Phenotypic heterogeneity in bacterial populations

—a mathematical study

Peter Czuppon

Dissertation zur Erlangung des Doktorgrades
der
Fakultät für Mathematik und Physik
der
Albert-Ludwigs-Universität Freiburg im Breisgau
Dekan / Dean  
Prof. Dr. Dietmar Kröner

Revisoren / Reviser  
Prof. Dr. Peter Pfaffelhuber  
Prof. Dr. Johannes Müller

Datum der Promotion / Date of defense  
20.09.2016
Abstract

Heterogeneity is ubiquitously observed in nature. From microbes to mammals, individuals established distinct life styles even given homogeneous environmental conditions and despite being genetically identical cells. This form of heterogeneity is called phenotypic or non-genetic heterogeneity in contrast to genetic heterogeneity which is caused by differences in the genome of individuals. Concerning phenotypic heterogeneity, first observations date back to the 70’s at least when Spudich and Koshland describe this phenomenon in their paper "Non-genetic heterogeneity: chance in a single cell", [SK76]. Since then scientists try to explain these findings and have so far identified several causes leading to this phenomenon. In the present work we study two of these explanatory approaches from a theoretical point of view. We focus on (i) random intracellular fluctuations leading to stochasticity in single cells and (ii) a specific model for cell-cell interactions which influence the cellular decisions regarding their phenotype.

(i) Studying random intracellular fluctuations, heterogeneity is observed in single cell studies where individuals appear to behave randomly. However, when looking at the population level colonies seem to reproduce certain patterns which appear to follow a deterministic rule. To resolve this putative paradox, we study a Markov process indexed by a binary tree where the dynamics along the two daughter nodes evolve independently. When rescaling time by $n$ and looking at the empirical measure process of the $[nt]$-th level we can derive a limit law for the distribution. This limit turns out to be deterministic (at least if the initial condition is fixed) although the individual itself evolved stochastically.

(ii) Concerning the second approach, phenotypic heterogeneity due to social interactions among cells, a spatial model with interacting particles seems well suited. Here, individuals make decisions, resulting in the expression of a certain phenotype, dependent on the type of neighboring cells. In the present work we model a public good producing colony. Therefore, we define a particle system consisting of two types, the contributing type (called cooperator) and the non-producing phenotype (called defector). The dynamics are described by a biased voter model favoring defectors extended by a cooperation mechanism. Both phenotypes are assumed to benefit from the public good, i.e. cooperators support both types equally. We study the limit behavior of this system on the lattice $\mathbb{Z}^d$. In one dimension we observe a phase transition from defectors taking over to cooperators dominating the colony dependent on the parameter configuration. Regarding dimensions $d \geq 3$ we describe the limit behavior in certain parameter regions given a weak interaction limit. Furthermore, by slightly altering the cooperation mechanism and considering a large cooperation rate we obtain a result for all dimensions $d \geq 1$ favoring cooperators.

The proofs of the results in the two approaches are based on weak convergence techniques, martingale theory, couplings with other interacting particles systems, approximations by reaction-diffusion equations and comparisons with oriented percolation processes.
Zusammenfassung


erhalten wir, nach einer kleinen Änderung des Kooperationsmechanismus', für große Kooperationsraten ein Resultat über die Überlebenschancen von Kooperierern für alle Dimensionen \( d \geq 1 \).

Die Beweise der obigen Resultate basieren auf Methoden der schwachen Konvergenz, Martingaltheorie, Kopplung mit weiteren interagierenden Teilchensystemen, Approximationen durch Reaktions-Diffusions-Gleichungen und Vergleichen mit orientierten Perkulationsprozessen.
Declaration of contributions as a co-author

In this dissertation I present my doctoral research from December 2012 until February 2016 on the theoretical study of heterogeneous behavior in bacterial populations. The research has led to the following publications and was conducted under the supervision of Prof. Dr. Peter Pfaffelhuber.

(a) Czuppon, P. and Pfaffelhuber, P.
Some limit results for Markov chains indexed by trees.
Electronic Communications in Probability (2014), 19, 1–11.

(b) Czuppon, P. and Pfaffelhuber, P.
A spatial model for selection and cooperation.

The results of (a) are shown in Chapter 2. I wrote parts of the manuscript. Reference (b) is presented in Chapters 3 and 4. The model studied there was developed and analyzed by me under the supervision of Prof. Dr. Peter Pfaffelhuber. I wrote main parts of the manuscript. The remaining contents of this doctoral thesis were developed by me, again under the supervision of Prof. Dr. Peter Pfaffelhuber. I wrote the manuscript.

Further publications
In addition, I contributed to the following papers:

(c) McIntosh M., Czuppon P., Best K., Becker A. and Pfaffelhuber P.
Modeling quorum sensing in *Sinorhizobium meliloti*.

(d) Schlüter J.P., Czuppon P., Schauer O., Pfaffelhuber P., McIntosh M. and Becker A.
Classification of phenotypic subpopulations in isogenic bacterial cultures by triple promoter probing at single cell level.

(e) Kumberger P., Kuttler C., Czuppon P. and Hense B.A.
Multiple regulation mechanisms of bacterial quorum sensing.
(2016), BIOMATH, 5 (1).

The papers listed in (c) and (d) emerged from collaborations with the experimental biologists group of Prof. Dr. Anke Becker at Synmikro at the Philipps-University of Marburg. (c) grew out of my Diploma thesis entitled "Probabilistic Tools in Chemical Reaction Network Theory and Applications" and deals with a stochastic description of a regulatory network. Concerning (d), I developed a statistical method for the classification of phenotypic subpopulations in the bacterium *Sinorhizobium meliloti*. 
Both works were completed under the supervision of Prof. Dr. Peter Pfaffelhuber. Additionally, I contributed to (e) by a stochastic treatment of the there studied model. This research was done in cooperation with Peter Kumberger under the supervision of Prof. Dr. Christina Kuttler and Dr. Burkhard Hense. The results from (c), (d) and (e) are not presented in this dissertation.
Acknowledgments

First of all I would like to thank Prof. Dr. Peter Pfaffelhuber who has been an excellent supervisor. I am especially grateful for all the time he spent discussing with me, sharing his knowledge and providing new ideas to the various topics I had the opportunity to work on. Without his advice and support this dissertation would not have been possible and by far not so heterogeneous.

I am grateful to Andrej Depperschmidt who never hesitated to discuss my questions about interacting particle systems with me and also to Franz Baumdicker for conversations on biological issues and the exceptionally great office atmosphere. For helping me out while constructing the figures I am indebted to Sebastian Bossert who besides that was a great companion during the last three years. Many thanks also to the rest of the Probability Group of the Albert-Ludwigs-University in Freiburg for the nice atmosphere and the joyful chats during all the coffee breaks.

Additionally, I would like to express my gratitude to the DFG for the financial support. It was a great experience being part of the SPP 1617 on “Phenotypic Heterogeneity and Sociobiology of Bacterial Populations”. I am especially grateful to our collaborators from the Philipps-University in Marburg, namely Prof. Dr. Anke Becker, Dr. Matthew McIntosh and Dr. Jan-Philip Schlüter. Thank you all for your patience during (Skype) discussions and your great effort (especially the cell-counting) when providing us your experimental results.

Now, turning to the non-scientific thanksgiving I switch to German. Zunächst einmal vielen lieben Dank Marco, Sus und Trixi für die unzähligen Kaffeepausen, Abende, Wochenenden, Sportveranstaltungen, etc., die für einen schönen Ausgleich zum Universitätsalltag gesorgt haben. Ohne euch wäre die Promotionszeit nur halb so schön gewesen. Außerdem auch ein großes Dankeschön an meine beiden langjährigen Mitbewohnerinnen Heike und Magda, ohne die ich mich sicherlich nicht so gesund und abwechslungsreich ernährt hätte und die einen wesentlichen Anteil an meinem unkomplizierten Wohnalltag hatten. Und zu guter Letzt noch ein Dank an alle bisher nicht namentlich erwähnten Freunde in Freiburg, Heidelberg, München, usw., die immer ein offenes Ohr und/oder eine freie Couch für mich hatten und sich meine Unigeschichten anhören durften oder mussten.
# Contents

1 Introduction

2 **Phenotypic heterogeneity due to random fluctuations**
   2.1 Markov chains indexed by trees
   2.2 A limit theorem for the empirical measure process
   2.3 Proof of the main result
      2.3.1 Auxiliary Results
      2.3.2 Convergence of the finite-dimensional distributions of \((\tilde{Z}^n)_{n\geq 1}\)
      2.3.3 Tightness of \((\tilde{Z}^n)_{n\geq 1}\)
   2.4 Discussion

3 **Cell-Cell Interactions**
   3.1 Introduction
   3.2 The model
      3.2.1 Existence and uniqueness of the process
   3.3 Unstructured populations

4 **The VMBC-process on \(\mathbb{Z}^d\)**
   4.1 Long-time behavior
   4.2 Preliminaries
      4.2.1 Comparison results
      4.2.2 A result on a jump process
   4.3 Proofs
      4.3.1 Proof of Theorem 4.2
      4.3.2 Proof of Theorem 4.3


5 Weak Interaction Limit
5.1 Voter model with perturbations
5.1.1 Hydrodynamic limit
5.2 The rescaled eVMBC
5.2.1 Limit behavior of the rescaled eVMBC
5.3 Proof of Theorem 5.8

6 Limit behavior for large cooperation rates
6.1 Definition and Result for the sVMBC
6.2 Preliminaries
6.2.1 Pure growth-process
6.2.2 Percolation
6.3 Proof of Theorem 6.3
6.4 Discussion
6.4.1 Justification of the sVMBC
6.4.2 Alternative models
6.4.3 Simulation results

7 Conclusion and Outlook

List of Figures

Bibliography
When thinking about heterogeneity and individual decision making one might not immediately think of bacterial cells. Yet, bacteria turn out to be very well suited to study these kinds of observations. This is not only due to the fact that bacteria are convenient to analyze under controllable environmental conditions in contrast to wild-life animals or humans. Furthermore, their genome is often enough sufficiently well understood and mutations are tractable. But, probably even more important there are a lot of examples where in bacterial colonies heterogeneous behavior is the rule and not only the exception, [Ave06]. In fact heterogeneity is ubiquitous in living individuals going from microbes to mammals, [BvOC11].

Talking about heterogeneity in bacteria, one usually distinguishes between genetic and phenotypic heterogeneity, [Ack15]. Here, genetic heterogeneity refers to diverse behavior of individuals due to differences in the genome. The term phenotypic or non-genetic heterogeneity on the other hand is used for diversity independent of genetic or environmental variation, i.e. heterogeneity amongst genetically identical cells facing the same external conditions.

In the following we focus on the latter type of heterogeneity. As already mentioned there are many bacteria known to behave heterogeneously despite being absolutely identical cells. The question which has arisen is about the mechanisms which are responsible for such phenotypic heterogeneity. By now at least four different mechanisms which lead to this non-genetic heterogeneity are identified:

(i) random intracellular fluctuations such as stochastic gene expression, [MA99, ELSS02], stochastic partitioning at cell division [HP11] or periodic oscillations, [LS11, LLE13];

(ii) cellular age, [MM08, BA11];
Introduction

(iii) cell-cell interactions together with intracellular regulatory mechanisms, [SP11, RE11];

(iv) epigenetic modification of chromatin through changes in DNA methylation, [MDSA11].

Several examples are today known to fall in one of the four categories; see the reviews [Ave06, Ack15]. Another question that results naturally from studying non-genetic heterogeneity is the subject of functionality and evolutionary impact of this phenomenon. Although there are a lot of interesting (theoretical) approaches to this question we will not deal with it in the following. For a review on recent research in this area see [Ack15].

Instead we focus on the causes leading to phenotypic heterogeneity in isogenic microbial colonies. Concerning the list from above we exclude the fourth point since epigenetic modifications are primarily observed in eukaryotic cells, see [RSP+12] for a review on mathematical models of dynamics in the epigenome. Hence, we concentrate on the first three causes in the following. Hereby, we treat cellular aging and random intracellular fluctuations as mathematically equivalent. Therefore, we are basically dealing with two mathematical frameworks, Markov processes evolving on trees in cases (i) and (ii) and interacting particle systems for a model relating to (iii).

In order to describe random fluctuations or cellular aging in a bacterial colony we consider a Markov process indexed by a tree. In the case of cells it is convenient to consider binary trees, i.e. each vertex has one ancestor node (except for the root - the founder cell) and two descendants. The stochasticity in gene expression or stochastic location of the old pole during cell division is then modeled by a (asymmetric) Markov process evolving along this binary tree. The object of interest is the limit behavior of the empirical measure process which resembles the observations when analyzing experimentally obtained data from different colonies of the same bacterial trait.

Concerning cell-cell interactions we rely on the framework of interacting particle systems. Over the past few decades very general particle systems have been introduced and classes of models have been identified. Hence, a lot of tools for the analysis of these kinds of models are available, see for example the monographs [Lig85] and [Lig99]. We are interested in the special case of a public good producing phenotype which simulates the behavior of the bacterium studied by our experimental collaborators from the group of Prof. Dr. Anke Becker at Synmikro at the Philipps-University of Marburg. In order to model this specific situation we extend the well known two-type biased voter model, see e.g. [BG81]. To be more concrete, some particles produce a public good and provide it to their environment (cooperators), whereas other particles focus on reproductive processes and do not participate in the public good production (defectors). Thus, defectors have a selective advantage over the cooperating phenotype. Again, the goal is to analyze the limit behavior of this system, namely the convergence to invariant laws under different parameter configurations and on different graph structures. The aim is
to identify parameter regions such that the public good producing type has a chance to survive and thus provide conditions for the evolution of altruistic or cooperative behavior in this setting.

The thesis is structured as follows:

Chapter 2 deals with the situation of intracellular random fluctuations. We define a bifurcating Markov process indexed by a binary tree, i.e. the two daughter cells evolve independently according to some Markov processes. By rescaling time we obtain a weak limit result for the empirical measure of the \([nt]-th\) level. This basically means that we may observe variability amongst the cells but the population itself turns out to behave deterministically (at least given a deterministic starting condition).

