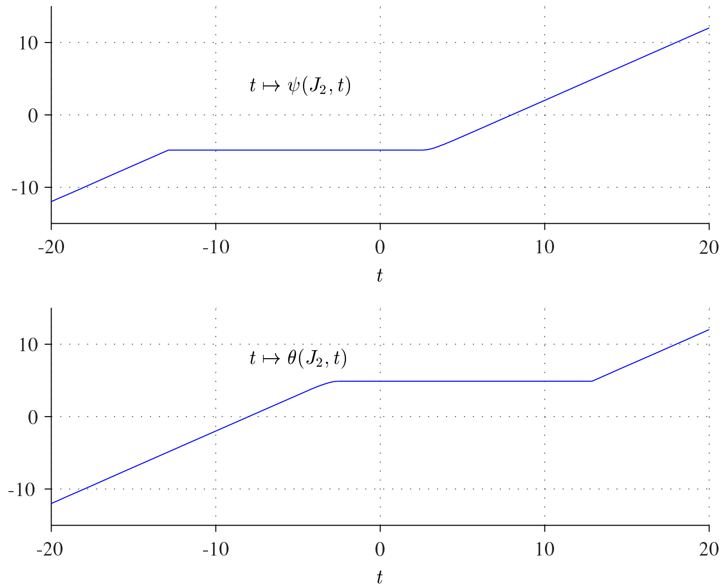

Figure 1.2: *The functions $t \mapsto \psi(J_2, t)$ and $t \mapsto \theta(J_2, t)$ in the case when $J_2 = 2$ and $d = 4$.*

Our next proposition deals with continuity properties of $f_{\pm}$ and $g_{\pm}$ with respect to the parameters J_1, J_2 and h_1. We will only discuss the function f_+, the other ones can be treated in a similar fashion.

Proposition 1.4.8. *Consider the Ising model on $\mathbb{T}^d$ and recall the notation from Theorem 1.4.7. Let*

$$a = a(J_1, J_2) = t_-(J_1, -h^*(J_1)) + |J_1 - J_2|$$
$$b = b(J_1, J_2) = t_+(J_1, -h^*(J_1)) + |J_1 - J_2|.$$

a) *Given J_1, $J_2 > 0$, the map $\mathbb{R} \ni h_1 \mapsto f_+(J_1, J_2, h_1)$ is continuous except possibly at $-h^*(J_1)$ depending on J_1 and J_2 in the following way:*

If $J_1 \leq J_c$ or $J_1 = J_2$ then it is continuous at $-h^*(J_1)$.

If $J_1 > J_c$ and $0 < J_2 \leq J_c$ then it is discontinuous at $-h^*(J_1)$.

If J_1, $J_2 > J_c$, $J_1 \neq J_2$ then it is discontinuous except when

$$t_-(J_2, -h^*(J_2)) \leq a < t^*(J_2) \quad \textit{and}$$
$$t_-(J_2, -h^*(J_2)) \leq b \leq t^*(J_2).$$

b) *Given $J_2 > 0$, $h_1 \in \mathbb{R}$, the map $(0, \infty) \ni J_1 \mapsto f_+(J_1, J_2, h_1)$ is continuous at J_1 if $0 < J_1 \leq J_c$ or $J_1 > J_c$ and $h_1 \neq -h^*(J_1)$. In the case when $h_1 = -h^*(J_1)$ it is discontinuous at J_1 except when*

$$t_-(J_2, -h^*(J_2)) \leq a < t^*(J_2) \ \textit{ and } \ t_-(J_2, -h^*(J_2)) \leq b \leq t^*(J_2).$$

c) *Given $J_1 > 0$, $h_1 \in \mathbb{R}$, the map $(0, \infty) \ni J_2 \mapsto f_+(J_1, J_2, h_1)$ is continuous for all $J_2 > 0$.*

Our last result for the Ising model is about how the measures $\{\mu_h^{J,+}\}_{J>0}$ are ordered with respect to J for fixed $h \in \mathbb{R}$.

Proposition 1.4.9. *Consider the Ising model on $\mathbb{T}^d$. The map $(0, \infty) \ni J \mapsto \mu_h^{J,+}$ is increasing in the following cases: a) $h \geq 0$ and $J \geq J_c$, b) $h < 0$ and $h^*(J) > -h$.*

Results for the fuzzy Potts model

For an infinite connected locally finite graph $G = (V, E)$ it is possible by a limiting procedure to define $q + 1$ basic examples of Gibbs measures for the Potts model, see [23] and the references therein. We denote these basic examples by $\pi_{q,J}^{G,i}$, $i \in \{0, \ldots, q\}$. (The measures $\{\pi_{q,J}^{G,i}\}_{i=1}^q$ are the analogs of the plus and minus states for the Ising model and $\pi_{q,J}^{G,0}$ is constructed by taking a free boundary condition outside a finite box and letting the box grow to infinity.) From them we can define new objects as follows: Fix $i \in \{0, \ldots, q\}$, suppose $r \in \{1, \ldots, q-1\}$ and pick a $\pi_{q,J}^{G,i}$-distributed random variable X and for $x \in V$ define

$$Y(x) = \begin{cases} -1 & \text{if } X(x) \in \{1, \ldots, r\} \\ \ \ 1 & \text{if } X(x) \in \{r+1, \ldots, q\}. \end{cases} \tag{6}$$

We write $\nu_{q,J,r}^{G,i}$ for the induced probability measure on $\{-1, 1\}^V$ and call it the fuzzy Potts measure with parameters q, J and r.

In words, the fuzzy Potts model is obtained from the ordinary q-state Potts model by identifying r states with a fuzzy spin denoted -1 and the remaining $q - r$ states with another fuzzy spin denoted 1. From this point of view, the fuzzy Potts model is one of the most basic examples of a so called hidden Markov field [27]. For earlier work on the fuzzy Potts model, see for example [21–23, 25, 35].

It is easy to see that when $G = \mathbb{Z}^d$ or $\mathbb{T}^d$ in the construction above it follows from symmetry that $\nu_{q,J,r}^{G,i} = \nu_{q,J,r}^{G,j}$ if $i, j \in \{1, \dots, r\}$ or $i, j \in \{r+1, \dots, q\}$, i.e. when the Potts spins i, j map to the same fuzzy spin. For that reason, we let $\nu_{q,J,r}^{G,-} := \nu_{q,J,r}^{G,1}$ and $\nu_{q,J,r}^{G,+} := \nu_{q,J,r}^{G,q}$ when $G = \mathbb{Z}^d$ or $\mathbb{T}^d$. (We stick to our earlier notation of $\nu_{q,J,r}^{G,0}$.)

Given a finite or countable set V and $p \in [0, 1]$, let γ_p denote the product measure on $\{-1, 1\}^V$ with $\gamma_p(\eta : \eta(x) = 1) = p$ for all $x \in V$. In [34], the authors proved the following results for the Ising model. (The second result was originally proved for $d = 2$ only but it trivially extends to all $d \geq 2$.)

Proposition 1.4.10 (Liggett, Steif). *Fix an integer $d \geq 2$ and consider the Ising model on $\mathbb{Z}^d$ with parameters $J > 0$ and $h = 0$. Then for any $p \in [0, 1]$, $\mu^{J,+} \geq \gamma_p$ if and only if $\mu^{J,-} \geq \gamma_p$.*

Proposition 1.4.11 (Liggett, Steif). *Let $d \geq 2$ be a given integer and consider the Ising model on $\mathbb{T}^d$ with paramteters $J > 0$ and $h = 0$. Moreover, let $\mu^{J,f}$ denote the Gibbs state obtained by using free boundary conditions. If $\mu^{J,+} \neq \mu^{J,-}$, then there exist $0 < p' < p$ such that $\mu^{J,+}$ dominates γ_p but $\mu^{J,f}$ does not dominate γ_p and $\mu^{J,f}$ dominates $\gamma_{p'}$ but $\mu^{J,-}$ does not dominate $\gamma_{p'}$.*

In words, on $\mathbb{Z}^d$ the plus and minus state dominate the same set of product measures while on $\mathbb{T}^d$ that is not the case except when the we have a unique phase. Our first result is a generalization of Proposition 1.4.10 to the fuzzy Potts model.

Proposition 1.4.12. *Let $d \geq 2$ be a given integer and consider the fuzzy Potts model on $\mathbb{Z}^d$ with parameters $q \geq 3$, $J > 0$ and $r \in \{1, \dots, q-1\}$. Then for any $k, l \in \{0, -, +\}$ and $p \in [0, 1]$, $\nu_{q,J,r}^{\mathbb{Z}^d,k} \geq \gamma_p$ if and only if $\nu_{q,J,r}^{\mathbb{Z}^d,l} \geq \gamma_p$.*

In the same way as for the Ising model, we believe that Proposition 1.4.12 fails completely on $\mathbb{T}^d$ except when we have a unique phase in the Potts model. Our last result is in that direction.

Proposition 1.4.13. *Let $d \geq 2$ be a given integer and consider the fuzzy Potts model on $\mathbb{T}^d$ with parameters $q \geq 3$, $J > 0$ and $r \in \{1, \dots, q-1\}$ where $e^{2J} \geq q - 2$. If the underlying Gibbs measures for the Potts model satisfy $\pi_{q,J}^{\mathbb{T}^d,1} \neq \pi_{q,J}^{\mathbb{T}^d,0}$, then there exists $0 < p < 1$ such that $\nu_{q,J,r}^{\mathbb{T}^d,0}$ dominates γ_p but $\nu_{q,J,r}^{\mathbb{T}^d,-}$ does not dominate γ_p.*

1.4.4 Paper IV

To model a pair spread the authors in [12] suggested the so called mean reverting Ornstein-Uhlenbeck process. In this paper, we generalize the model to also include possible jumps. More precisely, we let the difference $U = \{U_t\}_{t\geq 0}$ between the assets be the unique solution of the stochastic differential equation

$$dU_t = -\mu U_t\,dt + \sigma\,dW_t + dC_t^{\lambda,\varphi}, \quad t > 0, \tag{7}$$

where $\mu > 0$, $\sigma > 0$, $W = \{W_t\}_{t\geq 0}$ is a standard Brownian motion and $C^{\lambda,\varphi} = \{C_t^{\lambda,\varphi}\}_{t\geq 0}$ is a compound Poisson process with jump intensity $\lambda > 0$ and symmetric jump size distribution φ. Moreover, the support of φ is assumed to be contained in the interval $(-J, J)$ for some $J > 0$. (The solution to (7) is usually called a generalized Ornstein-Uhlenbeck process or an Ornstein-Uhlenbeck type process.) As discussed in [12] there is a large risk associated with a pair trading strategy. Indeed, if the market spread ceases to be mean reverting, the investor is exposed to substantial risk. Therefore, in practice the investor typically chooses in advance a stop-loss level $a < 0$, which corresponds to the level of loss above which the investor will close the pair trade.

Given such a stop-loss level $a < 0$, define

$$\tau_a = \inf\{t \geq 0 : U_t \leq a\}, \tag{8}$$

the first hitting time of the region $(-\infty, a]$, and the so called value function

$$V(x) = \sup_{\tau} \mathbf{E}_x[U_{\tau_a \wedge \tau}] \quad x \in \mathbb{R}, \tag{9}$$

where the supremum is taken over all stopping times with respect to U. (Here and in the sequel $\mathbf{E}_x$ means expected value when $U_0 = x$.) The major interest here is to characterize V, and perhaps more importantly, to describe the stopping time where the supremum is attained. Since the drift has the opposite sign as U, we have no reason to liquidate our position as long as U is negative. On the other hand, if U is positive, then the drift is working against the investor and for large values of U the size of the drift should overcome the possible benefits from random variations. Moreover, since the jumps are assumed to be symmetric, this indicates that there is a stopping barrier $b > 0$ with the property that we should keep our position when $U_t < b$ and liquidate as soon as $U_t \geq b$.

General optimal stopping theory (described for example in [37, Ch. 3]) leads us to believe that the value function is given by $V = u$, where (u, b) is the solution to the free boundary problem

$$\begin{aligned} \mathcal{G}_U u(x) &= 0, \quad x \in (a,b), \\ u(x) &= x, \quad x \notin (a,b), \\ u'(b) &= 1. \end{aligned} \tag{10}$$

Here $\mathcal{G}_U$ is the infinitesimal generator of U, which is defined on the space of twice continuously differentiable functions $f : \mathbb{R} \to \mathbb{R}$ with compact support:

$$\mathcal{G}_U f(x) = \frac{\sigma^2}{2} f''(x) - \mu x f'(x) + \lambda \int_{-\infty}^{\infty} (f(x+y) - f(x))\varphi(y)\,dy, \quad x \in \mathbb{R}. \tag{11}$$

Moreover, the stopping time where the supremum in (9) is attained should be

$$\tau_b = \inf\{t \geq 0 : U_t \geq b\}. \tag{12}$$

Indeed, our first result is a so called verification theorem.

Theorem 1.4.14. *Assume that (u, b) is a classical solution to (10) with*

a) $\mathcal{G}_U u(x) \leq 0$ *for* $x > b$,

b) $u(x) \geq x$ *for all* $x \in \mathbb{R}$.

Then $u(x) = V(x) = \mathbf{E}_x[U_{\tau_a \wedge \tau_b}]$, *for* $x \in \mathbb{R}$.

In the rest of the paper, we analyze the free boundary problem (10). By transforming to homogeneous boundary values and using the symmetry of φ, we get

$$\begin{aligned} \mathcal{L}v(x) - \mathcal{I}v(x) &= -\mu x, \quad && x \in (a,b), \\ v(x) &= 0, && x \notin (a,b), \\ v'(b) &= 0, \end{aligned} \tag{13}$$

where $v(x) = u(x) - x$ and

$$\begin{aligned} \mathcal{L}v(x) &= -\tfrac{1}{2}\sigma^2 v''(x) + \mu x v'(x), \\ \mathcal{I}v(x) &= \lambda \int_{-\infty}^{\infty} \big(v(x+y) - v(x)\big)\varphi(y)\, dy. \end{aligned}$$

We have not been able to give a rigorous proof of the existence and uniqueness of a solution (v, b) of the free boundary value problem (13). We therefore resort to a numerical solution by means of the finite element method. We begin to prove existence and uniqueness of solutions of the boundary value problem

$$\begin{aligned} \mathcal{L}v(x) - \mathcal{I}v(x) &= -\mu x, \quad && x \in (a,b), \\ v(x) &= 0, && x \notin (a,b), \end{aligned} \tag{14}$$

and the corresponding finite element equation. Next, we define the functions

$$F(b) = v'(b), \quad F_N(b) = v_N'(b), \quad b > 0,$$

where v_N denotes the finite element approximation of v when we use a uniform subdivision of the interval $[a, b]$ consisting of N number of points. Note that from the maximum principle proved in [13], there is no restriction to assume that $b > 0$. For $0 \leq b_1 < b_2$, we prove the error estimate

$$\|F - F_N\|_{L_\infty(b_1,b_2)} \leq C N^{-\frac{1}{2}}, \quad N \geq N_0, \tag{15}$$

where C and N_0 are constants depending on a, λ, μ, σ, b_1 and b_2.

From (15) and simulations we present strong evidence that there exists a unique $b > 0$ such that $F(b) = 0$, that there for N large exists unique b_N such that $F_N(b_N) = 0$ and that

$$\lim_{N\to\infty} b_N = b.$$

Finally, we do simulations to discuss the properties $a)$ and $b)$ from Theorem 1.4.14.

References

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

[2] M. Bramson, R. Durrett, and R. H. Schonmann, *The contact process in a random environment*, Ann. Probab. **19** (1991), 960–983.

[3] E. I. Broman, *Stochastic Domination for a Hidden Markov Chain with Applications to the Contact Process in a Randomly Evolving Environment*, Ann. Probab. **35** (2007), 2263–2293.

[4] P. Clifford and A. Sudbury, *A model for spatial conflict*, Biometrika **60** (1973), 581–588.

[5] R. L. Dobrushin, *The description of a random field by means of conditional probabilities and conditions of its regularity*, Theor. Prob. Appl. **13** (1968), 197–224.

[6] ______, *Gibbsian random fields for lattice systems with pairwise interactions*, Funct. Anal. Appl. **2** (1968), 292–301.

[7] ______, *The problem of uniqueness of a gibbs random field and the problem of phase transition*, Funct. Anal. Appl. **2** (1968), 302–312.

[8] ______, *Gibbsian random fields. The general case*, Funct. Anal. Appl. **3** (1969), 22–28.

[9] ______, *Prescribing a system of random variables by conditional distributions*, Theor. Prob. Appl. **15** (1970), 458–486.

[10] R. Durrett, *Lecture notes on paricle systems and percolation*, Wadsworth and Brooks/Cole Advances Books and Software, 1988.

[11] D. Ehrman, *The handbook of pairs trading*, John Wiley and Sons Inc, 2006.

[12] E. Ekström, C. Lindberg, and J. Tysk, *Optimal liquidation of a pair trade*, to appear.

[13] M. G. Garroni and J. L. Menaldi, *Second Order Elliptic Integro-Differential Problems*, Chapman and Hall/CRC, 2002.

[14] E. Gatev, W. Goetzmann, and G. Rouwenhorst, *Pairs trading: performance of a relative-value arbitrage*, Review of Financial Studies **19**, 797–827.

[15] H. O. Georgii, *Gibbs Measures and Phase Transitions*, Walter de Gruyter, 1988.

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

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

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

[19] R. Holley and T. M. Liggett, *Ergodic theorems for weakly interacting systems and the voter model*, Ann. Probab. **3** (1975), 643–663.

[20] O. Häggström, *Random-cluster representations in the study of phase transitions*, Markov Processes and Related Fields **4** (1998), 275–321.

[21] ______, *Positive correlations in the fuzzy Potts model*, Ann. Appl. Probab. **9** (1999), 1149–1159.

[22] ______, *Is the fuzzy Potts model gibbsian?*, Ann. Inst. H. Poincaré, Probab. Statist **39** (2003), 891–917.

[23] O. Häggström and C. Külske, *Gibbs properties of the fuzzy Potts model on trees and in mean field*, Markov Processes and Related Fields **10** (2004), 477–506.

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

[25] J. Kahn and N. Weininger, *Positive association in the fractional fuzzy Potts model*, Ann. Probab. **35** (2007), 2038–2043.

[26] A. Klein, *Extinction of contact and percolation processes in a random envionment*, Ann. Probab. **22** (1994), 1227–1251.

[27] H. Künsch, S. Geman, and A. Kehagias, *Hidden Markov random fields*, Ann. Appl. Probab. **5** (1995), 577–602.

[28] O. E. Lanford and D. Ruelle, *Observables at infinity and states with short range correlations in statistical mechanics*, Commun. Math. Phys. **13** (1969), 194–215.

[29] W. Lenz, *Beitrag zum Verständnis der magnetischen Erscheinungen in festen Körpern*, Physik. Zeitschr. **21** (1920), 613–615.

[30] T. M. Liggett, *Attractive nearest neighbor spin systems on the integers*, Ann. Probab. **6** (1978), 629–636.

[31] ______, *Interacting Particle Systems*, Springer, 1985.

[32] ______, *The survival of one-dimensional contact processes in random environments*, Ann. Probab. **20** (1992), 696–723.

[33] ______, *Stochastic interacting systems: contact, voter and exclusion processes*, Springer, 1999.

[34] T. M. Liggett and J. E. Steif, *Stochastic domination: the contact process, Ising models and FKG measures*, Ann. Inst. H. Poincaré, Probab. Statist **42** (2006), 223–243.

[35] C. Maes and K. Vande Velde, *The fuzzy Potts model*, J. Phys. A **28** (1995), 4261–4270.

[36] C. M. Newman and S. B. Volchan, *Persistent survival of one-dimensional contact processes in random environments*, Ann. Probab. **24** (1996), 411–421.

[37] G. Peskir and A. Shiryaev, *Optimal Stopping and Free-Boundary Problems*, Springer, 2006.

[38] G. Vidyamurthy, *Pairs Trading - Quantitative Methods and Analysis*, John Wiley and Sons Inc, 2004.

[39] M. Whistler, *Trading Pairs: Capturing Profits and Hedging Risk with Statistical Arbitrage Strategies*, John Wiley and Sons Inc, 2004.

Part II

PAPERS

PAPER I

The critical contact process in a randomly evolving environment dies out, Alea, **4** (2008), 337-357

Jeffrey E. Steif and Marcus Warfheimer

2

PAPER I

ABSTRACT

Bezuidenhout and Grimmett proved that the critical contact process dies out. Here, we generalize the result to the so called contact process in a random evolving environment (CPREE), introduced by Erik Broman. This process is a generalization of the contact process where the recovery rate can vary between two values. The rate which it chooses is determined by a background process, which evolves independently at different sites. As for the contact process, we can similarly define a critical value in terms of survival for this process. In this paper we prove that this definition is independent of how we start the background process, that finite and infinite survival (meaning nontriviality of the upper invariant measure) are equivalent and finally that the process dies out at criticality.

Key words and phrases: Contact process, varying environment.
Subject classification : 60K35.

2.1 Introduction and main results

The contact process, introduced by Harris [5], is a simple model for the spread of an infection on a lattice. The state at a certain time is described by a configuration, $\eta \in \{0,1\}^{\mathbb{Z}^d}$, where $\eta(x) = 0$ means that the individual at location x is healthy and $\eta(x) = 1$ means it is infected. The model is such that infected people recover at rate 1 and healthy people are infected with a rate proportional to the number of infected neighbors. In more mathematical language, the contact process is a Markov process,

$\{\eta_t\}_{t\geq 0}$, with state space $\{0,1\}^{\mathbb{Z}^d}$ where the configuration changes its state at site $x \in \mathbb{Z}^d$ as follows:

$$\begin{array}{llll} \eta \to \eta_x & \text{with rate} \quad 1 & \text{if} & \eta(x) = 1 \\ \eta \to \eta_x & \text{with rate} \quad \lambda \sum_{y \sim x} \eta(y) & \text{if} & \eta(x) = 0, \end{array}$$

where $y \sim x$ means that x and y are neighbors,

$$\eta_x(y) = \begin{cases} \eta(y) & \text{if } y \neq x \\ 1 - \eta(x) & \text{if } y = x \end{cases}$$

and λ is a positive parameter called the infection rate. See the standard references Liggett [7] and Durrett [4] for how these informal rates determine a Markov process and for much on the contact process as well as other interacting particle systems. Denote the distribution of this process when it starts with the configuration η by $\mathbf{P}_\lambda^\eta$. We say that the process *dies out at* λ if

$$\mathbf{P}_\lambda^{\{0\}}[\,\eta_t = \emptyset \text{ some } t \geq 0\,] = 1;$$

otherwise it is said to *survive at* λ. Here, the initial configuration $\{0\}$ means there is a single infection at the origin and the configuration $\emptyset$ means the element in $\{0,1\}^{\mathbb{Z}^d}$ consisting of all zeros. (As usual, we identify $\{0,1\}^{\mathbb{Z}^d}$ with subsets of $\mathbb{Z}^d$.) Using an easy monotonicity in λ, it is natural to define the critical value

$$\lambda_c := \inf\{\,\lambda : \ \mathbf{P}_\lambda^{\{0\}}[\,\eta_t \neq \emptyset \text{ for all } t \geq 0\,] > 0\,\}.$$

A fundamental first question concerning this model is whether it survives when λ is large and whether it dies out for small values of λ, i.e. whether $0 < \lambda_c < \infty$, and it is not very hard to show that this indeed is the case. Furthermore, since the contact process is attractive (see Liggett [7] for this definition), we can define

$$\lambda_c' := \inf\{\,\lambda : \ \bar{\nu}_\lambda \neq \delta_\emptyset\,\},$$

where $\bar{\nu}_\lambda$ is the so called upper invariant measure, defined to be the limiting distribution starting from all 1's. A self-duality equation (see [4] or [7]) easily leads to $\lambda_c = \lambda_c'$. A much harder question, and one which had been open for approximately 15 years, is whether the contact process survives or dies out *at* the critical value. A celebrated theorem by Bezuidenhout and Grimmett, [1], gives us the answer.

Theorem 2.1.1 (Bezuidenhout and Grimmett). *The critical contact process dies out.*

For a proof of this, see [1] or [9].

Note that changing λ to 1 and the recovery rate to δ corresponds to a trivial time scaling and so the process could have instead been defined in this way. We will denote the corresponding critical value by δ_c. This should be kept in mind in what follows.

In 1991, Bramson, Durrett and Schonmann [2] introduced the contact process in a random environment, in which the recovery rates are taken to be independently and identically distributed random variables and then fixed in time. For further results concerning this model see for example, Liggett [8], Klein [6] and Newman and Volchan [11]. Recently, Broman [3] introduced another variant where the environment changes in time in a simple Markovian way. More precisely, he considered the Markov process, $\{(B_t, C_t)\}_{t\geq 0}$ on $\{0,1\}^{\mathbb{Z}^d} \times \{0,1\}^{\mathbb{Z}^d}$ described by the following rates at a site x:

transition	rate
$(0,0) \to (0,1)$	$\sum_{y\sim x} C(y)$
$(1,0) \to (1,1)$	$\sum_{y\sim x} C(y)$
$(0,1) \to (0,0)$	δ_0
$(1,1) \to (1,0)$	δ_1
$(0,0) \to (1,0)$	γp
$(0,1) \to (1,1)$	γp
$(1,0) \to (0,0)$	$\gamma(1-p)$
$(1,1) \to (0,1)$	$\gamma(1-p)$

where $d \geq 1$, $\gamma, \delta_0, \delta_1 > 0$ with $\delta_1 \leq \delta_0$ and $p \in [0,1]$. In other words, at each site x independently, $\{B_t(x)\}_{t\geq 0}$ is a 2-state Markov chain with infinitesimal matrix

$$\begin{pmatrix} -\gamma p & \gamma p \\ \gamma(1-p) & -\gamma(1-p) \end{pmatrix}$$

which in turn determines the recovery rate of $\{C_t(x)\}_{t\geq 0}$ in the following way. For each t, the recovery rate at location x is δ_0 or δ_1 depending on whether $B_t(x) = 0$ or $B_t(x) = 1$. In addition, the infection rate is always taken to be the number of infected neighbors. (Actually, Broman did this on a more general graph, but here we will only consider $\mathbb{Z}^d$.) Broman referred to $\{B_t\}_{t\geq 0}$ as *the background process* and the whole process $\{(B_t, C_t)\}_{t\geq 0}$ as *the contact process in a randomly evolving environment* (CPREE). Let $\{C_t^\rho\}_{t\geq 0}$ denote the right marginal where the initial distribution of the whole process is ρ. In the case where $\rho = \mu \times \nu$ we write $\{C_t^{\mu,\nu}\}_{t\geq 0}$. Furthermore, let $\mathbf{P}_p$ denote the measure governing the process for the parameters p, γ, δ_0 and δ_1, where γ, δ_0 and δ_1 are considered fixed. Also, denote the product measure with density $q \in [0,1]$ by π_q. Broman defined the critical value

$$p_c := \inf\left\{p : \mathbf{P}_p[\, C_t^{\pi_p,\{0\}} \neq \emptyset \, \forall t > 0\,] > 0\right\}$$

(p_c is taken to be 1 if no p satisfies this) and proved that if $\delta_1 < \delta_c < \delta_0$ and $\gamma > \max(2d, \delta_c - \delta_1)$, then $p_c \in (0,1)$. At the end of his paper he asked whether the critical value is affected if we vary the initial distribution of the background process.

Our first result answers this question. Given $\gamma, \delta_0, \delta_1 > 0$ with $\delta_1 \leq \delta_0$, $q \in [0,1]$ and $A \subseteq \mathbb{Z}^d$ with $|A| < \infty$, define

$$p_c(q, A) := \inf \left\{ p : \mathbf{P}_p[\, C_t^{\pi_q, A} \neq \emptyset \, \forall t > 0\,] > 0 \right\}.$$

Theorem 2.1.2. *Given $A, A' \subseteq \mathbb{Z}^d$ with $|A|$, $|A'| < \infty$ and p, q, $q' \in [0,1]$,*

(1) $$\mathbf{P}_p[\, C_t^{\pi_q, A} \neq \emptyset \, \forall t > 0\,] > 0 \quad \Longleftrightarrow \quad \mathbf{P}_p[\, C_t^{\pi_{q'}, A'} \neq \emptyset \, \forall t > 0\,] > 0.$$

In particular, $p_c(q, A)$ is independent of both q and A.

We will let p_c denote this common value. (Recall, p_c of course depends on γ, δ_0 and δ_1.) Also, if $\mathbf{P}_p[\, C_t^{\pi_q, A} \neq \emptyset \, \forall t > 0\,] > 0$ holds (which we now know is independent of q and A), we say that $\{C_t\}$ *survives at p*; otherwise it is said to *die out at p*.

Later on, we will see that the process is attractive. (See Proposition 2.2.1.) This yields that the limiting distribution starting from all 1's exists and we will denote the limit by $\bar{\nu}_p$. Also, we will refer to this measure as the *upper invariant measure*. This measure gives us another natural way to define a critical value:

$$p_c' := \inf\{\, p : \bar{\nu}_p \neq \pi_p \times \delta_\emptyset \,\}.$$

For general attractive systems it might or might not be the case that these definitions coincide. However, for the ordinary contact process, this is the case (due to its self-duality) and our next result shows that this is also true in our situation.

Theorem 2.1.3. *$\{C_t\}$ survives at p if and only if $\bar{\nu}_p \neq \pi_p \times \delta_\emptyset$. In particular $p_c = p_c'$.*

Our final result is a generalization of Theorem 2.1.1.

Theorem 2.1.4. *If $\{C_t\}$ survives at $p > 0$, then there exists $\delta > 0$ so that it survives at $p - \delta$. In particular, if $p_c \in (0,1]$, then the critical contact process in a randomly evolving environment dies out.*

The rest of the paper is organized as follows. In Section 2, we provide some preliminaries, in Section 3, we prove Theorems 2.1.2 and 2.1.3 and in Section 4, we prove Theorem 2.1.4.

2.2 Some preliminaries

In this section we will present the basic construction of the CPREE via a graphical representation that is suitable for our situation. We will also prove the elementary fact that the CPREE is an attractive process. However, we will start off with some notation and basic definitions. When the initial distribution of the process is ρ, we will denote the distribution at time t by $\rho S_p(t)$, suppressing γ, δ_0 and δ_1 in the notation.

(Of course, ρ is a probability measure on $\{0,1\}^{\mathbb{Z}^d} \times \{0,1\}^{\mathbb{Z}^d}$.) When ρ is a product measure, $\rho = \mu \times \nu$, we will denote the process by $\{(B_t^{\mu}, C_t^{\mu,\nu})\}_{t\geq 0}$. In the case where $\mu = \delta_\beta$ and $\nu = \delta_\eta$ for some $\beta, \eta \in \{0,1\}^{\mathbb{Z}^d}$, we write $\{(B_t^{\beta}, C_t^{\beta,\eta})\}_{t\geq 0}$. To simplify notation, we freely interchange between talking about elements in $\{0,1\}^{\mathbb{Z}^d}$ and subsets of $\mathbb{Z}^d$. For $\eta, \eta' \in \{0,1\}^{\mathbb{Z}^d}$ we write $\eta \leq \eta'$ if $\eta(x) \leq \eta'(x)$ $\forall x \in \mathbb{Z}^d$. Furthermore, for $(\beta,\eta), (\beta',\eta') \in \{0,1\}^{\mathbb{Z}^d} \times \{0,1\}^{\mathbb{Z}^d}$ we write $(\beta,\eta) \leq (\beta',\eta')$ if both $\beta \leq \beta'$ and $\eta \leq \eta'$. These relations induce the concept of increasing function in the usual way.

Definition 2.2.1. We say that a function f on $\{0,1\}^{\mathbb{Z}^d}$ (or $\{0,1\}^{\mathbb{Z}^d} \times \{0,1\}^{\mathbb{Z}^d}$) is increasing if $f(\eta) \leq f(\eta')$ $(f(\beta,\eta) \leq f(\beta',\eta'))$ whenever $\eta \leq \eta'$ $((\beta,\eta) \leq (\beta',\eta'))$.

In our analysis we make extensive use of the concept of *stochastic domination*.

Definition 2.2.2. Given two probability measures μ_1 and μ_2 on $\{0,1\}^{\mathbb{Z}^d}$, we say that μ_1 is stochastically dominated by μ_2 if $\mu_1(f) \leq \mu_2(f)$ $\forall$ increasing continuous functions f and we denote this by $\mu_1 \leq \mu_2$. If μ_i is the distribution of X_i, $i = 1,2$, we also write $X_1 \leq_D X_2$.

It is well known (see for example [7]) that this is equivalent to the existence of random variables X_1, X_2 on a common probability space such that $X_1 \sim \mu_1$, $X_2 \sim \mu_2$ and $X_1 \leq X_2$ a.s. (The $\sim$ here means distributed according to.) Also, since we can identify $\{0,1\}^{\mathbb{Z}^d} \times \{0,1\}^{\mathbb{Z}^d}$ with $\{0,1\}^{\mathbb{Z}^d \times \{0,1\}}$ we have a similar result for measures on $\{0,1\}^{\mathbb{Z}^d} \times \{0,1\}^{\mathbb{Z}^d}$. (Of course, stochastic domination makes sense on any space of the form $\{0,1\}^S$ where S is countable.)

Now, we turn to the graphical representation from which our process will be defined. Let $\gamma, \delta_0, \delta_1 > 0$ with $\delta_1 \leq \delta_0$ and $p \in [0,1]$ be given parameters. Let $\{e_j\}_{j=1}^d$ denote the standard basis on $\mathbb{Z}^d$, i.e. for $i, j \in \{1, \ldots, d\}$

$$e_j(i) = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{if } i \neq j. \end{cases}$$

Define the following stochastic elements on a common probability space in such a way that they are independent:

- $M^{b,0\to 1} = \{M_t^{b,0\to 1}\}_{t\geq 0}$, a process with state space $\mathbb{N}^{\mathbb{Z}^d}$ where each marginal independently evolves as a Poisson process with intensity γp. (This process will correspond to the 0 to 1 flips in the background process, see below.)
- $M^{b,1\to 0} = \{M_t^{b,1\to 0}\}_{t\geq 0}$, a process with state space $\mathbb{N}^{\mathbb{Z}^d}$ where each marginal independently evolves as a Poisson process with intensity $\gamma(1-p)$. (This process will correspond to the 1 to 0 flips in the background process, see below.)
- $N^{\delta_1} = \{N_t^{\delta_1}\}_{t\geq 0}$, a process with state space $\mathbb{N}^{\mathbb{Z}^d}$ where each marginal independently evolves as a Poisson process with intensity δ_1.

- $N^{\delta_0-\delta_1} = \{N_t^{\delta_0-\delta_1}\}_{t\geq 0}$, a process with state space $\mathbb{N}^{\mathbb{Z}^d}$ where each marginal independently evolves as a Poisson process with intensity $\delta_0 - \delta_1$.
- $\vec{N}^j = \{\vec{N}_t^j\}_{t\geq 0}$, $j \in \{\pm e_1, \dots, \pm e_d\}$, independent processes with state space $\mathbb{N}^{\mathbb{Z}^d}$ where each marginal independently evolves as a Poisson process with intensity 1. (We think of the points in $\vec{N}^j(x)$ as being arrows from x to $x + e_j$ and will correspond to the potential spread of infection from x to $x + e_j$.)