Chapter 3 introduces a model for social interactions of bacterial cells in the setting of a public good producing type. Therefore, we apply tools provided by the general framework of interacting particle systems. The model essentially depends on two parameters, namely \(\alpha\) the selection parameter which represents the reproductive fitness advantage of defectors and \(\gamma\) which describes the benefit from having access to the public good. We mathematically define the model, prove the existence of a Markov process possessing these interaction dynamics and analyze the case of an unstructured population which simulates the case of a non-spatial system.

Chapter 4 presents limit theorems for the introduced interacting particle system on the lattice \(\mathbb{Z}^d\) for \(d \geq 1\). In particular, we study two different models. The first one is a purely altruistic model where it turns out that if \(\alpha > \frac{d-1}{d} \gamma\) cooperators do not survive in any dimension. On the other hand when considering a more cooperative (and thus less altruistic) form of the model dynamics we can prove that at least on \(\mathbb{Z}\) in the case \(\gamma > \alpha\) cooperators take over the colony, i.e. defectors die out in the limit for time \(t\) tending to infinity.

Chapter 5 analyzes the cooperative model on \(\mathbb{Z}^d\) for dimensions \(d \geq 3\). Therefore, we rescale the parameters by a factor \(\varepsilon > 0\) for \(\varepsilon\) small, such that we can approximate the system by a reaction-diffusion equation. To be more precise, the model then fits into the class of voter models with perturbations. Hence, in order to investigate the limit behavior of our interacting particle system we can apply known results about these systems. Unfortunately, we are limited to the case \(d \geq 3\), the dimensions where the voter model possesses non-trivial invariant laws. In this scenario, we obtain that for \(\alpha \geq \frac{4d^2-2d+1}{3d(2d-1)} \gamma\) cooperating individuals die out whereas for \(\alpha \leq \frac{1}{3d(2d-1)} \gamma\) defecting particles go extinct.

Chapter 6 studies a slightly altered model in contrast to the model introduced in Chapter 3. Here, the cooperator-defector interaction is changed. This allows us to study the limit behavior in all dimensions which finally also includes the
biologically most relevant case, $d = 2$. We apply methods from percolation theory on oriented graphs in order to show that for $\gamma$ large enough cooperators take over the colony. Furthermore, we briefly discuss the applicability of this method to other possible interaction forms for public good producing colonies and take a look at simulation results of the original model and compare them to the results obtained in this chapter.
In this part, we study the evolution of phenotypic heterogeneity in a microbial colony which arises due to random fluctuations of processes occurring inside cells. These processes can be chemical reactions taking place or uneven splitting of molecules during cell division or even the age of the cell pole which affects the cellular behavior. The underlying question of phenotypic heterogeneity caused by these circumstances is motivated by an observation in microbiology, where a population of bacteria is growing (along a binary tree, say), and every individual bacterial cell is in a certain state (e.g. some gene expression profile), which can be – at least partially – inherited. It has been observed for a long time that such populations tend to be heterogeneous although all cells carry the same genome; see [SK76] for an early reference.

Below, we analyze one consequence of intracellular random fluctuations, i.e. a law of large numbers. This result entails that the dynamics of single cells can be stochastic while the behavior of the whole population becomes deterministic.

In [BP94], Benjamini and Peres introduced the notion of a tree-indexed Markov chain which is a suitable framework for modeling cell divisions coupled with cell state inheritance. Since this first appearance of tree-indexed Markov chains, a lot of effort has been spent in studying weak and strong laws of large numbers for very general types of and even possibly random trees [YL02, Yan03, LW03, LY04, Tak06, Guy07, BdSG09, DM10, dSGM11].

Our approach consists of a Markov kernel dependent on some scaling parameter $n$ which will tend to infinity. We then look at the empirical measure process in the $[nt]$-th generation of the population, $t \geq 0$, which corresponds to a time-scaling of the empirical measure process. We will prove the weak convergence of these processes, which will be a deterministic limit (if the initial distribution is deterministic).

After introducing the general notions in Section 2.1, we present our main result in
2 Phenotypic heterogeneity due to random fluctuations

Theorem 2.5 in Section 2.2, together with two simple examples. Subsequently, we give the proof of Theorem 2.5 in Section 2.3 and discuss its applicability in Section 2.4.

The following results also appeared in [CP14].

2.1 Markov chains indexed by trees

We define by

\[ T = \bigcup_{k=0}^{\infty} T_k, \quad T_0 = \{\emptyset\}, \quad T_k = \{0,1\}^k, k = 1,2,\ldots \]

a complete binary tree, rooted at \( \emptyset \in T_0 \), where \( \sigma 0, \sigma 1 \in T_{k+1} \) are the two children of \( \sigma \in T_k \), \( k = 0,1,2,\ldots \). For \( \sigma \in T_k \) and \( j \leq k \), we denote by \( \pi_j \sigma \) the prefix of \( \sigma \) of length \( j \), e.g. for \( \sigma \in T_k \) the vertex \( \pi_{k-1} \sigma \) is the parental vertex of \( \sigma \). Furthermore, we set \( |\sigma| = k \) iff \( \sigma \in T_k \) and in addition, set \( \pi_{-1} \sigma := \pi_{|\sigma|-1} \sigma \), the immediate ancestor of \( \sigma \). We define the \( \leq \)-relation by writing

\[ \sigma \leq \tau \text{ iff there is } j \text{ such that } \pi_j \tau = \sigma \]

and

\[ \tau \land \tau' := \sup\{\sigma : \sigma \leq \tau, \sigma \leq \tau'\} \]

as the most recent common ancestor of \( \tau \) and \( \tau' \).

In the following, let \( (E,r) \) be a complete and separable metric space and denote by \( \mathcal{B}(E) \) the Borel-\( \sigma \)-field. We call a stochastic process \( X = (X_{\sigma})_{\sigma \in T} \) a time-homogeneous, tree-indexed Markov chain (extending a notion introduced in [BP94]), if there is a Markov transition kernel \( p \) from \( E \) to \( \mathcal{B}(E^2) \) such that for all \( \sigma \in T \) and \( A_0, A_1 \in \mathcal{B}(E) \),

\[ P(X_{\sigma_0} \in A_0, X_{\sigma_1} \in A_1 | X_{\tau} = x_{\sigma}, X_{\tau} = x_{\tau} \text{ for } \tau \in T \text{ with } \tau \land \sigma < \sigma) \]

\[ = P(X_{\sigma_0} \in A_0, X_{\sigma_1} \in A_1 | X_{\sigma} = x_{\sigma}) = p(x_{\sigma}, A_0 \times A_1). \quad (2.1) \]

We connect the Markov chain \( R = (R_n)_{n=0,1,2,\ldots} \) with \( X \) and define its transition kernel by

\[ p_R(x, A) := \frac{1}{2}(p(x,A \times E) + p(x,E \times A)). \]

Here, \( R \) arises from observing the state of \( X \) when walking along \( T \) starting from \( \sigma \) to \( \sigma 0 \) and \( \sigma 1 \) purely at random.

Another representation of \( R \) is as follows: let \( (\Sigma_k)_{k=0,1,2,\ldots} \) be a symmetric forward random walk on \( T \) (independent of \( X \)), i.e. \( \Sigma_k \in T_k \) almost surely and

\[ P(\Sigma_{k+1} = \sigma 0 | \Sigma_k = \sigma) = P(\Sigma_{k+1} = \sigma 1 | \Sigma_k = \sigma) = \frac{1}{2}. \]
Then, \( \mathcal{R} \overset{d}{=} (X_{\Sigma_k})_{k=0,1,2,...} \). For an illustration of a binary tree, its definitions and an exemplary realization of \( \mathcal{R} \) see Figure 2.1.

Additionally, we define the process of empirical measures \( Z = (Z_k)_{k=0,1,2,...} \) through

\[
Z_k := 2^{-k} \sum_{\sigma \in T_k} \delta_{X_{\sigma}},
\]

for \( (X_{\sigma})_{\sigma \in T} \) a (time-homogeneous) Markov chain as introduced above. Note that \( Z \) takes values in \( \mathcal{P}(E) \), the set of probability measures on \( \mathcal{B}(E) \) and that \( Z \) is a non-homogeneous Markov chain (indexed by \( k = 0,1,2,... \)).

**Remark 2.1** (Symmetric, tree-indexed Markov chains) The idea to consider different transition mechanisms to the two different children is due to the work of Guyon in 2007, [Guy07]. A special, classical case is that of a symmetric tree-indexed Markov chain which is defined as follows:

We call a time-homogeneous, (tree-indexed) Markov chain with transition kernel \( p \) (from \( E \) to \( \mathcal{B}(E^2) \)) symmetric, if there is a Markov transition kernel \( q \) (from \( E \) to \( \mathcal{B}(E) \)) such that for all \( x \in E, A_0, A_1 \in \mathcal{B}(E) \)

\[
p(x, A_0 \times A_1) = q(x, A_0) \cdot q(x, A_1).
\]

In other words, the transitions from \( X_{\sigma} \) to \( X_{\sigma_0} \) and to \( X_{\sigma_1} \) are independent and identically distributed. In this case, we have that \( R_k \overset{d}{=} X_{\sigma} \) for all \( \sigma \in \Sigma_k \). More general results for this symmetric case can be found in [BP94, YL02, Yan03, LW03, LY04, Tak06].
In the following, we consider the more general case of transition kernels like introduced in equation (2.1).

2.2 A limit theorem for the empirical measure process

Here, we state a limit theorem for the setup introduced in the last section. Therefore, let \((X^n)_{n=1,2,...}\) be a sequence of tree-indexed Markov chains with complete and separable metric state spaces \((E^n,r^n)_{n=1,2,...}\). We have a complete separable metric space \((E,r)\) as a limit state space and Borel-measurable maps \(\eta^n : E^n \to E\) mapping to it.

Let \(R^n\) be the process of observing \(X^n\) when moving randomly along the tree. We denote the corresponding transition kernels by \(p_n\) (for \(X^n\)) and \(p_{R^n}\), respectively. Moreover, let \(Z^n\) be the process of empirical measures based on \(X^n\), which has state space \(P(E^n)\), \(n = 1, 2, ...\). Our goal is to find sufficient conditions for \(X^n\) (via \(R^n\)), such that the process of empirical measures \(Z^n\) converges, and to characterize the limit process. We first recall some notation.

Remark 2.2 (Notation) Throughout this chapter, we will consider a complete and separable metric space \((E,r)\). The space of (continuous,) real-valued, bounded functions on \(E\) is denoted by \(B(E)(C_b(E))\). Weak convergence is indicated by \(\Rightarrow\). Given the functions \(f : [0, \infty) \to E_1\) and \(\eta : E_1 \to E_2\), we write (abusing notation) \(\eta \circ f = \eta(f)\) for the function \(\eta \circ f : t \mapsto \eta(f(t))\). If \((E_1,r_1),(E_2,r_2)\) are two metric spaces, \(\eta : E_1 \to E_2\) is measurable, and \(\nu \in P(E_1)\), we define the image measure of \(\nu\) under \(\eta\) by \(\eta_*\nu \in P(E_2)\), i.e. \(\eta_*\nu(A_2) = \nu(\eta^{-1}(A_2))\). Sometimes, we write \((z, \varphi) := \int \varphi dz\) for \(z \in P(E)\) and \(\varphi \in B(E)\). For \(f \in C_b(E)\) we write \(||f|| := \sup_{x \in E} |f(x)|\).

We need two more notions before we can proceed to the limit theorem.

Definition 2.3 (Feller property, compact containment condition) Recall that \((E,r)\) is complete and separable.

1. A Markov process \(X = (X_t)_{t \geq 0}\) with state space \(E\) and càdlàg-paths satisfies the Feller property, iff

   (i) \(X_t \xrightarrow{t \to 0} X_0\) and

   (ii) the map \(x \mapsto E[f(X_t)|X_0 = x]\) is continuous for all \(f \in C_b(E), t \geq 0\) and all \(x \in E\).

   Equivalently, let \((S_t)_{t \geq 0}\) be the semigroup of \(X\), i.e. \(S_tf(x) = E[f(X_t)|X_0 = x]\). Then, \(X\) is a Feller-process iff \((S_t)_{t \geq 0}\) is a Feller semigroup.

2. A Feller semigroup is a family of operators \((S_t)_{t \geq 0}, S_t : C_b(E) \to C_b(E)\) with

   (i) \(S_tf(x) \xrightarrow{t \to 0} f(x)\) for all \(x \in E\) and \(f \in C_b(E)\) and

   (ii) \(S_tf \in C_b(E)\) if \(f \in C_b(E)\).

We say that
(iii) \((S_t)_{t \geq 0}\) is a contraction iff \(||S_t f|| \leq ||f||\) and

(iv) \((S_t)_{t \geq 0}\) is strongly continuous iff \(||S_t f - f|| \xrightarrow{t \to 0} 0\).

We say that an operator \(G : \mathcal{D}(G) \subseteq C_b(E) \to C_b(E)\) generates a strongly continuous semigroup \((S_t)_{t \geq 0}\) if

\[
G f(x) := \lim_{t \to 0} \frac{1}{t} (S_t f(x) - f(x))
\]

for all \(f \in C_b(E)\) for which the limit in \((2.2)\) exists.

3. For a sequence \((X^n_t)_{t \geq 0}, (X^n_1)_{t \geq 0}, \ldots\) of \(E\)-valued stochastic processes, we say that the compact containment condition (in \(E\)) holds, if for every \(\varepsilon > 0\) and \(T \geq 0\) there is a compact set \(K_{\varepsilon,T} \subseteq E\) such that

\[
\sup_{n=1,2,\ldots} \mathbf{P}(X^n_t \in K_{\varepsilon,T} \text{ for all } 0 \leq t \leq T) < \varepsilon.
\]

Remark 2.4 (Properties of the generator) Recall that for the generator \(G\) of a Feller semigroup, the domain \(\mathcal{D}(G)\) is dense in \(C_b(E)\) (with respect to the topology of uniform convergence on compact sets) [EK86, Corollary 1.1.6], and if \((E,r)\) is locally compact, every Feller semigroup is a strongly continuous contraction semigroup [Kal02, Theorem 17.6] and is uniquely characterized by its generator [Kal02, Lemma 17.5].