For $s \geq 0$ and $\beta \in \{0,1\}^{\mathbb{Z}^d}$, we will begin to define a process $B^{\beta,s} = \{B_t^{\beta,s}\}_{t\geq s}$ where for each $x \in \mathbb{Z}^d$, $B^{\beta,s}(x)$ is a function of the arrivals of $M^{b,0\to 1}(x)$ and $M^{b,1\to 0}(x)$ in $[s,\infty)$. Assume for example that $\beta(x) = 0$; the case when $\beta(x) = 1$ can be handled in a similar fashion. We then define

$$\begin{array}{lll}
B_t^{\beta,s}(x) & = 0, & s \leq t < T_1 \\
B_t^{\beta,s}(x) & = 1, & T_1 \leq t < T_2 \\
B_t^{\beta,s}(x) & = 0, & T_2 \leq t < T_3 \\
B_t^{\beta,s}(x) & = 1, & T_3 \leq t < T_4 \\
 & \vdots &
\end{array}$$

where T_1 is the first arrival time of $M^{b,0\to 1}(x)$ after s, T_2 is the first arrival time of $M^{b,1\to 0}(x)$ after T_1, T_3 is the first arrival time of $M^{b,0\to 1}(x)$ after T_2, T_4 is the first arrival time of $M^{b,1\to 0}(x)$ after T_3 and so forth. In words, the points in $M^{b,0\to 1}$ are the times at which the background process switches to 1 (had it been in state 0) and similarily for $M^{b,1\to 0}$. Note importantly, we have all the processes $B^{\beta,s}$, as β and s vary, defined on the same probability space.

Given $B^{\beta,s}$, N^{δ_1} and $N^{\delta_0-\delta_1}$, define $X^{\beta,s} = \{X_t^{\beta,s}\}_{t\geq s}$, a point process on $\mathbb{Z}^d \times [s,\infty)$, in the following way:

$$X^{\beta,s} = \{\, (x,t) \in \mathbb{Z}^d \times [s,\infty) : \, (x,t) \in N^{\delta_1} \text{ or } \\ (x,t) \in N^{\delta_0-\delta_1} \text{ and } B_t^{\beta,s}(x) = 0 \,\}$$

In words, for each site x, we choose points in $[s,\infty)$ from $N^{\delta_1}(x)$ when the background process is in state 1 and from the union of $N^{\delta_1}(x)$ and $N^{\delta_0-\delta_1}(x)$ when the background process is in state 0.

Definition 2.2.3. Given space-time points (x,s) and (y,t) with $t > s$ and $\beta \in \{0,1\}^{\mathbb{Z}^d}$, we say that there is a β-active path from (x,s) to (y,t) if there is a sequence of times $s = s_0 < s_1 < \ldots < s_m < s_{m+1} = t$ and space points $x = x_0, x_1, \dots,$ $x_m = y$ so that for $i = 1, \dots, m$, there is an arrow from x_{i-1} to x_i at time s_i and there are no points in $X^{\beta,s}$ on the vertical segments $\{x_i\} \times (s_i, s_{i+1})$, $i = 0, \dots, m$.

Remark: Note importantly, that both $B^{\beta,s}$ and the existence of a β-active path from (x,s) to (y,t) are measurable with respect to the Poisson processes after time s and hence are independent of everything in the Poisson processes up to that time. The reason that these objects are introduced for $s > 0$ is that they are useful objects to which the original process can be usefully compared as will be done in the proof of Theorem 2.1.4.

To define the process $\{(B_t^\beta, C_t^{\beta,\eta})\}_{t\geq 0}$ for a given initial configuration $(\beta,\eta) \in \{0,1\}^{\mathbb{Z}^d} \times \{0,1\}^{\mathbb{Z}^d}$, we let $B_t^\beta = B_t^{\beta,0}$ and

$$C_t^{\beta,\eta} = \{\, y \in \mathbb{Z}^d : \text{ for some } x \in \mathbb{Z}^d \text{ with } \eta(x) = 1, \\ \text{there is a } \beta\text{-active path from } (x,0) \text{ to } (y,t) \,\}.$$

This is our formal definition of the CPREE. Note as β and η vary, we have all the processes $\{(B_t^\beta, C_t^{\beta,\eta})\}_{t\geq 0}$ defined on the same probability space.

Having defined $\{(B_t, C_t)\}_{t\geq 0}$ with initial configuration (β,η), it is a simple matter to extend the definition to an arbitrary initial distribution ρ. Just add to our probability space, independently of all the random variables already defined, two random variables on $\{0,1\}^{\mathbb{Z}^d}$ with joint distribution ρ. We will denote the probability measure governing all these variables by $\mathbf{P}_p$, suppressing γ, δ_0 and δ_1 in the notation.

The first easy fact about the CPREE we will show is that it is an attractive process.

Proposition 2.2.1. (B_t, C_t) *satisfies the attractivity condition:*

$$\rho \leq \sigma \quad \Longrightarrow \quad \rho S_p(t) \leq \sigma S_p(t) \quad \forall t > 0. \tag{2}$$

Proof. It is standard that (2) is equivalent to $(\delta_\beta \times \delta_\eta)S_p(t)$ being stochastically increasing in (β,η) for all $t \geq 0$. However, it is immediate from the construction that if $\beta_1 \leq \beta_2$ and $\eta_1 \leq \eta_2$, then for all $t \geq 0$

$$B_t^{\beta_1} \leq B_t^{\beta_2}$$

and

$$C_t^{\beta_1,\eta_1} \leq C_t^{\beta_2,\eta_2}.$$

This gives the stochastic domination (with an explicit coupling). $\square$

2.3 Proofs of Theorems 2.1.2 and 2.1.3

Recall, given γ, δ_0, $\delta_1 > 0$ with $\delta_1 \leq \delta_0$ and $q \in [0,1]$ we have defined

$$p_c(q,A) := \inf\left\{p : \mathbf{P}_p[\, C_t^{\pi_q,A} \neq \emptyset\, \forall t > 0\,] > 0\right\}$$

where $A \subseteq \mathbb{Z}^d$, $|A| < \infty$, and π_q denotes product measure with density q.

Proof of Theorem 2.1.2. We will prove the statements:

– For all $A \subseteq \mathbb{Z}^d$ with $|A| < \infty$ and p, $q \in [0,1]$,

$$\mathbf{P}_p[\, C_t^{\pi_q,A} \neq \emptyset\, \forall t > 0\,] > 0 \quad \Longleftrightarrow \quad \mathbf{P}_p[\, C_t^{\pi_q,\{0\}} \neq \emptyset\, \forall t > 0\,] > 0. \tag{3}$$

– For all $p \in [0,1]$,

$$\text{(4)} \quad \mathbf{P}_p[\, C_t^{\emptyset,\{0\}} \neq \emptyset \, \forall t > 0 \,] > 0 \quad \iff \quad \mathbf{P}_p[\, C_t^{\mathbb{Z}^d,\{0\}} \neq \emptyset \, \forall t > 0 \,] > 0.$$

Combining these two will yield the statement in Theorem 2.1.2. For (3), the left implication follows from translation invariance and the right implication follows easily from the additivity property of the process meaning

$$C_t^{\beta, A\cup B} = C_t^{\beta,A} \cup C_t^{\beta,B} \quad \forall A,B \subseteq \mathbb{Z}^d,\ \forall \beta \in \{0,1\}^{\mathbb{Z}^d}.$$

To prove (4), observe that the right implication is immediate from Proposition 2.2.1 and so we assume $\mathbf{P}_p[\, C_t^{\mathbb{Z}^d,\{0\}} \neq \emptyset \, \forall t > 0 \,] > 0$. Define

$$\varphi_t(x) = 1_{\{B_t^{\emptyset}(x) = B_t^{\mathbb{Z}^d}(x)\}} \quad x \in \mathbb{Z}^d,\ t \geq 0.$$

(Recall this is well defined since $\{B_t^{\emptyset}\}_{t\geq 0}$ and $\{B_t^{\mathbb{Z}^d}\}_{t\geq 0}$ are defined on the same probability space.) Note that φ_t has the property that for each site independently, after an exponentially distributed time with mean $\frac{1}{\gamma}$, the process flips to one and stays there. Therefore we have $\mathbf{P}_p[\, \varphi_t(x) = 1 \,] = 1 - e^{-\gamma t}$. For $A \subseteq \mathbb{Z}^d$, define $\{\tilde{C}_t^A\}_{t\geq 0}$ from the graphical representation in the same way as $\{C_t^{\cdot,A}\}_{t\geq 0}$ except that all recoveries are ignored. This is what is usually called the Richardson model, see Durrett [4].

Lemma 2.3.1. $\mathbf{P}_p[\, \tilde{C}_t^{\{0\}} \subseteq \varphi_t \,, \forall t \geq n \,] \to 1$ *as* $n \to \infty$.

Proof. Let $I_n = \{-n^2, \dots, n^2\}^d$ and for $x \in \mathbb{Z}^d$ define

$$t(x) = \inf\{\, t : \, x \in \tilde{C}_t^{\{0\}} \,\}.$$

From [4, p. 16], we get that there are constants $c_1, c_2, c_3 \in (0,\infty)$ such that

$$\mathbf{P}_p[\, t(x) < c_1 |x|_\infty \,] \leq c_2 e^{-c_3 |x|_\infty},$$

where $|\cdot|_\infty$ is the L^∞ norm. This easily gives us the estimate

$$\mathbf{P}_p[\, \tilde{C}_{c_1(n+1)}^{\{0\}} \nsubseteq I_n \,] \leq P(n) e^{-c_3 n},$$

where $P(n)$ is a polynomial in n, and from the Borel Cantelli lemma we can conclude

$$\text{(5)} \qquad \mathbf{P}_p[\, \exists N \geq 1 \text{ such that } \tilde{\mathrm{C}}_{\mathrm{c_1(n+1)}}^{\{0\}} \subseteq \mathrm{I_n} \,, \forall \mathrm{n} \geq \mathrm{N} \,] = 1.$$

Furthermore, independence gives

$$\mathbf{P}_p[\, I_n \subseteq \varphi_{c_1 n} \,] = (1 - e^{-\gamma c_1 n})^{(2n^2+1)^d}.$$

and since
$$\sum_{n=1}^{\infty} 1-(1-e^{-\gamma c_1 n})^{(2n^2+1)^d} < \infty,$$
the Borel Cantelli lemma again yields

$$\mathbf{P}_p[\,\exists N \geq 1 \text{ such that } \mathrm{I_n} \subseteq \varphi_{\mathrm{c_1 n}}\,, \forall \mathrm{n} \geq \mathrm{N}\,] = 1. \tag{6}$$

Combining (5) and (6), we obtain

$$\mathbf{P}_p[\,\exists N \geq 1 \text{ such that } \tilde{\mathrm{C}}_{\mathrm{t}}^{\{0\}} \subseteq \varphi_{\mathrm{t}}\,, \forall \mathrm{t} \geq \mathrm{N}\,] = 1,$$

as desired. □

Since $C_t^{\mathbb{Z}^d,\{0\}} \subseteq \tilde{C}_t^{\{0\}}$ $\forall t \geq 0$, the claim tells us that, with probability one, after some time and thereafter, the two background processes influence $C_t^{\emptyset,\{0\}}$ and $C_t^{\mathbb{Z}^d,\{0\}}$ in exactly the same way. Next, countable additivity gives us that for some $n \geq 1$ we have
$$\mathbf{P}_p[\,\tilde{C}_t^{\{0\}} \subseteq \varphi_t \;\forall t \geq n\,,\, C_t^{\mathbb{Z}^d,\{0\}} \neq \emptyset \;\forall t > 0\,] > 0$$
and then that for some m (depending on n)

$$\mathbf{P}_p[\,\tilde{C}_t^{\{0\}} \subseteq \varphi_t \;\forall t \geq n\,,\, \tilde{C}_t^{\{0\}} \subseteq [-m,m]^d \;\forall t \in [0,n],\, C_t^{\mathbb{Z}^d,\{0\}} \neq \emptyset \;\forall t > 0\,] > 0.$$

Denote the previous event by A and define the random set

$$U = \{\,(x,t) \in [-m,m]^d \times [0,n] : \; B_t^{\mathbb{Z}^d}(x) = 1\,\}$$

and let
$$B = \{\,\text{no arrivals in } N^{\delta_0-\delta_1} \text{ during } U\,\}.$$
It is clear that
$$A \cap B \subseteq \{\,C_t^{\emptyset,\{0\}} \neq \emptyset \;\forall t > 0\,\}$$
and so it remains to show that

$$\mathbf{P}_p[\,A \cap B\,] > 0.$$

However, if we condition on A and U, then we will not yield any information about the $N^{\delta_0-\delta_1}$ process on U and so

$$\mathbf{P}_p[\,B\,|A, U\,] = e^{-(\delta_0-\delta_1)\mathcal{L}(U)}$$

where $\mathcal{L}(U)$ is the "length" of U. This easily gives

$$\mathbf{P}_p[\,B|\,A\,] > 0$$

and the proof is complete. □

Remark: The same argument shows that strong survival does not depend on the initial distribution of the background process in the sense that

$$\mathbf{P}_p[0 \in C_t^{\emptyset,\{0\}} \text{ i.o.}] > 0 \quad \Longleftrightarrow \quad \mathbf{P}_p[0 \in C_t^{\mathbb{Z}^d,\{0\}} \text{ i.o.}] > 0.$$

This answers another question in [3].

Recall the definition of p_c' from the introduction:

$$p_c' := \inf\{\, p : \bar{\nu}_p \neq \pi_p \times \delta_\emptyset \,\}.$$

Here $\bar{\nu}_p = \lim_{t\to\infty}(\delta_{\mathbb{Z}^d} \times \delta_{\mathbb{Z}^d})S_p(t)$. (The limit exists due to Proposition 2.2.1.) To prove Theorem 2.1.3 we will use the next Lemma.

Lemma 2.3.2. *Given $p, q \in (0,1)$ with $q \geq p$ we have*

$$\lim_{t\to\infty}(\pi_q \times \delta_{\mathbb{Z}^d})S_p(t) = \bar{\nu}_p.$$

Proof. By simple stochastic comparison, it is enough to consider the case when $q = p$. We begin to establish the existence of that limit. Since π_p is the stationary distribution for the background process and the right marginal always occupies less than or equal to the whole $\{0,1\}^{\mathbb{Z}^d}$, we have

$$(\pi_p \times \delta_{\mathbb{Z}^d})S_p(t) \leq \pi_p \times \delta_{\mathbb{Z}^d} \quad \forall t > 0.$$

Using attractiveness and the Markov property yields

$$(\pi_p \times \delta_{\mathbb{Z}^d})S_p(s+t) \leq (\pi_p \times \delta_{\mathbb{Z}^d})S_p(s) \quad \forall s, t > 0,$$

and so the existence of the limit is clear from monotonicity. Denote this limit by ν_p' and observe it is necessarily stationary. It is clear that $\nu_p' \leq \bar{\nu}_p$ so we are done if $\bar{\nu}_p \leq \nu_p'$. For this, note that attractiveness again gives that the map

$$\mu \mapsto \mathbf{E}^\mu[f(\delta_t, \eta_t)]$$

is increasing whenever f is continuous and increasing. Using this, and the fact that any stationary distribution necessarily has as first marginal π_p, we can do the following calculation for any stationary distribution μ of (B_t, C_t) and $f : \{0,1\}^{\mathbb{Z}^d} \times \{0,1\}^{\mathbb{Z}^d} \to \mathbb{R}$ continuous and increasing:

$$\int f\, d\mu = \mathbf{E}^\mu[f(\delta_t, \eta_t)] \leq \mathbf{E}^{\pi_p \times \delta_{\mathbb{Z}^d}}[f(\delta_t, \eta_t)] \to \int f\, d\nu_p' \quad \text{as } t \to \infty.$$

Hence, $\mu \leq \nu_p'$ and we are done. □

Proof of Theorem 2.1.3. When the initial distribution of the background process is π_p, it is easy to see from the graphical representation that C_t is self-dual in the sense that

$$\mathbf{P}_p[\, C_t^{\pi_p, A} \cap B \neq \emptyset\,] = \mathbf{P}_p[\, C_t^{\pi_p, B} \cap A \neq \emptyset\,] \quad \forall t > 0,\ A, B \subseteq \mathbb{Z}^d. \tag{7}$$

If we take $A = \{0\}, B = \mathbb{Z}^d$ in this equation and let $t \to \infty$ using the previous lemma, we can easily conclude that

$$\mathbf{P}_p[\, C_t^{\pi_p,\{0\}} \neq \emptyset \, \forall t > 0 \,] > 0 \qquad \Longleftrightarrow \qquad \bar{\nu}_p \neq \pi_p \times \delta_\emptyset$$

and we are done. □

Remark: There is a weaker duality equation when the initial distribution of the background process differs from π_p, but this is less natural and seems less useful.

2.4 Proof of Theorem 2.1.4

We now turn to the proof of Theorem 2.1.4, that the critical CPREE dies out. Once Lemma 2.4.1 below is established, the rest follows similar lines as in the proofs of Theorem 2.1.1 carried out in [1] and [9]. Our main goal is to prove that if $\{C_t\}$ survives at $p > 0$, then there is a number $\delta > 0$ and integers n, a such that

(8) $$\mathbf{P}_{p-\delta}[\, C_t^{\emptyset,[-n,n]^d} \text{ survives in } \mathbb{Z} \times [-5a, 5a]^{d-1} \times [0, \infty) \,] > 0.$$

If $p_c \in (0, 1]$, this will immediately imply

$$\mathbf{P}_{p_c}[\, C_t^{\emptyset,\{0\}} \neq \emptyset \, \forall t \geq 0 \,] = 0.$$

To achieve (8), we begin by showing that if the CPREE survives, then it is very likely to have survival if the initial configuration is sufficiently large even if we start with all zeros in the background process.

Lemma 2.4.1. *If $\{C_t\}$ survives at $p > 0$ then*

$$\lim_{n \to \infty} \mathbf{P}_p[\, C_t^{\emptyset,[-n,n]^d} \neq \emptyset \, \forall t > 0 \,] = 1.$$

For the proof of this we use the following result.

Lemma 2.4.2. *For all $n \geq 1$, we have*

$$\lim_{\epsilon \to 0} \mathbf{P}_p[\, C_t^{\pi_{p-\epsilon},[-n,n]^d} \neq \emptyset \, \forall t > 0 \,] = \mathbf{P}_p[\, C_t^{\pi_p,[-n,n]^d} \neq \emptyset \, \forall t > 0 \,].$$

Proof. Fix $n \geq 1$. The probability on the left increases when ϵ decreases and so the limit exists and is clearly at most the right hand side. For the other inequality let $\delta > 0$ and define

$$\varphi_t^\epsilon(x) = 1_{\{B_t^{\pi_{p-\epsilon}}(x) = B_t^{\pi_p}(x)\}} \qquad x \in \mathbb{Z}^d, \ t \geq 0,$$

where $\pi_{p-\epsilon}$ and π_p are coupled in the usual monotone way. Recall the definition of φ_t from the proof of Theorem 2.1.2 and observe that

$$\varphi_t \subseteq \varphi_t^\epsilon \quad \forall t > 0, \ \forall \epsilon > 0.$$

Also, an easy modification of the proof of Lemma 2.3.1 yields

$$\lim_{T\to\infty} \mathbf{P}_p[\,\tilde{C}_t^{[-n,n]^d} \subseteq \varphi_t \,, \forall t \geq T\,] = 1.$$

(Recall that $\tilde{C}_t^A$ is the CPREE starting from the configuration A but with no recoveries.) This allows us to choose $T > 0$ such that

$$\begin{aligned}&\mathbf{P}_p[\,C_t^{\pi_p,[-n,n]^d} \neq \emptyset\,\forall t > 0\,]\\ &\quad\leq \mathbf{P}_p[\,\tilde{C}_t^{[-n,n]^d} \subseteq \varphi_t \,, \forall t \geq T,\, C_t^{\pi_p,[-n,n]^d} \neq \emptyset\,\forall t > 0\,] + \delta.\end{aligned}$$

Given this T, choose $m \geq 1$ such that

$$\mathbf{P}_p[\,\tilde{C}_t^{[-n,n]^d} \subseteq [-m,m]^d\,\forall\, 0 \leq t \leq T\,] > 1 - \delta$$

and for that m choose $\epsilon_0 > 0$ such that

$$\mathbf{P}_p[\,B_0^{\pi_{p-\epsilon}} = B_0^{\pi_p} \text{ on } [-m,m]^d\,] > 1-\delta, \quad \forall\, 0 < \epsilon \leq \epsilon_0.$$

Now since

$$\begin{aligned}&\{\,\tilde{C}_t^{[-n,n]^d} \subseteq \varphi_t \,, \forall t \geq T,\, \tilde{C}_t^{[-n,n]^d} \subseteq [-m,m]^d\,\forall\, 0 \leq t \leq T,\\ &\quad B_0^{\pi_{p-\epsilon}} = B_0^{\pi_p} \text{ on } [-m,m]^d,\, C_t^{\pi_p,[-n,n]^d} \neq \emptyset\,\forall t > 0\,\}\\ &\quad\subseteq \{\,C_t^{\pi_{p-\epsilon},[-n,n]^d} \neq \emptyset\,\forall t > 0\,\},\end{aligned}$$

we get

$$\begin{aligned}&\mathbf{P}_p[\,C_t^{\pi_p,[-n,n]^d} \neq \emptyset\,\forall t > 0\,]\\ &\quad\leq \mathbf{P}_p[\,C_t^{\pi_{p-\epsilon},[-n,n]^d} \neq \emptyset\,\forall t > 0\,] + 3\delta,\end{aligned}$$

whenever $0 < \epsilon \leq \epsilon_0$ and so the proof is complete. □

Proof of Lemma 2.4.1. Let $\delta > 0$. From the self-duality equation (7), Lemma 2.3.2 and the easily verified fact that the second marginal of $\bar{\nu}_p$ gives zero measure to $\emptyset$, we easily get that there is an $n \geq 1$ such that

$$\mathbf{P}_p[\,C_t^{\pi_p,[-n,n]^d} \neq \emptyset\,\forall t > 0\,] > 1 - \delta.$$

The previous lemma makes it possible to now choose an $\epsilon > 0$ such that

$$\mathbf{P}_p[\,C_t^{\pi_{p-\epsilon},[-n,n]^d} \neq \emptyset\,\forall t > 0\,] > 1 - \delta.$$

Denote the semigroup operator associated with the background process by $T(t)$ and note that for ϵ above there is a time s such that

$$\delta_\emptyset T(s) \geq \pi_{p-\epsilon}.$$

Now, let $B_{m,n}$ denote the box in $\mathbb{Z}^d$ with sidelength mn and write

$$B_{m,n} = \bigcup_{i=1}^{m^d} A_i,$$

where each A_i is a translation of the box with sidelength n and with the A_i's disjoint. Then, define

$$A^s_{m,n} = \{ \text{No arrivals in } N^{\delta_1} \text{ or } N^{\delta_0 - \delta_1} \text{ up to time } s \text{ in some } A_i \}.$$

Given n and s, we can choose m so large that

$$\mathbf{P}_p[A^s_{m,n}] > 1 - \delta.$$

The proof is finished by noting that monotonicity easily implies that

$$\mathbf{P}_p[C_t^{\emptyset,[-mn,mn]^d} \neq \emptyset \, \forall t > 0 \mid A^s_{m,n}] \geq \mathbf{P}_p[C_t^{\pi_{p-\epsilon},[-n,n]^d} \neq \emptyset \, \forall t > 0],$$

using the fact that $A^s_{m,n}$ is independent of the background process. □

Remark: A slightly more abstract but considerably shorter proof of Lemma 2.4.1 is found by Olle Häggström after submission of the paper and is as follows. For $x \in \mathbb{Z}^d$, let $Y_x^{\emptyset}$ be the indicator variable for survival when the process starts with only x infected and all zeros in the background process. By translation invariance, $\mathbf{P}_p[Y_x^{\emptyset} = 1]$ is independent of x and by Theorem 2.1.2 we know that it is positive. It follows from the graphical representation that the process $\{Y_x^{\emptyset}\}_{x \in \mathbb{Z}^d}$ is ergodic and hence a.s. there is some x for which $Y_x^{\emptyset} = 1$. Moreover, the event in Lemma 2.4.1 occurs as soon as some site in $[-n,n]^d$ has $Y_x^{\emptyset} = 1$ and so the lemma follows at once.

We have now set up the necessary ground work for our model in order to be able to follow the steps in [9]. For $L \geq 1$ and $A \subseteq (-L,L)^d$, let ${}_LC_t^{\emptyset,A}$ be the truncated process, using only $\emptyset$-active paths (recall Definition 2.2.3) which stay in $(-L,L)^d \times [0,t]$.

Lemma 2.4.3. *For all finite $A \subseteq \mathbb{Z}^d$ and $N \geq 1$, we have*

$$\lim_{t \to \infty} \lim_{L \to \infty} \mathbf{P}_p[|{}_LC_t^{\emptyset,A}| \geq N] = \mathbf{P}_p[C_t^{\emptyset,A} \neq \emptyset \, \forall t > 0]$$

Proof. Fix A and N. Since

$$C_t^{\emptyset,A} = \bigcup_{L=1}^{\infty} {}_LC_t^{\emptyset,A},$$

we easily get that for fixed t

$$\mathbf{P}_p[|C_t^{\emptyset,A}| \geq N] = \lim_{L \to \infty} \mathbf{P}_p[|{}_LC_t^{\emptyset,A}| \geq N],$$

and so we are done if

$$\lim_{t\to\infty} \mathbf{P}_p[\,|C_t^{\emptyset,A}| \geq N\,] = \mathbf{P}_p[\,C_t^{\emptyset,A} \neq \emptyset\,\forall t > 0\,].$$

For this, it is enough to check two things:

$$\lim_{t\to\infty} \mathbf{P}_p[\,|C_t^{\emptyset,A}| \geq N, C_s^{\emptyset,A} = \emptyset \text{ some } s > 0\,] = 0$$
$$\lim_{t\to\infty} \mathbf{P}_p[\,|C_t^{\emptyset,A}| \geq N, C_s^{\emptyset,A} \neq \emptyset\,\forall s > 0\,] = \mathbf{P}_p[\,C_t^{\emptyset,A} \neq \emptyset\,\forall t > 0\,]$$

The first equality follows easily by applying Fatou's Lemma. The second one follows if

$$\lim_{t\to\infty} |C_t^{\emptyset,A}| = \infty \quad \text{a.s} \quad \text{on} \quad \{\,C_t^{\emptyset,A} \neq \emptyset\,\forall t > 0\,\}.$$

Assume the contrary, i.e.

(9) $$\mathbf{P}_p[\,|C_t^{\emptyset,A}| \text{ does not converges to infinity, } C_s^{\emptyset,A} \neq \emptyset\,\forall s > 0\,] > 0.$$

From the martingale convergence theorem we get that

(10) $$\mathbf{P}_p[\,C_t^{\emptyset,A} \neq \emptyset\,\forall t \geq s \,|\, \mathcal{F}_s\,] \to 1_{\{\,C_t^{\emptyset,A} \neq \emptyset\,\forall t > 0\,\}} \quad \text{as } s \to \infty,$$

where $\mathcal{F}_s$ is the σ-algebra generated by the whole process up to time s. Equation (9) and (10) implies that with positive probability the following can happen:

$$\lim_{s\to\infty} \mathbf{P}^{(\beta_s, C_s)}[\,C_t \neq \emptyset\,\forall t > 0\,] = 1$$
$$\exists\, M > 0,\ \{\tau_i\}_{i\geq 1} \ni \tau_1 < \tau_2 < \ldots < \tau_i \to \infty\,, |C_{\tau_i}| \leq M\ \forall i.$$

However, using elementary facts about exponentially distributed variables, we get

$$\mathbf{P}^{(\beta_{\tau_i}, C_{\tau_i})}[\,C_t = \emptyset \text{ some } t > 0\,]$$
$$\geq \mathbf{P}^{(\mathbb{Z}^d, C_{\tau_i})}[\,C_t = \emptyset \text{ some } t > 0\,] \geq \left(\frac{\delta_1}{\delta_0 + \gamma + 2d}\right)^M \quad \forall i,$$

which yields a contradiction and the proof is complete. □

The next step is to take care of the sides of the space-time box. Define

$$S(L,T) = \{\,(x,t) \in \mathbb{Z}^d \times [0,T] :\, |x|_\infty = L\,\}.$$

Fix $A \subseteq (-L,L)^d$ and look at all points on $S(L,T)$ that can be reached from A by an $\emptyset$-active path using vertical segments where the space coordinate is in $(-L,L)^d$ and infection arrows from $(x,\cdot)$ to $(y,\cdot)$ with $x \in (-L,L)^d$. Define $N_\emptyset^A(L,T)$ to be the maximum number of such points with the following property: If (x,t_1) and (x,t_2) are any two points with the same spatial coordinate, then $|t_1 - t_2| \geq 1$.

Lemma 2.4.4. *Assume $L_j \nearrow \infty$ and $T_j \nearrow \infty$. Then for any $M, N \geq 1$ and finite $A \subseteq \mathbb{Z}^d$, we have*

$$\limsup_{j\to\infty} \mathbf{P}_p[\, N_\emptyset^A(L_j, T_j) \leq M\,]\, \mathbf{P}_p[\, |_{L_j} C_{T_j}^{\emptyset,A}| \leq N\,] \leq \mathbf{P}_p[\, C_t^{\emptyset,A} = \emptyset \text{ some } t > 0\,].$$

Proof. The proof follows the steps of Proposition 2.8 in [9] with some adjustments. Let $\mathcal{F}_{L,T}$ denote the σ-algebra generated by $M^{b,0\to 1}$, $M^{b,1\to 0}$, N^{δ_1}, $N^{\delta_0-\delta_1}$ and $\vec{N}^j$, $j \in \{\pm e_1, \ldots, \pm e_d\}$ in $(-L, L)^d \times [0, T]$. We first argue that

$$\begin{aligned} \mathbf{P}_p[\, C_t^{\emptyset,A} = \emptyset \text{ some } t > 0 \,|\, \mathcal{F}_{L,T}\,] &\geq \left(\frac{e^{-4d}\delta_1}{\delta_0+\gamma+2d}\right)^k \\ \text{a.s} \quad \text{on} \quad \{\, N_\emptyset^A(L,T) + |_L C_T^{\emptyset,A}| &\leq k\,\} \end{aligned} \tag{11}$$

For $x \in {}_L C_T^{\emptyset,A}$ there is a conditional probability of at least

$$\frac{\delta_1}{\delta_0+\gamma+2d}$$

that x becomes healthy before it infects any of its neighbors. So, if $|_L C_T^{\emptyset,A}| = m$, then the conditional probability that no $x \in {}_L C_T^{\emptyset,A}$ contributes to survival is at least

$$\left(\frac{\delta_1}{\delta_0+\gamma+2d}\right)^m .$$

For the sides of the box, consider a time line $\{x\} \times [0, T]$, where $|x|_\infty = L$ and let

$$(x, t_1), \ldots, (x, t_j)$$

be a maximal set of points that can be reached from A by an $\emptyset$-active path with the property that each pair is separated by at least distance 1. Let

$$I = \bigcup_{k=1}^{j} \{x\} \times (t_k - 1, t_k + 1)$$

and note that the probability that there are no arrows coming out from I is at least e^{-4dj}. Furthermore, for each interval of length y in the complement of I in $\{x\} \times [0, \infty)$, the probability of the event that if there is at least one arrival of the Poisson processes in the interval with the first one coming from N^{δ_1} or there is no arrivals at all is

$$\left(1 - e^{-(\delta_0+\gamma+2d)y}\right) \frac{\delta_1}{\delta_0+\gamma+2d} + e^{-(\delta_0+\gamma+2d)y} \geq \frac{\delta_1}{\delta_0+\gamma+2d}.$$

By independence, we get that the conditional probability that none of the points in the time line $\{x\} \times [0, T]$ contributes to survival is at least

$$\left(\frac{e^{-4d}\delta_1}{\delta_0+\gamma+2d}\right)^j .$$

Now, considering the contribution of different x's yields

$$\begin{aligned} &\mathbf{P}_p[\, C_t^{\emptyset,A} = \emptyset \text{ some } t > 0 \,|\, \mathcal{F}_{L,T} \,] \\ &\geq \left(\frac{\delta_1}{\delta_0 + \gamma + 2d}\right)^{|_L C_T^{\emptyset,A}|} \left(\frac{e^{-4d}\delta_1}{\delta_0 + \gamma + 2d}\right)^{N^A(L,T)} \end{aligned}$$

which implies (11). For the rest of the proof, one proceeds exactly as in the second half of Proposition 2.8 in [9, p. 48-49]. The needed inequality

$$\begin{aligned} &\mathbf{P}_p[\, N_\emptyset^A(L,T) \leq M, |_L C_T^{\emptyset,A}| \leq N \,] \\ &\quad \geq \mathbf{P}_p[\, N_\emptyset^A(L,T) \leq M \,] \mathbf{P}_p[\, |_L C_T^{\emptyset,A}| \leq N \,] \end{aligned}$$

is justified by the fact that $N_\emptyset^A(L,T)$ and $|_L C_T^{\emptyset,A}|$ are increasing functions of $\vec{N}^j$, $j \in \{\, \pm e_1, \ldots, \pm e_d \,\}$ and $M^{b,0\to 1}$, and decreasing in N^{δ_1}, $N^{\delta_0 - \delta_1}$ and $M^{b,1\to 0}$. This completes the proof. $\square$

We are soon ready to state and prove the so called finite space-time condition. However, we first need two more propositions. We just state them here since the proofs are exactly the same as for Propositions 2.6 and 2.11, pages 46-47 and 49 in [9].

Proposition 2.4.5. *For every $n, N \geq 1$ and $L \geq n$, we have*

$$\mathbf{P}_p[\, |_L C_t^{\emptyset,[-n,n]^d} \cap [0,L)^d| \leq N \,] \leq \left(\mathbf{P}_p[\, |_L C_t^{\emptyset,[-n,n]^d}| \leq 2^d N \,]\right)^{2^{-d}}$$

Let

$$S_+(L,T) = \{\, (x,t) \in \mathbb{Z}^d \times [0,T] :\, x_1 = L\,, x_i \geq 0,\ 2 \leq i \leq d \,\}$$

and define $N_{\emptyset,+}^A(L,T)$ in a similar manner as $N_\emptyset^A(L,T)$ using $S_+(L,T)$ instead of $S(L,T)$.

Proposition 2.4.6. *For any $L, M \geq 1$, $T > 0$ and $n < L$,*

$$\left(\mathbf{P}_p[\, N_{\emptyset,+}^{[-n,n]^d}(L,T) \leq M \,]\right)^{d2^d} \leq \mathbf{P}_p[\, N_\emptyset^{[-n,n]^d}(L,T) \leq M d 2^d \,]$$

The proof of these propositions requires certain random variables to be positively correlated. For Proposition 2.4.5, let $X_1 = |_L C_t^{[-n,n]^d} \cap [0,L)^d|$ and $X_2, \ldots, X_{2^d}$ be defined similarly with respect to the other orthants in $\mathbb{R}^d$. The needed positive correlation of $\{X_i\}_{i=1}^{2^d}$ is justified in the same way as in the end of the proof of Lemma 2.4.4. Similarly justification can be made in the proof of Proposition 2.4.6.