Now we can formulate our main result.

Theorem 2.5 (Convergence of \(Z^n\)) Let \(X^n, R^n, Z^n\) be as above, \(n = 1, 2, \ldots\) Moreover, let \(\tilde{R}_n := (\tilde{R}_n^0) \subseteq (\tilde{R}_n^0) \subseteq 0\), and \(\tilde{Z}_n := (\tilde{Z}_n^0) \subseteq (\tilde{Z}_n^0) \subseteq 0\), \(n = 1, 2, \ldots\) Assume that \(\eta_n(X^n_0) \xrightarrow{n \to \infty} \nu \in \mathcal{P}(E)\) and that the compact containment condition holds for \(\eta_1(\tilde{R}_1), \eta_2(\tilde{R}_2), \ldots\)

In addition, assume that there is a linear operator \(G_{R} : \mathcal{D}(G_{R}) \subseteq C_b(E) \to C_b(E)\), which generates a strongly continuous contraction semigroup, and such that \(\mathcal{D}(G_{R})\) contains an algebra \(\Pi\) that separates points. Furthermore, for each \(\varphi \in \mathcal{D}(G_{R})\), there is a sequence \(\varphi_1 \in \mathcal{B}(E^1), \varphi_2 \in \mathcal{B}(E^2), \ldots\) such that \(\sup_{n=1,2,\ldots} ||\varphi_n|| < \infty\) and

\[
\lim_{n \to \infty} \sup_{x \in E^n} |\varphi \circ \eta_n(x) - \varphi_n(x)| = 0, \tag{2.3}
\]

\[
\lim_{n \to \infty} \sup_{x \in E^n} |(G_{R}\varphi) \circ \eta_n(x) - G_{\tilde{R}_n}\varphi_n(x)| = 0, \tag{2.4}
\]

where

\[
G_{\tilde{R}_n}\varphi(x) := n \cdot \mathbf{E}[\varphi(\tilde{R}_n^{0:n}) - \varphi(\tilde{R}_n^{0:n})]|\tilde{R}_n^{0:n} = x].
\]

Then, there is an \(E\)-valued Feller process \(R = (R_t)_{t \geq 0}\) with \(R_0 \sim \nu\) and generator \(G_{R}\) with \(\eta_n(\tilde{R}_n^{0:n}) \xrightarrow{n \to \infty} R\), and a \(\mathcal{P}(E)\)-valued stochastic process \(Z = (Z_t)_{t \geq 0}\) such that \(\eta_1(Z_n) \xrightarrow{n \to \infty} Z\) with \(Z_0 \sim \delta_{R_0} \in \mathcal{P}(\mathcal{P}(E))\).

Moreover, if \(\nu = \delta_x\) for \(x \in E\), then \(Z_t = \delta_{E(x)}\).
Remark 2.6 (Convergence, Deterministic limit) 1. Under the assumptions given above, the convergence $\eta^n \circ R^n \rightarrow \mathbb{R}$ was shown in [EK86, Corollary 4.8.9].

2. As the Theorem shows, the limit process of empirical measures $Z$ is deterministic (if the initial distribution is a Dirac-measure). The heuristic behind this result is that two distinct values $X^n_\sigma, X^n_\tau$ with $\sigma, \tau \in \mathcal{T}_{[n]}$ have already evolved independently for $O(n)$ steps. Hence, $\tilde{Z}_i^n$ is approximately given by the empirical measure of $2^{[n]}$ independent processes, which leads to a deterministic limit. This argument will be made precise below.

We now give two simple examples for normal and Poisson convergence.

Example 2.7 1. Let $(Y_\sigma)_{\sigma \in \mathcal{\mathcal{T}}}$ be a family of real-valued, independent and identically distributed random variables with $E[Y_\sigma] = 0$, $\text{Var}[Y_\sigma] = 1$. Moreover, let $X^n_0 := 0$ and

$$(X^n_0, X^n_1) := (X^n_\sigma + \frac{1}{\sqrt{n}} Y_\sigma, X_\sigma - \frac{1}{\sqrt{n}} Y_\sigma).$$

In other words, the states of the two children of $\sigma$ are a pair of dependent random variables. Then, the process $R^n = (R^n_t)_{t=0,1,2,...}$ can be written as

$$R^n_t = \frac{1}{\sqrt{n}} \sum_{k=0}^{t-1} \tilde{Y}_k,$$

where $(\tilde{Y}_k)_{k=0,1,2,...}$ are independent and identically distributed with

$$(\tilde{Y}_k), P = \frac{1}{2}(Y_\sigma), P + \frac{1}{2}(-Y_\sigma), P,$$

a mixture of the distributions of $Y_\sigma$ and $-Y_\sigma$. Donsker’s Theorem (see e.g. [Kal02, Theorem 14.9]) yields the convergence $R^n \rightarrow \mathbb{W}$, the standard Brownian motion $\mathbb{W}$. Our Theorem now says that the limiting process $Z$ is the law of $\mathbb{W}$, so we find that $Z = (\mathcal{N}(0,t))_{t \geq 0}$, where $\mathcal{N}(0,t)$ is the normal distribution with mean 0 and variance $t$.

2. Let $(Y^n_\sigma)_{\sigma \in \mathcal{T}}$ be a family of independent, identically distributed random variables with values in $\mathbb{Z}_+$ and

$$P[Y^n_\sigma > 0] = 2\lambda/n + o(1/n), \quad P[Y^n_\sigma > 1] = o(1/n).$$

Moreover, let $X^n_0 := 0$ and $(X^n_0, X^n_2) := (X^n_\sigma, X^n_\sigma + Y^n_\sigma)$. In other words, the state of the left child equals the state of its parent while the state of the right child has a small probability of having increased by 1. Then, the process $\mathcal{R}^n = (R^n_t)_{t=0,1,2,...}$ can be written as

$$R^n_t = \frac{d}{\sqrt{n}} \sum_{k=0}^{t-1} \tilde{Y}_k, \quad \text{where } (\tilde{Y}_k^n)_{k=0,1,2,...} \text{ are independent and identically distributed with } (\tilde{Y}_k^n), P = \frac{1}{2} \delta_0 + \frac{1}{2} (Y^n_\sigma), P,$$

equiv

$$P[\tilde{Y}_k^n > 0] = \lambda/n + o(1/n), \quad P[\tilde{Y}_k^n > 1] = o(1/n).$$
Classical convergence results (see e.g. [Kal02, Theorem 5.7]) then show that \( \tilde{R}^n \) converges weakly to a Poisson process with rate \( \lambda \). Consequently, we then have that \( Z = (Z_t)_{t \geq 0} \) with \( Z_t = \text{Poi}(\lambda t) \) by the above Theorem.

2.3 Proof of the main result

Throughout this section, we build on the same assumptions as in Theorem 2.5. Furthermore, we will replace \( \eta^n(\tilde{R}^n) \) by \( \tilde{R}^n \) and \( \eta^n \tilde{Z}^n \) by \( \tilde{Z}^n \) in the sequel (and similarly for the processes without \( \sim \)) for the sake of readability.

The proof of Theorem 2.5 consists of showing two assertions which then yield the result by Theorem 3.7.8 from [EK86]:

1. The sequence \( (\tilde{Z}^n)_{n=1,2,...} \) is tight.

2. The finite-dimensional distributions of \( (\tilde{Z}^n)_{n=1,2,...} \) converge.

For 2. we will prove in Lemma 2.12 that \( \tilde{Z}^n \xrightarrow{n \to \infty} \delta_{L(R_t)} \) holds for all \( t \geq 0 \). We will see in Section 2.3.4 that this is sufficient for the finite-dimensional distributions to converge. As for 1. we use Jakubowski’s tightness criterion, which is recalled in Proposition 2.10 in the next section. For this criterion, we have to show that (i) \( \tilde{Z}^n \) satisfies the compact containment condition in \( \mathcal{P}(E) \) (see Definition 2.3) and (ii) that the sequence \( (\langle \tilde{Z}_t^1, \varphi \rangle)_{t \geq 0}, (\langle \tilde{Z}_t^2, \varphi \rangle)_{t \geq 0}, \ldots \) is tight for all \( \varphi \in \Pi \subset C_b(E) \), where \( \Pi \) is a vector space which separates points. (i) will be resolved in Lemma 2.13, while (ii) is a result in Lemma 2.14. Hence, we are done once we have shown Lemma 2.12, 2.13 and 2.14.

In the following section we state auxiliary results needed for the proof of the two central steps above. Thereafter, we prove the convergence of the finite-dimensional distributions of \( (\tilde{Z}^n)_{n=1,2,...} \), which is followed by a section dealing with the tightness of this sequence. In the last section we put things together and prove Theorem 2.5.

2.3.1 Auxiliary Results

At first we state some results about random measures. These will imply that the sequence \( (\tilde{Z}^n)_{n=1,2,...} \) has a deterministic limit.

Definition 2.8 (Moment measures) For a random variable \( Z \) taking values in \( \mathcal{P}(E) \) and \( k = 1,2,... \), there is a uniquely determined measure \( \mu^{(k)} \) on \( \mathcal{B}(E^k) \) such that

\[
\mathbb{E}[Z(A_1) \cdots Z(A_k)] = \mu^{(k)}(A_1 \times \cdots \times A_k)
\]

for \( A_1, \ldots, A_k \in \mathcal{B}(E) \). This measure is called the \( k \)-th moment measure. Equivalently, \( \mu^{(k)} \) is the unique measure such that \( \mathbb{E}[\langle Z, \varphi_1 \rangle \cdots \langle Z, \varphi_k \rangle] = \langle \mu^{(k)}, \varphi_1 \cdots \varphi_k \rangle \).
Lemma 2.9 (Characterization of deterministic random measures) Let $Z$ be a random variable taking values in $\mathcal{P}(E)$ with the first two moment measures $\mu := \mu^{(1)}$ and $\mu^{(2)}$. Then the following assertions are equivalent:

1. There is $\nu \in \mathcal{P}(E)$ with $Z = \nu$, almost surely.
2. The second moment measure has product-form, i.e. $\mu^{(2)} = \mu \otimes \mu$ which is equivalent to
   \[ \mathbb{E}[(Z, \varphi_1) \cdot (Z, \varphi_2)] = \langle \mu, \varphi_1 \rangle \cdot \langle \mu, \varphi_2 \rangle \]
   for all $\varphi_1, \varphi_2 \in B(E)$.
3. $\mathbb{E}[(Z, \varphi)^2] = \langle \mu, \varphi \rangle^2$ holds for all $\varphi \in B(E)$.

In either case, $\mu = \nu$.

Proof: 1. $\Rightarrow$ 2.: This is clear because of $\mathbb{E}[Z(A)] = \nu(A)$, i.e. $\mu = \nu$. Moreover, $\mathbb{E}[Z(A_1)Z(A_2)] = \nu(A_1)\nu(A_2) = \mu(A_1)\mu(A_2) = \mu \otimes \mu(A_1 \times A_2)$.

2. $\Rightarrow$ 1.: Since the second moment-measure has product form, for any measurable $A \subseteq E$, $\text{Var}[Z(A)] = \mathbb{E}[Z(A)Z(A)] - \mathbb{E}[Z(A)]^2 = \mu^{(2)}(A \times A) - (\mu(A))^2 = 0$, i.e. the random variable $Z(A)$ has zero variance and therefore is deterministic. In particular, $Z(A) = \mathbb{E}[Z(A)] = \mu(A)$ and the assertion follows with $\nu = \mu$.

3. $\Rightarrow$ 2.: For $\varphi_1, \varphi_2 \in B(E)$ when using 3. we have
   \[ \mathbb{E}[(Z, \varphi_1 + \varphi_2)^2] = \langle \mu, \varphi_1 + \varphi_2 \rangle^2 = \langle \mu, \varphi_1 \rangle^2 + 2\langle \mu, \varphi_1 \rangle \langle \mu, \varphi_2 \rangle + \langle \mu, \varphi_2 \rangle^2. \]

On the other hand applying linearity we obtain
   \[ \mathbb{E}[(Z, \varphi_1 + \varphi_2)^2] = \mathbb{E}[(Z, \varphi_1)^2 + 2(Z, \varphi_1)(Z, \varphi_2) + (Z, \varphi_2)^2] = \langle \mu, \varphi_1 \rangle^2 + 2\mathbb{E}[(Z, \varphi_1)(Z, \varphi_2)] + \langle \mu, \varphi_2 \rangle^2. \]

Comparing these two expressions yields the result.

Next, we recall Jakubowski’s tightness criterion from [Jak86]; see also [Daw93, Theorem 3.6.4]. We will make use of this characterization when proving the tightness of $(\widetilde{Z}^n)_{n=1,2,\ldots}$.