Theorem 2.4.7. *If $\{C_t\}$ survives at $p > 0$, then it satisfies the following condition: For all $\epsilon > 0$ there exist $n, L \geq 1$ and $T > 0$ such that*

$$\mathbf{P}_p[\, {}_{L+n}C_{T+1}^{\emptyset,[-n,n]^d} \supseteq x + [-n,n]^d \text{ some } x \in [0,L)^d \,] > 1 - \epsilon \tag{12}$$

$$\begin{aligned} \mathbf{P}_p[\, {}_{L+2n+1}C_{t+1}^{\emptyset,[-n,n]^d} &\supseteq x + [-n,n]^d \text{ some } 0 \leq t < T, \\ &\text{some } x \in \{L+n\} \times [0,L)^{d-1} \,] > 1 - \epsilon \end{aligned} \tag{13}$$

Proof. Again, we will follow the steps in [9] with some modifications. Let $0 < \delta < 1$. We will see at the end how to choose δ for a given $\epsilon > 0$. Lemma 2.4.1 gives us an n such that

$$\mathbf{P}_p[\, C_t^{\emptyset,[-n,n]^d} \neq \emptyset \,\forall t > 0\,] > 1 - \delta^2. \tag{14}$$

Given n, choose N' such that

$$\Big(1 - \mathbf{P}_p[\, {}_{n+1}C_1^{\emptyset,\{0\}} \supseteq [-n,n]^d\,]\Big)^{N'} < \delta$$

and then choose N so large such that if $A \subseteq \mathbb{Z}^d$ with $|A| \geq N$, then there exists $B \subseteq A$ with $|B| \geq N'$ and

$$|x - y|_\infty \geq 2n + 1 \quad \forall\, x, y \in B,\ x \neq y.$$

Let B_A be a fixed (deterministic) such choice for each A.

In a similar fashion, choose M' such that

$$(1 - a)^{M'} < \delta, \tag{15}$$

where

$$\begin{aligned} a = \mathbf{P}_p[\, &\text{There are } \emptyset\text{-active paths from the origin to every} \\ &\text{point in } [0,2n] \times [-n,n]^{d-1} \times \{1\} \text{ that} \\ &\text{stays in } [0,2n] \times [-n,n]^{d-1} \times [0,1]\,] \end{aligned}$$

Then choose M so large such that if $A \subseteq \mathbb{Z}^d \times [0, \infty)$ is a finite set with $|A| \geq M$, where the distance in time between points with the same spatial coordinate is at least 1, then there exists $B \subseteq A$ with $|B| \geq M'$ and with the property that for each pair of points $(x, s), (y, t) \in B$ we have either

$$x = y, \quad |s - t| \geq 1 \qquad \text{or} \qquad |x - y|_\infty \geq 2n + 1. \tag{16}$$

Let B_A be a fixed (deterministic) such choice for each A.

From Lemma 2.4.3, (14), the inequality $1 - \delta < 1 - \delta^2$ and the facts that for fixed L, n and N, the map $t \mapsto \mathbf{P}_p[\,|{}_LC_t^{\emptyset,[-n,n]^d}| > 2^dN\,]$ is continuous and that $\lim_{t\to\infty} \mathbf{P}_p[\,|{}_LC_t^{\emptyset,[-n,n]^d}| > 2^dN\,] = 0$, we can conclude that there exist $L_j \nearrow \infty$ and $T_j \nearrow \infty$ so that

$$\mathbf{P}_p[\,|{}_{L_j}C_{T_j}^{\emptyset,[-n,n]^d}| > 2^dN\,] = 1 - \delta \quad \forall j \geq 1.$$

Furthermore, Lemma 2.4.4 with M and N replaced by $Md2^d$ and 2^dN respectively and with $A = [-n,n]^d$, we get that for some j

$$\mathbf{P}_p[\, N_\emptyset^{[-n,n]^d}(L_j, T_j) > Md2^d\,] > 1 - \delta.$$

Let $L = L_j$ and $T = T_j$ for that specific j and apply Propositions 2.4.5 and 2.4.6 to get

$$\mathbf{P}_p[\, |{}_L C_T^{\emptyset,[-n,n]^d} \cap [0,L)^d| > N \,] > 1 - \delta^{2^{-d}} \tag{17}$$

$$\mathbf{P}_p[\, N_{\emptyset,+}^{[-n,n]^d}(L,T) > M \,] > 1 - \delta^{2^{-d}}/d. \tag{18}$$

To obtain (12), define for $B \subseteq \mathbb{Z}^d$ and $T > 0$

$$\begin{aligned} V_B^T = \{\, \exists\,(x,t) \in B \times \{T\} &\text{ such that there are } \emptyset\text{-active paths from} \\ &(x,t) \text{ to every } (y,s) \in \big(x + [-n,n]^d\big) \times \{T+1\} \\ &\text{that stays in } \big(x + [-n,n]^d\big) \times (T, T+1] \,\}, \end{aligned}$$

and note that

$$\begin{aligned} \bigcup_{A \subseteq [0,L)^d} &\{\, |{}_L C_T^{\emptyset,[-n,n]^d} \cap [0,L)^d| > N,\ {}_L C_T^{\emptyset,[-n,n]^d} \cap [0,L)^d = A,\ V_{B_A}^T \,\} \\ &\subseteq \{\, {}_{L+n} C_{T+1}^{\emptyset,[-n,n]^d} \supseteq x + [-n,n]^d \text{ some } x \in [0,L)^d \,\}. \end{aligned} \tag{19}$$

Let $\mathcal{F}_T$ be the σ-algebra generated by $M^{b,0\to 1}$, $M^{b,1\to 0}$, N^{δ_1}, $N^{\delta_0 - \delta_1}$, and $\vec{N}^j$, $j \in \{\, \pm e_1, \dots, \pm e_d \,\}$ up to time T and note that for given $A \subseteq [0,L)^d$ with $|A| \geq N$, $V_{B_A}^T$ is independent of $\mathcal{F}_T$ so

$$\begin{aligned} \mathbf{P}_p[\, V_{B_A}^T \,|\, \mathcal{F}_T \,] &= \mathbf{P}_p[\, V_{B_A}^T \,] \\ &\geq 1 - \Big(1 - \mathbf{P}_p[\, {}_{n+1} C_1^{\emptyset,\{0\}} \supseteq [-n,n]^d \,]\Big)^{N'} > 1 - \delta. \end{aligned}$$

By summing up over $A \subseteq [0,L)^d$ and using (17) and (19), we get

$$\begin{aligned} \mathbf{P}_p[\, &{}_{L+n} C_{T+1}^{\emptyset,[-n,n]^d} \supseteq x + [-n,n]^d \text{ some } x \in [0,L)^d \,] \\ &> (1-\delta)(1 - \delta^{2^{-d}}). \end{aligned}$$

This yields (12) when δ is chosen appropriately.

To obtain (13), define for each space-time point (x_i, t_i) we count in $N_{\emptyset,+}^{[-n,n]^d}(L,T)$ a variable $\tilde{Y}_i$ which is 1 if (x_i, t_i) infects all points in

$$\big(x_i + [0,2n] \times [-n,n]^{d-1}\big) \times \{t_i + 1\}$$

using $\emptyset$-active paths in

$$\big(x_i + [0,2n] \times [-n,n]^{d-1}\big) \times (t_i, t_i + 1]$$

only and 0 otherwise. If $N_{\emptyset,+}^{[-n,n]^d}(L,T) > M$, we can choose M' space-time points satisfying (16). Denote the corresponding variables by Y_i, $i = 1, \dots, M'$. Let $\mathcal{F}_{L,T}$ be as in the proof of Lemma 2.4.4 and note that conditioned on $\mathcal{F}_{L,T}$ restricted to the event $\{ N_{\emptyset,+}^{[-n,n]^d}(L,T) > M \}$, the M' space-time points are specified and $Y_1, Y_2, \dots, Y_{M'}$ are independent with the (conditional) probability of $Y_i = 1$ equal to a. This implies that

$$\mathbf{P}_p[\, Y_i = 1 \text{ some } i = 1, \dots, M' \,|\, \mathcal{F}_{L,T}\,] = 1 - (1-a)^{M'} \quad \text{on} \quad \{ N_{\emptyset,+}^{[-n,n]^d}(L,T) > M \},$$

which together with (15) and (18) yields

$$\begin{aligned}\mathbf{P}_p[\, {}_{L+2n+1}C_{t+1}^{\emptyset,[-n,n]^d} \supseteq x + [-n,n]^d \text{ some } 0 \le t < T,\\ \text{some } x \in \{L+n\} \times [0,L)^{d-1}\,]\\ > (1-\delta)(1-\delta^{2^{-d}/d}).\end{aligned}$$

This gives (13) when δ is chosen appropriately. □

The next part of the program is to carry out a comparison with oriented percolation. For this, we start to combine (12) and (13) into one.

Lemma 2.4.8. *If $\{C_t\}$ survives at $p > 0$, then it satisfies the following condition: For all $\epsilon > 0$ there exist $n, L \ge 1$ and $T > 0$ such that*

$$\begin{aligned}\mathbf{P}_p[\, {}_{2L+3n}C_t^{\emptyset,[-n,n]^d} \supseteq x + [-n,n]^d \textit{ some } T \le t < 2T,\\ \textit{some } x \in [L+n, 2L+n] \times [0,2L)^{d-1}\,] > 1 - \epsilon\end{aligned} \tag{20}$$

Proof. We follow Proposition 2.20 in [9]. Let (x,τ) be the first (in time) space-time point with the property appearing in the probability (13), where x is choosen according to some deterministic ordering of $\mathbb{Z}^d$ and restart (B_t, C_t) at time $\tau + 1$. From (12), (13) and the fact that these probabilities are increasing in the background process, it follows that

$$\begin{aligned}\mathbf{P}_p[\, {}_{2L+3n}C_t^{\emptyset,[-n,n]^d} \supseteq x + [-n,n]^d \text{ some } T+1 \le t < 2T+2,\\ \text{some } x \in [L+n, 2L+n] \times [0,2L)^{d-1}\,] > (1-\epsilon)^2.\end{aligned}$$

Replace $T+1$ with T and the proof is complete. □

Now we are ready for the fundamental step in the construction towards the comparison.

Lemma 2.4.9. *Assume $\{C_t\}$ survives at $p > 0$ and fix $\epsilon > 0$. Then there exist $\delta > 0$, n, a, b with $n < a$ such that for all $(x,t) \in [-a,a]^d \times [0,b]$*

$$\begin{aligned}\mathbf{P}_{p-\delta}[\, \exists\, (y,s) \in [a,3a] \times [-a,a]^{d-1} \times [5b,6b] \textit{ such that}\\ \textit{there are } \emptyset\textit{-active paths from } (x,t) + \big([-n,n]^d \times \{0\}\big)\\ \textit{to every point in } (y,s) + \big([-n,n]^d \times \{0\}\big)\\ \textit{that stays in } [-5a,5a]^d \times [0,6b]\,] > 1 - \epsilon.\end{aligned}$$

Proof. One can proceed exactly as in Proposition 2.22 in [9, p. 52-53] to first obtain the statement with $p-\delta$ replaced by p and therefore we only outline this part of the argument. The main idea is to use Lemma 2.4.8 (or a "reflected" version of it) repeatedly (between 4 to 10 times) to steer things properly so that the desired event occurs. The existence of $\delta > 0$ is a consequence of the fact that the event in question depends only on the graphical representation in $[-5a, 5a]^d \times [0, 6b]$ and hence is continuous in p. ☐

Repeated use of the previous lemma together with appropriate stopping times and monotonicity in the background process yields:

Lemma 2.4.10. *Assume $\{C_t\}$ survives at $p > 0$ and let $\epsilon > 0$ and $k \geq 1$ be fixed. Then there exist $\delta > 0$, n, a, b with $n < a$ such that the following holds: For all $(x,t) \in [-a,a]^d \times [0,b]$, with $\mathbf{P}_{p-\delta}$-probability at least $1-\epsilon$, there exists a translate $(y,s) + [-n,n]^d \times \{0\}$ of $[-n,n]^d \times \{0\}$ such that*

a) $(y,s) \in ([-a,a] + 2ka) \times [-a,a]^{d-1} \times [5kb, (5k+1)b]$

b) There are $\emptyset$-active paths from $(x,t) + [-n,n]^d \times \{0\}$ to every point in $(y,s) + [-n,n]^d \times \{0\}$ that stays in the region

$$\mathcal{A} = \bigcup_{j=0}^{k-1} ([-5a, 5a] + 2ja) \times [-5a, 5a]^{d-1} \times ([0, 6b] + 5jb).$$

Our final step towards (8) is to use the previous lemma in a so called renormalization argument. The set $\mathcal{A}$ from Lemma 2.4.10 (see Figure 2.1) and its reflection with respect to the t-axis will consist of our building blocks. Given the conditions in Lemma 2.4.10, the distance c in Figure 2.2 is well defined. (Define it to be zero if the dashed vertical line is to the right of the left corner of the rectangle R, see Figure 2.2.) It is easy to see that, if we choose $k > 5$, c will be bigger than $3a$, independent of the value of a. Fix such a k.

Theorem 2.4.11. *If $\{C_t\}$ survives at $p > 0$, then there are integers n,a and $\delta > 0$ such that*

$$\mathbf{P}_{p-\delta}[\, C_t^{\emptyset,[-n,n]^d} \textit{ survives in } \mathbb{Z} \times [-5a, 5a]^{d-1} \times [0,\infty)\,] > 0$$

Proof. The proof is a modification of Lemma 21 of [1]. Let $\eta > 0$ be given and take $\epsilon > 0$ such that $1-\epsilon > 1-\eta$ and let n, a, b and δ be as in Lemma 2.4.10. We will make an appropriate choice of η later. Construct a process $Z_n(i) = (X_n(i), Y_n(i))$, $i \geq 0, n \geq 0$, where $X_n(i) \in \{0,1\}$ and $Y_n(i)$ is a point in $\mathbb{Z}^d \times [0,\infty)$. $Y_n(i)$ will be undefined when $X_n(i) = 0$. Start with $Z_0(0) = (1,0)$, $X_0(i) = 0, i \neq 0$ and define inductively as follows: With $Z_k(i)$ already defined for $i \geq 0, 0 \leq k \leq n$ let $X_{n+1}(i) = 1$ if for either $j = i$ or $j = i-1$ it is the case that $X_n(j) = 1$ and there is a translation of $[-n,n]^d$ to the shaded area (see Figure 2.3 for the shaded regions) on the top of the

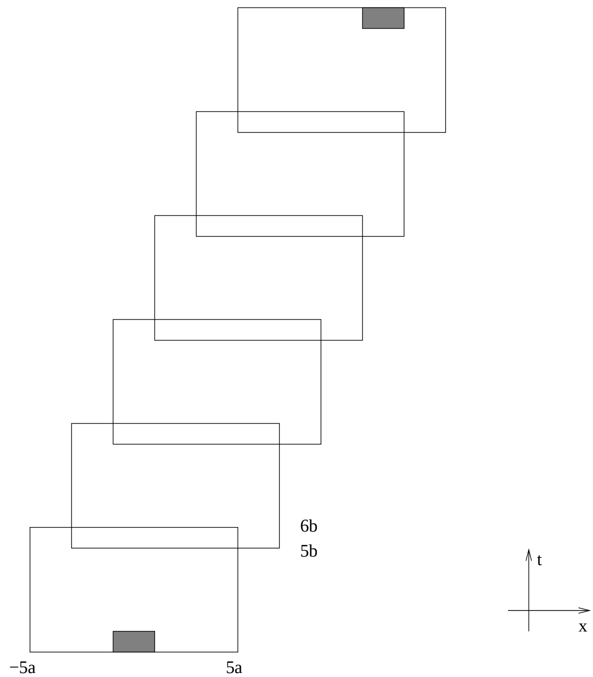

Figure 2.1: *The set $\mathcal{A}$.*

corresponding block such that $Y_n(j) + [-n,n]^d$ is connected with $\emptyset$-active paths to every point in that translation. Furthermore, define $Y_{n+1}(i) = (x_{n+1}(i), t_{n+1}(i))$, where $t_{n+1}(i)$ is the earliest center of such a translation and $x_{n+1}(i)$ is chosen according to some fixed ordering of $\mathbb{Z}^d$. Note that if $X_n(i) = 1$ for infinitely many pairs (i,n), then $C_t^{\emptyset,[-n,n]^d}$ survives in $\mathbb{Z} \times [-5a, 5a]^{d-1} \times [0, \infty)$ so it remains to prove that the former has positive probability. Let $\mathcal{F}_n$ be the σ-algebra generating by $Z_k(i)$, where $i \geq 0$, $0 \leq k \leq n$ and note that from Lemma 2.4.10 we get

$$\mathbf{P}_p[\, X_{n+1}(i) = 1 \,|\, \mathcal{F}_n \,] > 1 - \eta \quad \text{on} \quad \{\, X_n(i-1) = 1 \text{ or } X_n(i) = 1 \,\}.$$

Also, our choice of k and the fact that events that depend on disjoint parts of the graphical representation are independent, we have that, conditioned on $\mathcal{F}_n$, the collection of variables $\{\, X_{n+1}(i) \;:\; i \geq 0 \,\}$ is one-dependent. Now, we are ready to make the construction above for a specific choice of η. Take $1/4 \leq p < 1$ so large that an oriented percolation process, $\{A_n\}$, on $\mathbb{N}$ with parameter p survives with positive probability when it starts with a single infection at the origin and choose η such that $1 - \eta > 1 - (1 - \sqrt{p})^3$. A result of Liggett, Schonmann and Stacey [10] (see also Theorem B26 [9]) tells us that a one-dependent process with density $1 - \eta$ stochastically

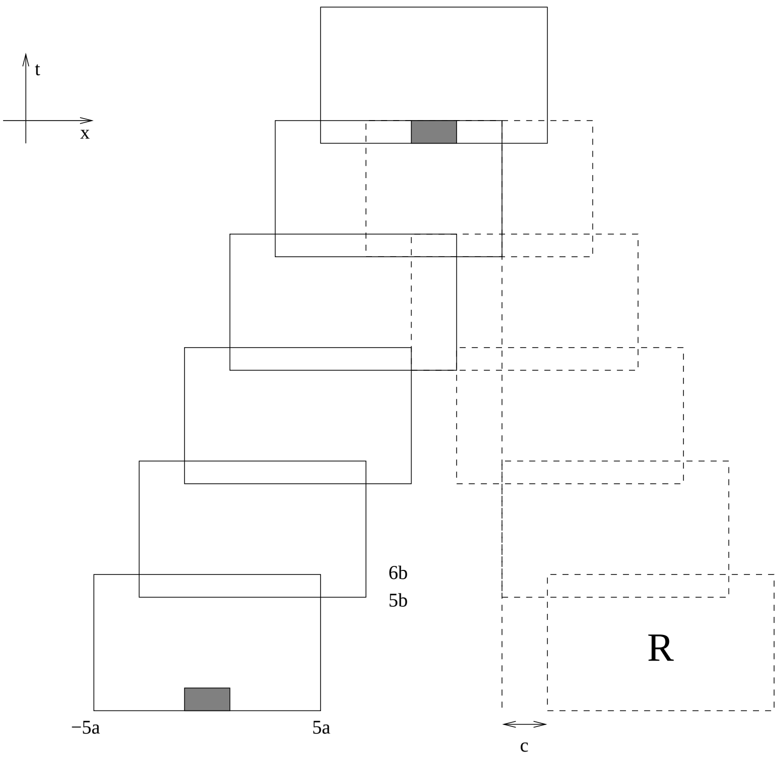

Figure 2.2: *The definition of c.*

dominates a product measure with density p on $\mathbb{N}$. We can then conclude that $\{X_n\}$ dominates $\{A_n\}$. This completes the proof. $\square$

We end with the following question:

Does the process obey a complete convergence theorem, i.e. is it the case that for all $p \in [0,1]$ and $\beta, \eta \in \{0,1\}^{\mathbb{Z}^d}$

$$(\delta_\beta \times \delta_\eta)S_p(t) \to \alpha_p(\beta,\eta)\bar{\nu}_p + (1-\alpha_p(\beta,\eta))\pi_p \times \delta_\emptyset \quad \text{as } t \to \infty,$$

where

$$\alpha_p(\beta,\eta) = \mathbf{P}_p[\, C_t^{\beta,\eta} \neq \emptyset \, \forall t \geq 0\,].$$

Contemporaneously and independently of our work, Remenik [12] has proved a complete convergence theorem for the special variant when $\delta_0 = \infty$. We strongly believe that a complete convergence theorem also holds in our case and plan to pursue some ideas that we have.

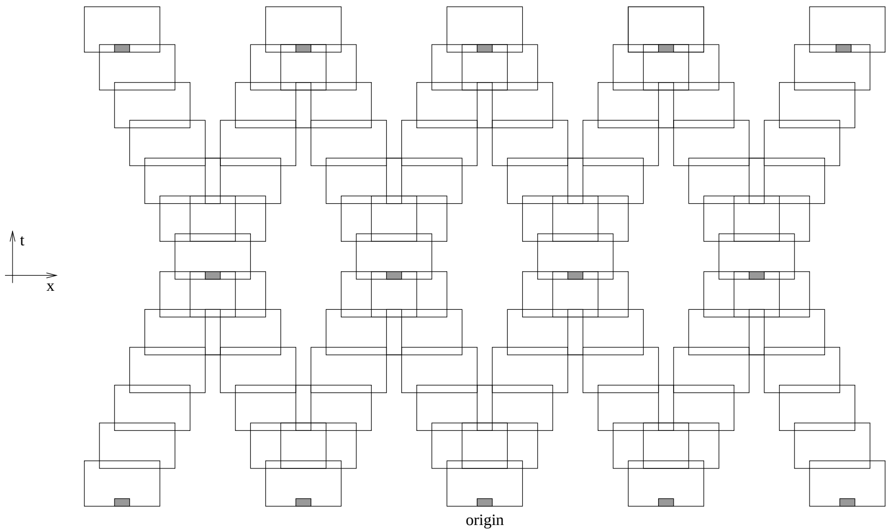

Figure 2.3: *Our building block $\mathcal{A}$ together with its reflection are translated in the x_1 and t direction. The shaded regions indicate where the paths start and stop in the definition of Z_n.*

Acknowledgement

The authors want to thank Olle Häggström for a careful reading of the manuscript and for valuable comments.

References

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

[2] M. Bramson, R. Durrett, and R. H. Schonmann, *The contact process in a random environment*, Ann. Probab. **19** (1991), 960–983.

[3] E. I. Broman, *Stochastic Domination for a Hidden Markov Chain with Applications to the Contact Process in a Randomly Evolving Environment*, Ann. Probab. **35** (2007), 2263–2293.

[4] R. Durrett, *Lecture notes on paricle systems and percolation*, Wadsworth and Brooks/Cole Advances Books and Software, 1988.

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

[6] A. Klein, *Extinction of contact and percolation processes in a random envionment*, Ann. Probab. **22** (1994), 1227–1251.

[7] T. M. Liggett, *Interacting Particle Systems*, Springer, 1985.

[8] ———, *The survival of one-dimensional contact processes in random environments*, Ann. Probab. **20** (1992), 696–723.

[9] ———, *Stochastic interacting systems: contact, voter and exclusion processes*, Springer, 1999.

[10] T. M. Liggett, R. H. Schonmann, and A. M Stacey, *Domination by product measures*, Ann. Probab. **25** (1997), 71–95.

[11] C. M. Newman and S. B. Volchan, *Persistent survival of one-dimensional contact processes in random environments*, Ann. Probab. **24** (1996), 411–421.

[12] D. Remenik, *The contact process in a dynamic random environment*, Ann. Appl. Probab. **18** (2008), 2392–2420.

PAPER II

Attractive nearest-neighbor spin systems on the integers in a randomly evolving environment

Marcus Warfheimer

3

PAPER II

ABSTRACT

We consider spin systems on $\mathbb{Z}$ (i.e. interacting particle systems on $\mathbb{Z}$ in which each coordinate only has two possible values and only one coordinate changes in each transition) whose rates are determined by another process, called a background process. A canonical example is the contact process in randomly evolving environment, introduced and analysed by Broman and further studied by Steif and the author, where the marginals of the background process independently evolve as 2-state Markov chains and determine the recovery rates for a contact process. We prove that, if the background process has a unique stationary distribution and if the rates satisfy a certain positivity condition, then there are at most two extremal stationary distributions. The proof follows closely the ideas of Liggett's proof of a corresponding theorem for spin systems on $\mathbb{Z}$ without a background process.

Key words and phrases: Spin systems, varying environment.
Subject classification : 60K35.

3.1 Introduction

The contact process in a random environment, in which the rates are taken to be random variables and then fixed in time, has been studied the last twenty years, see for example [1, 4, 7, 8]. However, recently Broman [2] introduced a variant where the environment changes in time in a Markovian way. (See also [9] for further analysis concerning that process.) More precisely, he considered the Markov process $\{(B_t, C_t)\}_{t\geq 0}$ on

$\{0,1\}^{\mathbb{Z}^d} \times \{0,1\}^{\mathbb{Z}^d}$ described by the following rates at a site x:

transition	rate
$(0,0) \to (0,1)$	$\sum_{y\sim x} C(y)$
$(1,0) \to (1,1)$	$\sum_{y\sim x} C(y)$
$(0,1) \to (0,0)$	δ_0
$(1,1) \to (1,0)$	δ_1
$(0,0) \to (1,0)$	γp
$(0,1) \to (1,1)$	γp
$(1,0) \to (0,0)$	$\gamma(1-p)$
$(1,1) \to (0,1)$	$\gamma(1-p)$

where $\gamma, \delta_0, \delta_1 > 0$ with $\delta_1 \leq \delta_0$ and $p \in [0,1]$. In other words, the background process evolves independently for each site and determines the recovery rate for the right marginal in the following way: At a given site x and time t, the rate is δ_0 or δ_1 depending on whether $B_t(x) = 0$ or $B_t(x) = 1$. Broman called $\{(B_t, C_t)\}$ the contact process in a randomly evolving environment, abbreviated CPREE. In this paper we study processes in one dimension with the same structure: a background process influencing another interacting particle system, but here both processes are more general. We prove, under certain conditions on the rates, that we have at most two extremal invariant distributions.

3.2 The model and main result

We consider the Markov process, $\{(\beta_t, \eta_t)\}_{t\geq 0}$ on $\{0,1\}^{\mathbb{Z}} \times \{0,1\}^{\mathbb{Z}}$ described by the following rates at a site x:

transition	rate
$(\beta, \eta) \to (\beta, \eta_x)$	$c_0(x,\eta)$ if $\beta(x) = 0$
$(\beta, \eta) \to (\beta, \eta_x)$	$c_1(x,\eta)$ if $\beta(x) = 1$
$(\beta, \eta) \to (\beta_x, \eta)$	$b(x,\beta)$

Here $c_0(x,\eta)$, $c_1(x,\eta)$ and $b(x,\beta)$ are given rate functions where the first two satisfy

$$
\begin{aligned}
c_0(x,\eta) &\leq c_1(x,\eta) \quad \text{if} \quad \eta(x) = 0, \\
c_1(x,\eta) &\leq c_0(x,\eta) \quad \text{if} \quad \eta(x) = 1,
\end{aligned} \tag{1}
$$

and all three satisfy the following attractivity condition:

Definition 3.2.1. A spin system on $\{0,1\}^{\mathbb{Z}}$, with rates $c(x,\eta)$ is said to be attractive if whenever $\eta \leq \eta'$,

$$
\begin{aligned}
c(x,\eta) &\leq c(x,\eta') \quad \text{if} \quad \eta(x) = \eta'(x) = 0, \\
c(x,\eta) &\geq c(x,\eta') \quad \text{if} \quad \eta(x) = \eta'(x) = 1.
\end{aligned} \tag{2}
$$

Here, $\leq$ refers to the usual partial ordering on $\{0,1\}^{\mathbb{Z}}$, i.e., $\eta \leq \eta'$ if and only if $\eta(x) \leq \eta'(x)$ for all $x \in \mathbb{Z}$. We also assume that the rate functions are translation invariant and that the rates $c_0(x,\eta)$, $c_1(x,\eta)$ only depend on η through

$$\{\eta(x-1), \eta(x), \eta(x+1)\}.$$

Moreover, to ensure that we have a well defined process we will assume that

$$\sum_{y \in \mathbb{Z}} \sup_{\beta \in \{0,1\}^{\mathbb{Z}}} |b(0,\beta) - b(0,\beta_y)| < \infty.$$

In other words, the rates for the system are completely described by $b(x,\beta)$ and the 16 parameters determining c_0 and c_1. To describe the values we will use the following notation:

$$c_i(001) = c_i(x,\eta) \quad \text{when} \quad \eta(x-1) = 0,\ \eta(x) = 0 \ \text{ and } \ \eta(x+1) = 1.$$

We always refer to the left marginal as the *background process*. Furthermore, note that we can equivalently view our process on $\{0,1\}^{\mathbb{Z}\times\{0,1\}}$ and that the conditions (1) and (2) then mean that the whole process is attractive on that space. (Definition 3.2.1 can of course be generalized to $\{0,1\}^S$ where S is countable.) The attractivity can be used to show (via monotonicity) the existence of two extremal stationary distributions ν_0 and ν_1 defined by

$$\nu_0 = \lim_{t\to\infty} \delta_0 S(t) \qquad \nu_1 = \lim_{t\to\infty} \delta_1 S(t),$$

where δ_0 and δ_1 denote the point masses corresponding to the elements $\eta \equiv 0$ and $\eta \equiv 1$ in $\{0,1\}^{\mathbb{Z}\times\{0,1\}}$ and $\{S(t)\}_{t\geq 0}$ denotes the semigroup associated to $\{(\beta_t, \eta_t)\}_{t\geq 0}$. The main result here is that, if the background process has a unique stationary distribution and the rates c_0, c_1 satisfy a certain positivity condition, then ν_0 and ν_1 are the only extremal stationary distributions. Let $\mathcal{I}$ denote the set of stationary distributions for the process and let $\mathcal{I}_e$ denote its extreme points. Furthermore, define

$$\begin{aligned} C_1 = \{\, & c_i(100) + c_j(110),\ c_i(001) + c_j(011), \\ & c_i(011) + c_j(110),\ c_i(100) + c_j(001),\ i = 0,1,\ j = 0,1 \,\} \end{aligned}$$

and let

$$C = \min\left(C_1\right).$$

Before we state our main result, we want to emphasize that the case with no background process has been studied before by Liggett. The proof of our main result follows closely the ideas of his proof. To state his result, let $c(x,\eta)$ be a rate function for an attractive, translation invariant, nearest-neighbor spin system $\{\eta_t\}_{t\geq 0}$ on $\{0,1\}^{\mathbb{Z}}$ and define $\mu_i = \lim_{t\to\infty} \delta_i T(t)$, $i = 0,1$, where δ_i is the point mass corresponding to the element $\eta \equiv i$ in $\{0,1\}^{\mathbb{Z}}$ and $\{T(t)\}_{t\geq 0}$ denotes the semigroup associated to $\{\eta_t\}_{t\geq 0}$. Moreover, let $\mathcal{J}_e$ denote the extreme points of the set of stationary distributions for $\{\eta_t\}_{t\geq 0}$.

Theorem 3.2.1 (Liggett). *Suppose*

$$c(x,\eta) + c(x,\eta_x) > 0 \quad \textit{whenever} \quad \eta(x-1) \neq \eta(x+1). \tag{3}$$

Then $\mathcal{J}_e = \{\mu_0, \mu_1\}$.

For a proof, see [5] or [6, p. 145-152]. In fact, he also proved that if condition (3) fails, then $\mathcal{J}_e$ contains infinitely many points, see [6, p. 145].

Theorem 3.2.2. *Suppose that the background process has a unique stationary distribution and assume* $C > 0$. *Then* $\mathcal{I}_e = \{\nu_0, \nu_1\}$.

Remarks:

(i) From [6, p. 152] we get that Theorem 3.2.1 is equivalent to the statement that (3) and

$$c(011) + c(110) > 0$$
$$c(100) + c(001) > 0$$

implies $\mathcal{J}_e = \{\mu_0, \mu_1\}$. By letting $c = c_0 = c_1$, it is now clear that Theorem 3.2.2 covers Theorem 3.2.1.

(ii) The hypotheses in Theorem 3.2.2 are true for the CPREE described in the introduction. Indeed, if c_0 and c_1 satisfy (1) and are symmetric under reflections, i.e.

$$c_i(100) = c_i(001)$$
$$c_i(110) = c_i(011), \quad i = 0, 1$$

then $C > 0$ if and only if $c_0(001) > 0$ and $c_1(011) > 0$.

(iii) Note that we are not assuming independence or even nearest-neighbor interaction between coordinates in the background process.

(iv) To see that the conclusion may fail if we drop the assumption about a unique stationary distribution for the background process, let $b(x,\beta)$, in addition to being attractive and translation invariant, be nearest-neighbor with $b(000) = b(111) = 0$ and satisfiy

$$b(x,\beta) + b(x,\beta_x) > 0 \quad \text{whenever} \quad \beta(x-1) \neq \beta(x+1).$$

Let $c_0 = c_1$ be the rates corresponding to a supercritical contact process on $\mathbb{Z}$. Then

$$\mathcal{I}_e = \{\, \delta_0 \times \delta_0,\ \delta_0 \times \bar{\nu},\ \delta_1 \times \delta_0,\ \delta_1 \times \bar{\nu} \,\},$$

where δ_0, δ_1 are the point masses corresponding to the elements $\eta \equiv 0$ and $\eta \equiv 1$ in $\{0,1\}^{\mathbb{Z}}$ respectively and $\bar{\nu}$ denotes the upper invariant measure for the contact process.

(v) If we take the same background process, but instead let $c_0 = c_1$ be the rates for a subcritical contact process, we see that the condition about a unique stationary distribution for the background process is not necessary for having only two extremal stationary distributions.

(vi) To see that the conclusion may fail if $C = 0$, let $b(x, \beta)$ be a rate function such that $\{\beta_t\}_{t\geq 0}$ has the point mass at $\beta \equiv 1$ as its unique stationary distribution and let c_1 satisify

$$c_1(001) + c_1(011) = 0.$$

It is easy to check that for each $n \in \mathbb{Z}$, $\delta_1 \times \delta_{\eta^n}$ is an extremal stationary distribution where

$$\eta^n(x) = \begin{cases} 1 & \text{if } x \geq n \\ 0 & \text{if } x < n. \end{cases}$$

A natural next step is to ask when there is a unique stationary distribution, i.e. when $\nu_0 = \nu_1$. In the case of no background process, Gray proved in [3] that there can only be one stationary distribution provided that the rates are strictly positive. We conjecture an analogous statement in our situation.

Theorem 3.2.3 (Gray). *If $c(x, \eta) > 0$ for all $x \in \mathbb{Z}$ and $\eta \in \{0, 1\}^{\mathbb{Z}}$, then $\mu_0 = \mu_1$.*

Conjecture 3.2.4. *Suppose that the background process has a unique stationary distribution and assume that $c_i(x, \eta) > 0$ for all x, η, $i = 1, 2$. Then $\nu_0 = \nu_1$.*

The rest of the paper is organized as follows. In Section 3.3 we prove Theorem 3.2.2 and in Section 3.4 we discuss Conjecture 3.2.4.

3.3 Proof of Theorem 3.2.2

In the proof, we make extensive use of a maximal type coupling which we now describe. Denote

$$U = \{0, 1\}^{\mathbb{Z}}, \qquad V = \{\, (\eta, \gamma, \xi) \in U^3 : \eta \leq \gamma \leq \xi \,\} \quad \text{and} \quad W = U \times V.$$

The coupled process $(\beta_t, \eta_t, \gamma_t, \xi_t)$, which we now define, lives on W and its flip rates are described as follows: First, let flips of the type

$$(\beta, \eta, \gamma, \xi) \to (\beta_x, \eta, \gamma, \xi)$$

occur at rate $b(x, \beta)$.

Then, let the other three marginals flip according to Tables 3.1 and 3.2. These tables should be interpreted as follows. For example, when $\beta_t(x) = 0$, $\eta_t(x) = 0$, $\gamma_t(x) = 0$ and $\xi_t(x) = 1$, $\xi_t(x)$ will flip alone at rate $c_0(x, \xi_t)$, $\gamma_t(x)$ will flip alone at rate $c_0(x, \gamma_t) - c_0(x, \eta_t)$ and $\eta_t(x)$ and $\gamma_t(x)$ flip together at rate $c_0(x, \eta_t)$. Note that

	(0,0,0,0)	(0,0,0,1)	(0,0,1,1)	(0,1,1,1)
(0,0,0,0)	–	$c_0(x,\xi) - c_0(x,\gamma)$	$c_0(x,\gamma) - c_0(x,\eta)$	$c_0(x,\eta)$
(0,0,0,1)	$c_0(x,\xi)$	–	$c_0(x,\gamma) - c_0(x,\eta)$	$c_0(x,\eta)$
(0,0,1,1)	$c_0(x,\xi)$	$c_0(x,\gamma) - c_0(x,\xi)$	–	$c_0(x,\eta)$
(0,1,1,1)	$c_0(x,\xi)$	$c_0(x,\gamma) - c_0(x,\xi)$	$c_0(x,\eta) - c_0(x,\gamma)$	–

Table 3.1: *Transition rates when the background process is in state 0.*

	(1,0,0,0)	(1,0,0,1)	(1,0,1,1)	(1,1,1,1)
(1,0,0,0)	–	$c_1(x,\xi) - c_1(x,\gamma)$	$c_1(x,\gamma) - c_1(x,\eta)$	$c_1(x,\eta)$
(1,0,0,1)	$c_1(x,\xi)$	–	$c_1(x,\gamma) - c_1(x,\eta)$	$c_1(x,\eta)$
(1,0,1,1)	$c_1(x,\xi)$	$c_1(x,\gamma) - c_1(x,\xi)$	–	$c_1(x,\eta)$
(1,1,1,1)	$c_1(x,\xi)$	$c_1(x,\gamma) - c_1(x,\xi)$	$c_1(x,\eta) - c_1(x,\gamma)$	–

Table 3.2: *Transition rates when the background process is in state 1.*

the pairs $\{(\beta_t,\eta_t)\}$, $\{(\beta_t,\gamma_t)\}$, $\{(\beta_t,\xi_t)\}$ each evolve as the original Markov process and that the second, third and fourth marginals try to flip together as much as possible. Also, observe that the background process is not allowed to flip together with any of the other processes.

As in the proof of Theorem 3.2.1, the proof of Theorem 3.2.2 consists of several lemma concerning certain functionals of the process. For $m \le n$, let $f_{m,n}(\beta,\eta,\gamma,\xi)$ be the number of intervals of zeros and ones in γ between m and n (including m and n), counted only where η and ξ differ. Furthermore, let

$$m \le x_1 < x_2 < \ldots < x_k \le n,$$

be all those x's between m and n for which $\eta(x) = 0$ and $\xi(x) = 1$. For $l \ge 1$, define

$$\begin{aligned} g^l_{m,n}(\beta,\eta,\gamma,\xi) &= \text{number of } i \text{ such that } i \ge 1,\ i+l+1 \le k \text{ and} \\ &\quad \gamma(x_i) \ne \gamma(x_{i+1}) = \gamma(x_{i+2}) = \ldots = \gamma(x_{i+l}) \ne \gamma(x_{i+l+1}). \end{aligned}$$

In other words, $g^l_{m,n}(\beta,\eta,\gamma,\xi)$ is the number of interior intervals of zeros and ones of length l in γ between m and n, counted only where η and ξ differ. For example if,

$$\begin{array}{cccccccccccc|c} \cdots & 1 & 0 & 1 & 1 & 1 & 1 & 1 & 0 & 1 & 1 & 1 & \cdots & \xi \\ \cdots & 1 & 0 & 1 & 1 & 0 & 0 & 1 & 0 & 1 & 1 & 0 & \cdots & \gamma \\ \cdots & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \cdots & \eta \\ \cdots & 1 & 0 & 1 & 0 & 0 & 1 & 1 & 1 & 0 & 1 & 1 & \cdots & \beta \\ & m & & & & & & & & & & n & & \end{array}$$

then $f_{m,n} = 4$, $g^2_{m,n} = 1$, $g^3_{m,n} = 1$ and $g^l_{m,n} = 0$ when $l \notin \{2,3\}$. Let

$$K = \max\left(\max_\eta c_0(x,\eta),\ \max_\eta c_1(x,\eta)\right)$$

and denote the set of stationary distributions and the generator of the coupled process by $\tilde{\mathcal{I}}$ and $\tilde{\Omega}$ respectively. Furthermore, for a given set $\mathcal{A}$, denote the set of extreme points by $\mathcal{A}_e$. The first lemma concerns certain basic properties of $f_{m,n}$ and $g^l_{m,n}$.

Lemma 3.3.1.

a) $f_{m,n}, g^l_{m,n}$ *are increasing when* n *increases or* m *decreases.*

b) $f_{m,n} \leq 2 + \sum_{l=1}^{\infty} g^l_{m,n}$.

c) $\sum_{l=1}^{\infty} l g^l_{m,n} \leq n - m + 1$.

If $\nu \in \tilde{\mathcal{I}}$,

d) $C \int g^1_{m,n}\, d\nu \leq K \int [f_{m-1,n} + f_{m,n+1} - 2f_{m,n}]\, d\nu$, *for* $m \leq n$

e) $C \int g^{l+1}_{m,n}\, d\nu \leq 12Kl \int g^l_{m,n}\, d\nu$, *for* $m \leq n$, $l \geq 1$.

Proof. a), b) and c) follow directly from the definitions. For d) and e) assume $\nu \in \tilde{\mathcal{I}}$. Note that $f_{m,n}$ and $g^l_{m,n}$ are cylinder functions so that

$$\int \tilde{\Omega} f_{m,n}\, d\nu = \int \tilde{\Omega} g^l_{m,n}\, d\nu = 0. \tag{4}$$

For cylinder function f, the generator has the form

$$\begin{aligned}\tilde{\Omega} f(\beta,\eta,\gamma,\xi) = &\sum_{(\beta,\bar{\eta},\bar{\gamma},\bar{\xi})} c(\beta,\eta,\gamma,\xi,\bar{\eta},\bar{\gamma},\bar{\xi})\left(f(\beta,\bar{\eta},\bar{\gamma},\bar{\xi}) - f(\beta,\eta,\gamma,\xi)\right)\\ &+ \sum_x b(x,\beta)\left(f(\beta_x,\eta,\gamma,\xi) - f(\beta,\eta,\gamma,\xi)\right)\end{aligned} \tag{5}$$

where the first sum is over all possible transitions when the second, third or fourth marginal flip. (Recall that the first marginal is not allowed to flip together with any of the others.) Here, since both $f_{m,n}$ and $g^l_{m,n}$ do not depend on β, the second sum is zero, so our task is to calculate the first part. For this, we follow the approach in [6, Lemma 3.7]. The argument given here is almost the same as in [6], we supply it for the sake of completeness. Let (β,η,γ,ξ) be fixed and note that the only way $f_{m,n}$ can increase because of a flip is if $f_{m-1,n} = f_{m,n} + 1$ or $f_{m,n+1} = f_{m,n} + 1$. In the first case the flip must occur at $x = m$ and in the second at $x = n$. The rate for such a flip is at most K so the positive terms in (5) are bounded above by

$$K\left[f_{m-1,n} + f_{m,n+1} - 2f_{m,n}\right].$$

Furthermore, there are $g^1_{m,n}$ sites x where a flip decreases $f_{m,n}$ by two. At such an x, $\gamma(x) = 0$ or $\gamma(x) = 1$. Assume $\gamma(x) = 1$. Then we necessarely have $\gamma(x-1) =$

$\eta(x-1)$ and $\gamma(x+1) = \eta(x+1)$. Therefore, the flip rate at x becomes

$$c_0(x,\gamma) + c_0(x,\eta) = \begin{cases} c_0(010) + c_0(000) & \text{if } \gamma(x-1) = 0, \gamma(x+1) = 0, \\ c_0(011) + c_0(001) & \text{if } \gamma(x-1) = 0, \gamma(x+1) = 1, \\ c_0(110) + c_0(100) & \text{if } \gamma(x-1) = 1, \gamma(x+1) = 0, \\ c_0(111) + c_0(101) & \text{if } \gamma(x-1) = 1, \gamma(x+1) = 1, \end{cases}$$

when $\beta(x) = 0$ and

$$c_1(x,\gamma) + c_1(x,\eta) = \begin{cases} c_1(010) + c_1(000) & \text{if } \gamma(x-1) = 0, \gamma(x+1) = 0, \\ c_1(011) + c_1(001) & \text{if } \gamma(x-1) = 0, \gamma(x+1) = 1, \\ c_1(110) + c_1(100) & \text{if } \gamma(x-1) = 1, \gamma(x+1) = 0, \\ c_1(111) + c_1(101) & \text{if } \gamma(x-1) = 1, \gamma(x+1) = 1, \end{cases}$$

when $\beta(x) = 1$. Also the attractivity condition gives

$$\begin{aligned} c_i(010) &\geq \max\{\, c_i(011),\, c_i(110)\,\} \\ c_i(101) &\geq \max\{\, c_i(001),\, c_i(100)\,\}, \quad i = 0, 1 \end{aligned}$$

and so the rates above are bounded below by $C/2$. The same argument works if $\gamma(x) = 0$ and so we can conclude that the negative terms in (5) are bounded above by $-Cg^1_{m,n}$. We get the estimate

$$\tilde{\Omega} f_{m,n} \leq K\left[f_{m-1,n} + f_{m,n+1} - 2f_{m,n}\right] - Cg^1_{m,n}$$

which via (4) gives d). For e), note that $g^l_{m,n}$ can only decrease via flips at no more than $lg^l_{m,n}$ sites or their neighbors, i.e. in total at most $3lg^l_{m,n}$ sites. The rate for such a flip is bounded by $2K$ and $g^l_{m,n}$ can at most decrease by two. The negative terms in the generator are therefore bounded below by $-12Klg^l_{m,n}$. Furthermore, $g^l_{m,n}$ can increase at no fewer than $g^{l+1}_{m,n}$ pair of sites. These pair of sites are the endpoints of an interval of length $l+1$. To get a lower bound on the flip rate for such endpoints, let $x < y$ denote such a pair and suppose $\gamma(x) = \gamma(y) = 1$. Then we have $\gamma(x-1) = \eta(x-1)$ and $\gamma(y+1) = \eta(y+1)$. The flip rate at x is at least $c_i(100)$ if $\gamma(x-1) = \eta(x-1) = 1$, $\beta(x) = i$ and at least $c_i(011)$ if $\gamma(x-1) = \eta(x-1) = 0$, $\beta(x) = i$. In a similar fashion, the flip rate at y is at least $c_i(001)$ if $\gamma(y+1) = \eta(y+1) = 1$, $\beta(y) = i$ and at least $c_i(110)$ if $\gamma(y+1) = \eta(y+1) = 0$, $\beta(y) = i$. In either case the sum of the flip rates for the pair is always at least C. The same statement holds if $\gamma(x) = \gamma(y) = 0$ and so we obtain that the positive terms in the generator expression are bounded below by $Cg^{l+1}_{m,n}$. Hence, we get the estimate

$$\tilde{\Omega} g_{m,n} \geq Cg^{l+1}_{m,n} - 12Klg^l_{m,n}.$$

Equation (4) then finally gives us

$$C \int g^{l+1}_{m,n}\, d\nu \leq 12Kl \int g^l_{m,n}\, d\nu$$

and the proof is complete. □

Denote

$$\begin{aligned}
A_1 &= \{\, (\beta,\eta,\gamma,\xi) \in W \,:\, \gamma \equiv \eta \,\},\\
A_2 &= \{\, (\beta,\eta,\gamma,\xi) \in W \,:\, \gamma \equiv \xi \,\},\\
A_3 &= \{\, (\beta,\eta,\gamma,\xi) \in W \setminus A_1 \cup A_2 \,:\, \exists x \in \mathbb{Z} \text{ such that}\\
&\qquad \gamma(y) = \eta(y) \text{ when } y \le x \text{ and } \gamma(y) = \xi(y) \text{ when } y > x \,\},\\
A_4 &= \{\, (\beta,\eta,\gamma,\xi) \in W \setminus A_1 \cup A_2 \,:\, \exists x \in \mathbb{Z} \text{ such that}\\
&\qquad \gamma(y) = \xi(y) \text{ when } y \le x \text{ and } \gamma(y) = \eta(y) \text{ when } y > x \,\},
\end{aligned}$$

Lemma 3.3.2. *Assume $C > 0$. Then*

$$\begin{aligned}
&a)\ \nu \in \tilde{\mathcal{I}} &&\Longrightarrow &&\nu\,(A_1 \cup A_2 \cup A_3 \cup A_4) = 1,\\
&b)\ \nu \in \tilde{\mathcal{I}}_e &&\Longrightarrow &&\nu\,(A_i) = 1 \ \text{ for some } i.
\end{aligned}$$

Proof. $b)$ follows from a) since A_i is closed for the coupled process in the sense that

$$\mathbf{P}^{(\beta,\eta,\gamma,\xi)}[\,(\beta_t,\eta_t,\gamma_t,\xi_t) \in A_i\,] = 1 \quad \forall t > 0$$

whenever $(\beta,\eta,\gamma,\xi) \in A_i$. To prove $a)$, suppose $\nu \in \tilde{\mathcal{I}}$. Since

$$\bigcup_{i=1}^{4} A_i = \{\, g^l_{m,n} = 0 \,\forall m \le n,\, l \ge 1 \,\}$$

we obtain that

$$\int g^l_{m,n} d\nu = 0 \text{ for all } m \le n,\, l \ge 1 \tag{6}$$

is equivalent to

$$\nu\,(A_1 \cup A_2 \cup A_3 \cup A_4) = 1.$$

To see that (6) holds, we proceed as in [6, Lemma 3.10]. Note that

$$f_{m-1,n} \le f_{m,n} + 1 \qquad \text{and} \qquad f_{m,n+1} \le f_{m,n} + 1$$

and so parts $d)$ and $e)$ of Lemma 3.3.1 gives us

$$M = \sup_{m \le n} \int g^l_{m,n}\, d\nu < \infty, \quad \forall l \ge 1. \tag{7}$$

Let $L \ge 1$. From part $b)$ of the same lemma, we get

$$\frac{1}{n-m} \int f_{m,n}\, d\nu \le \frac{2}{n-m} + \frac{1}{n-m} \int \sum_{l \ge 1} g^l_{m,n}\, d\nu.$$

Split the sum and now use part $c)$ of the lemma together with (7) to obtain that for any L

$$\frac{1}{n-m}\int f_{m,n}\,d\nu \le \frac{2}{n-m} + \frac{ML}{n-m} + \frac{1}{L}\left(1+\frac{1}{n-m}\right),$$

and so

$$\limsup_{n-m\to\infty} \frac{1}{n-m}\int f_{m,n}\,d\nu \le \frac{1}{L}.$$

Since $L \ge 1$ was arbitrary we can conclude

$$\lim_{n-m\to\infty} \frac{1}{n-m}\int f_{m,n}\,d\nu = 0. \tag{8}$$

Now, for $N \ge 1$, part $d)$ of Lemma 3.3.1 gives us

$$\begin{aligned} &C \sum_{m=-N+1}^{0}\sum_{n=0}^{N-1}\int g^1_{m,n}\,d\nu \\ &\quad\le K \sum_{m=-N+1}^{0}\sum_{n=0}^{N-1}\int [f_{m-1,n}+f_{m,n+1}-2f_{m,n}]\,d\nu. \end{aligned} \tag{9}$$

After some cancellations in the sum to the right, we get

$$\begin{aligned} &\sum_{m=-N+1}^{0}\sum_{n=0}^{N-1}\int [f_{m-1,n}+f_{m,n+1}-2f_{m,n}]\,d\nu \\ &\quad\le \sum_{m=-N+1}^{0}\int f_{m,N}\,d\nu + \sum_{n=0}^{N-1}\int f_{-N,n}\,d\nu \end{aligned}$$

and together with (8) and (9) we obtain

$$\lim_{N\to\infty}\frac{1}{N^2}\sum_{m=-N+1}^{0}\sum_{n=0}^{N-1}\int g^1_{m,n}\,d\nu = 0.$$

Using the monotonicity property of $g^1_{m,n}$ this implies $\int g^1_{m,n}\,d\nu = 0$ for all $m \le n$ and part $e)$ of the lemma gives $\int g^l_{m,n}\,d\nu = 0$ for all $l \ge 1$ and we are done with the proof. □

We are soon ready for the proof of Theorem 3.2.2. However, in the proof we make use of a 5-variant coupling $\{(\beta_t,\eta_t,\gamma_{1,t},\gamma_{2,t},\xi_t)\}$ of the one used so far. This coupling is also of maximal type and evolves on

$$X = \left\{ (\beta,\eta,\gamma_1,\gamma_2,\xi) \in U^5 : \eta \le \gamma_1 \le \xi,\ \eta \le \gamma_2 \le \xi \right\}$$

in a way such that $\{(\beta_t,\eta_t,\gamma_{1,t},\xi_t)\}$ and $\{(\beta_t,\eta_t,\gamma_{2,t},\xi_t)\}$ evolve exactly as the previous described coupling. We can therefore apply all we have done so far to each of these processes. The last tool we need is to have existence of an extremal stationary distribution for the 5-variant coupled process, given extremal stationary distributions for the $\{(\beta_t, \eta_t)\}$ process. For a stochastic variable X and a distribution μ, let $X \sim \mu$ denote that X is distributed according to μ. Also, let $\mathcal{I}^5$ denote the set of stationary distributions for the 5-variant coupled process on X.

Lemma 3.3.3. *Given μ, $\mu' \in \mathcal{I}_e$ there exists $\nu((\beta, \eta, \gamma_1, \gamma_2, \xi) \in \cdot) \in \mathcal{I}_e^5$ such that $(\beta, \eta) \sim \nu_0$, $(\beta, \gamma_1) \sim \mu$, $(\beta, \gamma_2) \sim \mu'$ and $(\beta, \xi) \sim \nu_1$.*

Proof. For any measure μ let μ_{ij} denote the projection to the *i*th and *j*th coordinate. Construct a coupling on $(\{0,1\}^{\mathbb{Z}} \times \{0,1\}^{\mathbb{Z}})^4$ of four $\{\beta_t, \eta_t\}$-processes such that the background processes agree as much as possible as well as the right marginals. Note that our 5-variant coupling above can be identified with such a coupling started with all the background processes equal. Starting the coupling with

$$\delta_{(\emptyset,\emptyset)} \times \mu \times \mu' \times \delta_{(\mathbb{Z},\mathbb{Z})}$$

and taking a suitable subsequence of Cesaro averages gives us a stationary distribution ρ for the coupling and by projecting to the first, second, fourth, sixth and eighth coordinate we get a probability measure $\tilde{\nu} \in \mathcal{I}^5$ with

$$\tilde{\nu}((\beta, \eta, \gamma_1, \gamma_2, \xi) \in U^5 : \eta \leq \gamma_1 \leq \xi, \eta \leq \gamma_2 \leq \xi) = 1.$$

Here it is important to note that the set

$$\{ (\beta_1, \eta, \beta_2, \gamma_1, \beta_3, \gamma_2, \beta_4, \xi) \in U^8 : \beta_1 \leq \beta_2 \leq \beta_4, \beta_1 \leq \beta_3 \leq \beta_4,$$
$$\eta \leq \gamma_1 \leq \xi, \eta \leq \gamma_2 \leq \xi \}$$

is closed under the evolution of the coupling and that the first, third, fifth and seventh coordinate are equal under ρ. Furthermore, it is clear that $\tilde{\nu}$ satisfies

$$\tilde{\nu}_{12} = \nu_0, \quad \tilde{\nu}_{13} = \mu \quad \tilde{\nu}_{14} = \mu' \quad \text{and} \quad \tilde{\nu}_{15} = \nu_1.$$

Define

$$\mathcal{B} = \{ \nu \in \mathcal{I}^5 : \nu_{12} = \nu_0, \nu_{13} = \mu, \nu_{14} = \mu', \nu_{15} = \nu_1 \}.$$

$\mathcal{B}$ is non-empty by the above and is compact and convex. Hence, by the Krein-Milman theorem, $\mathcal{B}$ can be written as the closed convex hull of its extreme points. Therefore, since $\mathcal{B} \neq \emptyset$, we have $\mathcal{B}_e \neq \emptyset$. Hence, the proof is complete if $\mathcal{B}_e \subset \mathcal{I}_e^5$. Assume $\nu \in \mathcal{B}_e$ and let $\nu = \alpha\rho + (1-\alpha)\sigma$, where $0 < \alpha < 1$ and $\rho, \sigma \in \mathcal{I}^5$. If $\rho, \sigma \in \mathcal{B}$ we get $\nu = \rho = \sigma$ and we are done. In order to see this, let (i,j) be one of the pairs $(1,2)$, $(1,3)$, $(1,4)$ or $(1,5)$. Since $\nu_{ij} = \alpha\rho_{ij} + (1-\alpha)\sigma_{ij}$, where $\rho_{ij}, \sigma_{ij} \in \mathcal{I}$, and the left hand side is an element of $\{\nu_0, \mu, \mu', \nu_1\} \subseteq \mathcal{I}_e$, we obtain

$$\nu_0 = \rho_{12} = \sigma_{12} \qquad \mu = \rho_{13} = \sigma_{13}$$
$$\mu' = \rho_{14} = \sigma_{14} \qquad \nu_1 = \rho_{15} = \sigma_{15}$$

and so $\rho, \sigma \in \mathcal{B}$. □

Proof of Theorem 3.2.2. We follow the steps in [6, Theroem 3.13]. Let $\mu_1 \in \mathcal{I}_e$. Since $\nu_0 \le \mu \le \nu_1$ for every stationary distribution μ, we can assume $\nu_0 \ne \nu_1$. Let $\mu_2 = \mu_1 \circ \theta_x^{-1}$, where θ_x is a translation by $x \in \mathbb{Z}$. Since the dynamics are translation invariant and $\mu_1 \in \mathcal{I}_e$, we get that $\mu_2 \in \mathcal{I}_e$. Let ρ be an extremal stationary distribution for the 5-variant coupling mentioned above with

$$\begin{aligned} (\beta, \eta) &\sim \nu_0 & \qquad (\beta, \gamma_1) &\sim \mu_1 \\ (\beta, \gamma_2) &\sim \mu_2 & \qquad (\beta, \xi) &\sim \nu_1 \end{aligned}$$

Such a measure exists by Lemma 3.3.3. Let ρ_1 and ρ_2 be the distributions obtained from the projections

$$\begin{aligned} (\beta, \eta, \gamma_1, \gamma_2, \xi) &\to (\beta, \eta, \gamma_1, \xi) \\ (\beta, \eta, \gamma_1, \gamma_2, \xi) &\to (\beta, \eta, \gamma_2, \xi) \end{aligned}$$

respectively. Since $\rho_1, \rho_2 \in \tilde{\mathcal{I}}_e$, Lemma 3.3.2 gives

$$\rho_1(A_i) = 1 \quad \text{some } 1 \le i \le 4 \quad \text{and} \quad \rho_2(A_i) = 1 \quad \text{some } 1 \le i \le 4.$$

However, γ_1 and γ_2 are just translations of each other so there is an i such that $\rho_1(A_i) = \rho_2(A_i) = 1$. It follows that

$$\rho\Big((\beta, \eta, \gamma_1, \gamma_2, \xi) : \sum_x |\gamma_1(x) - \gamma_2(x)| < \infty\Big) = 1.$$

Also, $(\gamma_{1,t}, \gamma_{2,t})$ has the property that

$$\mathbf{P}^{(\gamma,\gamma)}[\gamma_{1,t} = \gamma_{2,t}] = 1 \quad \text{and} \quad \mathbf{P}^{(\gamma_1,\gamma_2)}[\gamma_{1,t} = \gamma_{2,t}] > 0$$

whenever $\sum_x |\gamma_1(x) - \gamma_2(x)| < \infty$ and so since ρ is stationary, we must in fact have

$$\rho\Big((\beta, \eta, \gamma_1, \gamma_2, \xi) : \gamma_1 = \gamma_2\Big) = 1.$$

This implies $\mu_1 = \mu_2$, i.e. μ_1 is translation invariant. Therefore i equals 1 or 2 (recall $\nu_0 \ne \nu_1$). If $i = 1$, $\mu_1(U \times (\cdot)) = \nu_0(U \times (\cdot))$ and since the background process has a unique stationary distribution we must also have $\mu_1((\cdot) \times U) = \nu_0((\cdot) \times U)$. But since $\nu_0 \le \mu_1$ this yields $\mu_1 = \nu_0$. If $i = 2$ we get in a similar way that $\mu_1 = \nu_1$. $\square$

3.4 Discussion of Conjecture 3.2.4

We begin by describing a graphical representation which may be useful for a possible proof of Conjecture 3.2.4. The representation is similar as in [3] and we will explain it in a quite informal way. For simplicity, we will assume that the rates for the background process, in addition to attractive and translation invariant, also are uniformly bounded. (Of course, our assumptions on c_0 and c_1 from Section 3.2 imply that they are also uniformly bounded.) For $x \in \mathbb{Z}$, define

$$\begin{aligned}
\bar{b}_x &= \sup_{\beta:\, \beta(x)=0} b(x,\beta) + \sup_{\beta:\, \beta(x)=1} b(x,\beta) \\
\bar{c}^0_x &= \sup_{\eta:\, \eta(x)=0} c_0(x,\eta) + \sup_{\eta:\, \eta(x)=1} c_0(x,\eta) \\
\bar{c}^1_x &= \sup_{\eta:\, \eta(x)=0} c_1(x,\eta) + \sup_{\eta:\, \eta(x)=1} c_1(x,\eta) \\
\bar{c}_x &= \bar{c}^0_x + \bar{c}^1_x.
\end{aligned}$$

Define the following collection of independent random variables on some probability space $(\Omega, \mathcal{F}, \mathbf{P})$:

- $B_j(x)$ exponentially distributed with mean $1/\bar{b}_x$, $j \geq 1$, $x \in \mathbb{Z}$. (Define $B_j(x) = \infty$ if $\bar{b}_x = 0$.)
- $D_n(x)$ uniformly distributed on $[0, \bar{b}_x]$, $n \geq 1$, $x \in \mathbb{Z}$.
- $S_j(x)$ exponentially distributed with mean $1/\bar{c}_x$, $j \geq 1$, $x \in \mathbb{Z}$.
- $U^0_n(x)$ uniformly distributed on $[0, \bar{c}^0_x]$, $n \geq 1$, $x \in \mathbb{Z}$.
- $U^1_n(x)$ uniformly distributed on $[0, \bar{c}^1_x]$, $n \geq 1$, $x \in \mathbb{Z}$.

Moreover, for $n \geq 1$ and $x \in \mathbb{Z}$ define

$$C_n(x) = \sum_{j=1}^{n} B_j(x) \quad \text{and} \quad T_n(x) = \sum_{j=1}^{n} S_j(x).$$

For a given initial configuration $\beta \in \{0,1\}^{\mathbb{Z}}$, define a process $\{\beta^\beta_t\}_{t \geq 0}$ from $\{C_n(x)\}$ and $\{D_n(x)\}$ as follows:

- $\beta^\beta_0 = \beta$,
- $\beta^\beta_s(x)$ flips from 0 to 1 iff $\beta^\beta_{s-}(x) = 0$ and there exists an $n \geq 1$ such that $s = C_n(x)$ and $D_n(x) \geq \bar{b}_x - b(x, \beta^\beta_{s-})$,
- $\beta^\beta_s(x)$ flips from 1 to 0 iff $\beta^\beta_{s-}(x) = 1$ and there exists an $n \geq 1$ such that $s = C_n(x)$ and $D_n(x) < b(x, \beta^\beta_{s-})$.

By an approximation procedure, it is possible to prove that there exists a process with those properties and that such a process has flip rates $b(x, \beta)$.

Given $\beta, \eta \in \{0,1\}^{\mathbb{Z}}$, we now define a process $\{\eta^{\beta,\eta}_t\}_{t \geq 0}$ from $\{\beta^\beta_t\}$, $\{T_n(x)\}$, $\{U^0_n(x)\}$ and $\{U^1_n(x)\}$ in the following way:

- $\eta_0^{\beta,\eta} = \eta$,
- if $\beta_s^\beta(x) = 0$, then $\eta_s^{\beta,\eta}(x)$ flips from 0 to 1 iff $\eta_{s-}^{\beta,\eta}(x) = 0$ and there exists an $n \geq 1$ such that $s = T_n(x)$ and $U_n^0(x) \geq \bar{c}_x^0 - \frac{\bar{c}_x^0}{\bar{c}_x} c_0(x, \eta_{s-}^{\beta,\eta})$ and $\eta_s^{\beta,\eta}(x)$ flips from 1 to 0 iff $\eta_{s-}^{\beta,\eta}(x) = 1$ and there exists an $n \geq 1$ such that $s = T_n(x)$ and $U_n^0(x) < \frac{\bar{c}_x^0}{\bar{c}_x} c_0(x, \eta_{s-}^{\beta,\eta})$,
- if $\beta_s^\beta(x) = 1$, then $\eta_s^{\beta,\eta}(x)$ flips from 0 to 1 iff $\eta_{s-}^{\beta,\eta}(x) = 0$ and there exists an $n \geq 1$ such that $s = T_n(x)$ and $U_n^1(x) \geq \bar{c}_x^1 - \frac{\bar{c}_x^1}{\bar{c}_x} c_1(x, \eta_{s-}^{\beta,\eta})$ and $\eta_s^{\beta,\eta}(x)$ flips from 1 to 0 iff $\eta_{s-}^{\beta,\eta}(x) = 1$ and there exists an $n \geq 1$ such that $s = T_n(x)$ and $U_n^1(x) < \frac{\bar{c}_x^1}{\bar{c}_x} c_1(x, \eta_{s-}^{\beta,\eta})$.

It is clear that the process $\{(\beta_t^\beta, \eta_t^{\beta,\eta})\}$ has the correct flip rates. Moreover, the graphical representation gives us a coupling for all possible initial states and this coupling is exactly the maximal type coupling used in Section 3.3. If we want to start the process at a random state with distribution ρ, we just add, independent of everything else, two random variables with joint distribution ρ. We then write $\{\beta_t^{\rho_1}, \eta_t^{\rho_1,\rho_2}\}$ where ρ_i denotes the ith marginal of ρ.

A possible proof of Conjecture 3.2.4 may be based on the following lemma.

Lemma 3.4.1. *If*

$$\liminf_{k\to\infty} \liminf_{t\to\infty} \mathbf{P}\left[\eta_t^{\beta,\emptyset}(x) = \eta_t^{\beta,\mathbb{Z}}(x),\ -k \leq x \leq k\right] > 0 \tag{10}$$

for all $\beta \in \{0,1\}^{\mathbb{Z}}$, then $\nu_0 = \nu_1$.

Proof. From Lemma 3.3.3 (or more precisely from the version of it with three processes) there exists a probability measure γ on

$$\{\ (\beta, \eta, \xi) \in U^3 : \ \eta \leq \xi\}$$

which is stationary for $\{(\beta_t, \eta_t, \xi_t\}_{t\geq 0}$ and satisfies

$$\gamma_{12} = \nu_0, \quad \gamma_{13} = \nu_1 \quad \text{and} \quad \gamma_1 = \mu,$$

where μ is the unique stationary distribution for the background process. (Here, we use the same notation as in Lemma 3.3.3.) Our goal is to show that

$$\gamma\,(\eta = \xi) = 1.$$

For given $k \geq 1$ and $t \geq 0$, we get

$$\begin{aligned} &\gamma\,(\eta(x) = \xi(x),\ -k \leq x \leq k) = \gamma\,(\eta = \xi) \\ &\quad + \mathbf{P}\,[\,\eta_t^{\mu,\gamma_2}(x) = \eta_t^{\mu,\gamma_3}(x),\ -k \leq x \leq k \,|\, \eta_0^{\mu,\gamma_2} \neq \eta_0^{\mu,\gamma_3}\,]\,(1 - \gamma\,(\eta = \xi)). \end{aligned} \tag{11}$$

Here, we have used that γ is stationary and the fact that

$$\mathbf{P}\left[\,\eta_t^{\mu,\gamma_2}(x)=\eta_t^{\mu,\gamma_3}(x),\ -k\le x\le k\,|\,\eta_0^{\mu,\gamma_2}=\eta_0^{\mu,\gamma_3}\,\right]=1.$$

From the inequalities

$$\eta_t^{\mu,\emptyset}\le\eta_t^{\mu,\gamma_2}\le\eta_t^{\mu,\gamma_3}\le\eta_t^{\mu,\mathbb{Z}},\quad t\ge 0,$$

we get,

$$\begin{aligned}&\mathbf{P}\left[\,\eta_t^{\mu,\gamma_2}(x)=\eta_t^{\mu,\gamma_3}(x),\ -k\le x\le k\,|\,\eta_0^{\mu,\gamma_2}\ne\eta_0^{\mu,\gamma_3}\,\right]\\ &\quad\ge\mathbf{P}\left[\eta_t^{\mu,\emptyset}(x)=\eta_t^{\mu,\mathbb{Z}}(x),\ -k\le x\le k\,|\,\eta_0^{\mu,\gamma_2}\ne\eta_0^{\mu,\gamma_3}\right].\end{aligned}\tag{12}$$

Moreover, from the graphical representation, we get that the events

$$\{\,\eta_t^{\mu,\emptyset}(x)=\eta_t^{\mu,\mathbb{Z}}(x),\ -k\le x\le k\,\}\quad\text{and}\quad\{\,\eta_0^{\mu,\gamma_2}\ne\eta_0^{\mu,\gamma_3}\,\}$$

are conditionally independent given the initial state of the background process and so we can write

$$\begin{aligned}&\mathbf{P}\left[\eta_t^{\mu,\emptyset}(x)=\eta_t^{\mu,\mathbb{Z}}(x),\ -k\le x\le k,\ \eta_0^{\mu,\gamma_2}\ne\eta_0^{\mu,\gamma_3}\right]\\ &\int\mathbf{P}\left[\eta_t^{\mu,\emptyset}(x)=\eta_t^{\mu,\mathbb{Z}}(x),\ -k\le x\le k\,|\,\beta_0^{\mu}=\beta\right]\gamma\,(\eta\ne\xi\,|\,\beta)\ d\mu(\beta).\end{aligned}\tag{13}$$

Now, let us assume that

$$\gamma\,(\eta\ne\xi)>0.$$

Then

$$\gamma\,(\eta\ne\xi\,|\,\beta)>0.$$

on a set of positive μ-measure. By using (10), (13) together with Fatou's Lemma and then (12), we can conclude that

$$\liminf_{k\to\infty}\liminf_{t\to\infty}\mathbf{P}\left[\,\eta_t^{\mu,\gamma_2}(x)=\eta_t^{\mu,\gamma_3}(x),\ -k\le x\le k\,|\,\eta_0^{\mu,\gamma_2}\ne\eta_0^{\mu,\gamma_3}\,\right]>0.$$

However, by taking limits in (11) we arrive at a contradiction and so we are done with the proof. □

The question now is if it is possible to prove (10). A natural first try is to fix the initial state of the background process and then proceed as in [3, p. 393] and define so called left and right edge processes. The properties on p. 394 and Proposition 2 on p. 395 are then easily verified. For the correlation property between the left and right edge processes, we can use [6, Ch. II, Corollary 2.21] and since the Lemma in the proof of [3, Theorem 1] relies on the properties on [3, p. 394], it may be possible to prove a version of it for our process. Having succeeded so far, there is some hard work left which we at the moment are not able to decide on if it is possible to do something similar or not. The only thing we can say is that the argument given in [3, p. 399-403] is based on a very similar construction as we have and if all the preliminary work go through, then there may be a quite good chance to get a full proof of Conjecture 3.2.4.