Proposition 2.10 (Jakubowski’s tightness criterion) Assume the family $\Pi \subseteq C_b(E)$ is a vector space that separates points. A sequence $\mathcal{Z}^1 = (Z^1_t)_{t \geq 0}, \mathcal{Z}^2 = (Z^2_t)_{t \geq 0}, \ldots$ of $\mathcal{P}(E)$-valued processes with càdlàg-paths is tight if the following holds:

1. $(\mathcal{Z}^n)_{n=1,2,\ldots}$ satisfies the compact containment condition.
2. For every $f \in \Pi$, the sequence $(f(\mathcal{Z}^n))_{n=1,2,\ldots}$ with $f(\mathcal{Z}^n) = (\langle Z^n, f \rangle_t)_{t \geq 0}$ is tight.
We end this section by giving some basic relationships between the processes $\tilde{R}_n$ and $\tilde{Z}_n$, which will be used frequently. Some more refined relationships will be given in the proof of Lemma 2.12. Let $\varphi \in C_b(E)$. Then,

$$
E[\langle \tilde{Z}_n^t, \varphi \rangle] = E[\langle Z_n^t, \varphi \rangle] = E [\frac{1}{2^{\lfloor nt \rfloor}} \sum_{\sigma \in \mathcal{T}_{\lfloor nt \rfloor}} \langle \delta_{X^n_\sigma}, \varphi \rangle] = E [\frac{1}{2^{\lfloor nt \rfloor}} \sum_{\sigma \in \mathcal{T}_{\lfloor nt \rfloor}} \varphi(X^n_\sigma)] \quad (2.5)
$$

Similarly, we write

$$
\langle Z^n_k, \varphi \rangle = 2^{-k} \sum_{\sigma \in \mathcal{T}_k} \varphi(X^n_\sigma) = E[\varphi(R^n_k)|Z^n_k],
$$

and $\langle Z^n_k, \varphi \rangle = E[\varphi(R^n_k)|Z^n_k, Z^n_{k-1}]$ holds by the same argument, such that

$$
E[\langle Z^n_k, \varphi \rangle|Z^n_{k-1}] = E[\varphi(R^n_k)|Z^n_k, Z^n_{k-1}] = E[\varphi(R^n_k)|Z^n_{k-1}]. \quad (2.7)
$$

### 2.3.2 Convergence of the finite-dimensional distributions of $(\tilde{Z}_n)_{n \geq 1}$

We start with a fundamental observation, which is based on the fact that two random leaves from $\mathcal{T}_n$ have a most recent common ancestor node which is close to the root. The heuristics is that due to the scaling they have evolved independently for such a long time that their most recent common ancestor is approximated by the root.

Recall that by [EK86, Corollary 4.8.9] we already have that $\tilde{R}_n \xrightarrow{n \to \infty} R$ for a Feller-(hence càdlàg)-process $R$.

**Lemma 2.11** (Convergence to the initial state) Assume that $\tilde{R}_n \xrightarrow{n \to \infty} R$ for a càdlàg-process $R = (R_t)_{t \geq 0}$ with state space $E$. Then, the following holds:

1. Let $\sigma_1, \ldots, \sigma_k \in \mathcal{T}$ be fixed. Then,

   $$(X^n_{\sigma_i})_{i=1, \ldots, k} \xrightarrow{n \to \infty} (R_0)_{i=1, \ldots, k}$$

   in probability.

2. Let $\Sigma_1^n, \Sigma_2^n$ be two vertices, chosen uniformly at random from $\mathcal{T}_{\lfloor nt \rfloor}$. Then,

   $$(X^n_{\Sigma_1^n \wedge \Sigma_2^n}, X^n_{(\Sigma_1^n \wedge \Sigma_2^n)0}, X^n_{(\Sigma_1^n \wedge \Sigma_2^n)1}) \xrightarrow{n \to \infty} (R_0, R_0, R_0)$$

   in probability.
Proof: Recall that for the (independent) random walk \((\Sigma_k)_{k=0,1,...}\) on \(\mathcal{T}\) we have that \(R^n_k = X^n_{\Sigma_k}\). It suffices to prove the result for deterministic \(R_0 \in E\). By assumption, for all \(m \in \mathbb{N}\),
\[
P(r(R^n_m, R_0) > \varepsilon) = P(r(\widetilde{R}^n_{m/n}, R_0) > \varepsilon) \xrightarrow{n \to \infty} 0, \tag{2.8}
\]
since \(\mathcal{R}\) has càdlàg-paths.
1. Let \(\sigma \in \mathcal{T}_m\). We have for all \(\varepsilon > 0\) that
\[
P(r(R^n_m, X^n_{\sigma}) \leq \varepsilon) \geq P(r(R^n_m, X^n_{\sigma}) \leq \varepsilon, R^n_m = X^n_{\sigma}) = P(R^n_m = X^n_{\sigma}) \geq 2^{-m} \tag{2.9}
\]
since the random walk \((\Sigma_m)_{m=0,1,2,...}\) along \(\mathcal{T}\) has a chance of \(2^{-m}\) to pass through vertex \(\sigma \in \mathcal{T}_m\). Now, assume that the assertion does not hold, i.e. \(X^n_{\sigma}\) does not converge to \(R_0\) in probability. Let \(\varepsilon > 0\) such that \(P(r(X^n_{\sigma}, R_0) > \varepsilon) > \varepsilon\) for all \(n\). Then, for such an \(\varepsilon > 0\) we obtain
\[
P(r(R^n_m, R_0) > \varepsilon) \geq P(r(R^n_m, R_0) > \varepsilon, R^n_m = X^n_{\sigma}) \geq P(r(X^n_{\sigma}, R_0) > \varepsilon, \Sigma_m = \sigma) = P(r(X^n_{\sigma}, R_0) > \varepsilon) \cdot P(\Sigma_m = \sigma) \geq \varepsilon 2^{-m},
\]
where we used inequality (2.9) in the last line. This result is in contradiction to (2.8). Hence, 1. follows.
2. Let \(\varepsilon > 0\) and \(m\) be large enough such that \(2^{-m} < 2\varepsilon\) holds. From 1., we have that \((X^n_{\sigma})_{\sigma \in \bigcup_{k=0}^n \mathcal{T}_k} \xrightarrow{n \to \infty} (R_0)_{\sigma \in \bigcup_{k=0}^n \mathcal{T}_k}\) in probability. Moreover, for \(n > m\) and \(\Sigma^n_1, \Sigma^n_2\) two randomly chosen vertices of \(\mathcal{T}_{[nt]}\), we obtain
\[
P(\Sigma^n_1 \land \Sigma^n_2 \in \bigcup_{k=0}^m \mathcal{T}_k) = \sum_{k=0}^m 2^{-(k+1)} = 1 - 2^{-(m+1)} > 1 - \varepsilon.
\]
Hence, we can write
\[
P(r(X^n_{\Sigma^n_1 \land \Sigma^n_2}, R_0) > \varepsilon) \leq P\left(r(X^n_{\Sigma^n_1 \land \Sigma^n_2}, R_0) > \varepsilon, \Sigma^n_1 \land \Sigma^n_2 \in \bigcup_{k=0}^m \mathcal{T}_k\right) + P\left(\Sigma^n_1 \land \Sigma^n_2 \notin \bigcup_{k=0}^m \mathcal{T}_k\right)
\leq P\left(\sup_{\sigma \in \bigcup_{k=0}^m \mathcal{T}_k} r(X^n_{\sigma}, R_0) > \varepsilon\right) + P\left(\Sigma^n_1 \land \Sigma^n_2 \notin \bigcup_{k=0}^m \mathcal{T}_k\right) \xrightarrow{n \to \infty} 2^{-(m+1)} < \varepsilon
\]
by the first assertion. Thus, we have shown that \(X^n_{\Sigma^n_1 \land \Sigma^n_2} \xrightarrow{n \to \infty} R_0\) in probability. By the same arguments, we also find that \(X^n_{(\Sigma^n_1 \land \Sigma^n_2)_1} \xrightarrow{n \to \infty} R_0\) in probability for \(i = 0, 1\) and we are done. \(\square\)
2.3 Proof of the main result

**Lemma 2.12** (Convergence of $\tilde{Z}^n$ at fixed times) Consider the same situation as in Theorem 2.5 and let $t \geq 0$. If $\nu = \delta_x$ for some $x \in E$, we have that $\tilde{Z}^n \overset{n \to \infty}{\longrightarrow} \delta_{E(R_t)}$.

**Proof:** Note that the assertion holds once we show that

$$\langle \tilde{Z}^n_t, \varphi \rangle \overset{n \to \infty}{\longrightarrow} E[\varphi(R_t)]$$

(2.10)

for all $\varphi \in C_b(E)$ because this set of functions is separating. Indeed, the family $((\tilde{Z}^n_t, \varphi))_{n=1,2,\ldots}$ is tight by the boundedness of $\varphi$.

Furthermore, we already know from (2.5) that $E[\langle \tilde{Z}^n_t, \varphi \rangle] = E[\varphi(\tilde{R}^n_t)] \overset{n \to \infty}{\longrightarrow} E[\varphi(R_t)]$. Hence, it remains to show that

$$\text{Var}[\langle \tilde{Z}^n_t, \varphi \rangle] \overset{n \to \infty}{\longrightarrow} 0$$

(2.11)

which then implies (2.10) due to Lemma 2.9. For this, consider two randomly picked vertices $\Sigma_1, \Sigma_2 \in T_{[nt]}$ with $\Sigma_1 \neq \Sigma_2$. Then, without loss of generality we assume that

$$\pi_{[\Sigma_1 \land \Sigma_2]+1}X_{\Sigma_1} = X_{(\Sigma_1 \land \Sigma_2)0} \quad \text{and} \quad \pi_{[\Sigma_1 \land \Sigma_2]+1}X_{\Sigma_2} = X_{(\Sigma_1 \land \Sigma_2)1}$$

such that (setting $\delta^n_t := 2^{-[nt]}(E[\varphi^2(X^n_{\Sigma_1})] - \varphi(X^n_{\Sigma_1})\varphi(X^n_{\Sigma_2})))$

$$E[\langle \tilde{Z}^n_t, \varphi \rangle^2] = \frac{1}{2^{[nt]}} \sum_{\sigma_1, \sigma_2 \in T_{[nt]}} E[\varphi(X^n_{\Sigma_1})\varphi(X^n_{\Sigma_2})]
\begin{align*}
&= \frac{2^{[nt]} - 1}{2^{[nt]}(2^{[nt]} - 1)} \sum_{\sigma_1 \neq \sigma_2} E[\varphi(X^n_{\Sigma_1})\varphi(X^n_{\Sigma_2})] + \frac{1}{2^{[nt]}} \sum_{\sigma_1 \in T_{[nt]}} E[\varphi(X^n_{\Sigma_1})] \\
&= E[\varphi(X^n_{\Sigma_1})\varphi(X^n_{\Sigma_2})] + \delta^n_t \\
&= E[E[\varphi(X^n_{\Sigma_1})\varphi(X^n_{\Sigma_2})]X^n_{(\Sigma_1 \land \Sigma_2)0}, X^n_{(\Sigma_1 \land \Sigma_2)1}] + \delta^n_t \\
&= E[E[\varphi(X^n_{\Sigma_1})X^n_{\pi_{[\Sigma_1 \land \Sigma_2]+1}\Sigma_1}] \cdot E[\varphi(X^n_{\Sigma_2})X^n_{\pi_{[\Sigma_1 \land \Sigma_2]+1}\Sigma_2}]] + \delta^n_t \\
&= E[E[\varphi(R^n_{[nt]-[\Sigma_1 \land \Sigma_2]-1})R^n_0 = X^n_{\pi_{[\Sigma_1 \land \Sigma_2]+1}\Sigma_1}] \\
&\cdot E[E[\varphi(R^n_{[nt]-[\Sigma_1 \land \Sigma_2]-1})R^n_0 = X^n_{\pi_{[\Sigma_1 \land \Sigma_2]+1}\Sigma_2}]] + \delta^n_t \\
&= E[\varphi(\tilde{R}^n_t)]^2 + \varepsilon^n_t = E[\langle \tilde{Z}^n_t, \varphi \rangle^2] + \varepsilon^n_t
\end{align*}$$

for

$$\varepsilon^n_t := E[E[\varphi(R^n_{[nt]-[\Sigma_1 \land \Sigma_2]-1}) \cdot E[\varphi(R^n_{[nt]-[\Sigma_1 \land \Sigma_2]-1})]R^n_0 = X^n_{\pi_{[\Sigma_1 \land \Sigma_2]+1}\Sigma_1}] \\
\cdot E[E[\varphi(R^n_{[nt]-[\Sigma_1 \land \Sigma_2]-1}) \cdot E[\varphi(R^n_{[nt]-[\Sigma_1 \land \Sigma_2]-1})]R^n_0 = X^n_{\pi_{[\Sigma_1 \land \Sigma_2]+1}\Sigma_2}]] + \delta^n_t.$$  

(2.12)
Hence, we must show $\varepsilon_t^n \xrightarrow{n \to \infty} 0$ for (2.11), which is implied by the boundedness of $\varphi$ (showing convergence to 0 of the last term of (2.12)) and by the Cauchy-Schwarz inequality once we show that for $i \in \{0, 1\}$

$$E[\varphi(R^n_t) | R_0 = X^n_{\pi_{[\Pi, \Sigma_2]} + 1} | R_0 = x] \xrightarrow{n \to \infty} E[\varphi(R_t) | R_0 = x]$$

(2.13)

in probability. Indeed, we already know from Lemma 2.11 that $X^n_{\pi_{[\Pi, \Sigma_2]} + 1} \xrightarrow{n \to \infty} x$ in probability. This and since $R$ has càdlàg-paths, the convergence of semigroups, Theorem 1.6.1 from [EK86] (see also Theorem 1.6.5 and Remark 4.8.8) and the strong continuity of the semigroup for $R$ yields

$$\left| E[\varphi(R^n_{nt} - [\Pi, \Sigma_2] - 1)] | R^n_0 = X^n_{\pi_{[\Pi, \Sigma_2]} + 1} | - E[\varphi(R_t) | R_0 = x] \right|$$

$$\leq \left| E[\varphi(R^n_{nt} - [\Pi, \Sigma_2] - 1)] | R^n_0 = X^n_{\pi_{[\Pi, \Sigma_2]} + 1} | - E[\varphi(R_n) | R_0 = X^n_{\pi_{[\Pi, \Sigma_2]} + 1} ] \right|$$

$$+ \left| E[\varphi(R^n_t) | R^n_0 = X^n_{\pi_{[\Pi, \Sigma_2]} + 1} ] - E[\varphi(R_t) | R_0 = X^n_{\pi_{[\Pi, \Sigma_2]} + 1} ] \right|$$

$$\xrightarrow{n \to \infty} 0$$

in probability, which shows (2.13). This completes the proof.

2.3.3 Tightness of $(\tilde{Z}^n)_{n \geq 1}$

It remains to prove the tightness of the sequence $(\tilde{Z}^n)_{n \geq 1}$. We therefore check the conditions of Jakubowski’s tightness criterion from Proposition 2.10. We start with the proof of the compact containment condition for $(\tilde{Z}^n)_{n \geq 1}$.