Acknowledgement

The author wants to thank Jeffrey Steif for valuable comments and especially for helping me with Lemma 3.3.3.

References

[1] M. Bramson, R. Durrett, and R. H. Schonmann, *The contact process in a random environment*, Ann. Probab. **19** (1991), 960–983.

[2] E. I. Broman, *Stochastic Domination for a Hidden Markov Chain with Applications to the Contact Process in a Randomly Evolving Environment*, Ann. Probab. **35** (2007), 2263–2293.

[3] L. F. Gary, *The Positive Rates Problem for Attractive Nearest Neighbor Spin Systems on $\mathbb{Z}$*, Probab. Th. Rel. Fields **61** (1982), 389–404.

[4] A. Klein, *Extinction of contact and percolation processes in a random envionment*, Ann. Probab. **22** (1994), 1227–1251.

[5] T. M. Liggett, *Attractive nearest neighbor spin systems on the integers*, Ann. Probab. **6** (1978), 629–636.

[6] ———, *Interacting Particle Systems*, Springer, 1985.

[7] ———, *The survival of one-dimensional contact processes in random environments*, Ann. Probab. **20** (1992), 696–723.

[8] C. M. Newman and S. B. Volchan, *Persistent survival of one-dimensional contact processes in random environments*, Ann. Probab. **24** (1996), 411–421.

[9] Jeffrey E. Steif and Marcus Warfheimer, *The contact process in a randomly evolving environment dies out*, Alea **4** (2008), 337–355.

PAPER III

Stochastic domination for the Ising and fuzzy Potts models

Marcus Warfheimer

4

PAPER III

ABSTRACT

We discuss various aspects concerning stochastic domination for the Ising model and the fuzzy Potts model. We begin by considering the Ising model on the homogeneous tree of degree d, $\mathbb{T}^d$. For given interaction parameters J_1, $J_2 > 0$ and external field $h_1 \in \mathbb{R}$, we compute the smallest external field $\tilde{h}$ such that the plus measure with parameters J_2 and h dominates the plus measure with parameters J_1 and h_1 for all $h \geq \tilde{h}$. Moreover, we discuss continuity of $\tilde{h}$ with respect to the three parameters J_1, J_2, h and also how the plus measures are stochastically ordered in the interaction parameter for a fixed external field. Next, we consider the fuzzy Potts model and prove that on $\mathbb{Z}^d$ the fuzzy Potts measures dominate the same set of product measures while on $\mathbb{T}^d$, for certain parameter values, the free and minus fuzzy Potts measures dominate different product measures. For the Ising model, Liggett and Steif proved that on $\mathbb{Z}^d$ the plus measures dominate the same set of product measures while on $\mathbb{T}^2$ that statement fails completely except when there is a unique phase.

Key words and phrases: Stochastic domination, Ising model, fuzzy Potts model.
Subject classification : 60K35.

4.1 Introduction and main results

The concept of stochastic domination has played an important role in probability theory over the last couple of decades, for example in interacting particle systems and statistical mechanics. In [13], various results were proved concerning stochastic domination for the Ising model with no external field on $\mathbb{Z}^d$ and on the homogeneous binary tree $\mathbb{T}^2$ (i.e. the unique infinite tree where each site has 3 neighbors). As an example, the following distinction between $\mathbb{Z}^d$ and $\mathbb{T}^2$ was shown: On $\mathbb{Z}^d$, the plus and minus states dominate the same set of product measures, while on $\mathbb{T}^2$ that statement fails completely except in the case when we have a unique phase. In this paper we study stochastic domination for the Ising model in the case of nonzero external field and also for the so called fuzzy Potts model.

Let V be a finite or countable set and equip the space $\{-1,1\}^V$ with the following natural partial order: For η, $\eta' \in \{-1,1\}^V$, we write $\eta \leq \eta'$ if $\eta(x) \leq \eta'(x)$ for all $x \in V$. Moreover, whenever we need a topology on $\{-1,1\}^V$ we will use the product topology. We say that a function $f : \{-1,1\}^V \to \mathbb{R}$ is increasing if $f(\eta) \leq f(\eta')$ whenever $\eta \leq \eta'$. We will use the following usual definition of stochastic domination.

Definition 4.1.1 (Stochastic domination). Given a finite or countable set V and probability measures μ_1, μ_2 on $\{-1,1\}^V$, we say that μ_2 dominates μ_1 (written $\mu_1 \leq \mu_2$ or $\mu_2 \geq \mu_1$) if

$$\int f \, d\mu_1 \leq \int f \, d\mu_2$$

for all real-valued, continuous and increasing functions f on $\{-1,1\}^V$.

It is well known that a necessary and sufficient condition for two probability measures μ_1, μ_2 to satisfy $\mu_1 \leq \mu_2$ is that there exists a coupling measure ν on $\{-1,1\}^V \times \{-1,1\}^V$ with first and second marginals equal to μ_1 and μ_2 respectively and

$$\nu(\,(\eta,\xi) : \eta \leq \xi\,) = 1.$$

(For a proof, see for example [12, p. 72-74].) Given any set $S \subseteq \mathbb{R}$ and a family of probability measures $\{\mu_s\}_{s\in S}$ indexed by S, we will say that the map $S \ni s \mapsto \mu_s$ is increasing if $\mu_{s_1} \leq \mu_{s_2}$ whenever $s_1 < s_2$.

4.1.1 The Ising model

The ferromagnetic Ising model is a well studied object in both physics and probability theory. For a given infinite, locally finite (i.e. each vertex has a finite number of neighbors), connected graph $G = (V,E)$, it is defined from the nearest-neighbor potential

$$\Phi_A^{J,h}(\eta) = \begin{cases} -J\eta(x)\eta(y) & \text{if } A = \{x,y\}, \text{ with } \langle x,y\rangle \in E, \\ -h\eta(x) & \text{if } A = \{x\}, \\ 0 & \text{otherwise} \end{cases}$$

where $A \subseteq V$, $\eta \in \{-1,1\}^V$, $J > 0$, $h \in \mathbb{R}$ are two parameters called the coupling strength and the external field respectively and $\langle x,y\rangle$ denotes the edge connecting x

and y. A probability measure μ on $\{-1,1\}^V$ is said to be a Gibbs measure (or sometimes Gibbs state) for the ferromagnetic Ising model with parameters $h \in \mathbb{R}$ and $J > 0$ if it admits conditional probabilities such that for all finite $U \subseteq V$, all $\sigma \in \{-1,1\}^U$ and all $\eta \in \{-1,1\}^{V \setminus U}$

$$\begin{aligned}\mu(X(U) = \sigma \,|\, X(V \setminus U) = \eta) \\ = \frac{1}{Z^{U,\eta}_{J,h}} \exp\Bigg[J \Bigg(\sum_{\langle x,y\rangle \in E, x,y \in U} \sigma(x)\sigma(y) + \sum_{\langle x,y\rangle \in E, x \in U, y \in \partial U} \sigma(x)\eta(y) \Bigg) \\ + h \sum_{x \in U} \sigma(x) \Bigg]\end{aligned}$$

where $Z^{U,\eta}_{J,h}$ is a normalizing constant and

$$\partial U = \{\, x \in V \setminus U : \exists y \in U \text{ such that } \langle x,y\rangle \in E \,\}.$$

For given $J > 0$ and $h \in \mathbb{R}$, we will denote the set of Gibbs measures with parameters J and h by $\mathcal{G}(J,h)$ and we say that a phase transition occurs if $|\mathcal{G}(J,h)| > 1$, i.e. if there exist more than one Gibbs state. (From the general theory described in [2] or [12], $\mathcal{G}(J,h)$ is always nonempty.) At this stage one can ask, for fixed $h \in \mathbb{R}$, is it the case that the existence of multiple Gibbs states is increasing in J? When $h = 0$ it is possible from the so called random-cluster representation of the Ising model to show a positive answer to the last question (see [5] for the case when $G = \mathbb{Z}^d$ and [7] for more general G). However, when $h \neq 0$ there are graphs where the above monontonicity property no longer holds, see [15] for an example of a relatively simple such graph.

Furthermore, still for fixed $J > 0$, $h \in \mathbb{R}$, standard monotonicity arguments can be used to show that there exists two particular Gibbs states $\mu^{J,+}_h$, $\mu^{J,-}_h$, called the plus and the minus state, which are extreme with respect to the stochastic ordering in the sense that

$$\mu^{J,-}_h \leq \mu \leq \mu^{J,+}_h \quad \text{for any other } \mu \in \mathcal{G}(J,h). \tag{1}$$

To simplify the notation, we will write $\mu^{J,+}$ for $\mu^{J,+}_0$ and $\mu^{J,-}$ for $\mu^{J,-}_0$. (Of course, most of the things we have defined so far are also highly dependent on the graph G, but we suppress that in the notation.)

In [13] the authors studied, among other things, stochastic domination between the plus measures $\{\mu^{J,+}\}_{J>0}$ in the case when $G = \mathbb{T}^2$. For example, they showed that the map $(0,\infty) \ni J \mapsto \mu^{J,+}$ is increasing when $J > J_c$ and proved the existence of and computed the smallest $J > J_c$ such that $\mu^{J,+}$ dominates $\mu^{J',+}$ for all $0 < J' \leq J_c$. (On $\mathbb{Z}^d$, the fact that $\mu^{J_1,+}$ and $\mu^{J_2,+}$ are not stochastically ordered when $J_1 \neq J_2$ gives that such a J does not even exist in that case.) Our first result deals with the following question: Given J_1, $J_2 > 0$, $h_1 \in \mathbb{R}$, can we find the smallest external field $\tilde{h} = \tilde{h}(J_1, J_2, h_1)$ with the property that $\mu^{J_2,+}_h$ dominates $\mu^{J_1,+}_{h_1}$ for all $h \geq \tilde{h}$?

To clarify the question a bit more, note that an easy application of Holley's theorem (see [3]) tells us that for fixed $J > 0$, the map $\mathbb{R} \ni h \mapsto \mu_h^{J,+}$ is increasing. Hence, for given J_1, J_2 and h_1 as above the set

$$\{ h \in \mathbb{R} : \mu_h^{J_2,+} \geq \mu_{h_1}^{J_1,\pm} \}$$

is an infinite interval and we want to find the left endpoint of that interval (possibly $-\infty$ or $+\infty$ at this stage). For a general graph not much can be said, but we have the following easy bounds on $\tilde{h}$ when G is of bounded degree.

Proposition 4.1.1. *Consider the Ising model on a general graph $G = (V, E)$ of bounded degree. Define*

$$\tilde{h} = \tilde{h}(J_1, J_2, h_1) = \inf\{ h \in \mathbb{R} : \mu_h^{J_2,+} \geq \mu_{h_1}^{J_1,+} \}.$$

Then

$$h_1 - N(J_1 + J_2) \leq \tilde{h} \leq h_1 + N|J_1 - J_2|,$$

where $N = \sup_{x \in V} N_x$ and N_x is the number of neighbors of the site $x \in V$.

For the Ising model, we will now consider the case when $G = \mathbb{T}^d$, the homogeneous d-ary tree, defined as the unique infinite tree where each site has exactly $d + 1 \geq 3$ neighbors. The parameter d is fixed in all that we will do and so we suppress that in the notation. For this particular graph it is well known that for given $h \in \mathbb{R}$, the existence of multiple Gibbs states is increasing in J and so as a consequence there exists a critical value $J_c(h) \in [0, \infty]$ such that when $J < J_c(h)$ we have a unique Gibbs state whereas for $J > J_c(h)$ there are more than one Gibbs states. In fact, much more can be shown in this case. As an example it is possible to derive an explicit expression for the phase transition region

$$\{ (J, h) \in \mathbb{R}^2 : |\mathcal{G}(J, h)| > 1 \},$$

in particular one can see that $J_c(h) \in (0, \infty)$ for all $h \in \mathbb{R}$. Moreover,

$$J_c := J_c(0) = \operatorname{arccoth} d = \frac{1}{2} \log \frac{d+1}{d-1},$$

see [2] for more details. (Here and in the sequel, := will mean definition.)

To state our results for the Ising model on $\mathbb{T}^d$, we need to recall some more facts, all of which can be found in [2, p. 247-255]. To begin, we just state what we need very briefly and later on we will give some more details. Given $J > 0$ and $h \in \mathbb{R}$, there is a one-to-one correspondence $t \mapsto \mu$ between the real solutions of a certain equation (see (7) and the function ϕ_J in (6) below) and the completely homogeneous Markov chains in $\mathcal{G}(J, h)$ (to be defined in Section 4.2). Let $t_\pm(J, h)$ denote the real numbers which correspond to the plus and minus measure respectively. (It is easy to see that the plus and minus states are completely homogeneous Markov chains, see Section 4.2.) We will write $t_\pm(J)$ instead of $t_\pm(J,0)$. Furthermore, let

$$h^*(J) = \max_{t\geq 0}\left(d\,\phi_J(t) - t\right)$$

and denote by $t^*(J)$ the $t \geq 0$ where the function $t \mapsto d\,\phi_J(t) - t$ attains its unique maximum. In [2], explicit expressions for both h^* and t^* are derived:

$$h^*(J) = \begin{cases} 0 & \text{if } J \leq J_c \\ d\,\text{arctanh}\left(\frac{d\tanh(J)-1}{d\coth(J)-1}\right)^{1/2} - \text{arctanh}\left(\frac{d-\coth(J)}{d-\tanh(J)}\right)^{1/2} & \text{if } J > J_c \end{cases}$$

$$t^*(J) = \begin{cases} 0 & \text{if } J \leq J_c \\ \text{arctanh}\left(\frac{d-\coth(J)}{d-\tanh(J)}\right)^{1/2} & \text{if } J > J_c \end{cases}$$

In particular one can see that both h^* and t^* are continuous functions of J and by computing derivatives one can show that they are strictly increasing for $J > J_c$.

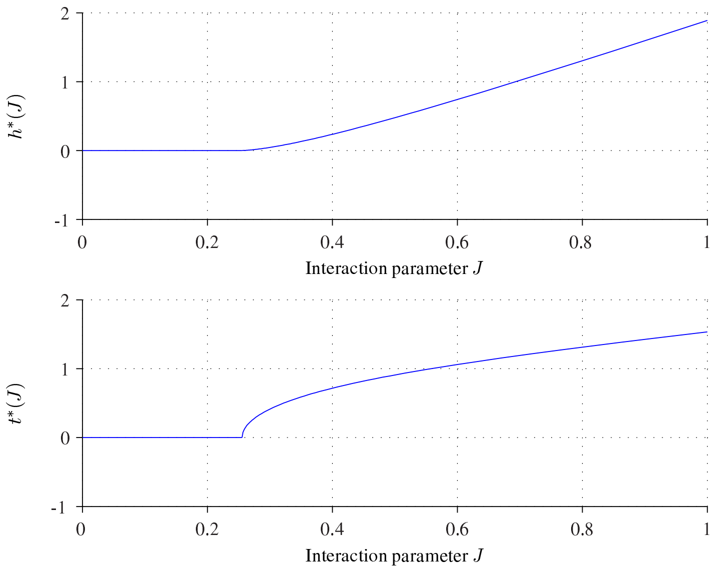

Figure 4.1: *The functions h^* and t^* in the case when $d = 4$.*

Theorem 4.1.2. *Consider the Ising model on $\mathbb{T}^d$ and let J_1, $J_2 > 0$, $h_1 \in \mathbb{R}$ be given. Define*

$$f_\pm(J_1,J_2,h_1) = \inf\{\, h\in\mathbb{R} : \mu_h^{J_2,+} \geq \mu_{h_1}^{J_1,\pm}\,\}$$
$$g_\pm(J_1,J_2,h_1) = \inf\{\, h\in\mathbb{R} : \mu_h^{J_2,-} \geq \mu_{h_1}^{J_1,\pm}\,\}$$

and denote $\tau_\pm = \tau_\pm(J_1, J_2, h_1) = t_\pm(J_1, h_1) + |J_1 - J_2|$. *Then the following holds:*

$$(2)\quad f_\pm(J_1, J_2, h_1) = \begin{cases} -h^*(J_2) & \text{if } t_-(J_2, -h^*(J_2)) \le \tau_\pm < t^*(J_2) \\ \tau_\pm - d\,\phi_{J_2}(\tau_\pm) & \text{if } \tau_\pm \ge t^*(J_2) \text{ or } \tau_\pm < t_-(J_2, -h^*(J_2)) \end{cases}$$

$$(3)\quad g_\pm(J_1, J_2, h_1) = \begin{cases} h^*(J_2) & \text{if } -t^*(J_2) < \tau_\pm \le t_+(J_2, h^*(J_2)) \\ \tau_\pm - d\,\phi_{J_2}(\tau_\pm) & \text{if } \tau_\pm \le -t^*(J_2) \text{ or } \tau_\pm > t_+(J_2, h^*(J_2)) \end{cases}$$

Remarks:

(i) Note that if $0 < J_2 \le J_c$, then $h^*(J_2) = 0$ and

$$t_-(J_2, -h^*(J_2)) = t^*(J_2) = t_+(J_2, h^*(J_2)) = 0$$

and hence the first interval disappears in the formulas and we simply get

$$\begin{aligned} f_\pm(J_1, J_2, h_1) &= g_\pm(J_1, J_2, h_1) \\ &= \tau_\pm(J_1, J_2, h_1) - d\phi_{J_2}(\tau_\pm(J_1, J_2, h_1)). \end{aligned}$$

(ii) By looking at the formulas (2) and (3), we see that there are functions ψ, θ : $(0, \infty) \times \mathbb{R} \mapsto \mathbb{R}$ such that

$$\begin{aligned} &f_\pm(J_1, J_2, h_1) = \psi(J_2, \tau_\pm(J_1, J_2, h_1)) \quad \text{and} \\ &g_\pm(J_1, J_2, h_1) = \theta(J_2, \tau_\pm(J_1, J_2, h_1)). \end{aligned}$$

(Of course, $\psi(J_2, t)$ and $\theta(J_2, t)$ are just (2) and (3) with t instead of $\tau_\pm$.) It is easy to check that for fixed $J_2 > 0$, the maps $t \mapsto \psi(J_2, t)$ and $t \mapsto \theta(J_2, t)$ are continuous. A picture of these functions when $J_2 = 2$, $d = 4$ can be seen in Figure 4.2.

(iii) It is not hard to see by direct computations that f_+ satisfies the bounds in Proposition 4.1.1. We will indicate how this can be done after the proof of Theorem 4.1.2.

(iv) We will see in the proof that if

$$t_-(J_2, -h^*(J_2)) \le \tau_\pm(J_1, J_2, h_1) < t^*(J_2),$$

then

$$\{\, h \in \mathbb{R} : \mu_h^{J_2,+} \ge \mu_{h_1}^{J_1,\pm} \,\} = [-h^*(J_2), \infty),$$

and if $-t^*(J_2) < \tau_\pm(J_1, J_2, h_1) \le t_+(J_2, h^*(J_2))$, then

$$\{\, h \in \mathbb{R} : \mu_h^{J_2,-} \ge \mu_{h_1}^{J_1,\pm} \,\} = (h^*(J_2), \infty).$$

Hence in the first case the left endpoint belongs to the interval, while in the second case it does not.

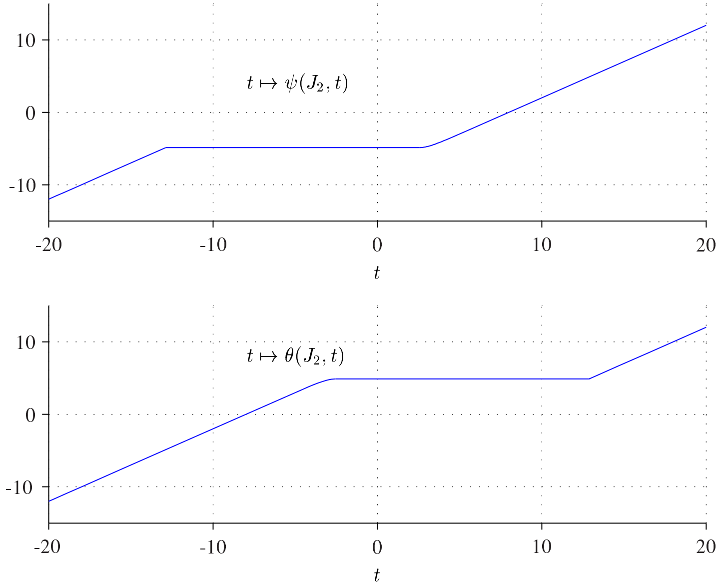

Figure 4.2: *The functions $t \mapsto \psi(J_2, t)$ and $t \mapsto \theta(J_2, t)$ in the case when $J_2 = 2$ and $d = 4$.*

Our next proposition deals with continuity properties of $f_\pm$ and $g_\pm$ with respect to the parameters J_1, J_2 and h_1. We will only discuss the function f_+, the other ones can be treated in a similar fashion.

Proposition 4.1.3. *Consider the Ising model on $\mathbb{T}^d$ and recall the notation from Theorem 4.1.2. Let*

$$a = a(J_1, J_2) = t_-(J_1, -h^*(J_1)) + |J_1 - J_2|$$
$$b = b(J_1, J_2) = t_+(J_1, -h^*(J_1)) + |J_1 - J_2|$$

a) Given J_1, $J_2 > 0$, the map $\mathbb{R} \ni h_1 \mapsto f_+(J_1, J_2, h_1)$ is continuous except possibly at $-h^*(J_1)$ depending on J_1 and J_2 in the following way:

If $J_1 \le J_c$ or $J_1 = J_2$ then it is continuous at $-h^*(J_1)$.

If $J_1 > J_c$ and $0 < J_2 \le J_c$ then it is discontinuous at $-h^*(J_1)$.

If J_1, $J_2 > J_c$, $J_1 \ne J_2$ then it is discontinuous except when

$$t_-(J_2, -h^*(J_2)) \le a < t^*(J_2) \text{ and } t_-(J_2, -h^*(J_2)) \le b \le t^*(J_2).$$

b) Given $J_2 > 0$, $h_1 \in \mathbb{R}$, the map $(0, \infty) \ni J_1 \mapsto f_+(J_1, J_2, h_1)$ is continuous at J_1 if $0 < J_1 \le J_c$ or $J_1 > J_c$ and $h_1 \ne -h^*(J_1)$. In the case when

$h_1 = -h^*(J_1)$ *it is discontinuous at* J_1 *except when*

$$t_-(J_2, -h^*(J_2)) \leq a < t^*(J_2) \text{ and } t_-(J_2, -h^*(J_2)) \leq b \leq t^*(J_2).$$

c) *Given* $J_1 > 0$, $h_1 \in \mathbb{R}$, *the map* $(0, \infty) \ni J_2 \mapsto f_+(J_1, J_2, h_1)$ *is continuous for all* $J_2 > 0$.

We conclude this section with a result about how the measures $\{\mu_h^{J,+}\}_{J>0}$ are ordered with respect to J for fixed $h \in \mathbb{R}$.

Proposition 4.1.4. *Consider the Ising model on* $\mathbb{T}^d$. *The map* $(0, \infty) \ni J \mapsto \mu_h^{J,+}$ *is increasing in the following cases: a)* $h \geq 0$ *and* $J \geq J_c$, *b)* $h < 0$ *and* $h^*(J) > -h$.

4.1.2 The fuzzy Potts model

Next, we consider the so called fuzzy Potts model. To define the model, we first need to define the perhaps more familiar Potts model. Let $G = (V, E)$ be an infinite locally finite graph and suppose that $q \geq 3$ is an integer. Let U be a finite subset of V and consider the finite graph H with vertex set U and edge set consisting of those edges $\langle x, y \rangle \in E$ with $x, y \in U$. In this way, we say that the graph H is induced by U. The finite volume Gibbs measure for the q-state Potts model at inverse temperature $J \geq 0$ with free boundary condition is defined to be the probability measure $\pi_{q,J}^H$ on $\{1, 2, \ldots, q\}^U$ which to each element σ assigns probability

$$\pi_{q,J}^H(\sigma) = \frac{1}{Z_{q,J}^H} \exp\left(2J \sum_{\langle x,y \rangle \in E, x,y \in U} I_{\{\sigma(x)=\sigma(y)\}}\right),$$

where $Z_{q,J}^H$ is a normalizing constant.

Now, suppose $r \in \{1, \ldots, q-1\}$ and pick a $\pi_{q,J}^H$ - distributed object X and for $x \in U$ let

$$Y(x) = \begin{cases} -1 & \text{if } X(x) \in \{1, \ldots, r\} \\ 1 & \text{if } X(x) \in \{r+1, \ldots, q\}. \end{cases} \tag{4}$$

We write $\nu_{q,J,r}^H$ for the resulting probability measure on $\{-1, 1\}^U$ and call it the finite volume fuzzy Potts measure on H with free boundary condition and parameters q, J and r.

We also need to consider the case when we have a boundary condition. For finite $U \subseteq V$, consider the graph H induced by the vertex set $U \cup \partial U$ and let $\eta \in \{1, \ldots, q\}^{V \setminus U}$. The finite volume Gibbs measure for the q-state Potts model at inverse temperature $J \geq 0$ with boundary condition η is defined to be the probability measure on $\{1,\ldots,q\}^U$ which to each element assigns probability

$$\pi_{q,J}^{H,\eta}(\sigma) = \frac{1}{Z_{q,J}^{H,\eta}} \exp\Bigg(2J \sum_{\langle x,y\rangle\in E, x,y\in U} I_{\{\sigma(x)=\sigma(y)\}} + 2J \sum_{\langle x,y\rangle\in E, x\in U, y\in\partial U} I_{\{\sigma(x)=\eta(y)\}}\Bigg),$$

where $Z_{q,J}^{H,\eta}$ is a normalizing constant. In the case when $\eta \equiv i$ for some $i \in \{1,\ldots,q\}$, we replace η with i in the notation.

Furthermore, we introduce the notion of infinite volume Gibbs measure for the Potts model. A probability measure μ on $\{1,\ldots,q\}^V$ is said to be an infinite volume Gibbs measure for the q-state Potts model on G at inverse temperature $J \geq 0$, if it admits conditional probabilities such that for all finite $U \subseteq V$, all $\sigma \in \{1,\ldots,q\}^U$ and all $\eta \in \{1,\ldots,q\}^{V\setminus U}$

$$\mu(X(U) = \sigma \,|\, X(V\setminus U) = \eta) = \pi_{q,J}^{H,\eta}(\sigma)$$

where H is the graph induced by $U \cup \partial U$. Let $\{V_n\}_{n\geq 1}$ be a sequence of finite subsets of V such that $V_n \subseteq V_{n+1}$ for all n, $V = \bigcup_{n\geq 1} V_n$ and for each n, denote by G_n the induced graph by $V_n \cup \partial V_n$. Furthermore, for each $i \in \{1,\ldots,q\}$, extend $\pi_{q,J}^{G_n,i}$ (and use the same notation for the extension) to a probability measure on $\{1,\ldots,q\}^V$ by assigning with probability one the spin value i outside V_n. It is well known (and independent of the sequence $\{V_n\}$) that there for each spin $i \in \{1,\ldots,q\}$ exists a infinite volume Gibbs measure $\pi_{q,J}^{G,i}$ which is the weak limit as $n \to \infty$ of the corresponding measures $\pi_{q,J}^{G_n,i}$. Moreover, there exists another infinite volume Gibbs measure denoted $\pi_{q,J}^{G,0}$ which is the limit of $\pi_{q,J}^{G_n}$ in the sense that the probabilities on cylinder sets converge. The existence of the above limits as well as the independence of the choice of the sequence $\{V_n\}$ when constructing them follows from the work of Aizenman et al. [1].

Given the infinite volume Gibbs measures $\{\pi_{q,J}^{G,i}\}_{i\in\{0,\ldots,q\}}$, we define the corresponding infinite volume fuzzy Potts measures $\{\nu_{q,J,r}^{G,i}\}_{i\in\{0,\ldots,q\}}$ using (4).

In words, the fuzzy Potts model can be thought of arising from the ordinary q-state Potts model by looking at a pair of glasses that prevents from distinguishing some of the spin values. From this point of view, the fuzzy Potts model is one of the most basic examples of a so called hidden Markov field [11]. For earlier work on the fuzzy Potts model, see for example [6, 8–10, 14].

Given a finite or countable set V and $p \in [0,1]$, let γ_p denote the product measure on $\{-1,1\}^V$ with $\gamma_p(\eta : \eta(x) = 1) = p$ for all $x \in V$. In [13] the authors proved the following results for the Ising model. (The second result was originally proved for $d = 2$ only but it trivially extends to all $d \geq 2$.)

Proposition 4.1.5 (Liggett, Steif). *Fix an integer $d \geq 2$ and consider the Ising model on $\mathbb{Z}^d$ with parameters $J > 0$ and $h = 0$. Then for any $p \in [0,1]$, $\mu^{J,+} \geq \gamma_p$ if and only if $\mu^{J,-} \geq \gamma_p$.*

Proposition 4.1.6 (Liggett, Steif). *Let $d \geq 2$ be a given integer and consider the Ising model on $\mathbb{T}^d$ with paramteters $J > 0$ and $h = 0$. Moreover, let $\mu^{J,f}$ denote the Gibbs state obtained by using free boundary conditions. If $\mu^{J,+} \neq \mu^{J,-}$, then there exist $0 < p' < p$ such that $\mu^{J,+}$ dominates γ_p but $\mu^{J,f}$ does not dominate γ_p and $\mu^{J,f}$ dominates $\gamma_{p'}$ but $\mu^{J,-}$ does not dominate $\gamma_{p'}$.*

In words, on $\mathbb{Z}^d$ the plus and minus state dominate the same set of product measures while on $\mathbb{T}^d$ that is not the case except when the we have a unique phase.

To state our next results we will take a closer look at the construction of the infinite volume fuzzy Potts measures when $G = \mathbb{Z}^d$ or $G = \mathbb{T}^d$. In those cases it follows from symmetry that $\nu^{G,i}_{q,J,r} = \nu^{G,j}_{q,J,r}$ if $i,j \in \{1,\ldots,r\}$ or $i,j \in \{r+1,\ldots,q\}$, i.e. when the Potts spins i,j map to the same fuzzy spin. For that reason, we let $\nu^{G,-}_{q,J,r} := \nu^{G,1}_{q,J,r}$ and $\nu^{G,+}_{q,J,r} := \nu^{G,q}_{q,J,r}$ when $G = \mathbb{Z}^d$ or $\mathbb{T}^d$. (Of course, we stick to our earlier notation of $\nu^{G,0}_{q,J,r}$.) Our first result is a generalization of Proposition 4.1.5 to the fuzzy Potts model.

Proposition 4.1.7. *Let $d \geq 2$ be a given integer and consider the fuzzy Potts model on $\mathbb{Z}^d$ with parameters $q \geq 3$, $J > 0$ and $r \in \{1,\ldots,q-1\}$. Then for any $k,l \in \{0,-,+\}$ and $p \in [0,1]$, $\nu^{\mathbb{Z}^d,k}_{q,J,r} \geq \gamma_p$ if and only if $\nu^{\mathbb{Z}^d,l}_{q,J,r} \geq \gamma_p$.*

In the same way as for the Ising model, we believe that Proposition 4.1.7 fails completely on $\mathbb{T}^d$ except when we have a unique phase in the Potts model. Our last result is in that direction.

Proposition 4.1.8. *Let $d \geq 2$ be a given integer and consider the fuzzy Potts model on $\mathbb{T}^d$ with parameters $q \geq 3$, $J > 0$ and $r \in \{1,\ldots,q-1\}$ where $e^{2J} \geq q-2$. If the underlying Gibbs measures for the Potts model satisfy $\pi^{\mathbb{T}^d,1}_{q,J} \neq \pi^{\mathbb{T}^d,0}_{q,J}$, then there exists $0 < p < 1$ such that $\nu^{\mathbb{T}^d,0}_{q,J,r}$ dominates γ_p but $\nu^{\mathbb{T}^d,-}_{q,J,r}$ does not dominate γ_p.*

4.2 Proofs

We start to recall some facts from [2] concerning the notion of completely homogeneous Markov chains on $\mathbb{T}^d$. Denote the vertex set and the edge set of $\mathbb{T}^d$ with $V(\mathbb{T}^d)$ and $E(\mathbb{T}^d)$ respectively. Given a directed edge $\langle x,y\rangle \in E(\mathbb{T}^d)$ define the "past" sites by

$$]-\infty,\langle x,y\rangle[= \{\, z \in V(\mathbb{T}^d) : z \text{ is closer to } x \text{ than to } y \,\}.$$

For $A \subseteq V(\mathbb{T}^d)$ denote by $\mathcal{F}_A$ the σ-algebra generated by the spins in A. A probability measure μ on $\{-1,1\}^{V(\mathbb{T}^d)}$ is called a Markov chain if

$$\mu(\,\eta(y) = 1 \mid \mathcal{F}_{]-\infty,\langle x,y\rangle[}\,) = \mu(\,\eta(y) = 1 \mid \mathcal{F}_{\{x\}}\,) \quad \mu\text{-a.s.}$$

for all $\langle x, y\rangle \in E(\mathbb{T}^d)$. Furthermore, a Markov chain μ is called completely homogeneous with transition matrix $P = \{\, P(i,j) : i,j \in \{\,-1,1\}\,\}$ if

$$\mu(\,\eta(y) = u \,|\, \mathcal{F}_{\{x\}}\,) = P(\eta(x), u) \quad \mu\text{-a.s.} \tag{5}$$

for all $\langle x, y\rangle \in E(\mathbb{T}^d)$ and $u \in \{-1,1\}$. Observe that such a P necessarily is a stochastic matrix and if it in addition is irreducible denote its stationary distribution by ν. In that situation, we get for each finite connected set $C \subseteq V(\mathbb{T}^d)$, $z \in C$ and $\xi \in \{-1,1\}^C$ that

$$\mu(\eta = \xi) = \nu(\xi(z)) \prod_{\langle x,y\rangle \in D} P(\xi(x), \xi(y))$$

where D is the set of directed edges $\langle x, y\rangle$, where $x, y \in C$ and x is closer to z than y is. In particular, it follows that every completely homogeneous Markov chain which arise from an irreducible stochastic matrix is invariant under all graph automorphisms.

Next, we give a short summary from [2] of the Ising model on $\mathbb{T}^d$. For $J > 0$, define

$$\phi_J(t) = \frac{1}{2} \log \frac{\cosh(t+J)}{\cosh(t-J)}, \quad t \in \mathbb{R}. \tag{6}$$

The function ϕ_J is trivially seen to be odd. Moreover, ϕ_J is concave on $[0, \infty)$, increasing and bounded. (In fact, $\phi_J(t) \to J$ as $t \to \infty$.) Furthermore, there is a one-to-one correspondence $t \mapsto \mu_t$ between the completely homogeneous Markov chains in $\mathcal{G}(J,h)$ and the numbers $t \in \mathbb{R}$ satisfying the equation

$$t = h + d\,\phi_J(t). \tag{7}$$

In addition, the transition matrix P_t of μ_t is given by

$$\begin{pmatrix} P_t(-1,-1) & P_t(-1,1) \\ P_t(1,-1) & P_t(1,1) \end{pmatrix} = \begin{pmatrix} \frac{e^{J-t}}{2\cosh(J-t)} & \frac{e^{t-J}}{2\cosh(J-t)} \\ \frac{e^{-J-t}}{2\cosh(J+t)} & \frac{e^{J+t}}{2\cosh(J+t)} \end{pmatrix}. \tag{8}$$

Given $h \in \mathbb{R}$ and $J > 0$ the fixed point equation (7) has one, two or three solutions. In fact Lemma 4.2.1 below tells us exactly when the different situations occur. The largest solution, denoted $t_+(J,h)$, corresponds to the plus measure $\mu_h^{J,+}$ and the smallest, denoted $t_-(J,h)$, to the minus measure $\mu_h^{J,-}$. To see why the last statement is true, let $\mu_\pm = \mu_{t_\pm(J,h)}$ and note that Lemma 4.2.2 from Section 4.2.2 implies that $\mu_- \le \mu \le \mu_+$ for any $\mu \in \mathcal{G}(J,h)$ which is also a completely homogeneous Markov chain on $\mathbb{T}^d$. Moreover, equation (1) implies that $\mu_h^{J,-} \le \mu_\pm \le \mu_h^{J,+}$ and so $\mu_\pm = \mu_h^{J,\pm}$ will follow if $\mu_h^{J,\pm}$ are completely homogeneous Markov chains. To see that, note that equation (1) also implies that $\mu_h^{J,\pm}$ are extremal in $\mathcal{G}(J,h)$ which in turn (see Theorem 12.6 in [2]) gives us that they are Markov chains on $\mathbb{T}^d$. Finally, from the fact that $\mu_h^{J,\pm}$ are invariant under all graph automorphisms on $\mathbb{T}^d$, we obtain the completely homogeneous property (5).

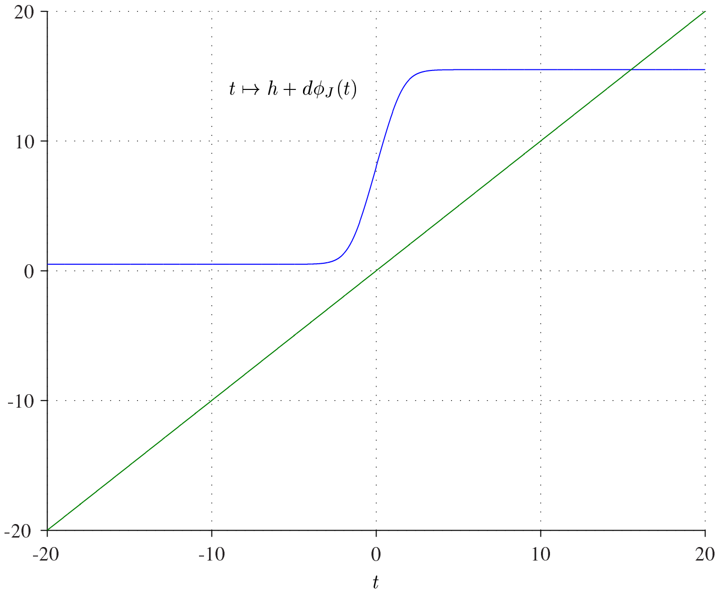

Figure 4.3: *A picture of the fixed point equation* (7) *when* $d = 5$, $h = 8$ *and* $J = 3/2$. *In this particular case we have a unique solution.*

Lemma 4.2.1 (Georgii). *The fixed point equation* (7) *has*

a) a unique solution when $|h| > h^*(J)$ *or* $h = h^*(J) = 0$,

b) two distinct solutions $t_-(J,h) < t_+(J,h)$ *when* $|h| = h^*(J) > 0$,

c) three distinct solutions $t_-(J,h) < t_0(J,h) < t_+(J,h)$ *when* $|h| < h^*(J)$.

4.2.1 Proof of Proposition 4.1.1

For the upper bound, just invoke Proposition 4.16 in [3] which gives us that $\mu_h^{J_2,+} \geq \mu_{h_1}^{J_1,+}$ if $h \geq h_1 + N|J_1 - J_2|$.

For the lower bound, we argue by contradiction as follows. Assume

$$\tilde{h} < h_1 - N(J_1 + J_2)$$

and pick h_0 such that

$$\tilde{h} < h_0 < h_1 - N(J_1 + J_2). \tag{9}$$

The right inequality of (9) is equivalent to

$$2(h_0 + NJ_2) < 2(h_1 - NJ_1)$$

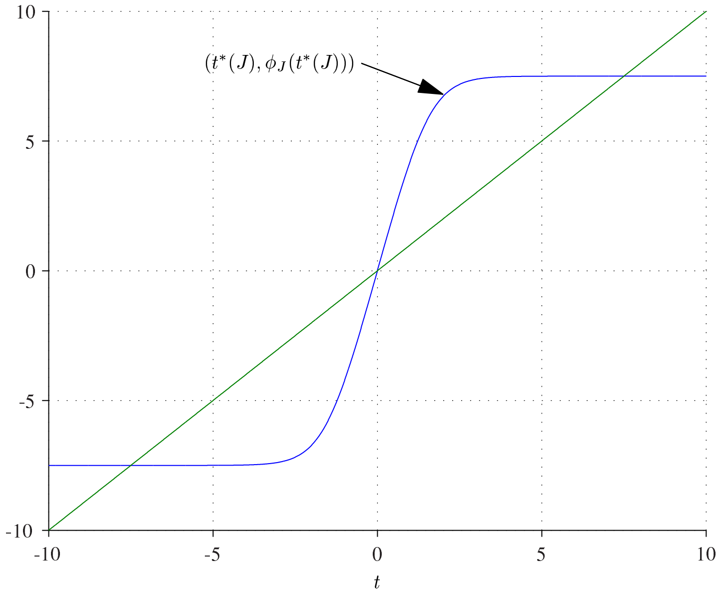

Figure 4.4: *A picture of the fixed point equation* (7) *when* $d = 5$, $h = 0$ *and* $J = 3/2$.

and so we can pick $0 < p_1 < p_2 < 1$ such that

$$2(h_0 + NJ_2) < \log(\frac{p_1}{1-p_1}) < \log(\frac{p_2}{1-p_2}) < 2(h_1 - NJ_1).$$

By using the last inequalities together with Proposition 4.16 in [3], we can conclude that

$$\begin{aligned} \mu_{h_0}^{J_2,+} &\leq \gamma_{p_1} \\ \mu_{h_1}^{J_1,+} &\geq \gamma_{p_2}. \end{aligned}$$

Since $p_1 < p_2$ this tells us that $\mu_{h_0}^{J_2,+} \not\geq \mu_{h_1}^{J_1,+}$. On the other hand we have $h_0 > \tilde{h}$ which by definition of $\tilde{h}$ implies that $\mu_{h_0}^{J_2,+} \geq \mu_{h_1}^{J_1,+}$. Hence, we get a contradiction and the proof is complete. $\square$

4.2.2 Proof of Theorem 4.1.2

We will make use of the following lemma from [13] concerning stochastic domination for completely homogeneous Markov chains on $\mathbb{T}^d$.

Lemma 4.2.2 (Liggett, Steif). *Given two 2-state transition matrices P and Q, let μ_P and μ_Q denote the corresponding completely homogeneous Markov chains on $\mathbb{T}^d$. Then μ_P dominates μ_Q if and only if $P(-1,1) \geq Q(-1,1)$ and $P(1,1) \geq Q(1,1)$.*

Proof of Theorem 4.1.2. To prove (2), let $J_1,\, J_2 > 0$ and $h_1 \in \mathbb{R}$ be given and note that we get from Lemma 4.2.2 and equation (8) that $\mu_h^{J_2,+} \geq \mu_{h_1}^{J_1,\pm}$ if and only if

$$\frac{e^{t_+(J_2,h)-J_2}}{2\cosh(t_+(J_2,h)-J_2)} \geq \frac{e^{t_\pm(J_1,h_1)-J_1}}{2\cosh(t_\pm(J_1,h_1)-J_1)}$$

and

$$\frac{e^{t_+(J_2,h)+J_2}}{2\cosh(t_+(J_2,h)+J_2)} \geq \frac{e^{t_\pm(J_1,h_1)+J_1}}{2\cosh(t_\pm(J_1,h_1)+J_1)}.$$

Since the map $\mathbb{R} \ni x \mapsto \frac{e^x}{2\cosh(x)}$ is strictly increasing this is equivalent to

$$t_+(J_2,h) \geq t_\pm(J_1,h_1) + J_2 - J_1$$

and

$$t_+(J_2,h) \geq t_\pm(J_1,h_1) + J_1 - J_2$$

which in turn is equivalent to

$$t_+(J_2,h) \geq t_\pm(J_1,h_1) + |J_1 - J_2| = \tau_\pm(J_1,J_2,h_1), \tag{10}$$

and so we want to compute the smallest $h \in \mathbb{R}$ such that (10) holds. Note that since the map $h \mapsto t_+(J_2,h)$ is strictly increasing and $t_+(J_2,h) \to \pm\infty$ as $h \to \pm\infty$ there always exists such an $h \in \mathbb{R}$. If $\tau_\pm \geq t^*(J_2)$ or $\tau_\pm < t_-(J_2,-h^*(J_2))$, then the equation

$$h + d\phi_{J_2}(\tau_\pm) = \tau_\pm$$

is equivalent to

$$t_+(J_2,h) = \tau_\pm$$

and so in that case the smallest $h \in \mathbb{R}$ such that (10) holds is equal to

$$\tau_\pm - d\phi_{J_2}(\tau_\pm).$$

If $t_-(J_2,-h^*(J_2)) \leq \tau_\pm < t^*(J_2)$, then since $t_+(J_2,h) \geq t^*(J_2)$ whenever $h \geq -h^*(J_2)$ and $t_+(J_2,h) < t_-(J_2,-h^*(J_2))$ whenever $h < -h^*(J_2)$, we have in this case that

$$\{\, h \in \mathbb{R} : \mu_h^{J_2,+} \geq \mu_{h_1}^{J_1,\pm} \,\} = [-h^*(J_2), \infty),$$

and so the h we are looking for is given by $-h^*(J_2)$.

For (3), we note as above that $\mu_h^{J_2,-} \geq \mu_{h_1}^{J_1,\pm}$ if and only if

$$t_-(J_2,h) \geq \tau_\pm(J_1,J_2,h_1). \tag{11}$$

If $\tau_\pm \leq -t^*(J_2)$ or $\tau_\pm > t_+(J_2, h^*(J_2))$ then we can proceed exactly as in the first case above. If $-t^*(J_2) < \tau_\pm \leq t_+(J_2, h^*(J_2))$, then $t_-(J_2, h) < \tau_\pm$ whenever $h \leq h^*(J_2)$ and $t_-(J_2, h) > \tau_\pm$ whenever $h > h^*(J_2)$ and so in that case we have

$$\{\, h \in \mathbb{R} : \mu_h^{J_2,-} \geq \mu_{h_1}^{J_1,\pm} \,\} = (h^*(J_2), \infty),$$

which yields (3) and the proof is complete. □

We will now indicate how to compute the bounds in Proposition 4.1.1 in the special case when $G = \mathbb{T}^d$. By looking at the formula for f_+ and using the definition of h^* we get that

$$f_+(J_1, J_2, h_1) \leq \tau_+(J_1, J_2, h_1) - d\phi_{J_2}(\tau_+(J_1, J_2, h_1)).$$

Substituting τ_+ and using the bounds $-J \leq \phi_J(t) \leq J$ for all $t \in \mathbb{R}$ we get the upper bound in Proposition 4.1.1 with $N = d + 1$. For the lower bound, first note that

$$\begin{aligned} \tau_+ - d\phi_{J_2}(\tau_+) &= h_1 + d\big(\phi_{J_1}(t_+(J_1, h_1)) - \phi_{J_2}(t_+(J_1, h_1))\big) + |J_1 - J_2| \\ &\geq h_1 - (d+1)(J_1 + J_2). \end{aligned}$$

Moreover it is easy to check that

$$-h^*(J_2) \geq h_1 - (d+1)(J_1 + J_2)$$

when

$$t_-(J_2, -h^*(J_2)) \leq \tau_+ \leq t^*(J_2) = t_+(J_2, -h^*(J_2))$$

and so the lower bound follows at once.

4.2.3 Proof of Proposition 4.1.3

Before we prove anything we would like to recall the fact that we can write (see Remark (ii) after Theorem 4.1.2)

$$f_+(J_1, J_2, h_1) = \psi(J_2, \tau_+(J_1, J_2, h_1)) \quad J_1, J_2 > 0, h_1 \in \mathbb{R},$$

where

$$\tau_+(J_1, J_2, h_1) = t_+(J_1, h_1) + |J_1 - J_2|$$

and the map $t \mapsto \psi(J_2, t)$ is continuous (see Figure 4.2 for a picture). In the rest of the proof, we will use this fact without further notification. For example, the above immediately gives that $h_1 \mapsto f_+(J_1, J_2, h_1)$ is continuous at a point $h_1 \in \mathbb{R}$ if $h_1 \mapsto t_+(J_1, h_1)$ is so.

Proof of Proposition 4.1.3. We will only prove part $a)$ and $c)$. The proof of part $b)$ follows the same type of argument as the proof of part $a)$.

To prove part a), we start to argue that for given $J_1 > 0$ the map $h_1 \mapsto t_+(J_1, h_1)$ is right-continuous at every point $h_1 \in \mathbb{R}$. To see that, take a sequence of reals $\{h_n\}$ such that $h_n \downarrow h_1$ as $n \to \infty$ and note that since the map $h_1 \mapsto t_+(J_1, h_1)$ is increasing, the sequence $\{t_+(J_1, h_n)\}$ converges to a limit $\tilde{t}$ with $\tilde{t} \geq t_+(J_1, h_1)$. Moreover, by taking the limit in the fixed point equation we see that

$$\tilde{t} = h_1 + d\phi_{J_1}(\tilde{t}) \tag{12}$$

and since $t_+(J_1, h_1)$ is the largest number satisfying (12) we get $\tilde{t} = t_+(J_1, h_1)$.

Next, assume $h_1 \neq -h^*(J_1)$ and $h_n \uparrow h_1$ as $n \to \infty$. As before, the limit of $\{t_+(J_1, h_n)\}$ exists, denote it by T. The number T will again satisfy (12). By considering different cases described in Figure 4.5, we easily conclude that $T = t_+(J_1, h_1)$. Hence, the function $h_1 \mapsto t_+(J_1, h_1)$ is continuous for all $h_1 \neq -h^*(J)$ and so we get that $h_1 \mapsto f_+(J_1, J_2, h_1)$ is also continuous for all $h_1 \neq -h^*(J_1)$.

Now assume $h_1 = -h^*(J_1)$. By considering sequences $h_n \downarrow -h^*(J_1)$ and $h_n \uparrow -h^*(J_1)$ we can similarly as above see that

$$\tau_+(J_1, J_2, -h^*(J_1)+) := \lim_{h \downarrow -h^*(J_1)} \tau_+(J_1, J_2, h) = t_+(J_1, -h^*(J_1)) + |J_1 - J_2|$$
$$\tau_+(J_1, J_2, -h^*(J_1)-) := \lim_{h \uparrow -h^*(J_1)} \tau_+(J_1, J_2, h) = t_-(J_1, -h^*(J_1)) + |J_1 - J_2|$$

and so

$$\tau_+(J_1, J_2, -h^*(J_1)+) = \tau_+(J_1, J_2, -h^*(J_1)-) \quad \iff \quad h^*(J_1) = 0.$$

Since $h^*(J_1) = 0$ if and only if $0 < J_1 \leq J_c$ the continuity of $h_1 \mapsto f_+(J_1, J_2, h_1)$ at $-h^*(J_1)$ follows at once in that case. If $J_1 = J_2$, then

$$\tau_+(J_1, J_2, -h^*(J_1)+) = t_+(J_2, -h^*(J_2))$$
$$\tau_+(J_1, J_2, -h^*(J_1)-) = t_-(J_2, -h^*(J_2))$$

and since

$$\psi(J_2, t_+(J_2, -h^*(J_2))) = \psi(J_2, t_-(J_2, -h^*(J_2))),$$

the continuity is clear also in that case. If $J_1 > J_c$ and $0 < J_2 \leq J_c$, then

$$\tau_+(J_1, J_2, -h^*(J_1)+) \neq \tau_+(J_1, J_2, -h^*(J_1)-)$$

and the map $t \mapsto \psi(J_2, t)$ becomes strictly increasing, hence $h_1 \mapsto f_+(J_1, J_2, h_1)$ is discontinuous at $-h^*(J_1)$. For the case when $J_1 > J_c$, $J_2 > J_c$, $J_1 \neq J_2$ just note that $h_1 \mapsto f_+(J_1, J_2, h_1)$ is continuous at $-h^*(J_1)$ if and only if a and b (defined in the statement of the proposition) are in the flat region in the upper graph of Figure 4.2.

To prove part c) we take a closer look at the map $(J_2, t) \mapsto \psi(J_2, t)$. By definition, this map is

$$\psi(J_2, t) = \begin{cases} -h^*(J_2) & \text{if} \quad t_-(J_2, -h^*(J_2)) \leq t < t^*(J_2) \\ t - d\,\phi_{J_2}(t) & \text{if} \quad t \geq t^*(J_2) \text{ or } t < t_-(J_2, -h^*(J_2)). \end{cases}$$

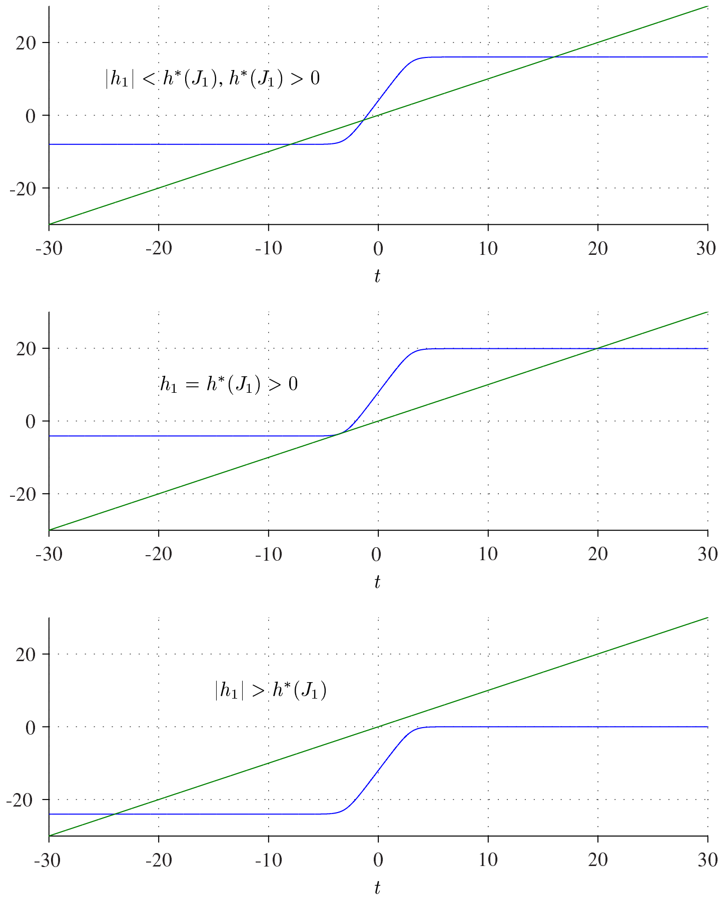

Figure 4.5: *A picture of the different cases in the fixed point equation that can occur when $h_1 \neq -h^*(J_1)$. Here, $d = 4$ and $J_1 = 3$.*

From the continuity of $t \mapsto \psi(J_2, t)$ for fixed J_2 and the facts that $J_2 \mapsto t^*(J_2)$, $J_2 \mapsto t_-(J_2, -h^*(J_2))$, $J_2 \mapsto -h^*(J_2)$ and $(J_2, t) \mapsto t - d\phi_{J_2}(t)$ are all continuous, we get that ψ is (jointly) continuous and so the result follows. $\square$

4.2.4 Proof of Proposition 4.1.4

To prove the statement, we will show that the inequality

$$\frac{\partial}{\partial J} t_+(J, h) \geq 1 \tag{13}$$

holds if $a)$ $h \geq 0$ and $J \geq J_c$ or $b)$ $h < 0$ and $h^*(J) > -h$. By integrating equation (13) the statement follows. The proof of equation (13) will be an easy modification of the proof of Lemma 5.2 in [13]. The proof is quite short and so we give a full proof here, even though it is more or less the same as the proof in [13].

Write $\phi(J,t)$ for $\phi_J(t)$ and use subscripts to denote partial derivatives. By differentiating the relation

$$h + d\,\phi(J, t_+(J,h)) = t_+(J,h)$$

with respect to J and solving, we get

$$\frac{\partial}{\partial J} t_+(J,h) = \frac{d\,\phi_1(J, t_+(J,h))}{1 - d\,\phi_2(J, t_+(J,h))}.$$

To get the left hand side bigger or equal to one, we need

$$d\,\phi_2(J, t_+(J,h)) < 1 \tag{14}$$

and

$$\phi_1(J, t_+(J,h)) + \phi_2(J, t_+(J,h)) \geq \frac{1}{d}. \tag{15}$$

The first inequality is immediate since in the cases $a)$ and $b)$ above, the function $t \mapsto h + d\,\phi(J,t)$ crosses the line $t \mapsto t$ from above to below. For (15), note that

$$\begin{aligned} \phi_1(J,t) &= \frac{1}{2}\big(\tanh(J+t) - \tanh(J-t)\big) \\ \phi_2(J,t) &= \frac{1}{2}\big(\tanh(J+t) + \tanh(J-t)\big) \end{aligned}$$

and so

$$\phi_1(J,t) + \phi_2(J,t) = \tanh(J+t),$$

which yields that $\phi_1 + \phi_2$ is increasing in both variables. Moreover, since $\tanh(J_c) = \frac{1}{d}$ (see [2]), we get

$$\phi_1(J_c, 0) + \phi_2(J_c, 0) = \frac{1}{d}$$

and so

$$\phi_1(J,t) + \phi_2(J,t) \geq \frac{1}{d} \quad \text{if } J \geq J_c,\, t \geq 0. \tag{16}$$

To complete the proof, observe that in the cases $a)$ and $b)$, we have $J \geq J_c$ and $t_+(J,h) \geq 0$. $\square$

4.2.5 Proof of Proposition 4.1.7

In the proof we will use the following results from [13] concerning domination of product measures.

Definition 4.2.1 (Downward FKG, Liggett, Steif). Given a finite or countable set V, a measure μ on $\{-1,1\}^V$ is called downward FKG if for any finite $A \subseteq V$, the conditional measure $\mu(\,\cdot\,|\,\eta \equiv 0$ on $A\,)$ has positive correlations.

Here, as usual, positive correlations is defined as follows:

Definition 4.2.2 (Positive correlations). A probability measure μ on $\{-1,1\}^V$ where V is a finite or countable set is said to have positive correlations if

$$\int fg\,d\mu \geq \int f\,d\mu \int g\,d\mu$$

for all real-valued, continuous and increasing functions f, g on $\{-1,1\}^V$.

Theorem 4.2.3 (Liggett, Steif). *Let μ be a translation invariant measure on $\{-1,1\}^{\mathbb{Z}^d}$ which also is downward FKG and let $p \in [0,1]$. Then the following are equivalent:*

a) $\mu \geq \gamma_p$.

b) $\limsup_{n\to\infty} \mu(\,\eta \equiv -1 \text{ on } [1,n]^d\,)^{1/n^d} \leq 1-p$.

Remarks:

(i) In particular, Theorem 4.2.3 gives us that if two translation invariant, downward FKG measures have the same above limsup, then they dominate the same set of product measures.

(ii) In [13], it is a third condition in Theorem 4.2.3 which we will not use and so we simply omit it.

Before we state the next lemma we need to recall the following definition.

Definition 4.2.3 (FKG lattice condition). Suppose V is a finite set and let μ be a probability measure on $\{-1,1\}^V$ which assigns positive probabilty to each element. For $\eta, \xi \in \{-1,1\}^V$ define $\eta \vee \xi$ and $\eta \wedge \xi$ by

$$(\eta \vee \xi)(x) = \max(\eta(x), \xi(x)),\ (\eta \wedge \xi)(x) = \min(\eta(x), \xi(x)),\ x \in V.$$

We say that μ satisfies the FKG lattice condition if

$$\mu(\eta \wedge \xi)\mu(\eta \vee \xi) \geq \mu(\eta)\mu(\xi)$$

for all $\eta, \xi \in \{-1,1\}^V$

Given a measure μ on $\{-1,1\}^{\mathbb{Z}^d}$ we will denote its projection on $\{-1,1\}^T$ for finite $T \subseteq \mathbb{Z}^d$ by μ_T.

Lemma 4.2.4. *The measures $\nu^{\mathbb{Z}^d,\pm}_{q,J,r}$ are FKG in the sense that $\nu^{\mathbb{Z}^d,\pm}_{T,q,J,r}$ satisfies the FKG lattice condtion for each finite $T \subseteq \mathbb{Z}^d$.*

Proof. For $n \geq 2$, let $\Lambda_n = \{-n,\dots,n\}^d$ and denote the finite volume Potts measures on $\{-1,1\}^{\Lambda_n}$ with boundary condition $\eta \equiv 1$ and $\eta \equiv q$ by $\pi^{n,1}_{q,J}$ and $\pi^{n,q}_{q,J}$. Furthermore, let $\nu^{n,-}_{q,J,r}$ and $\nu^{n,+}_{q,J,r}$ denote the corresponding fuzzy Potts measures. Given the convergence in the Potts model, it is clear that $\nu^{n,\pm}_{T,q,J,r}$ converges weakly to $\nu^{\mathbb{Z}^d,\pm}_{T,q,J,r}$ as $n \to \infty$ for each finite $T \subseteq \mathbb{Z}^d$. Since the FKG lattice condition is closed under taking projections (see [4, p. 28]) and weak limits we are done if we can show that $\nu^{n,\pm}_{q,J,r}$ satisfies the FKG lattice condition for each $n \geq 2$. In [6] it is proved that for an arbitrary finite graph $G = (V, E)$ the finite volume fuzzy Potts measure with free boundary condition and parameters q, J, r is monotone in the sense that

$$\nu^G_{q,J,r}(Y(x) = 1 \,|\, Y(V \setminus \{x\}) = \eta) \leq \nu^G_{q,J,r}(Y(x) = 1 \,|\, Y(V \setminus \{x\}) = \eta') \tag{17}$$

for all $x \in V$ and η, $\eta' \in \{-1,1\}^{V\setminus\{x\}}$ with $\eta \leq \eta'$. We claim that it is possible to modify the argument given there to prove that $\nu^{n,\pm}_{q,J,r}$ are monotone for each $n \geq 2$. (Recall from [4] the fact that if V is finite and μ is a probabilty measure on $\{-1,1\}^V$ that assigns positive probabilty to each element, then monotone is equivalent to the FKG lattice condition.) The proof of (17) is quite involved. However, the changes needed to prove our claim are quite straightforward and so we will only give an outline for how that can be done. Furthermore, we will only consider the minus case, the plus case is similar.

By considering a sequence $\eta = \eta_1 \leq \eta_2 \leq \cdots \leq \eta_m = \eta'$ where η_i and η_{i+1} differ only at a single vertex, it is easy to see that it is enough to prove that for all x, $y \in \Lambda_n$ and $\eta \in \{-1,1\}^{\Lambda_n\setminus\{x,y\}}$ we have

$$\begin{aligned} &\nu^{n,-}_{q,J,r}(Y(x) = 1, Y(y) = 1 \,|\, Y(\Lambda_n \setminus \{x,y\}) = \eta) \\ &\quad\geq \nu^{n,-}_{q,J,r}(Y(x) = 1 \,|\, Y(\Lambda_n \setminus \{x,y\}) = \eta) \\ &\quad\quad\cdot \nu^{n,-}_{q,J,r}(Y(y) = 1 \,|\, Y(\Lambda_n \setminus \{x,y\}) = \eta). \end{aligned} \tag{18}$$

Fix $n \geq 2$, x, y and η as above. We will first consider the case when x and y are not neighbors. At the end we will see how to modify the argument to work when x, y are neighbors as well. Define $V_- = \{z \in \Lambda_n \setminus \{x,y\} : \eta(z) = -1\}$ and $V_+ = \{z \in \Lambda_n \setminus \{x,y\} : \eta(z) = 1\}$. Furthermore, denote by E_n the set of edges $\langle u,v\rangle$ with either u, $v \in \Lambda_n$ or $u \in \Lambda_n$, $v \in \partial\Lambda_n$ and let $\mathbf{P}$ denote the probability measure on $W = \{1,\dots,q\}^{\Lambda_n\cup\partial\Lambda_n} \times \{0,1\}^{E_n}$ which to each site $u \in \Lambda_n \cup \partial\Lambda_n$ chooses a spin value uniformly from $\{1,\dots,q\}$, to each edge $\langle u,v\rangle$ assigns value 1 or 0 with probabilities p and $1-p$ respectively and which does those things independently for all sites and edges. Define the following events on W

$$\begin{aligned}
A &= \{(\sigma,\xi) : (\sigma(u)-\sigma(v))\xi(e) = 0, \forall e = \langle u,v\rangle \in E_n\},\\
B &= \{(\sigma,\xi) : \sigma(z) \in \{1,\dots,r\}\, \forall z \in V_-, \sigma(z) \in \{r+1,\dots,q\}\, \forall z \in V_+\},\\
C &= \{(\sigma,\xi) : \sigma(z) = 1, \forall z \in \partial\Lambda_n\},
\end{aligned}$$

and let $\mathbf{P}'$ and $\mathbf{P}''$ be the probability measures on $\{1,\dots,q\}^{\Lambda_n} \times \{0,1\}^{E_n}$ obtained from $\mathbf{P}$ by conditioning on $A \cap C$ and $A \cap B \cap C$ respectively. ($\mathbf{P}'$ is usually referred to as the Edward-Sokal coupling, see [3].) It is well known (and easy to check) that the spin marginal of $\mathbf{P}'$ is $\pi^{n,1}_{q,J}$ and that the edge marginal is the so called random-cluster measure defined as the probability measure on $\{0,1\}^{E_n}$ which to each $\xi \in \{0,1\}^{E_n}$ assigns probability proportional to

$$q^{k(\xi)} \prod_{e\in E_n} p^{\xi(e)}(1-p)^{1-\xi(e)},$$

where $k(\xi)$ is the number of connected components in ξ not reaching $\partial\Lambda_n$. In a similar way it is possible (by counting) to compute the spin and edge marginal of $\mathbf{P}''$: The spin marginal π'' is simply $\pi^{n,1}_{q,J}$ conditioned on B and the edge marginal ϕ'' assigns probability to a configuration $\xi \in \{0,1\}^{E_n}$ proportional to

$$1_D r^{k_0(\xi)}(q-r)^{k_1(\xi)} q^{k_x(\xi)+k_y(\xi)} \prod_{e\in E_n} p^{\xi(e)}(1-p)^{1-\xi(e)},$$

where $k_0(\xi)$ is the number of clusters intersecting V_- but not reaching $\partial\Lambda_n$, $k_1(\xi)$ is the number of clusters intersecting V_+, $k_x(\xi)$ (resp $k_y(\xi)$) is 1 if x (resp y) is in a singleton connected component and 0 otherwise and D is the event that no connected component in ξ intersects both V_- and V_+. Observe that (18) is the same as

$$\begin{aligned}
&\pi''(X(x) \in \{r+1,\dots,q\}, X(y) \in \{r+1,\dots,q\})\\
&\quad \geq \pi''(X(x) \in \{r+1,\dots,q\})\,\pi''(X(y) \in \{r+1,\dots,q\}).
\end{aligned} \tag{19}$$

An important feature of the coupling $\mathbf{P}''$ is that it gives a way to obtain a spin configuration $X \in \{1,\dots,q\}^{\Lambda_n}$ distributed as π'':

1. Pick an edge configuration ξ according to ϕ''.

2. Assign $X = 1$ to the connected components of ξ that intersect $\partial\Lambda_n$ and denote the union of those components by $\tilde{C}$.

3. Assign independently spins to a connected component $C \neq \tilde{C}$ of ξ where the spin is taken according to the uniform distribution on

$$\begin{array}{ll}
\{1,\dots,r\} & \text{if } C \text{ intersects } V_-,\\
\{r+1,\dots,q\} & \text{if } C \text{ intersects } V_+,\\
\{1,\dots,q\} & \text{if } C \text{ is a singleton vertex } x \text{ or } y.
\end{array}$$

By defining the functions $f_x, f_y : \{0,1\}^{E_n} \to \mathbb{R}$ as

$$f_x(\xi) = \begin{cases} 0, \text{ if } C_x = \tilde{C} \text{ or } C_x \text{ intersects } V_-, \\ \frac{q-r}{q}, \text{ if } C_x \text{ is a singleton}, \\ 1, \text{ otherwise}, \end{cases}$$

where C_x is the connected component of ξ containing x (f_y defined analogously), we see as in [6] that (19) follows if

$$\int f_x f_y \, d\phi'' \geq \int f_x \, d\phi'' \int f_y \, d\phi''. \tag{20}$$

The significance of f_x and f_y is that $f_x(\xi)$ is the conditional probability that $X(x) \in \{r+1, \ldots, q\}$ given ξ and similarly for f_y, and that the events $X(x) \in \{r+1, \ldots, q\}$ and $X(y) \in \{r+1, \ldots, q\}$ are conditionally independent given ξ. With all this setup done it is a simple task to see that to prove (20) we can proceed exactly as in [6, p. 1154-1155].