**Lemma 2.13** (Compact containment condition for $\tilde{Z}^n$) If $(\tilde{R}^n)_{n=1,2,...}$ satisfies the compact containment condition (in $E$), then $(\tilde{Z}^n)_{n=1,2,...}$ satisfies the compact containment condition (in $P(E)$) as well.

**Proof:** For all $\delta > 0$, let $K_\delta \subseteq E$ be compact and such that

$$\sup_{n=1,2,...} P(\tilde{R}^n_t \notin K_\delta \text{ for some } 0 \leq t \leq T) < \delta.$$ 

Since there are only $([nT] + 1)$ different states of $\tilde{R}^n$ in $[0, T]$ (including the root state) this is the same as saying that for the discrete set $S_n$ of $\{0, \frac{1}{n}, ..., \frac{Tn}{n}\}$-valued random times,

$$\sup_{n=1,2,...} \sup_{\tau_n \in S_n} P(\tilde{R}^n_{\tau_n} \notin K_\delta) < \delta.$$ 

For $\delta, \varepsilon > 0$, set

$$L_\delta := \{ \mu \in P(E) : \mu(K^c_\delta) < \delta \}, \quad L := \bigcap_{n=1}^{\infty} L_{\varepsilon 2^{-n}}.$$
Then, the closure of $L$ is a compact subset of $\mathcal{P}(E)$ by Prohorov’s Theorem. Moreover, by Markov’s inequality and equality (2.5) we obtain
\[
\sup_{n=1,2,\ldots} \sup_{\tau_n \in \mathcal{S}_n} P(\tilde{Z}^n_{\tau_n}(K_{2,4-t}^c) > \varepsilon 2^{-t}) \leq \sup_{n=1,2,\ldots} \sup_{\tau_n \in \mathcal{S}_n} \frac{1}{\varepsilon^{2-t}} E[(\tilde{Z}^n_{\tau_n}, 1_{K_{2,4-t}^c})]
\]
\[
= \sup_{n=1,2,\ldots} \sup_{\tau_n \in \mathcal{S}_n} \frac{1}{\varepsilon^{2-t}} E[1_{K_{2,4-t}^c}(\tilde{R}^n_{\tau_n})] = \sup_{n=1,2,\ldots} \sup_{\tau_n \in \mathcal{S}_n} \frac{1}{\varepsilon^{2-t}} P(\tilde{R}^n_{\tau_n} \notin K_{2,4-t}) < \varepsilon 2^{-t},
\]
and hence,
\[
\sup_{n=1,2,\ldots} P(\tilde{Z}^n_t \notin L \text{ for some } 0 \leq t \leq T) = \sup_{n=1,2,\ldots} \sup_{\tau_n \in \mathcal{S}_n} P(\tilde{Z}^n_{\tau_n} \notin L) \leq \sup_{n=1,2,\ldots} \sup_{\tau_n \in \mathcal{S}_n} \sum_{\ell=1}^{\infty} P(\tilde{Z}^n_{\tau_n}(K_{2,4-t}^c) > \varepsilon 2^{-t}) < \varepsilon,
\]
and the statement is proved.

The last step is to prove the second assertion in Jakubowski’ tightness criterion, namely the tightness of sequences of the form $(f(\tilde{Z}^n))_{n=1,2,\ldots}$ for $f \in \Pi$, where $\Pi$ is as in Theorem 2.5. This will be done by proving the weak convergence of an auxiliary martingale containing this sequence which then implies its tightness.

Lemma 2.14 (Martingale convergence) Consider the same situation as in Theorem 2.5 with $\nu = \delta_x$ for some $x \in E$. Let $\varphi \in \Pi$ and $\varphi_n \in B(E)$ such that $\|\varphi_n - \varphi\|_{n \rightarrow \infty} \rightarrow 0$. For $f_n(z) := \langle z, \varphi_n \rangle$, consider the mean-zero martingale $M^{n,\varphi_n} = (M^{n,\varphi_n}_t)_{t \geq 0}$, given by
\[
M^{n,\varphi_n}_t := f_n(Z^n_{[nt]}) - f_n(Z^n_0) - \sum_{k=1}^{[nt]} E[f_n(Z^n_k) - f_n(Z^n_{k-1})|Z^n_{k-1}] - \sum_{k=1}^{[nt]} E[\varphi(Z^n_k - Z^n_{k-1}, \varphi_n)|Z^n_{k-1}].
\]

Then, $M^{n,\varphi_n} \xrightarrow{n \rightarrow \infty} 0$ and $f_n(\tilde{Z}^n) \xrightarrow{n \rightarrow \infty} E[\varphi(R)]$. In particular, $(f_n(\tilde{Z}^n))_{n=1,2,\ldots}$ is tight.

Proof: We start by reformulating the martingale, using (2.6),
\[
M^{n,\varphi_n}_t = \langle Z^n_{[nt]}, \varphi_n \rangle - \langle Z^n_0, \varphi_n \rangle - \sum_{k=1}^{[nt]} E[\langle Z^n_k, \varphi_n \rangle - \langle Z^n_{k-1}, \varphi_n \rangle|Z^n_{k-1}] - \sum_{k=1}^{[nt]} E[\varphi_n(R^n_{[nt]})|Z^n_{[nt]}]
\]
\[
= \int_0^{t-1/n} n \cdot E[\varphi_n(R^n_{[nt]}) - \varphi_n(R^n_{[nt]})|Z^n_{[nt]}] ds
\]
\[
= \int_0^{t-1/n} n \cdot E[\varphi_n(\tilde{R}^n_{[nt]}) - \varphi_n(\tilde{R}^n_{[nt]})|Z^n_{[nt]}] ds.
\]

We now show that $M^{n,\varphi_n} \xrightarrow{n \rightarrow \infty} 0$. From Lemma 2.12, we already know that for fixed $t \geq 0$
\[
\tilde{Z}^n_t \xrightarrow{n \rightarrow \infty} \delta_{\mathcal{L}(R_t)}.
\]
We complement this by showing that for all $s \geq 0$

$$n \cdot \mathbb{E}\left[\varphi_n(\tilde{R}_{s+1/n}) - \varphi_n(\tilde{R}_s)\right] \xrightarrow{n \to \infty} \mathbb{E}[G_R \varphi(R_s)].$$

Note that the right hand side is deterministic. Indeed,

$$\mathbb{E}\left[n \cdot \mathbb{E}\left[\varphi_n(\tilde{R}_{s+1/n}) - \varphi_n(\tilde{R}_s)\right] - \mathbb{E}[G_R \varphi(R_s)]\right]$$

$$\leq \mathbb{E}\left[n \cdot \mathbb{E}\left[\varphi_n(\tilde{R}_{s+1/n}) - \varphi_n(\tilde{R}_s)\right] - \mathbb{E}[G_R \varphi(\tilde{R}_s)]\right]$$

$$+ \mathbb{E}\left[\mathbb{E}[G_R \varphi(\tilde{R}_s)] - \mathbb{E}[G_R \varphi(R_s)]\right] \xrightarrow{n \to \infty} 0$$

in probability, by (2.4) and Lemma 2.12, which shows that – using (2.6) –

$$\mathbb{E}[G_R \varphi(\tilde{R}_s)|\tilde{Z}_s^n] = \langle \tilde{Z}_s^n, G_R \varphi \rangle \xrightarrow{n \to \infty} \mathbb{E}[G_R \varphi(R_s)].$$

For every $t \geq 0$ and setting $\varphi(R_0) = 0$ without loss of generality, we now have that

$$M_t^{n,\varphi_n} \xrightarrow{n \to \infty} \mathbb{E}[\varphi(R_t)] - \int_0^t \mathbb{E}[G_R \varphi(R_s)] ds = 0. \quad (2.16)$$

Hence, we can write by Doob’s inequality

$$\mathbb{P}\left(\sup_{0 \leq s \leq t} |M_s^{n,\varphi_n}| > \varepsilon\right) \leq \frac{1}{\varepsilon} \mathbb{E}[|M_t^{n,\varphi_n}|] \xrightarrow{n \to \infty} 0, \quad (2.17)$$

since $M_t^{n,\varphi_n}$ is bounded in $n$ and convergence in (2.16) also holds in probability which proves $M_t^{n,\varphi_n} \xrightarrow{n \to \infty} 0$. For the second convergence, $f_n(\tilde{Z}^n) \xrightarrow{n \to \infty} \mathbb{E}[\varphi(\mathcal{R})]$, we use (2.14) and obtain

$$\mathbb{P}\left(\sup_{0 \leq s \leq t} |\langle \tilde{Z}_s^n, \varphi_n \rangle - \mathbb{E}[\varphi(R_s)]| > 3\varepsilon\right)$$

$$\leq \mathbb{P}\left(\sup_{0 \leq s \leq t} |M_s^{n,\varphi_n}| > \varepsilon\right)$$

$$+ \mathbb{P}\left(\int_0^{t-1/n} n \cdot \mathbb{E}[\varphi_n(\tilde{R}_{s+1/n}) - \varphi_n(\tilde{R}_s)|\tilde{Z}_s^n] ds - \mathbb{E}[G_R \varphi(R_s)] ds > \varepsilon\right)$$

$$+ \mathbb{P}\left(\int_{t-1/n}^t \mathbb{E}[G_R \varphi(R_s)] ds > \varepsilon\right) \xrightarrow{n \to \infty} 0$$

by (2.17) and (2.15). This finishes the proof. \qed
2.3.4 Proof of Theorem 2.5

Putting things together, we have shown

(i) $\tilde{Z}_t^n \Rightarrow \delta_{L(R_0)}$ by Lemma 2.12,

(ii) the compact containment condition of $\tilde{Z}^n$ in $\mathcal{P}(E)$ by Lemma 2.13 and

(iii) the tightness of $(f(\tilde{Z}^n))_{n=1,2,...}$ for any $f \in \Pi$ as in Theorem 2.5 by Lemma 2.14.

By (ii) and (iii) we can deduce that the sequence $(\tilde{Z}^n)_{n=1,2,...}$ is tight by Jakubowski’s tightness criterion given in Proposition 2.10.

As already mentioned (i) already implies the convergence of the finite-dimensional distributions of the sequence $(\tilde{Z}^n)_{n=1,2,...}$ because the right-hand side is deterministic by Lemma 2.9. By mixing over deterministic initial conditions $\nu = \delta_x$ and thus assuming arbitrary initial measures $\nu$ as claimed in Theorem 2.5 we obtain the general result of this finite-dimensional convergence.

Now, the tightness of the stochastic process $\tilde{Z}^n$ and the convergence of its finite-dimensional distributions implies the statement in Theorem 2.5 by applying [EK86, Theorem 3.7.8] and we are done.

2.4 Discussion

Briefly summarizing the results of this chapter we aimed to explain the paradox of stochasticity on the single cell level versus determinism on the population level. Therefore, we introduced a Markov process evolving on a binary tree. We allowed for independent dynamics along the two descendants of a vertex which represent randomly evolving individuals. Now, turning to the population level we analyzed the limit behavior of the empirical measure process of the $[nt]$—th level. Here, interestingly we obtain a deterministic limit distribution for $n \to \infty$ when started deterministically.

This result gives a theoretical explanation for the often observed randomness in single cells combined with reproducible population behavior. Furthermore, it allows for the analysis of high-throughput data where cell lineages are typically not identifiable. Here, our result provides a justification for the analysis of this kind of data in cases where inheritance of phenotypes is observed and seems to be crucial for the development of phenotypes. However, the concrete dynamics underlying the heterogeneous behavior cannot be identified by this method. Still, the stationary distribution of the population composition concerning the phenotypes is derivable from these information.
In this chapter we study the emergence of phenotypic heterogeneity due to social interactions amongst individuals which trigger regulation processes inside cells and thus lead to phenotypically heterogeneous populations. Several examples of this kind of non-genetic heterogeneity are known, one of the most popular being quorum sensing, [WB05]. In the following we study the scenario of a public good producing type in a bacterial colony. This can be related to the phenotypic behavior of the bacterium *Sinorhizobium meliloti*, the model organism of our collaborators at the Becker Lab at Synmikro associated to the Philipps-University of Marburg. Some individuals of *S. meliloti* produce so called Exopolysaccharides, a kind of sugar, and release it to their environment. There it works as a public good since it increases the local fitness and is available to all cells in a certain environment. We view this process as a form of altruism and/or cooperation. For this reason we are interested in the development of a population containing cooperatively behaving individuals.

The chapter is structured as follows. In Section 3.1 we first give an introduction to cooperation and review existing models based on spatial and stochastic interactions describing cooperative behavior. Thereafter, in Section 3.2 we define a interacting particle system which models a population with a public good producing type. Here, we also prove the existence of a Markov process possessing the described dynamics and study special cases of the interactions. In particular, we consider the case of type-dependent cooperation as opposed to type-independent cooperation. Furthermore, we distinguish between a altruistic and a cooperative model dependent on the concrete structure of the cooperation mechanism. The last part of this chapter, Section 3.3, analyzes the limit behavior of our model in the non-spatial case. The spatial setting will be investigated in the following Chapters 4-6.

The following setup is part of [CP16].
3.1 Introduction

In nature cooperative behavior amongst individuals is widely spread. It is observed in animals, e.g. [GW03, CB09] as well as in microorganisms, e.g. [Cre01, WL06]. While understanding this phenomenon by models, theoretical approaches introduced different interpretations and forms of cooperation, mostly within the area of game theory [Now06]. In all such approaches, a defector (or selfish) type tends to have more offspring, but there are cases in which defecting particles are outcompeted by the cooperator type. Although, in all models describing cooperation the question of extinction and survival of a type or the coexistence of several types are main subjects of the mathematical analysis, the frameworks for the theoretical studies may vary. While (stochastic) differential equations are mainly used for non-spatial systems (see for example [AS12, HJM15]), the theory of interacting particle systems provides a suitable setup for the analysis of models with local interactions between particles, [BK11, SS15, EL16]. In the following we define a model using the latter structure and terminology.