To take care of the case when x and y are neighbors, observe that everything we have done so far also works for the graph with one edge deleted, i.e. the graph with vertex set Λ_n and edge set $E_n \setminus \{\langle x, y\rangle\}$. Hence we can get (19) for that graph. However the observation in [6, 1156] gives us (19) even in the case when we reinsert the edge $\langle x, y\rangle$. □

Proof of Proposition 4.1.7. Let $k, l \in \{0, -, +\}$ be given and let $A_n = [1, n]^d$, $n \geq 2$. We are done if there exists $0 < c < 1$ (independent of k, l and n) such that

$$\nu_{q,J,r}^{\mathbb{Z}^d,k}(\,\eta \equiv -1 \text{ on } A_n\,) \geq c^{|\partial A_n|} \nu_{q,J,r}^{\mathbb{Z}^d,l}(\,\eta \equiv -1 \text{ on } A_n\,) \text{ for all } n.$$

As for the Ising model, it is well known that the infinite volume Potts measures satisfy the so called uniform nonnull property (sometimes called uniform finite energy property), which means that for some $c > 0$, the conditional probability of having a certain spin at a given site given everything else is at least c. (See for example [8] for a more precise definition.) We get for arbitrary $\sigma \in \{1, \ldots, q\}^{\partial A_n}$

$$\nu_{q,J,r}^{\mathbb{Z}^d,k}(\,\eta \equiv -1 \text{ on } [1,n]^d\,) \geq c^{|\partial A_n|} \pi_{q,J}^{A_n,\sigma}(\,Y \equiv -1 \text{ on } A_n\,). \tag{21}$$

Since $\nu_{q,J,r}^{\mathbb{Z}^d,l}(\,\eta \equiv -1 \text{ on } [1,n]^d\,)$ can be written as a convex combination of the terms in the far right side of (21) the result follows at once. □

4.2.6 Proof of Proposition 4.1.8

Let ρ denote the root of $\mathbb{T}^d$ and let V_n be the set of all sites in $\mathbb{T}^d$ with distance at most n from ρ. If x is on the unique self-avoiding path from ρ to y, we say that y is a descendant of x. Given $x \in \mathbb{T}^d$, let S_x denote the set of vertices of all descendants of x (including x). Moreover, let T_x denote the subtree of $\mathbb{T}^d$ whose vertex set is S_x and edge set consisting of all edges $\langle u, v\rangle \in E(\mathbb{T}^d)$ with $u, v \in S_x$. In the proof of Proposition 4.1.8, we will use the following Lemma from [13]:

Proposition 4.2.5 (Liggett, Steif). *Let $p \in [0,1]$, $\{ P(i,j) : i,j \in \{ -1,1\} \}$ be a transition matrix for an irreducible 2-state Markov chain with $P(-1,1) \leq P(1,1)$ and let μ be the distribution of the corresponding completely homogeneous Markov chain on $\mathbb{T}^d$. Then the following are equivalent:*

a) $\mu \geq \gamma_p$.

b) $\limsup_{n\to\infty} \mu(\, \eta \equiv -1 \text{ on } V_n\,)^{1/|V_n|} \leq 1 - p$.

c) $P(-1,1) \geq p$.

Proof of Proposition 4.1.8. Fix $J > 0$, $q \geq 3$ and $r \in \{1, \ldots, q-1\}$ with $e^{2J} \geq q-2$. In [9], it is proved that $\nu^{\mathbb{T}^d,0}_{q,J,r}$ is a completely homogeneous Markov chain on $\mathbb{T}^d$ for all values of the parameters with transition matrix

$$\begin{pmatrix} \frac{e^{2J}+r-1}{e^{2J}+q-1} & \frac{q-r}{e^{2J}+q-1} \\ \frac{r}{e^{2J}+q-1} & \frac{e^{2J}+q-r-1}{e^{2J}+q-1} \end{pmatrix}.$$

Hence, from Proposition 4.2.5 we get that $\nu^{\mathbb{T}^d,0}_{q,J,r} \geq \gamma_p$ if and only if

$$p \leq \frac{q-r}{e^{2J}+q-1}. \tag{22}$$

Furthermore, in [9, p. 10] the authors also derive the transition matrix for $\pi^{\mathbb{T}^d,1}_{q,J}$ from which we can compute the following:

$$\begin{aligned}
\nu^{\mathbb{T}^d,-}_{q,J,r}(\, \eta \equiv -1 \text{ on } V_n\,) &\geq \sum_{i=1}^{r} \pi^{\mathbb{T}^d,1}_{q,J}(\, X \equiv i \text{ on } V_n\,) \\
&= \frac{b}{b+q-1}\left(\frac{ce^{2J}}{ce^{2J}+q-1}\right)^{|V_n|-1} \\
&\quad + \frac{r-1}{b+q-1}\left(\frac{e^{2J}}{c+e^{2J}+q-2}\right)^{|V_n|-1}
\end{aligned}$$

where

$$b = \frac{\pi^{\mathbb{T}^d,1}_{q,J}(\, X(\rho) = 1\,)}{\pi^{\mathbb{T}^d,1}_{q,J}(\, X(\rho) = 2\,)}$$

$$c = \frac{\pi^{T_x,1}_{q,J}(\, X(x) = 1\,)}{\pi^{T_x,1}_{q,J}(\, X(x) = 2\,)}, \quad x \neq \rho.$$

(Of course, homogeneity gives that the last quotient is independent of x.) We get that

$$\limsup_{n\to\infty} \nu^{\mathbb{T}^d,-}_{q,J,r}(\,\eta \equiv -1 \text{ on } V_n\,)^{1/|V_n|} \geq \frac{ce^{2J}}{ce^{2J}+q-1} + \frac{e^{2J}}{c+e^{2J}+q-2}. \tag{23}$$

Now, assume that the underlying Gibbs measures for the Potts model satisfy $\pi^{\mathbb{T}^d,1}_{q,J} \neq \pi^{\mathbb{T}^d,0}_{q,J}$. It is known [1] that this is equivalent to having

$$\pi^{\mathbb{T}^d,1}_{q,J}(\,X(x)=1\,) > \frac{1}{q}, \quad \forall x \in \mathbb{T}^d.$$

In [9], the authors observed that if $a = \pi^{\mathbb{T}^d,1}_{q,J}(\,X(\rho)=1\,)$, then from symmetry reasons

$$b = \frac{(q-1)a}{1-a}.$$

Hence, if $a > \frac{1}{q}$ we get $b > 1$. Moreover, from the recursion formula in [9, p. 9] we obtain

$$b = \frac{(ce^{2J}+q-1)^{d+1}}{(c+e^{2J}+q-2)^{d+1}}. \tag{24}$$

It is easy to see from (24) that if $b > 1$ then $c > 1$. Hence, we can choose $p \in (0,1)$ such that

$$\frac{q-r}{ce^{2J}+q-1} < p \leq \frac{q-r}{e^{2J}+q-1}. \tag{25}$$

Moreover, an easy calculation gives us that

$$\frac{ce^{2J}}{ce^{2J}+q-1} + \frac{e^{2J}}{c+e^{2J}+q-2} \geq \frac{ce^{2J}+q-2}{ce^{2J}+q-1}$$

and since

$$1-p < \frac{ce^{2J}+r-1}{ce^{2J}+q-1} \leq \frac{ce^{2J}+q-2}{ce^{2J}+q-1}$$

we get from (23)

$$\limsup_{n\to\infty} \nu^{\mathbb{T}^d,-}_{q,J,r}(\,\eta \equiv -1 \text{ on } V_n\,)^{1/|V_n|} > 1-p.$$

It is now clear that for p as in (25) we have that $\nu^{\mathbb{T}^d,0}_{q,J,r}$ dominates γ_p but $\nu^{\mathbb{T}^d,-}_{q,J,r}$ does not dominate γ_p. $\square$

Remark: By deriving the transition matrix for $\pi^{\mathbb{T}^d,q}_{q,J}$ it is probably possible to prove that there exists $p \in (0,1)$ such that $\nu^{\mathbb{T}^d,0}_{q,J,r}$ dominates γ_p but $\nu^{\mathbb{T}^d,+}_{q,J,r}$ does not dominate γ_p.

4.3 Conjectures

We end with the following conjectures concerning the fuzzy Potts model. The corresponding statements for the Ising model are proved in [13].

Conjecture 4.3.1. *Let $q \geq 3$, $r \in \{1, \ldots, q-1\}$ and consider the fuzzy Potts model on $\mathbb{Z}^d$. If J_1, $J_2 > 0$ with $J_1 \neq J_2$, then $\nu^{\mathbb{Z}^d,+}_{q,J_1,r}$ and $\nu^{\mathbb{Z}^d,+}_{q,J_2,r}$ are not stochastically ordered.*

Conjecture 4.3.2. *Let $q \geq 3$, $r \in \{1, \ldots, q-1\}$ and consider the fuzzy Potts model on $\mathbb{Z}^d$. If $0 < J_1 < J_2$, then*

$$\sup\{\, p \in [0,1] : \nu^{\mathbb{Z}^d,+}_{q,J_1,r} \geq \gamma_p \,\} > \sup\{\, p \in [0,1] : \nu^{\mathbb{Z}^d,+}_{q,J_2,r} \geq \gamma_p \,\}.$$

Conjecture 4.3.3. *Let $J > 0$, $q \geq 3$, $r \in \{1, \ldots, q-1\}$ and consider the fuzzy Potts model on $\mathbb{T}^d$. Define the sets:*

$$\begin{aligned} D_+ &= \{\, p \in [0,1] : \nu^{\mathbb{Z}^d,+}_{q,J,r} \geq \gamma_p \,\}, \\ D_- &= \{\, p \in [0,1] : \nu^{\mathbb{Z}^d,-}_{q,J,r} \geq \gamma_p \,\}, \\ D_0 &= \{\, p \in [0,1] : \nu^{\mathbb{Z}^d,0}_{q,J,r} \geq \gamma_p \,\}, \end{aligned} \tag{26}$$

If the underlying Gibbs measures for the Potts model satisfy $\pi^{\mathbb{T}^d,1}_{q,J} \neq \pi^{\mathbb{T}^d,0}_{q,J}$, then the sets in (26) *are all different from each other.*

Conjecture 4.3.4. *Let $q \geq 3$, $r \in \{1, \ldots, q-1\}$ and consider the fuzzy Potts model on $\mathbb{T}^d$. Denote the critical value corresponding to non-uniqueness of Gibbs states for the Potts model by J_c. If $J_c < J_1 < J_2$ then $\nu^{\mathbb{T}^d,+}_{q,J_1,r} \leq \nu^{\mathbb{T}^d,+}_{q,J_2,r}$.*

Remark: If $J_1 < J_2 < J_c$, then

$$\nu^{\mathbb{T}^d,+}_{q,J_1,r}(\,\eta(x) = 1\,) = \nu^{\mathbb{T}^d,+}_{q,J_2,r}(\,\eta(x) = 1\,) = \frac{q-r}{q}$$

and so in that case, $\nu^{\mathbb{T}^d,+}_{q,J_1,r}$ and $\nu^{\mathbb{T}^d,+}_{q,J_2,r}$ can not be stochastically ordered.

Acknowledgement

The author wants to thank Jeffrey Steif for presenting the problems, a careful reading of the manuscript and for valuable comments and discussions.

References

[1] M. Aizenman, J. T. Chayes, L. Chayes, and C. M. Newman, *Discontinuity of the magnetization in one-dimensional* $1/|x-y|^2$ *Ising and Potts models*, J. Statist. Phys. **50** (1988), 1–40.

[2] H. O. Georgii, *Gibbs Measures and Phase Transitions*, Walter de Gruyter, 1988.

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

[4] G. Grimmett, *The random-cluster model*, Springer, 2006.

[5] O. Häggström, *Random-cluster representations in the study of phase transitions*, Markov Processes and Related Fields **4** (1998), 275–321.

[6] ______, *Positive correlations in the fuzzy Potts model*, Ann. Appl. Probab. **9** (1999), 1149–1159.

[7] ______, *Markov random fields and percolation on general graphs*, Adv. Appl. Prob. **32** (2000), 39–66.

[8] ______, *Is the fuzzy Potts model gibbsian?*, Ann. Inst. H. Poincaré, Probab. Statist **39** (2003), 891–917.

[9] O. Häggström and C. Külske, *Gibbs properties of the fuzzy Potts model on trees and in mean field*, Markov Processes and Related Fields **10** (2004), 477–506.

[10] J. Kahn and N. Weininger, *Positive association in the fractional fuzzy Potts model*, Ann. Probab. **35** (2007), 2038–2043.

[11] H. Künsch, S. Geman, and A. Kehagias, *Hidden Markov random fields*, Ann. Appl. Probab. **5** (1995), 577–602.

[12] T. M. Liggett, *Interacting Particle Systems*, Springer, 1985.

[13] T. M. Liggett and J. E. Steif, *Stochastic domination: the contact process, Ising models and FKG measures*, Ann. Inst. H. Poincaré, Probab. Statist **42** (2006), 223–243.

[14] C. Maes and K. Vande Velde, *The fuzzy Potts model*, J. Phys. A **28** (1995), 4261–4270.

[15] R. H. Schonmann and N.I. Tanaka, *Lack of monotonicity in ferromagnetic Ising model phase diagrams*, Ann. Appl. Probab. **8** (1998), 234–245.

PAPER IV

Optimal closing of a pair trade with a model containing jumps

Stig Larsson, Carl Lindberg and Marcus Warfheimer

5
PAPER IV

ABSTRACT
A pair trade is a portfolio consisting of a long position in one asset and a short position in another, and it is a widely applied investment strategy in the financial industry. Recently, Ekström, Lindberg and Tysk studied the problem of optimally closing a pair trading strategy when the difference of the two assets is modelled by an Ornstein-Uhlenbeck process. In this paper we study the same problem, but the model is generalized to also include jumps. More precisely we assume that the above difference is an Ornstein-Uhlenbeck type process, driven by a Lévy process of finite activity. We prove a verification theorem and analyze a numerical method for the associated free boundary problem. We prove rigorous error estimates, which are used to draw some conclusions from numerical simulations.

Key words and phrases: Pairs trading, Ornstein-Uhlenbeck type process, finite element method, error estimate.
Subject classification : 91B28, 65N30, 45J05.

5.1 Introduction

A portfolio which consists of a positive position in one asset, and a negative position in another is called a pair trade. Pairs trading was developed at Morgan Stanley in the late 1980's, and today it is one of the most common investment strategies in the financial industry. The idea behind pairs trading is quite intuitive: the investor finds two assets, for which the prices have moved together historically. When the price spread widens, the investor takes a short position in the outperforming asset, and a long position in the underperforming one with the hope that the spread will converge again, generating a profit. A main advantage of pairs trading is that the short position can, in principle, remove any exposure to market risk. For a historical evaluation of pairs trading we refer to [6].

To model the pair spread the authors in [3] proposed a mean reverting Gaussian Markov chain which they considered to be observed in Gaussian noise. Recently, in [2] the authors suggested the continuous time analogue, the so called mean reverting Ornstein-Uhlenbeck process. In this paper we generalize the model of the spread to also include possible jumps. Let $(\Omega, \mathcal{F}, \mathbf{P})$ be a complete probability space where the following processes are defined in such a way that they are independent:

– A standard Brownian motion $W = \{W_t\}_{t\geq 0}$.

– A Possion process $N^\lambda = \{N_t^\lambda\}_{t\geq 0}$ with intensity $\lambda > 0$.

– A sequence of independent random variables $\{X_k^\varphi\}_{k=1}^\infty$ with common continuous symmetric density φ. Moreover, the support of φ is contained in the interval $(-J, J)$ for some $J > 0$.

Define the compound Poisson process $C^{\lambda,\varphi} = \{C_t^{\lambda,\varphi}\}_{t\geq 0}$ in the usual way as

$$C_t^{\lambda,\varphi} = \sum_{k=1}^{N_t^\lambda} X_k^\varphi$$

and denote the filtration generated by W, $C^{\lambda,\varphi}$ and the null sets of $\mathcal{F}$ by $\mathbb{F} = \{\mathcal{F}_t\}_{t\geq 0}$. It is well known that this filtration satisfies the usual hypotheses (see for example [10]). From now on, when we say that a process is a martingale, submartingale or supermartingale we mean that this is with respect to $\mathbb{F}$.

Let the difference $U = \{U_t\}_{t\geq 0}$ between the assets be the unique solution of the stochastic differential equation

$$dU_t = -\mu U_t\, dt + \sigma\, dW_t + dC_t^{\lambda,\varphi}, \quad t > 0, \tag{1}$$

where $\mu > 0$, $\sigma > 0$. (The solution of equation (1) is usually called a generalized Ornstein-Uhlenbeck process or an Ornstein-Uhlenbeck type process.) Sometimes we will denote the driving Lévy process in (1) by $Z^{\sigma,\lambda,\varphi}$, i.e.

$$Z_t^{\sigma,\lambda,\varphi} = \sigma W_t + C_t^{\lambda,\varphi}, \quad t \geq 0.$$

As discussed in [2], there is a large risk associated with a pair trading strategy. Indeed, if the market spread ceases to be mean reverting, the investor is exposed to substantial risk. Therefore, in practice the investor typically chooses in advance a stop-loss level $a < 0$, which corresponds to the level of loss above which the investor will close the pair trade. For a given stop-loss level $a < 0$ define

$$\tau_a = \inf\{t \geq 0 : U_t \leq a\}, \tag{2}$$

the first hitting time of the region $(-\infty, a]$, and the so called value function

$$V(x) = \sup_{\tau} \mathbf{E}_x[U_{\tau_a \wedge \tau}] \quad x \in \mathbb{R}, \tag{3}$$

where the supremum is taken over all stopping times with respect to U. (Here and in the sequel $\mathbf{E}_x$ means expected value when $U_0 = x$.) The major interest here is to characterize V, and perhaps more importantly, to describe the stopping time where the supremum is attained. Since the drift has the opposite sign as U, we have no reason to liquidate our position as long as U is negative. On the other hand, if U is positive, then the drift is working against the investor and for large values of U the size of the drift should overcome the possible benefits from random variations. Moreover, since the jumps are assumed to be symmetric, this indicates that there is a stopping barrier $b > 0$ with the property that we should keep our position when $U_t < b$ and liquidate as soon as $U_t \geq b$. We note that we cannot be sure to close the pair trade at any of the boundaries a or b, because the spread can exhibit jumps. This was not the case in [2] and it is the major reason for the additional difficulties encountered in the present paper.

General optimal stopping theory (described for example in [9, Ch. 3]) leads us to believe that the value function is given by $V = u$, where (u, b) is the solution of the free boundary problem

$$\begin{aligned} \mathcal{G}_U u(x) &= 0, \quad x \in (a, b), \\ u(x) &= x, \quad x \notin (a, b), \\ u'(b) &= 1. \end{aligned} \tag{4}$$

Here $\mathcal{G}_U$ is the infinitesimal generator of U, which is defined on the space of twice continuously differentiable functions $f : \mathbb{R} \to \mathbb{R}$ with compact support:

$$\mathcal{G}_U f(x) = \frac{\sigma^2}{2} f''(x) - \mu x f'(x) + \lambda \int_{-\infty}^{\infty} (f(x+y) - f(x)) \varphi(y)\, dy, \quad x \in \mathbb{R}. \tag{5}$$

Moreover, the stopping time where the supremum in (3) is attained should be

$$\tau_b = \inf\{t \geq 0 : U_t \geq b\}. \tag{6}$$

Indeed, our first result is a so called verification theorem.

Theorem 5.1.1. *Assume that* (u, b) *is a classical solution of* (4) *with*

a) $\mathcal{G}_U u(x) \leq 0$, *for* $x > b$,

b) $u(x) \geq x$, *for* $x \in \mathbb{R}$.

Then $u(x) = V(x) = \mathbf{E}_x[U_{\tau_a \wedge \tau_b}]$, *for* $x \in \mathbb{R}$, *where* V *is given by* (3).

Remark: As seen from the assumptions on φ, we are assuming that the absolute value of the jumps of the process $\{U_t\}_{\geq 0}$ are bounded. The reason is that on the financial market, an asset cannot jump to arbitrarily large levels. If nothing else, the jumps are bounded by all the money in the world.

The rest of the paper is organized as follows. In Section 5.2 we prove Theorem 5.1.1 and in Section 5.3 we discuss a numerical solution of the free boundary problem (4). We also present strong evidence for the existence and uniqueness of a solution of (4).

5.2 Proof of Theorem 5.1.1

Before we start to prove Theorem 5.1.1 we need to recall some facts. From the general theory in [5] we get that the boundary value problem

$$\begin{aligned} \mathcal{G}_U u(x) &= 0, \quad x \in (a,b), \\ u(x) &= x, \quad x \notin (a,b), \end{aligned} \tag{7}$$

has a unique classical solution and that such a solution belongs to the space

$$C^2(\mathbb{R} \setminus \{a,b\}) \cap C^1(\mathbb{R} \setminus \{a,b\}) \cap C(\mathbb{R}).$$

Moreover, the finite left and right limits of u' and u'' exist at a and b. Although these facts follow from [5], we present in Theorem 5.3.1 a self-contained proof for the simpler situation that we consider here. Hence, if (u,b) is a classical solution of (4), then necessarily

$$u \in C^2(\mathbb{R} \setminus \{a,b\}) \cap C^1(\mathbb{R} \setminus \{a\}) \cap C(\mathbb{R})$$

with finite left and right limits of u' and u'' everywhere. Furthermore, recall a generalized version of Itô's formula for convex functions (see for example [10, Ch. 4]):

Theorem 5.2.1 (Meyer-Itô formula). *Let $X = \{X_t\}_{\geq 0}$ be a semimartingale and let f be the difference of two convex functions. Then*

$$\begin{aligned} f(X_t) = f(X_0) &+ \int_{0+}^t D^- f(X_{s-})\, dX_s \\ &+ \sum_{0<s\leq t} \big(f(X_s) - f(X_{s-}) - D^- f(X_{s-})\, \Delta X_s\big) \\ &+ \frac{1}{2}\int_{-\infty}^{\infty} L_t^y(X)\, d\mu(y), \end{aligned}$$

where $D^- f$ is the left derivative of f, μ is a signed measure which is the second generalized derivative of f and $\{L_t^a(X)\}_{t\geq 0}$ is the local time process of X at a.

Due to the regularity of u it can be written as a difference of two convex functions (see Problem 6.24 in [7, Ch. 3]). Moreover, the second derivative measure μ of u can be split into two parts $\mu = \mu_c + \mu_d$, where the continuous part μ_c is given by $d\mu_c = u''\,dx$ and the discrete part $\mu_d = \delta_a$ is a point mass at a. Here, $u''(x)$ denotes the second derivative of u at x except at the points a and b, where it denotes the right second derivative (which we know is finite). By Corollary 1 of the Meyer-Itô formula in [10], we can now write

$$\begin{aligned}\frac{1}{2}\int_{-\infty}^{\infty} L_t^y(U)\,d\mu(y) &= \frac{1}{2}\int_0^t u''(U_{s-})\,d[U,U]_s^c + \frac{1}{2}L_t^a(U)\big(u'(a+) - u'(a-)\big)\\ &= \frac{\sigma^2}{2}\int_0^t u''(U_{s-})\,ds + \frac{1}{2}L_t^a(U)\big(u'(a+) - u'(a-)\big),\end{aligned} \tag{8}$$

where $[U,U]^c$ denotes the continuous part of the quadratic variation $[U,U]$.

Furthermore, by using (1) and the compensated Poisson random measure

$$\tilde{N}_Z(dt,dy) = N_Z(dt,dy) - \lambda\,dt\,\varphi(y)\,dy,$$

where N_Z denotes the jump measure associated with $Z^{\sigma,\lambda,\varphi}$, we get

$$\begin{aligned}&\int_{0+}^t D^- u(U_{s-})\,dU_s + \sum_{0<s\le t}\big(u(U_s) - u(U_{s-}) - D^- u(U_{s-})\,\Delta U_s\big)\\ &\quad= -\mu\int_0^t U_{s-}D^- u(U_{s-})\,ds + \sigma\int_0^t D^- u(U_{s-})\,dW_s\\ &\qquad+ \int_{0+}^t\int_{\mathbb{R}}\big(u(U_{s-}+y) - u(U_{s-})\big)\,\tilde{N}_Z(ds,dy)\\ &\qquad+ \lambda\int_0^t\int_{\mathbb{R}}\big(u(U_{s-}+y) - u(U_{s-})\big)\varphi(y)\,dy\,ds.\end{aligned} \tag{9}$$

Summing up, we now have for $t \ge 0$

$$\begin{aligned}u(U_t) = u(U_0) &+ \int_0^t\Big(\frac{\sigma^2}{2}u''(U_{s-}) - \mu U_{s-}D^- u(U_{s-})\Big)\,ds\\ &+ \lambda\int_0^t\int_{\mathbb{R}}\big(u(U_{s-}+y) - u(U_{s-})\big)\varphi(y)\,dy\,ds\\ &+ \frac{1}{2}L_t^a(U)\big(u'(a+) - u'(a-)\big) + M_t,\end{aligned} \tag{10}$$

where

$$M_t = \sigma\int_0^t D^- u(U_{s-})\,dW_s + \int_{0+}^t\int_{\mathbb{R}}\big(u(U_{s-}+y) - u(U_{s-})\big)\,\tilde{N}_Z(ds,dy).$$

Since u is Lipschitz, has a bounded left derivative and since the jumps density φhas a finite swe get that $\{M_t\}_{t\ge 0}$ is a martingale.

Lemma 5.2.2. *Assume $a \in \mathbb{R}$ and $U_0 > a$. Then a.s. $L^a_{\tau_a \wedge t}(U) = 0$ for all $t \geq 0$.*

Proof. Fix $a \in \mathbb{R}$ and assume $U_0 > a$. Since the local time process $\{L^a_t\}_{t\geq 0}$ is continuous in t it is enough to prove that for fixed $t \geq 0$ we have $L^a_{\tau_a\wedge t}(U) = 0$ a.s. From [10, p. 217], we get that

$$\begin{aligned}\frac{1}{2}L^a_{\tau_a\wedge t}(U) = (U_{\tau_a\wedge t} - a)^- - \sum_{0<s\leq \tau_a\wedge t} 1_{\{U_{s-}>a\}}(U_s - a)^- \\ + \int_{0+}^{\tau_a\wedge t} 1_{\{U_{s-}\leq a\}}\, dU_s - \sum_{0<s\leq\tau_a\wedge t} 1_{\{U_{s-}\leq a\}}(U_s - a)^+.\end{aligned}$$

Futhermore, from the fact that $U_s > a$ for all $0 < s < \tau_a \wedge t$, we get that $U_{s-} \geq a$ for all $0 < s < \tau_a \wedge t$ and from the left continuity of U_{s-}, we can conclude that we also have $U_{\tau_a\wedge t-} \geq a$. From that and by splitting the integral and the sum, we obtain

$$\begin{aligned}\frac{1}{2}L^a_{\tau_a\wedge t}(U) &= 1_{\{U_{\tau_a\wedge t-}=a\}}(U_{\tau_a\wedge t} - a)^- + 1_{\{U_{\tau_a\wedge t-}=a\}}(U_{\tau_a\wedge t} - a) \\ &\quad - 1_{\{U_{\tau_a\wedge t-}=a\}}(U_{\tau_a\wedge t} - a)^+ + \int_{0+}^{\tau_a\wedge t-} 1_{\{U_{s-}=a\}}\, dU_s \\ &\quad - \sum_{0<s<\tau_a\wedge t} 1_{\{U_{s-}=a\}}(U_s - a)^+ \\ &= \int_{0+}^{\tau_a\wedge t-} 1_{\{U_{s-}=a\}}\, dU_s - \sum_{0<s<\tau_a\wedge t} 1_{\{U_{s-}=a\}}(U_s - a)^+.\end{aligned}$$

From the observation that if $U_{s-} = a$ for some $0 < s < \tau_a \wedge t$, then s is a jump time and the jump must be in the up direction, we conclude that the right hand side of the last expression is zero and so we are done. □

Remark: In a similar way one can show that, if $a < U_0 < b$, then

$$L^a_{\tau_a\wedge\tau_b\wedge t}(U) = 0 \text{ and } L^b_{\tau_a\wedge\tau_b\wedge t}(U) = 0 \text{ for } t \geq 0.$$

Proof of Theorem 5.1.1. Since $u(x) = V(x) = \mathbf{E}_x[U_{\tau_a\wedge\tau_b}] = x$, when $x \leq a$, we can assume that $x > a$. Define $Y_t = u(U_{\tau_a\wedge t})$, $t \geq 0$. By using (10), Lemma 5.2.2, the expression (5) for the generator of U, and (4), we get

$$\begin{aligned}Y_t &= u(x) - \int_0^{\tau_a\wedge t} \mu U_{s-} 1_{\{U_{s-}\geq b\}}\, ds \\ &\quad + \lambda \int_0^{\tau_a\wedge t}\int_{\mathbb{R}} \big(u(U_{s-} + y) - u(U_{s-})\big)\varphi(y) 1_{\{U_{s-}\geq b\}}\, dy\, ds + M_{\tau_a\wedge t}.\end{aligned} \tag{11}$$

Property $a)$ and the martingale property of $\{M_{\tau_a\wedge t}\}$ give that $\{Y_t\}_{t\geq 0}$ is a supermartingale. Furthermore, from property $b)$ we get that $Y_t \geq U_{\tau_a\wedge t}$, for $t \geq 0$, and since

$$U_{\tau_a\wedge t} \geq a - J, \quad t \geq 0, \tag{12}$$

we can apply the optional sampling theorem (see [7]) and obtain

$$\mathbf{E}_x[U_{\tau_a\wedge\tau}] \le \mathbf{E}_x[Y_\tau] \le \mathbf{E}_x[Y_0] = u(x),$$

where τ is an arbitrary stopping time with respect to U. Hence, $V(x) \le u(x)$ for $x > a$. In particular, if $x \ge b$ then $x \le V(x) \le u(x) = x$ and so $u(x) = V(x) = \mathbf{E}_x[U_{\tau_a\wedge\tau_b}]$ when $x \ge b$.

For the case when $a < x < b$, note that from (11) we get for $t \ge 0$ that

$$Y_{\tau_b\wedge t} = M_{\tau_a\wedge\tau_b\wedge t} + u(x)$$

and since

$$a - J \le Y_{\tau_b\wedge t} \le b + J, \quad t \ge 0,$$

the optional sampling theorem applies again and we obtain $u(x) = \mathbf{E}_x[Y_{\tau_b}]$. Finally, the fact that $Y_{\tau_b} = U_{\tau_a\wedge\tau_b}$ gives us $u(x) = \mathbf{E}_x[U_{\tau_a\wedge\tau_b}] \le V(x)$ and the proof is complete. □

5.3 Numerical solution of the equation (4)

We have not been able to give a rigorous proof of the existence and uniqueness of the solution (u, b) of the free boundary value problem (4). We therefore resort to a numerical solution by means of the finite element method. However, at the end of this section we will show that we have strong computational evidence for both existence and uniqueness for (4). In order to achieve this we first show rigorous existence and regularity results for the boundary value problem (7) and rigorous convergence estimates with explicit constants for the finite element approximation.

5.3.1 The boundary value problem

We begin by transforming the free boundary value problem (4) to a problem with homogeneous boundary values. Set $v(x) = u(x) - x$ and use $\int_{-\infty}^{\infty} y\varphi(y)\,dy = 0$ to get

$$\begin{aligned} &-\tfrac{1}{2}\sigma^2 v''(x) + \mu x v'(x) \\ &\qquad -\lambda\int_{-\infty}^{\infty}\big(v(x+y)-v(x)\big)\varphi(y)\,dy = -\mu x, \quad && x \in (a,b), \\ &\qquad\qquad v(x) = 0, && x \notin (a,b), \\ &\qquad\qquad v'(b) = 0. \end{aligned} \tag{13}$$

Introducing the operators

$$\begin{aligned} \mathcal{L}v(x) &= -\tfrac{1}{2}\sigma^2 v''(x) + \mu x v'(x), \\ \mathcal{I}v(x) &= \lambda\int_{-\infty}^{\infty}\big(v(x+y)-v(x)\big)\varphi(y)\,dy, \end{aligned}$$

our approach will be to first solve the boundary value problem

$$\begin{aligned} \mathcal{L}v - \mathcal{I}v &= f, \quad x \in (a,b), \\ v(x) &= 0, \quad x \notin (a,b), \end{aligned} \tag{14}$$

with $f(x) = -\mu x$, and then for fixed $a < 0$ find $b > a$ such that $v'(b) = 0$.

To solve (14) we follow a standard approach based on a weak formulation and Fredholm's alternative. We denote by $(\cdot,\cdot)$ and $\|\cdot\|$ the standard scalar product and norm in $L_2(a,b)$, and we denote by $H^k(a,b)$ and $H_0^1(a,b) = \{v \in H^1(a,b) : v(a) = v(b) = 0\}$ the standard Sobolev spaces. We denote the derivative $Dv = dv/dx$. We choose $v \mapsto \|Dv\|$ to be the norm in $H_0^1(a,b)$, which is equivalent to the standard H^1-norm. We extend functions $v \in L_2(a,b)$ by zero outside (a,b) so that $\mathcal{I}v$ is properly defined. We define bilinear forms

$$\begin{aligned} A_{\mathcal{L}}(u,v) &= \int_a^b \big(\tfrac{1}{2}\sigma^2 u'(x)v'(x) + \mu x u'(x)v(x)\big)\,dx, \quad u,v \in H_0^1(a,b), \\ A_{\mathcal{I}}(u,v) &= \int_a^b \mathcal{I}u(x)v(x)\,dx, \quad u,v \in L_2(a,b) \\ A(u,v) &= A_{\mathcal{L}}(u,v) - A_{\mathcal{I}}(u,v). \end{aligned} \tag{15}$$

Since $\int_{-\infty}^{\infty} \varphi(y)\,dy = 1$, $\varphi(-y) = \varphi(y)$, and $v(x) = 0$ for $x \notin (a,b)$, we also have

$$\mathcal{I}v(x) = \lambda \int_a^b \varphi(x-y)v(y)\,dy - \lambda v(x), \quad v \in L_2(a,b). \tag{16}$$

The convolution operator $\mathcal{I}_1 v(x) = \int_{-\infty}^{\infty} \varphi(x-y)v(y)\,dy$ is bounded in $L_2(a,b)$ with constant $c = \int_{-\infty}^{\infty} \varphi(y)\,dy = 1$ by Young's inequality. Hence,

$$\|\mathcal{I}v\| \le 2\lambda\|v\|, \quad v \in L_2(a,b), \tag{17}$$

$$\|D\mathcal{I}v\| \le 2\lambda\|Dv\|, \quad v \in H_0^1(a,b), \tag{18}$$

and

$$-A_{\mathcal{I}}(v,v) \ge \lambda\big(\|v\|^2 - \|\mathcal{I}_1 v\|\|v\|\big) \ge 0, \quad v \in L_2(a,b).$$

Hence,

$$\begin{aligned} |A(u,v)| &\le \tfrac{1}{2}\sigma^2\|Du\|\|Dv\| + \mu\max(|a|,|b|)\|Du\|\|v\| + 2\lambda\|u\|\|v\| \\ &\le c_1\|Du\|\|Dv\|, \quad u,v \in H_0^1(a,b), \\ c_1 &= \tfrac{1}{2}\sigma^2 + c_2(\mu\max(|a|,|b|) + 2\lambda c_2), \end{aligned}$$

where we also used Poincaré's inequality

$$\|v\| \le c_2\|Dv\|, \quad v \in H_0^1(a,b), \quad c_2 = (b-a)/\pi. \tag{19}$$

By integration by parts we obtain

$$A_{\mathcal{L}}(v,v) = \tfrac{1}{2}\sigma^2\|Dv\|^2 - \tfrac{1}{2}\mu\|v\|^2, \quad v \in H_0^1(a,b),$$

so that $A(\cdot,\cdot)$ is bounded and coercive on $H_0^1(a,b)$:

$$|A(u,v)| \le c_1\|Du\|\|Dv\|, \qquad u,v \in H_0^1(a,b), \tag{20}$$

$$A(v,v) \ge \tfrac{1}{2}\sigma^2\|Dv\|^2 - \tfrac{1}{2}\mu\|v\|^2, \qquad v \in H_0^1(a,b). \tag{21}$$

We say that $v \in H_0^1(a,b)$ is a weak solution of (14) if

$$A(v,\phi) = (f,\phi) \quad \forall \phi \in H_0^1(a,b). \tag{22}$$

We also use the adjoint problem: find $w \in H_0^1(a,b)$ such that

$$A(\phi,w) = (\phi,g) \quad \forall \phi \in H_0^1(a,b). \tag{23}$$

The strong form is (note that $\mathcal{I}$ is self-adjoint in $L_2(a,b)$)

$$\begin{aligned} \mathcal{L}^* w(x) - \mathcal{I}w(x) &= g(x), \quad && x \in (a,b), \\ w(x) &= 0, && x \notin (a,b), \end{aligned} \tag{24}$$

where

$$\mathcal{L}^* w(x) = -\tfrac{1}{2}\sigma^2 w''(x) - \mu x w'(x) - \mu w(x).$$

We may now prove the existence and uniqueness of a classical solution of (14). In principle this follows from the general theory in [5], but we present a self-contained proof, with explicit constants, for the simpler situation that we consider here. The theorem also provides results necessary for the analysis of the finite element method.

Theorem 5.3.1. *The boundary value problem* (14) *has a unique weak solution $v \in H_0^1(a,b)$ for every $f \in L_2(a,b)$. The solution belongs to $H^2(a,b)$ and there is a constant c_3 such that*

$$\|D^2 v\| \le c_3\|f\|. \tag{25}$$

Moreover, if $f(x) = -\mu x$, then the solution is classical, $v \in C^2([a,b])$. Similarly, the adjoint problem (24) *has a unique weak solution $w \in H_0^1(a,b)$ for each $g \in L_2(a,b)$, which belongs to $H^2(a,b)$ and*

$$\|D^2 w\| \le c_3\|g\|. \tag{26}$$

Proof. The proof is a standard argument as presented, for example, in [4, Ch. 6] for elliptic PDEs. The only difference is that that the lowest order term in $A(\cdot,\cdot)$ is defined by means of an integral operator, but the crucial properties (20), (21) are the same.

We first show that weak solutions are regular. We use a regularity result for elliptic problems (see [4, p. 323]): If v is a weak solution of

$$\mathcal{L}v(x) = g(x),\ x \in (a,b); \quad v(a) = v(b) = 0,$$

and if $g \in H^k(a,b)$ for some $k \geq 0$, then $v \in H^{k+2}(a,b)$. A weak solution $v \in H_0^1(a,b)$ of (14) satisfies this with $g = f + \mathcal{I}v$, where by (17), (18) $\mathcal{I}v \in H^1(a,b)$. For $f \in L_2(a,b)$ we conclude that $v \in H^2(a,b)$. If $f \in H^1(a,b)$, then we have $v \in H^3(a,b)$ and by Sobolev's inbedding $v \in C^2([a,b])$. In particular, a weak solution is classical when $f(x) = 0$ and $f(x) = -\mu x$. Analogous regularity results hold for the adjoint problem.

Now we can prove existence. Let

$$A_\mu(u,v) = A(u,v) + \tfrac{1}{2}\mu(u,v).$$

By the Lax-Milgram lemma we know that the shifted problem

$$A_\mu(u,\phi) = (g,\phi) \quad \forall \phi \in H_0^1(a,b),$$

has a unique solution $u \in H_0^1(a,b)$ for each $g \in L_2(a,b)$. This defines the bounded linear operator $\mathcal{A}_\mu^{-1} : L_2(a,b) \to H_0^1(a,b)$ by $u = \mathcal{A}_\mu^{-1}g$. The equation (22) is now equivalent to

$$v = \mathcal{A}_\mu^{-1}f + \tfrac{1}{2}\mu\mathcal{A}_\mu^{-1}v,$$

or $v - Kv = h$, where $h = \mathcal{A}_\mu^{-1}f$ and where $K = \frac{1}{2}\mu\mathcal{A}_\mu^{-1} : L_2(a,b) \to L_2(a,b)$ is a compact operator, because $H_0^1(a,b)$ is compactly inbedded in $L_2(a,b)$.

By the Fredholm alternative we know that the latter equation is uniquely solvable for every $h \in L_2(a,b)$ if and only if the corresponding homogeneous equation has no non-trivial solution. But a non-trivial solution of $v - Kv = 0$ would be a weak solution, and hence a classical solution, of (14) with $f = 0$.

Then we can apply the maximum principle for classical solutions of (14), see [5, Theorem 3.1.3]. It says that if a classical function satisfies $(\mathcal{L} - \mathcal{I})u \leq 0$ in (a,b), then $\max_{[a,b]} u = \max_{\mathbb{R}\setminus(a,b)} u$. (The maximum principle for the integro-differential equation is proved in the same way as for the differential equation after noting that $-\mathcal{I}u(x_0) \geq 0$ if u has a maximum at x_0.) We conclude that that the homogeneous equation has no non-trivial solution and therefore (14) has a unique weak solution for every $f \in L_2(a,b)$. By the Fredholm theory the adjoint problem (24) is then also uniquely solvable for all $g \in L_2(a,b)$.

Finally, we prove the bounds (25) and (26). Let $v = \mathcal{A}^{-1}f$ and $w = (\mathcal{A}^*)^{-1}g$ denote the solution operators of (14) and (24), respectively.

Let $f \in H_0^1(a,b)$. Then $v = \mathcal{A}^{-1}f$ is classical and the maximum principle gives

$$\|v\|_{L_\infty(a,b)} \leq c_4\|f\|_{L_\infty(a,b)}. \tag{27}$$

In order to compute the explicit constant we briefly recall the proof. Let

$$\phi(x) = \begin{cases} e^{\gamma(b-a)} - e^{\gamma(x-a)}, & x \le b, \\ 0, & x \ge b, \end{cases}$$

where $\gamma > 0$ is chosen so that that $\mathcal{A}\phi \ge 1$ in (a,b). Then $u(x) = \|f\|_{L_\infty(a,b)}\phi(x)$ satisfies $\mathcal{A}u \ge \|f\|_{L_\infty(a,b)} \ge f = \mathcal{A}v$ in (a,b) and $u \ge 0 = v$ outside (a,b), so that the maximum principle gives $\max_{[a,b]}(v-u) = \max_{\mathbb{R}\setminus(a,b)}(v-u) = 0$, that is, $u \ge v$ in $[a,b]$. Hence $v \le \|\phi\|_{L_\infty(a,b)}\|f\|_{L_\infty(a,b)}$ in $[a,b]$. The lower bound $v \ge -\|\phi\|_{L_\infty(a,b)}\|f\|_{L_\infty(a,b)}$ is obtained in a similar way and so we get

$$\|v\|_{L_\infty(a,b)} \le \|\phi\|_{L_\infty(a,b)}\|f\|_{L_\infty(a,b)} \le e^{\gamma(b-a)}\|f\|_{L_\infty(a,b)}.$$

To determine γ, let $x \in (a,b)$ and compute

$$\begin{aligned} -\mathcal{I}\phi(x) &= \lambda e^{\gamma(x-a)} \int_{-\infty}^{b-x} (e^{\gamma y} - 1)\varphi(y)\,dy \\ &\quad + \lambda(e^{\gamma(b-a)} - e^{\gamma(x-a)}) \int_{b-x}^{\infty} \varphi(y)\,dy \\ &\ge -\lambda e^{\gamma(x-a)} \int_{-\infty}^{\infty} \varphi(y)\,dy = -\lambda e^{\gamma(x-a)}. \end{aligned}$$

Hence,

$$\mathcal{A}\phi(x) \ge (\tfrac{1}{2}\sigma^2\gamma^2 - \mu b\gamma - \lambda)e^{\gamma(x-a)} \ge 1, \quad x \in (a,b),$$

if $\frac{1}{2}\sigma^2\gamma^2 - \mu b\gamma - \lambda \ge 1$, that is, if

$$\gamma = \hat{\gamma} = \frac{\mu b}{\sigma^2} + \sqrt{\frac{2(\lambda+1)}{\sigma^2}}.$$

Then we conclude that (27) holds with $c_4 = e^{\hat{\gamma}(b-a)}$.

Hence, since $\|v\| \le (b-a)^{\frac{1}{2}}\|v\|_{L_\infty(a,b)}$ and $\|f\|_{L_\infty(a,b)} \le (b-a)^{\frac{1}{2}}\|Df\|$, we obtain the bound

$$\|v\| = \|\mathcal{A}^{-1}f\| \le c_5\|Df\| \quad \forall f \in H_0^1(a,b),\ c_5 = (b-a)c_4.$$

By duality we conclude

$$\|(\mathcal{A}^{-1})^*\|_{B(L_2,H^{-1})} = \|\mathcal{A}^{-1}\|_{B(H_0^1,L_2)} \le c_5.$$

Hence

$$\|w\|_{H^{-1}} = \|(\mathcal{A}^*)^{-1}g\|_{H^{-1}} = \|(\mathcal{A}^{-1})^*g\|_{H^{-1}} \le c_5\|g\| \quad \forall g \in L_2(a,b), \tag{28}$$

where $H^{-1}(a,b) = (H_0^1(a,b))^*$ and

$$\|w\|_{H^{-1}} = \sup_{\phi \in H_0^1} \frac{(\phi, w)}{\|D\phi\|}.$$

Recall that $v \mapsto \|Dv\|$ is the chosen norm in $H_0^1(a,b)$. By using $\phi = w \in H_0^1(a,b)$ here we obtain

$$\|w\|^2 \le \|w\|_{H^{-1}} \|Dw\|. \tag{29}$$

We take $\phi = w$ in the adjoint equation (23) and use coercivity (21), the inequality $2ab \le \epsilon a^2 + \epsilon^{-1} b^2$, and (29) to get

$$\begin{aligned}
\tfrac{1}{2}\sigma^2\|Dw\|^2 &\le A(w,w) + \tfrac{1}{2}\mu\|w\|^2 \le \|g\|\|w\| + \tfrac{1}{2}\mu\|w\|^2 \\
&\le \tfrac{1}{2}\mu^{-1}\|g\|^2 + \mu\|w\|^2 \le \tfrac{1}{2}\mu^{-1}\|g\|^2 + \mu\|w\|_{H^{-1}}\|Dw\| \\
&\le \tfrac{1}{2}\mu^{-1}\|g\|^2 + \mu^2\sigma^{-2}\|w\|_{H^{-1}}^2 + \tfrac{1}{4}\sigma^2\|Dw\|^2.
\end{aligned}$$

With (28) this leads to

$$\begin{aligned}
\|Dw\|^2 &\le 2\sigma^{-2}\mu^{-1}\|g\|^2 + 4\sigma^{-4}\mu^{-2}\|w\|_{H^{-1}}^2 \\
&\le (2\sigma^{-2}\mu^{-1} + 4\sigma^{-4}\mu^{-2}c_5^2)\|g\|^2
\end{aligned}$$

and with Poincaré's inequality (19),

$$\|w\| \le c_2\|Dw\| \le c_2(2\sigma^{-2}\mu^{-1} + 4\sigma^{-4}\mu^{-2}c_5^2)^{\frac{1}{2}}\|g\|.$$

Hence

$$\begin{aligned}
&\|(\mathcal{A}^*)^{-1}g\| = \|w\| \le c_6\|g\| \quad \forall g \in L_2(a,b), \\
&c_6 = c_2(2\sigma^{-2}\mu^{-1} + 4\sigma^{-4}\mu^{-2}c_5^2)^{\frac{1}{2}}.
\end{aligned} \tag{30}$$

By duality in L_2 we also have

$$\|v\| = \|\mathcal{A}^{-1}f\| \le c_6\|f\| \quad \forall f \in L_2(a,b). \tag{31}$$

In order to bound D^2v we recall that $v \in H^2(a,b)$. Hence it satisfies (14) strongly, so that with (17) we obtain

$$\begin{aligned}
\tfrac{1}{2}\sigma^2\|D^2v\| &\le \mu\|xDv\| + \|\mathcal{I}v\| + \|f\| \\
&\le \mu\max(|a|,|b|)\|Dv\| + 2\lambda\|v\| + \|f\| \\
&\le \mu\max(|a|,|b|)\|D^2v\|^{\frac{1}{2}}\|v\|^{\frac{1}{2}} + 2\lambda\|v\| + \|f\| \\
&\le \tfrac{1}{4}\sigma^2\|D^2v\| + (2\lambda + \sigma^{-2}\mu^2\max(|a|,|b|)^2)\|v\| + \|f\|.
\end{aligned}$$

Hence,

$$\begin{aligned}&\|D^2v\| \le c_7\|f\| + c_8\|v\|,\\ &\quad c_7 = 4\sigma^{-2},\ c_8 = 4\sigma^{-2}(2\lambda + \mu + \sigma^{-2}\mu^2\max(|a|,|b|)^2).\end{aligned}$$

In the last step we replaced 2λ by $2\lambda + \mu$ in c_8, so that the same result holds also for the adjoint equation (24). Using also (30) and (31) we finally conclude

$$\begin{aligned}&\|D^2v\| \le c_3\|f\|, \quad \|D^2w\| \le c_3\|g\|,\\ &\quad c_3 = c_7 + c_6c_8.\end{aligned}$$

This completes the proof. □

5.3.2 The finite element method

The finite element method is based on a family of subdivisions $\mathcal{T}_h$ of the interval $[a,b]$ parametrized by the maximal mesh size h. Each mesh is of the form

$$\mathcal{T}_h : a = x_0 < x_1 < \cdots < x_{j-1} < x_j < \cdots < x_N = b, \quad h = \max_{j=1,\ldots,N}(x_j - x_{j-1}).$$

We introduce the space $V_h \subset H_0^1(a,b)$ consisting of all continuous functions that reduce to piecewise polynomials of degree ≤ 1 with respect to $\mathcal{T}_h$. See [8, Ch. 5] or [1, Ch. 1]. Then there is an interpolator $I_h : C([a,b]) \to V_h$ such that $I_hu(x_j) = u(x_j)$, $j = 1,\ldots,N$, and

$$(32)\quad \|D(u - I_hu)\|_{L_p(a,b)} \le h^{\frac{1}{2}+\frac{1}{p}}\|D^2u\|, \quad u \in H^2(a,b) \cap H_0^1(a,b),\ p = 2,\infty.$$

To prove this we use the identity

$$D(u - I_hu)(x) = h_j^{-1}\int_{x_{j-1}}^{x_j}\big(u'(x) - u'(y)\big)\,dy = h_j^{-1}\int_{x_{j-1}}^{x_j}\int_y^x u''(z)\,dz\,dy,$$

for $x \in (x_{j-1}, x_j)$ and with $h_j = x_j - x_{j-1}$, which yields

$$|D(u - I_hu)(x)| \le h_j^{\frac{1}{2}}\|D^2u\|_{L_2(x_{j-1},x_j)} \le h^{\frac{1}{2}}\|D^2u\|, \quad x \in (x_{j-1}, x_j).$$

This proves the case $p = \infty$ and for $p = 2$ we have

$$\|D(u - I_hu)\|^2 \le \sum_{j=1}^N h_j^2\|D^2u\|^2_{L_2(x_{j-1},x_j)} \le h^2\|D^2u\|^2.$$

The finite element problem is based on the weak formulation in (22): find $v_h \in V_h$ such that

$$(33)\qquad A(v_h, \phi_h) = (f, \phi_h) \quad \forall \phi_h \in V_h,$$

where $A(\cdot,\cdot)$ is defined in (15) with the integral operator computed as in (16). In the following theorem we prove convergence estimates with explicit constants.

Theorem 5.3.2. *Let v be the solution of* (14) *as in Theorem 5.3.1. There is $h_0 = \sigma/(2^{\frac{1}{2}}\mu^{\frac{1}{2}}c_1c_3)$ such that, for $h \le h_0$,* (33) *has a unique solution $v_h \in V_h$ and*

$$\|v - v_h\| \le 4c_1^2c_3^2\sigma^{-2}h^2\|f\|, \quad \|D(v - v_h)\| \le 4c_1c_3\sigma^{-2}h\|f\|. \tag{34}$$

Proof. We adapt an argument from [11]. Let $e = v - v_h$ denote the error. By subtraction of (33) and (22) with $\phi = \phi_h \in V_h \subset H_0^1(a, b)$ we get

$$A(e, \phi_h) = 0 \quad \forall \phi_h \in V_h. \tag{35}$$

Consider the adjoint problem (23) with $g = e$ and solution $w = (\mathcal{A}^*)^{-1}e$. With $\phi = e$ this yields

$$\begin{aligned}\|e\|^2 = A(e, w) &= A(e, w - I_hw) \le c_1\|De\|\|D(w - I_hw)\|\\ &\le c_1\|De\|h\|D^2w\| \le c_1c_3h\|De\|\|e\|.\end{aligned}$$

Here we used (35), (20), (32), and (26). We conclude

$$\|e\| \le c_1c_3h\|De\|. \tag{36}$$

In view of (35) we have $A(e, e) = A(e, v - v_h) = A(e, v)$, so that by (21) and (36),

$$\begin{aligned}\tfrac{1}{2}\sigma^2\|De\|^2 &\le A(e, e) + \tfrac{1}{2}\mu\|e\|^2 = A(e, v) + \tfrac{1}{2}\mu\|e\|^2\\ &\le c_1\|De\|\|Dv\| + \tfrac{1}{2}\mu c_1^2c_3^2h^2\|De\|^2.\end{aligned} \tag{37}$$

Hence, for $h \le h_0$ sufficiently small ($h_0^2 = \sigma^2/(2\mu c_1^2c_3^2)$), we have

$$\|De\| \le c_9\|Dv\|, \quad c_9 = 4c_1\sigma^{-2}.$$

Now if $f = 0$ in (22) and (33), then $v = 0$ by uniqueness, and hence $e = 0$, so that $v_h = 0$. This means that we have uniqueness for the finite element problem (33). But this is an equation in a finite dimensional space so existence also follows. Therefore, (33) has a unique solution for all $f \in L_2(a, b)$ if $h \le h_0$.

In order to prove the error estimate (34) we return to (37) but use $A(e, e) = A(e, v - v_h) = A(e, v - I_hv)$ instead:

$$\begin{aligned}\tfrac{1}{2}\sigma^2\|De\|^2 &\le A(e, e) + \tfrac{1}{2}\mu\|e\|^2 = A(e, v - I_hv) + \tfrac{1}{2}\mu\|e\|^2\\ &\le c_1\|De\|\|D(v - I_hv)\| + \tfrac{1}{2}\mu c_1^2c_3^2h^2\|De\|^2,\end{aligned}$$

and conclude, for $h \le h_0$,

$$\|De\| \le c_9\|D(v - I_hv)\|, \quad c_9 = 4c_1\sigma^{-2}.$$

Hence, by (32), (25), and (36),

$$\begin{aligned}\|De\| &\le c_9h\|D^2v\| \le c_9c_3h\|f\| = 4c_1c_3\sigma^{-2}h\|f\|,\\ \|e\| &\le c_1c_3h\|De\| \le 4c_1^2c_3^2\sigma^{-2}h^2\|f\|,\end{aligned}$$

which is (34). ∎

We finish by proving the pointwise convergence of the derivative.

Corollary 5.3.3. *Assume that each finite element mesh $\mathcal{T}_h$ is uniform, that is, $x_j - x_{j-1} = h$ for $j = 1, \dots, N$. Then, for $h \le h_0$ as in Theorem 5.3.2, we have*

$$|v'(b) - v_h'(b)| \le c_{10} h^{\frac{1}{2}} \|f\|, \quad c_{10} = 2 + 4c_1 c_3 \sigma^{-2}.$$

Proof. We use the inverse inequality

$$\|D\phi_h\|_{L_\infty(a,b)} \le h^{-\frac{1}{2}} \|D\phi_h\|, \quad \phi_h \in V_h.$$

To prove this we note that

$$D\phi_h(x) = h^{-1} \int_{x_{j-1}}^{x_j} D\phi_h(y)\, dy, \quad x \in (x_{j-1}, x_j),\ h = x_j - x_{j-1},$$

which yields

$$|D\phi_h(x)| \le h^{-\frac{1}{2}} \|D\phi_h\|_{L_2(x_{j-1},x_j)} \le h^{-\frac{1}{2}} \|D\phi_h\|, \quad x \in (x_{j-1}, x_j).$$

Hence, by (32) and (34),

$$\begin{aligned}
\|De\|_{L_\infty(a,b)} &\le \|D(v - I_h v)\|_{L_\infty(a,b)} + \|D(I_h v - v_h)\|_{L_\infty(a,b)} \\
&\le \|D(v - I_h v)\|_{L_\infty(a,b)} + h^{-\frac{1}{2}} \|D(I_h v - v_h)\| \\
&\le \|D(v - I_h v)\|_{L_\infty(a,b)} + h^{-\frac{1}{2}} \|D(I_h v - v)\| + h^{-\frac{1}{2}} \|D(v - v_h)\| \\
&\le 2h^{\frac{1}{2}} \|D^2 v\| + h^{-\frac{1}{2}} \|D(v - v_h)\| \le (2 + 4c_1 c_3 \sigma^{-2}) h^{\frac{1}{2}} \|f\|.
\end{aligned}$$

Therefore

$$|v'(b) - v_h'(b)| \le (2 + 4c_1 c_3 \sigma^{-2}) h^{\frac{1}{2}} \|f\|.$$

□

In particular, with $f(x) = -\mu x$, Corollary 5.3.3 gives

$$|v'(b) - v_h'(b)| \le c_{11} h^{\frac{1}{2}}, \quad c_{11} = c_{10} \mu \sqrt{\frac{b^3 - a^3}{3}}. \tag{38}$$

Given numerical values for the parameters $a, b, \sigma, \mu, \lambda$ we may now compute numerical values for h_0 and c_{11}. Alternatively, we may conclude that there are uniform bounds $h_0 \ge \hat{h}_0$, $c_{11} \le \hat{c}_{11}$ for $b \in [b_1, b_2]$ and with the other parameters fixed.

5.3.3 The free boundary value problem

We use uniform meshes $\mathcal{T}_h$ with

$$x_j - x_{j-1} = h = \frac{b-a}{N}, \quad j = 1, \ldots, N.$$

Since we want to vary b, we parametrize by N instead of h. Let $f(x) = -\mu x$, fix $a < 0$ and let v, v_N denote the solutions of (22) and (33) for $b > a$. Define the functions

$$F(b) = v'(b), \quad F_N(b) = v_N'(b).$$

From (38), we get for $a < b_1 < b_2$

$$\begin{aligned} \|F - F_N\|_{L_\infty(b_1,b_2)} &\leq \hat{c}_{12} N^{-\frac{1}{2}}, \quad N \geq \hat{N}_0, \\ \hat{c}_{12} = \hat{c}_{11}(b_2 - a)^{\frac{1}{2}}&, \quad \hat{N}_0 = \frac{b_2 - a}{\hat{h}_0}. \end{aligned} \tag{39}$$

By writing down the matrix equation for solving the finite element problem (33), it is easy to see that, for fixed N, the function $b \mapsto F_N(b)$ is continuous on (a, ∞). From (39) we conclude that $b \mapsto F(b)$ is also continuous on (a, ∞). Moreover, by a direct consequence of the strong maximum principle and the Hopf boundary point principle for our equation (see [5, Theorem 3.1.4-3.1.5]), we get the following:

Lemma 5.3.4. *If $a < b \leq 0$, then $F(b) < 0$. In particular, if (u, b) is a solution to the free boundary problem* (4)*, then $b > 0$.*

We believe that there exists a unique $b > 0$ such that $F(b) = 0$. We are not able to provide a rigorous proof of this, but numerical simulations present strong evidence in the following way. Assign numerical values to the parameters a, σ, μ, λ and fix a jump density φ. In all our computations, we took φ to be the truncated normal distribution with mean zero, variance $\gamma > 0$ and support $[-J, J]$, i.e.

$$\varphi(y) = \begin{cases} \dfrac{e^{-\frac{y^2}{2\gamma^2}}}{\gamma\sqrt{2\pi}\,(2\Phi(J/\gamma) - 1)} & \text{if} \quad -J < y < J, \\ 0 & \text{otherwise,} \end{cases}$$

where

$$\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-\frac{y^2}{2}}\, dy, \quad x \in \mathbb{R}.$$

From computations of the boundary value problem (33) (see Figures 5.1 and 5.2), we can find $0 \leq b_1 < b_2$ and $\tilde{N} \geq \hat{N}_0$ such that

$$F_{\tilde{N}}(b_1) \leq -\frac{1}{2}, \quad F_{\tilde{N}}(b_2) \geq \frac{1}{2}, \quad \text{and} \quad \hat{c}_{12}\tilde{N}^{-\frac{1}{2}} < \frac{1}{4}.$$

(The $1/2$ and $1/4$ may vary if we change the parameters.) From (39), we can then conclude that

$$F(b_1) < 0, \qquad F(b_2) > 0,$$
$$F_N(b_1) < 0, \quad F_N(b_2) > 0 \quad \text{for all } N \geq \tilde{N}.$$

Hence, there exists $b \in (b_1, b_2)$ such that $F(b) = 0$ and for each $N \geq \tilde{N}$ there exists $b_N \in (b_1, b_2)$ such that $F_N(b_N) = 0$. Moreover, (39) gives us that

$$\lim_{N\to\infty} F(b_N) = 0.$$

Of course, we cannot conclude that b is unique and $b_N \to b$ as $N \to \infty$. However, Figure 5.1 suggests that b is unique and from computations with increasing N, it seems like b_N converges, see Table 5.1.

We now discuss whether the properties $a)$ and $b)$ in the statement of Theorem 5.1.1 hold for a solution (u, b) of (4). We have no rigorous proof, but computational evidence. The properties $a)$ and $b)$ boil down to

$$\lambda \int_a^b v(y)\varphi(y-x)\,dy \leq \mu x, \quad \text{for } x > b, \tag{40}$$

and $v \geq 0$ respectively, where (v, b) solves (13). We believe that $v \geq 0$ holds for all values of the parameters, but computations suggests that (40) may fail for certain parameter values, typically when σ is small and λ is three or four times larger than μ. See Figures 5.3 and 5.4, where we check (40) for (v_N, b_N) instead of (v, b).

N	b_N
2000	0.0572939
4000	0.0572743
6000	0.0572678
8000	0.0572653

Table 5.1: $a = -0.1$, $\lambda = 10$, $\sigma = 0.2$, $\mu = \frac{\sigma^2}{0.005}$, $\gamma = 0.02$ *and* $J = 0.05$.

References

[1] S. C. Brenner and L. R. Scott, *The Mathematical Theory of Finite Element Methods*, Springer, 2008.

[2] E. Ekström, C. Lindberg, and J. Tysk, *Optimal liquidation of a pair trade*, to appear.

[3] R. Elliot, J. van der Hoek, and W. Malcolm, *Pairs trading*, Quant. Finance **5** (2005), 271–276.

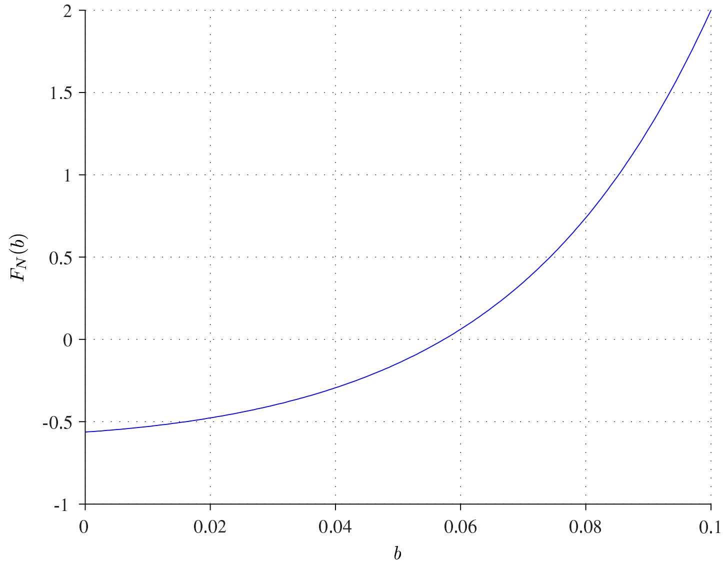

Figure 5.1: *The function* F_N *when* $a = -0.1$, $\lambda = 10$, $\sigma = 0.2$, $\mu = \frac{\sigma^2}{0.005}$, $\gamma = 0.02$ *and* $J = 0.05$.

[4] L. C. Evans, *Partial Differential Equations*, American Mathematical Society, 1998.

[5] M. G. Garroni and J. L. Menaldi, *Second Order Elliptic Integro-Differential Problems*, Chapman and Hall/CRC, 2002.

[6] E. Gatev, W. Goetzmann, and G. Rouwenhorst, *Pairs trading: performance of a relative-value arbitrage*, Review of Financial Studies **19**, 797–827.

[7] I. Karatzas and S. E. Shreve, *Brownian Motion and Stochastic Calculus*, Springer, 1998.

[8] S. Larsson and V. Thomée, *Partial Differential Equations with Numerical Methods*, Springer, 2003.

[9] G. Peskir and A. Shiryaev, *Optimal Stopping and Free-Boundary Problems*, Springer, 2006.

[10] P. E. Protter, *Stochastic Integration and Differential Equations*, Springer, 2005.

[11] A. H. Schatz, *An observation concerning Ritz-Galerkin methods with indefinite bilinear forms*, Math. Comp. **28** (1974), 959–962.

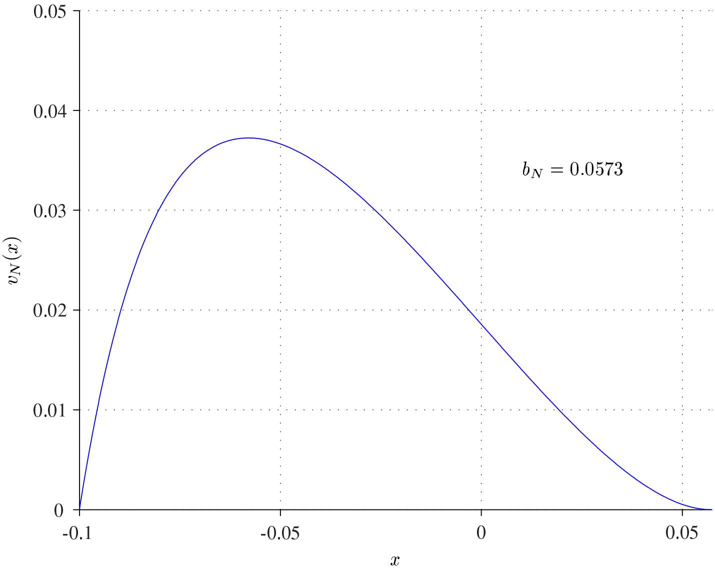

Figure 5.2: *The solution* (v_N, b_N) *when* $a = -0.1$, $\lambda = 10$, $\sigma = 0.2$, $\mu = \frac{\sigma^2}{0.005}$, $\gamma = 0.02$ *and* $J = 0.05$.

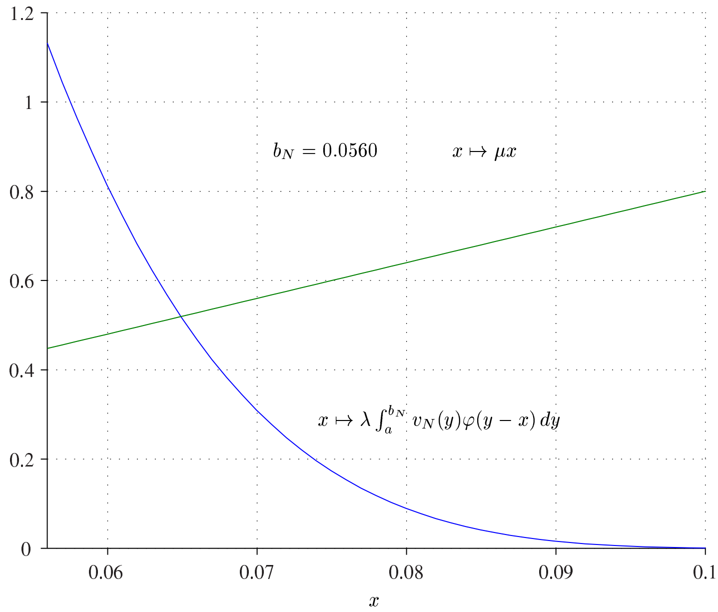

Figure 5.3: *A simulation of* (40) *when* $a = -0.1$, $\lambda = 30$, $\sigma = 0.2$, $\mu = \frac{\sigma^2}{0.005}$, $\gamma = 0.02$ *and* $J = 0.05$. *The condition fails.*

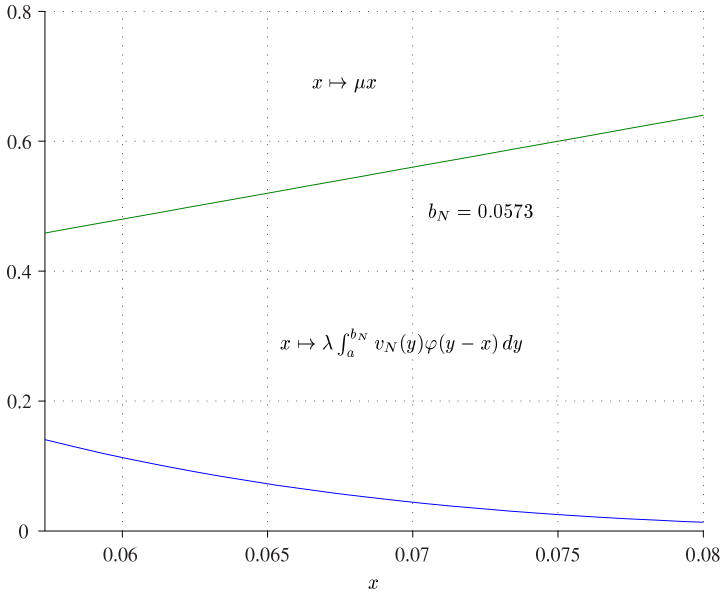

Figure 5.4: *A simulation of* (40) *when* $a = -0.1$, $\lambda = 10$, $\sigma = 0.2$, $\mu = \frac{\sigma^2}{0.005}$, $\gamma = 0.02$ *and* $J = 0.05$. *The condition holds.*