Investigations of models incorporating cooperation are interesting because of the following dichotomy: in non-spatial (well-mixed) situations, the whole population benefits from the cooperative behavior. If defectors have a higher fitness than cooperators, defectors always outcompete cooperators in the long run. However, if the system is truly spatial, cooperators can form clusters and thus use the benefit of cooperation in order to defend themselves against defectors, even though those might have higher (individual) fitness. This heuristic implies that only structured models can help to understand cooperative behavior in nature.

Due to the variety of interpretations of cooperative behavior there are different ways of implementing these mechanisms in a spatial context. In the field of population dynamics, Sturm and Swart [SS15] study an interacting particle system containing a cooperative-branching mechanism which can be understood as a sexual reproduction event. In [BK11], a branching-annihilating random walk is studied and again, a cooperation mechanism is interpreted as sexual reproduction. Different to that, the model introduced in [EL16] originates from the game-theoretical study of a two player game with different strategies where the strategies can be altruistic or selfish. Here, the altruistic strategies represent the cooperator type.

Various interacting particle systems which appear in the literature are monotone, i.e. two versions of the system, which start in configurations where one dominates the other, can be coupled such that this property holds for all times; see e.g. [SS15] for a monotone model mentioned above. For such processes, there exist several general results (cf. [Lig85]) which provide some useful techniques helping in the analysis. However, cooperation often leads to non-monotone interacting particle systems; compare with [BK11, EL16] and the one presented here. In our case the reason is that cooperators (or altruists) do not distinguish between non-cooperators and their own type which then contradicts the monotonicity condition.
The motivation for our model, as already mentioned above, came from studies of bacterial cells whose division (reproduction) rate is dependent on the types of neighboring particles. The idea is that there are two phenotypes (defector=0, cooperator=1), where only cooperators produce some public good which helps neighboring cells to reproduce. However, this production is costly which means that defectors will have a selective advantage over the cooperator type. The resulting model is a biased voter model with an additional cooperation mechanism. The main objective of our analysis is to study the long-time-behavior of such a model dependent on the parameters of the system which will be done in the following chapters.

In particular, in Chapter 4 we prove that the system clusters in one dimension independently of the parameter configuration. In Chapter 5 we study limit results in dimensions \( d \geq 3 \) in case of weakly interacting particles and finally, in Chapter 6 we derive a condition on the model such that cooperators survive in all dimensions \( d \geq 1 \). But, first of all we now introduce our model.

### 3.2 The model

Let \( V \) be a countable vertex set, and \( (a(u,v))_{u,v \in V} \) be a (not necessarily symmetric) Markov kernel from \( V \) to \( V \). Additionally, \( (b(u,(v,w)))_{u \in V, (v,w) \in V \times V} \) is a second Markov kernel mapping from \( V \) to \( V \times V \). In the following we are interested in a stochastic process \( X = ((X_t(u))_{u \in V})_{t \geq 0} \) with state space \( \{0,1\}^V \), where \( X_t(u) \in \{0,1\} \) is the type at site \( u \) at time \( t \). We call a particle in state 0 defector and a particle in state 1 cooperator. The dynamics of the interacting particle system, which is a Markov process, are (informally) as follows: for some \( \alpha, \gamma \geq 0 \) we have

**Reproduction:** A particle at site \( u \in V \) reproduces with rate \( a(u,v) \) to site \( v \), i.e. \( X(v) \) changes to \( X(u) \). (This mechanism is well-known from the voter model.)

**Selection:** If \( X(u) = 0 \) (i.e. there is a defector at site \( u \in V \)), it reproduces with additional rate \( \alpha \cdot a(u,v) \) to site \( v \), i.e. \( X(v) \) changes to 0. (A defector has a fitness advantage over cooperators by this additional chance to reproduce. This mechanism is well-known from the biased voter model.)

**Cooperation:** If \( X(u) = 1 \) (i.e. there is a cooperator at site \( u \in V \)), the individual at site \( v \) (no matter which state it has) reproduces to site \( w \) at rate \( \gamma \cdot b(u,(v,w)) \geq 0 \). (A cooperator at site \( u \) helps an individual at site \( v \) to reproduce to site \( w \).)

**Remark 3.1** (Interpretation) 1. **Selection:** Since cooperation imposes an energetic cost on cooperators, the non-cooperating individuals can use these free resources for reproduction processes. This leads to a fitness advantage which we describe with the parameter \( \alpha \).
2. **Cooperation**: The idea of the cooperation mechanism in our model is that each cooperator supports a neighboring individual, independent of its type, to reproduce to an adjacent site according to the Markov kernel $b$. A biological interpretation for this supportive interaction is a public good produced by cooperators and released to the environment helping the colony to expand. The corresponding interaction parameter is $\gamma$.

Below, we will deal with two situations, depending on whether $b(u,(v,u)) > 0$ or $b(u,(v,u)) = 0$. In the former case, we speak of an altruistic system, since a cooperator at site $u$ can help the particle at site $v$ to kill itself. In the latter case, we speak of a cooperative system because the effect of the supportive behavior cannot be used against the cooperator itself.

In order to uniquely define a Markov process, we will need the following assumption.

**Assumption 3.2** (Markov kernels) The Markov kernels $a(.,.)$ and $b(.,(.,.))$ satisfy

$$\sum_{u \in V} a(u,v) < \infty \text{ for all } v \in V$$  

(3.1)

and

$$\sum_{u,v \in V} b(u,(v,w)) < \infty \text{ for all } w \in V.$$  

(3.2)

**Remark 3.3** (Some special cases) A special case (the altruistic model) is when

$$b(u,(v,w)) = a(u,v) \cdot a(v,w).$$  

(3.3)

Then, (3.2) is implied by the assumption

$$\sup_{v \in V} \sum_{u \in V} a(u,v) < \infty,$$

which is stronger than (3.1). Furthermore, we study a similar case where $u$ cannot help $v$ to replace $u$ (the cooperative model), i.e. when $b(u,(v,u)) = 0$. More precisely, if $a(v,u) < 1$,

$$b(u,(v,w)) = a(u,v) \cdot \frac{a(v,w)1_{\{w \neq u\}}}{\sum_{w' \neq u} a(v,w')}.$$  

(3.4)

### 3.2.1 Existence and uniqueness of the process

In order to become more formal, we define the (pre-)generator of the process $X$ via its transition rates. Given $X \in \{0,1\}^V$, the transition rates $c(u,X)$ from $X$ to

$$X^u(v) = \begin{cases} 
X(v), & v \in V\{u\}; \\
1 - X(u), & v = u;
\end{cases}$$
are as follows:
If \( X(u) = 0 \), then
\[
c(u, X) = \sum_v a(v, u)X(v) + \gamma \sum_v X(v) \sum_w X(w)b(w, (v, u)). \tag{3.5}
\]
If \( X(u) = 1 \), then
\[
c(u, X) = \sum_v a(v, u)(1 - X(v))(1 + \alpha) + \gamma \sum_v (1 - X(v)) \sum_w X(w)b(w, (v, u)). \tag{3.6}
\]
Here, the first sum in \( c(u, X) \) represents the rates triggered by reproduction and selection whereas the last terms emerge from the cooperation mechanism.

The existence of a unique Markov process corresponding to the transition rates \( c(u, X) \) is guaranteed by standard existence results, see for example [Lig85, Chapter 1]. Precisely, we define the (pre-)generator \( \Omega \) of the process through
\[
(\Omega f)(X) = \sum_{u \in G} c(u, X)(f(X^u) - f(X)),
\]
where \( f \in \mathcal{D}(\Omega) \). Here, the domain of \( \Omega \) is given by
\[
\mathcal{D}(\Omega) := \{ f \in C(\{0,1\}^V) \text{ depends only on finitely many coordinates} \},
\]
with \( C(\{0,1\}^V) \) the set of continuous functions on \( \{0,1\}^V \). We note that \( \mathcal{D}(\Omega) \) is dense in \( C_b(\{0,1\}^V) \), the set of bounded continuous functions on \( \{0,1\}^V \), because of the Stone-Weierstrass-Theorem. We find the following general statement.

**Proposition 3.4** (Existence of unique Markov process) *If Assumption 3.2 holds, the transition rates \( c(\ldots) \) given in (3.5) and (3.6) define a unique Markov process \( X \) on \( \{0,1\}^V \). Moreover, the closure \( \overline{\Omega} \) of \( \Omega \) is the generator of \( X \).*

**Proof:** We need to show that the closure of \( \Omega \) in \( C(\{0,1\}^V) \) is a generator of a semigroup which then uniquely defines a Markov process (see for example [Lig85, Theorem 1.1.5]). In order to show this we follow [Lig85, Theorem 1.3.9] and check the following two conditions:

\[
sup_{u \in V} \sup_{X \in \{0,1\}^V} c(u, X) < \infty, \tag{3.7}
\]
\[
sup_{u \in V} \sum_{v \neq u} \tilde{c}_u(v) < \infty, \tag{3.8}
\]
where
\[
\tilde{c}_u(v) := \sup\{\|c(u, X_1) - c(u, X_2)\| : X_1(w) = X_2(w) \text{ for all } w \neq v\}
\]
measures the dependence of the transition rate $c(u, X)$ of the site $v \in V$ and $\| \cdot \|_T$ denotes the total variation norm.

Both inequalities follow from Assumption 3.2 and the definition of the transition rates $c(\cdot, \cdot)$. Using these we obtain for any $X \in \{0, 1\}^V$ and $u \in V$

$$c(u, X) \leq (1 + \alpha) \sum_{v \in V} a(v, u) + \gamma \sum_{v, w \in V} b(w, (v, u)) < \infty$$

showing (3.7). For (3.8), we note that $\tilde{c}_u(v) \neq 0$ only when either $a(v, u) > 0$ or $b(w, (v, u)) > 0$ or $b(v, (w, u)) > 0$ for some $w \in V$, i.e. whenever site $v$ has an influence on site $u$. Hence, for all $u \in V$ we obtain

$$\sum_{v \neq u} \tilde{c}_u(v) \leq \sum_{v \neq u} \left( (1 + \alpha)a(v, u) + \gamma \sum_{w \in V} b(w, (v, u)) + b(v, (w, u)) \right)$$
$$\leq \sum_{v \in V} (1 + \alpha)a(v, u) + 2\gamma \sum_{v, w \in V} b(v, (w, u)) < \infty,$$

where we used the inequalities (3.1) and (3.2) again and we have proved (3.8).

Now, using [Lig85, Theorem 1.3.9] we see that the closure of $\Omega$ in $C(\{0, 1\}^V)$ is a Markov generator of a Markov semigroup. This finishes the proof.

We can now define the voter model with bias and cooperation.

**Definition 3.5** ((Cooperative/Altruistic) Voter Model with Bias and Cooperation)
Let $a(\cdot, \cdot)$ be a Markov kernel from $V$ to $V$ satisfying (3.1) and $b(\cdot, (\cdot, \cdot))$ be a Markov kernel from $V$ to $V \times V$ satisfying (3.2).

1. The (unique) Markov process with transition rates given by (3.5) and (3.6) is called *Voter Model with Bias and Cooperation* (VMBC).

2. If (3.4) holds, the VMBC is called *cooperative Voter Model with Bias and Cooperation* (cVMBC).

3. If (3.3) holds, the VMBC is called *altruistic Voter Model with Bias and Cooperation* (aVMBC).

**Remark 3.6** (Graphical Construction) It is also possible to construct the interacting particle systems described above via Poisson processes corresponding to the model dynamics. We do this exemplarily for the nearest-neighbor case in one dimension, the higher dimensional cases follow analogously. To start with, let $u \in \mathbb{Z}$ and $v \in \mathbb{Z} + 1/2$ and let $\tilde{\omega}_r(u), \tilde{\omega}_r(u), \tilde{\omega}_r(u), \tilde{\omega}_r(u), \tilde{\omega}_r(v)$ and $\tilde{\omega}_r(v)$ be independent Poisson subsets of $\mathbb{R}$.

The Poisson sets $\tilde{\omega}_r(u)$ and $\tilde{\omega}_r(u)$ have intensity $1/2$. These represent the type-independent reproduction events. The processes $\tilde{\omega}_r(u)$ and $\tilde{\omega}_r(u)$ jump with rate $\alpha/2$. 

26
3.2 The model

\[ Z_t \]

\[ \sigma \]

\[ \omega_c(u) \]

\[ u - 1 \quad u \quad u + 1 \]

cooporation event

\[ \omega_r(u) \]

\[ u \quad u + 1 \]

reproduction event

\[ \omega_s(u) \]

\[ u \quad u + 1 \]

selective reproduction event

Figure 3.1: Example of a graphical representation of the cVMBC in one dimension and explanation of the dynamical mechanisms. Time is given vertically whereas space is plotted horizontally. Bold lines indicate the presence of a cooperator (labeled by a 1), thin lines the presence of a defector (labeled by a 0). Solid arrows are those which can be used by the underlying particle configuration, dashed arrows will not be used and therefore entail no update.
These are connected to the selective advantage of defectors. Finally, $\omega_c(v)$ and $\omega_s^r(v)$ have intensity $\gamma/2$. For an illustration see Figure 3.1.

For $u \in \mathbb{Z}/2$ the presence of a point $t \in \omega_{r,s,c}(u)$ is indicated by an arrow at time $t \in \mathbb{R}$ from $u$ to $u - 1$. The arrows pointing to the right correspond to the Poisson sets $\omega_{r,s,c}(u)$. Additionally, to distinguish between the sets $\omega_r$ and $\omega_s$ the arrows corresponding to the latter sets are marked by $\sigma$.

In order to describe our dynamical behavior of the cVMBC we need to specify what happens at the different arrows. For the unmarked arrows pointing from $u \in \mathbb{Z}$ to $u + 1$ in one of the two directions the type of the particle at $u + 1$ is replaced by the type of the particle at $u$. The arrows having a $\sigma$ and pointing from $u \in \mathbb{Z}$ to $u + 1$ only overwrite the type at $u + 1$ if the particle at $u$ is a defector (labeled by a 0). Otherwise we do nothing. Concerning the arrows emerging from some $u \in \mathbb{Z} + 1/2$ and pointing to $u + 1$ the particle type at site $u + 3/2$ will be updated by the type of the particle at $u + 1/2$ if the particle at $u - 1/2$ is a cooperator (labeled by a 1). In all other cases there is no update. Of course, the same rules hold for the arrows looking to the left.

It can be checked that this construction yields the existence of a Markov process corresponding to the model dynamics from the cVMBC. For a general result on this, see for example [Dur95, Theorem 2.1]. The main idea is to prove that for a given time $s \leq t$ the set of sites whose state could influence a particle at site $u \in \mathbb{Z}$ is almost surely finite.

We introduce the graphical representation additionally to the already proved existence by semigroup theory because we will need this form of description in Chapter 6. There we use it for a comparison with a oriented percolation process. For the comparison to work, it is necessary to see that sites are grouped in finite time-space boxes which do not influence each other, which is exactly the result of the graphical representation of the cVMBC.

### 3.3 Unstructured populations

As a first result, we show that cooperators always die out (for $\alpha > 0$) if $V$ is a large complete graph. Here, we consider the special case of an unstructured population in the sense that $|V^N| = N$ and

$$a^N(u, v) = \frac{1}{N - 1}$$

for $u, v \in V^N$ with $u \neq v$. Due to the global neighborhood it is equally likely to find configurations of the form "101" and "110". Hence, cooperation events favoring a defector or a cooperator happen with the same rate and thus cancel out when looking at the mean field behavior of the system. We will show that defectors always take over the system for large $N$. It can easily be seen that the aVMBC is dominated by the
cVMBC, so it suffices to show extinction of cooperators for the cVMBC, i.e. we have
\[ b^N(u, (v, w)) = \frac{1_{\{u \neq v\}} 1_{\{v \neq w\}} 1_{\{w \neq u\}}}{N-1} = \frac{1}{(N-1)(N-2)} 1_{\{u, v, w \text{ different}\}}. \]

We prove that in the limit for large \( N \) the frequency of cooperators follows a logistic equation with negative drift, hence cooperators die out. See also [EK86, Chapter 11].

**Proposition 3.7** (Convergence in the unstructured case) Let \( X^N \) be a cVMBC on \( V^N \) and \( S^N := \frac{1}{N} \sum_u X^N(u) \) the frequency of cooperators. If \( S^N_0 \xrightarrow{N \to \infty} s_0 \), then
\[ S^N \xrightarrow{N \to \infty} S, \]
where \( S \) solves the ODE
\[ \frac{dS}{dt} = -\alpha S(1 - S) \]
with \( S_0 = s_0 \), independently of \( \gamma \).

**Proof:** In order to prove the limiting behavior for \( N \to \infty \), we observe that \( S^N \) is a Markov process. A calculation of the generator \( \Omega^N \) applied to some smooth function \( f \) yields
\[
\Omega^N f(s) = N s \frac{1 - s}{1 - \frac{1}{N}} (f(s + \frac{1}{N}) - f(s)) + (1 + \alpha) N (1 - s) \frac{s}{1 - \frac{1}{N}} (f(s - \frac{1}{N}) - f(s)) \\
+ \gamma N s \frac{s - \frac{1}{N}}{1 - \frac{1}{N}} \frac{1 - s}{1 - \frac{2}{N}} (f(s + \frac{1}{N}) - f(s)) \\
+ \gamma N s \frac{1 - s}{1 - \frac{1}{N}} \frac{s - \frac{1}{N}}{1 - \frac{2}{N}} (f(s - \frac{1}{N}) - f(s)) \\
\xrightarrow{N \to \infty} -\alpha s(1 - s)f'(s).
\]
Applying standard weak convergence results, see for example [EK86, Theorem 4.8.2], this shows the claimed convergence. \( \square \)
Having proved a non-spatial result, this chapter is devoted to the spatial setting. We study limit results for the aVMBC on $\mathbb{Z}^d$ for $d \geq 1$ and for the cVMBC on $\mathbb{Z}$. In case of the cVMBC we prove that, when starting in a translation invariant configuration, for $\alpha > \gamma$ defectors take over the colony, whereas for $\gamma > \alpha$ cooperators win; see Theorem 4.2. Additionally, we also show that a finite number of cooperators dies out if $\alpha > \gamma$, but may survive if $\gamma > \alpha$. The converse holds true for defectors; see Theorem 4.5. Concerning the aVMBC we identify a parameter region dependent on the dimension such that defecting particles win; compare with Theorem 4.3.

The following results also appeared in [CP16].

4.1 Long-time behavior

Our goal is to derive the long-time behavior of the VMBC with $V = \mathbb{Z}^d$. In spin-flip systems, results on the ergodic behavior can be obtained by general principles if the process is monotone. Thereby, a spin-system is called monotone if for two configurations $X, Y \in \{0, 1\}^V$ with $X \leq Y$ componentwisely, the corresponding transition rates $c$ satisfy the following two relations for all $u \in V$

$$X(u) = Y(u) = 0 \Rightarrow c(u, X) \leq c(u, Y), \quad (4.1)$$
$$X(u) = Y(u) = 1 \Rightarrow c(u, X) \geq c(u, Y). \quad (4.2)$$

However, the VMBC is not monotone for $\gamma > 0$. Indeed, consider the simple case when $V = \{u, v, w\}$ with Markov kernels

$$a(u, v) = a(v, w) = a(w, u) = 1$$
and \( b(u, (v, w)) = a(u, v) a(v, w) \). Then, let \( X = (001) \) and \( Y = (101) \) (i.e. \( X(u) = 0, Y(u) = 1, X(v) = Y(v) = 0, X(w) = Y(w) = 1 \)) and note that \( X \leq Y \), but
\[
c(w, X) = 1 + \alpha < 1 + \alpha + \gamma = c(w, Y).
\]
This shows that (4.2) is not satisfied at \( w \in V \). Hence, proofs for the long-time behavior require other strategies which do not rely on the monotonicity of the process.

Before we state our main results we define what we mean by extinction, survival and clustering.

**Definition 4.1 (Extinction, Survival, Clustering)**

1. We say that in the VMBC-process \( (X_t)_{t \geq 0} \) type \( i \in \{0,1\} \) dies out if
\[
P \left( \lim_{t \to \infty} X_t = 1 - i \right) = 1,
\]
where 0 and 1 are the all-zero (exclusively defectors) and all-one (exclusively cooperators) configurations, respectively.

2. We say that type \( i \in \{0,1\} \) survives, if for all \( u \in V \)
\[
\lim_{t \to \infty} \inf P(X_t(u) = i) > 0.
\]

3. We say that the VMBC-process clusters if for all \( u, v \in V \)
\[
\lim_{t \to \infty} P(X_t(u) = X_t(v)) = 1.
\]

Throughout the rest of the chapter we set \( V = \mathbb{Z}^d \) and use uniform nearest neighbor interaction for the kernels \( a \) and \( b \). In this case we have that for all \( u, v, w \in \mathbb{Z}^d \) with \( |u - v| = |w - v| = 1 \)
\[
a(u, v) = \frac{1}{2d}, \quad b(u, (v, w)) = \frac{1}{(2d)^2}
\]
for the aVMBC and
\[
a(u, v) = \frac{1}{2d}, \quad b(u, (v, w)) = \frac{1}{2d(2d-1)} 1_{\{u \neq w\}}
\]
for the cVMBC. Furthermore, we say that (the distribution of) a \( \{0,1\}^{\mathbb{Z}^d} \)-valued random configuration \( X \) is non-trivial if \( P(X = 0) + P(X = 1) < 1 \). This implies that, with positive probability, \( X \) is not mono-typic. Furthermore, we call \( X \) translation invariant if \( (X(u_1), ..., X(u_n)) \overset{d}{=} (X(u_1 + v), ..., X(u_n + v)) \) for all \( n \in \mathbb{N}, u_1, ..., u_n, v \in \mathbb{Z}^d \). If the VMBC model is started in a translation invariant configuration \( X_0 \in \{0,1\}^{\mathbb{Z}^d} \), the configuration \( X_t \) is translation invariant due to the homogeneous model dynamics.

Now, we can state our main results. For cVMBC, we distinguish between the case \( \alpha > \gamma \) where we can state a convergence result in all dimensions \( d \geq 1 \), the case \( \gamma > \alpha \) and the case \( \gamma = \alpha \). In the last two cases, the method of proof is only applicable in dimension \( d = 1 \).
4.1 Long-time behavior

Figure 4.1: Relative frequencies of cooperators after 100,000 transitions of the cVMBC on a 1,000 sites torus in one dimension (dashed line), a 40 × 40 sites torus in two dimensions (solid line) and a 12 × 12 × 12 sites torus in three dimensions (dotted line). The initial configuration was a Bernoulli-product measure with probability 0.5. The selection rate \( \alpha \) was set to 0.5 and the reproduction rate was equal to 1.

Theorem 4.2 (cVMBC-limits) Let \( V = \mathbb{Z}^d \) and \( a(.,.) \) be the nearest neighbor random walk kernel and \( X \) be the cVMBC with \( \alpha, \gamma \geq 0 \) starting in some non-trivial translation invariant initial configuration.

(i) If \( d \geq 1 \) and \( \alpha > \gamma \), cooperators die out.

(ii) If \( d = 1 \) and \( \gamma > \alpha \), defectors die out.

(iii) If \( d = 1 \) and \( \gamma = \alpha \), the process clusters.

The proof of Theorem 4.2 is given in Section 4.3.1. Briefly, for \( \alpha > \gamma \), we will use a comparison argument with a biased voter model, see Definition 4.7. For \( \gamma > \alpha \), however, the proof relies on the help of a cluster size process which takes the special form of a jump process in one dimension. In particular, this argument cannot be extended to higher dimensions. Resorting to some simulation results for \( d = 2 \) and \( d = 3 \), we see a similar behavior (with a different threshold) like in \( d = 1 \), see Figure 4.1. For higher dimensions, cooperators are less likely to cluster, leading to a reduced chance of survival.

The next Theorem deals with the aVMBC and its proof can be found in Section 4.3.2. Here, we can only state a threshold when cooperators die out.
Theorem 4.3 (aVMBC-limits) Let $V = \mathbb{Z}^d$ and $a(.,.)$ be the nearest neighbor random walk kernel and $X$ be the aVMBC with $\alpha, \gamma \geq 0$ starting in some non-trivial translation invariant initial configuration.

1. If $d \geq 1$ and $\alpha > \gamma \frac{d-1}{d}$, the cooperators die out. In particular, for $d = 1$, the cooperators die out if $\alpha > 0$ independently of $\gamma$.

2. If $d = 1$, the process equals the cVMBC with parameters $\alpha + \gamma/2$ and $\gamma/2$ in distribution. In particular, if $\gamma > \alpha = 0$, the process clusters.

Remark 4.4 (Cooperation only among cooperators) Another cooperation mechanism one might consider arises if cooperators only help other cooperators, i.e. cells recognize related cells. In ecological literature this behavior is called kin-recognition or kin-discrimination, see [PF10] for an overview. As to the dynamical behavior of the model this changes the transition rate in (3.6), i.e. if $X(u) = 1$ then

$$c(u, X) = (1 + \alpha) \sum_{v} a(v, u)(1 - X(v)).$$

Here, cooperators are less likely to die and hence, this process dominates the cVMBC. In particular, for translation invariant initial conditions, defectors die out for $\gamma > \alpha$ in one dimension. Moreover, as can be seen from a calculation similar as in the proof of Lemma 4.9, a biased voter model, where type 0 is favored, still dominates this process for $\alpha > \gamma$. Hence, we also have that cooperators die out in this case. Furthermore, in order to show clustering of this process one can carry out a similar calculation as in the proof of Theorem 4.2 (iii) such that the same results as in Theorem 4.2 hold.

In what follows we want to study the emergence of cooperation in a population. Therefore, we consider initial configurations which are not translation invariant and contain a finite number of cooperators or defectors. Since cooperators always die out in $d = 1$ for the aVMBC (as long as $\alpha > 0$), we focus on the cVMBC in the sequel. We obtain the following results.

Theorem 4.5 (Finite initial configurations) Let $V = \mathbb{Z}$ and $a(.,.)$ be the nearest neighbor random walk kernel and $X$ be the cVMBC with $\alpha, \gamma \geq 0$. Let $X_0$ contain either finitely many defectors or finitely many cooperators (i.e. $X_0 = 1_A$ or $X_0 = 1 - 1_A$ for some finite $A \subseteq V$). Then the following holds:

(i) The process clusters.

(ii) If $\alpha \geq \gamma$ and $X_0$ contains finitely many cooperators, cooperators die out.

(iii) If $\gamma \geq \alpha$ and $X_0$ contains finitely many defectors, defectors die out.

Remark 4.6 (Starting with a single particle) A particularly simple initial condition is given if $|A| = 1$. In case there is only a single cooperator, we note that the size of
the cluster of cooperators \((C_t)_{t \geq 0}\) is a birth-death process which jumps from \(x\) to
\[
\begin{align*}
x + 1 & \text{ at rate } I_{\{x>0\}} + \gamma \cdot I_{\{x \geq 2\}}, \\
x - 1 & \text{ at rate } (1 + \alpha) \cdot I_{\{x>0\}}.
\end{align*}
\]
Conversely, if there is only a single defector, the size of the cluster of defectors \((D_t)_{t \geq 0}\) is a birth-death process which jumps from \(y\) to
\[
\begin{align*}
y + 1 & \text{ at rate } (1 + \alpha) \cdot I_{\{y>0\}} + \gamma \cdot I_{\{y=1\}}, \\
y - 1 & \text{ at rate } (1 + \gamma) \cdot I_{\{y>0\}}.
\end{align*}
\]
Hence, either cooperators or defectors die out, depending on whether \((C_t)_{t \geq 0}\) (or \((D_t)_{t \geq 0}\)) hits 0 or not.

The proof of Theorem 4.5 is given in Section 4.3.3. Note, that the only situations where the process does not converge to a deterministic configuration in this setting are the cases where \(\gamma > \alpha\) (\(\alpha > \gamma\)) and the process starts with finitely many cooperators (defectors). Here, the limit distribution is a linear combination of the invariant measures \(\delta_0\) and \(\delta_1\) which basically means that we observe clustering, which is statement (i) above.

### 4.2 Preliminaries

Here we provide some useful results for the proofs of the Theorems. In particular, we give a comparison with a biased voter model in Section 4.2.1 and study a jump process in Section 4.2.2.

#### 4.2.1 Comparison results

In cases where \(\alpha > \gamma\), it is possible to prove a stochastic domination of the VMBC by a biased voter model, i.e. the existence of a coupling such that the VMBC is componentwisely smaller than a biased voter model for all times. The precise statements will be given below. But first, we define the process called biased voter model, which was introduced in [WB72] and first rigorously studied in [BG81].

**Definition 4.7 (Biased Voter Model)** The biased voter model with bias \(\beta \geq -1\) and \(\delta \geq -1\) is a spin system \(\tilde{X}\) with state space \(\{0, 1\}^V\) and transition rates as follows: If \(\tilde{X}(u) = 0\), then
\[
\tilde{c}(u, \tilde{X}) = (1 + \beta) \sum_v a(v, u)X(v).
\]
If \(\tilde{X}(u) = 1\), then
\[
\tilde{c}(u, \tilde{X}) = (1 + \delta) \sum_v a(v, u)(1 - X(v)).
\]
Remark 4.8 (Long-time behavior of the biased voter model) The long-time behavior of the biased voter model is quite simple. In [BG81], the process on $V = \mathbb{Z}^d$ with nearest neighbor interactions is studied. Generalizations to the case of $d$-regular trees for $d \geq 3$ can be found in [LTV14]. We restate the results for $V = \mathbb{Z}^d$:

Let $\tilde{X}$ be a biased voter model with bias $\beta > -1$ and $\delta > -1$ as introduced in Definition 4.7. For any configuration $X_0 \in \{0, 1\}^{\mathbb{Z}^d}$ with infinitely many particles of both types it holds that the type with less bias dies out, i.e.:

1. If $\beta > \delta$, type 1 dies out (i.e. $P(\lim_{t \to \infty} \tilde{X}_t = 0) = 1$).
2. If $\delta > \beta$, type 0 dies out (i.e. $P(\lim_{t \to \infty} \tilde{X}_t = 1) = 1$).

Lemma 4.9 (cVMBC ≤ biased voter model) Let $X$ be a cVMBC with bias $\alpha$ and cooperation coefficient $\gamma$ and $\tilde{X}$ a biased voter model with bias $\gamma$ and $\alpha$. Then, if $b(., (., .))$ satisfies $\sum_u b(u, (v, w)) \leq a(v, w)$, and $X_0 \leq \tilde{X}_0$, it is possible to couple $X$ and $\tilde{X}$ such that $X_t \leq \tilde{X}_t$ for all $t \geq 0$.

Proof: We need to show (see [Lig85, Theorem 3.1.5]) that for $X \leq \tilde{X}$

\begin{align}
\text{if } X(u) = \tilde{X}(u) = 0, \quad & c(u, X) \leq \tilde{c}(u, \tilde{X}), \\
\text{if } X(u) = \tilde{X}(u) = 1, \quad & c(x, X) \geq \tilde{c}(u, \tilde{X}).
\end{align}

These inequalities then yield the existence of a basic coupling with the claimed properties (see also [Lig85, Chapter 3.1]). We start with the first assertion and write

\[ c(u, X) = \sum_v a(v, u)X(v) + \gamma \sum_v X(v) \sum_w X(w)b(w, (v, u)) \]
\[ \leq \sum_v a(v, u)X(v) + \gamma \sum_v X(v)a(v, u) \]
\[ \leq (1 + \gamma) \sum_v a(v, u)\tilde{X}(v) = \tilde{c}(u, \tilde{X}), \]

and for the second inequality we have

\[ c(u, X) = (1 + \alpha) \sum_v a(v, u)(1 - X(v)) + \gamma \sum_v (1 - X(v)) \sum_w X(w)b(w, (v, u)) \]
\[ \geq (1 + \alpha) \sum_v a(v, u)(1 - \tilde{X}(v)) \geq (1 + \alpha) \sum_v a(v, u)(1 - \tilde{X}(v)) = \tilde{c}(u, \tilde{X}). \]

This shows the result and we are done.

Next, we focus on the aVMBC in the case $V = \mathbb{Z}^d$ and a symmetric, nearest-neighbor random walk kernel.
4.2 Preliminaries

Lemma 4.10 (aVMBC≤biased voter model) Let $V = \mathbb{Z}^d$, $a(\cdot,\cdot)$ be a nearest-neighbor random walk kernel, $X$ be an aVMBC with bias $\alpha$ and cooperation coefficient $\gamma$ and $\tilde{X}$ a biased voter model with bias $\gamma(2d-1)/(2d)$ and $\alpha + \gamma/(2d)$. Then, if $X_0 \leq \tilde{X}_0$, it is possible to couple $X$ and $\tilde{X}$ such that $X_t \leq \tilde{X}_t$ for all $t \geq 0$.

Proof: Again, we need to show that for $X \leq \tilde{X}$ the inequalities in (4.3) hold. We start with the first assertion and write

$$c(u,X) = \sum_v a(v,u)X(v) + \gamma \sum_v X(v) \sum_w X(w)a(w,v)a(v,u)$$

$$\leq \sum_v a(v,u)X(v) + \gamma \sum_v X(v)a(v,u) \sum_{w \neq u} a(w,v)$$

$$\leq \left(1 + \gamma \frac{2d-1}{2d}\right) \sum_v a(v,u)\tilde{X}(v) = \tilde{c}(u, \tilde{X}),$$

and for the second inequality we have

$$c(u,X) = (1 + \alpha) \sum_v a(v,u)(1 - X(v)) + \gamma \sum_v (1 - X(v)) \sum_w X(w)a(w,v)a(v,u)$$

$$\geq (1 + \alpha) \sum_v a(v,u)(1 - X(v)) + \gamma \sum_v (1 - X(v))a(u,v)a(v,u)$$

$$= \left(1 + \alpha + \frac{\gamma}{2d}\right) \sum_v a(v,u)(1 - X(v))$$

$$\geq \left(1 + \alpha + \frac{\gamma}{2d}\right) \sum_v a(v,u)(1 - \tilde{X}(v)) = \tilde{c}(u, \tilde{X}).$$

This proves the second inequality and we are done.

4.2.2 A result on a jump process

In the proof of Theorem 4.2, we will use the dynamics of the size of a cluster of cooperators and rely on a comparison of this cluster size process with a certain jump process (which jumps downwards by at most one and upwards by at most two). The following lemma will be of use.

Lemma 4.11 (A jump process) Let $(C_t)_{t \geq 0}$ be a $\mathbb{Z}$-valued jump process, which jumps from $x$ to

$$x - 1 \text{ at rate } \mu,$$

$$x + 1 \text{ at rate } \lambda_1,$$

$$x + 2 \text{ at rate } \lambda_2.$$

If $\lambda_1 + 2\lambda_2 > \mu$, then $C_t \xrightarrow{t \to \infty} \infty$ almost surely.
4 The VMBC-process on $\mathbb{Z}^d$

**Proof:** We assume without loss of generality that $\lambda_1 + \lambda_2 + \mu = 1$. Otherwise, use a time-rescaling.

We first show that there exists an $a_c > 0$ such that for all $a \in (0, a_c)$, the process $(\exp(-a C_t))_{t \geq 0}$ is a positive super-martingale. For this, consider the generator $G^C$ of the process $(C_t)_{t \geq 0}$ applied to the function $f(x) = \exp(-ax)$ which yields

$$(G^C f)(x) = \lambda_1 \exp(-a(x + 1)) + \lambda_2 \exp(-a(x + 2)) + \mu \exp(-a(x - 1)) - \exp(-ax)$$

$$= \exp(-ax)(\lambda_1 \exp(-a) + \lambda_2 \exp(-2a) + \mu \exp(a) - 1).$$

We define $g(x) := \lambda_1 \exp(-x) + \lambda_2 \exp(-2x) + \mu \exp(x) - 1$. Noting that $g(0) = 0$ and $g'(0) < 0$ we find $a_c > 0$ such that $a_c := \inf(a \in \mathbb{R}_{>0} : g(a) = 0)$. This yields that for all $a \in (0, a_c)$ we have that $g(a) < 0$ which means that $(\exp(-a C_t))_{t \geq 0}$ is a super-martingale. By the martingale convergence theorem, it converges almost surely and the only possible almost sure limit is 0. The result now follows since $C_t \xrightarrow{t \to \infty} \infty$ if and only if $\exp(-a C_t) \xrightarrow{t \to \infty} 0$ for some $a > 0$. $\square$

**Corollary 4.12** $(C$ does not need to hit 1) Let $(C_t)_{t \geq 0}$ be the $\mathbb{Z}$-valued jump process from above with $C_0 = 2$ and $T_1 := \inf\{t : C_t = 1\}$. Then, $P(T_1 = \infty) > 0$.

**Proof:** For some $a > 0$, the process $(\exp(-a C_{t \wedge T_1}))_{t \geq 0}$ is a non-negative super-martingale. Let us assume that $T_1 < \infty$ almost surely, which occurs if and only if $C_{t \wedge T_1} \xrightarrow{t \to \infty} 1$ almost surely. Then we obtain with the dominated convergence theorem and $C_0 = 2$

$$\exp(-2a) = E[\exp(-a C_0)]$$

$$\geq \lim_{t \to \infty} E[\exp(-a C_{t \wedge T_1})] = E[\lim_{t \to \infty} \exp(-a C_{t \wedge T_1})] = \exp(-a),$$

a contradiction. Thus, we have that $P(T_1 = \infty) > 0$. $\square$

**Corollary 4.13** (Time-dependent jump rates) Let $(\mu(t))_{t \geq 0}, (\lambda_1(t))_{t \geq 0}, (\lambda_2(t))_{t \geq 0}$ be càdlàg-stochastic processes which satisfy $\lambda_1(t) + 2\lambda_2(t) - \mu(t) > \varepsilon > 0$ for some $\varepsilon > 0$. In addition, let $(C_t)_{t \geq 0}$ be a $\mathbb{Z}$-valued jump process with $C_0 = 2$, which jumps at time $t$ from $x$ to

$$x - 1 \text{ at rate } \mu(t),$$

$$x + 1 \text{ at rate } \lambda_1(t),$$

$$x + 2 \text{ at rate } \lambda_2(t)$$

and $T_1 := \inf\{t : C_t = 1\}$. Then, $P(T_1 = \infty) > 0$.

**Proof:** By the same arguments as in the proof of Lemma 4.11, we obtain that $(\exp(-a C_t))_{t \geq 0}$ is a non-negative super-martingale for $a > 0$ small enough and thus $C_t \xrightarrow{t \to \infty} \infty$ almost surely. As in the proof of Corollary 4.12, the assertion then follows. $\square$
4.3 Proofs

Here, we give the proofs of Theorems 4.2, 4.3 and 4.5.

4.3.1 Proof of Theorem 4.2

For (i), we have $\alpha > \gamma$. The assertion is a consequence of the coupling with the biased voter model from Lemma 4.9 (with biases $\gamma$ and $\alpha$). Since the biased voter model dominates the cVMBC and type 1 dies out in the biased voter model (4.8), the same is true for the cVMBC.

The proof of (ii) is more involved. We have to show that cooperators survive almost surely when started in a non-trivial translation invariant configuration. Therefore, we analyze an arbitrary cluster of cooperators and show that the size of such a cluster has a positive probability to diverge off to infinity. Note that the flanking regions of a cluster of cooperators can have three different forms:

<table>
<thead>
<tr>
<th>Case A</th>
<th>Case B</th>
<th>Case C</th>
</tr>
</thead>
<tbody>
<tr>
<td>001...100,</td>
<td>101...101,</td>
<td>001...101 or 101...100</td>
</tr>
<tr>
<td>cluster of cooperators</td>
<td>cluster of cooperators</td>
<td>cluster of cooperators</td>
</tr>
</tbody>
</table>

These are the only possible environments a cluster of cooperators can encounter in one dimension. Note, that a cluster can also only consist of a single cooperator. The dynamics of the cluster size depends on the environment. Precisely, by the dynamics of the process, we obtain the following. A cluster of size $x > 1$

- in case A jumps to $y = x + 1$ at rate $1 + \gamma$
- jumps to $y = x - 1$ at rate $1 + \alpha$
- in case B jumps to $y \geq x + 2$ at rate at least $2 + \gamma$
- jumps to $y = x - 1$ at rate $1 + \alpha + \gamma$
- in case C jumps to $y \geq x + 2$ at rate at least $1 + \frac{\gamma}{2}$
- jumps to $y = x + 1$ at rate $1 + \frac{\gamma}{2}$
- jumps to $y = x - 1$ at rate $1 + \alpha + \frac{\gamma}{2}$.

(4.4)

Under the assumptions of Theorem 4.2, let $(V_t)_{t \geq 0}$ be the cluster of cooperators at time 0 which is closest to the origin and contains at least two cooperators. (If there is no such cluster at time 0, wait for some time $\varepsilon > 0$ and pick the cluster then.) We will show that

$$P(V_t \uparrow \mathbb{Z}) > 0.$$  (4.5)

For this, we compare $|V| = (|V_t|)_{t \geq 0}$ with a jump process $(\tilde{V}_t)_{t \geq 0}$ as in Corollary 4.13, where $\lambda_1(t) = 1 + \gamma$ at times $t$ when the cluster is in case A, $\lambda_1(t) = 0$ and $\lambda_1(t) = 1 + \frac{\gamma}{2}$.
for cases B and C. Moreover, $\lambda_2(t) = 0, 2 + \gamma, 1 + \frac{\gamma}{2}$ for the three cases and the decrease rate is given by $\mu(t) = 1 + \alpha, 1 + \alpha + \gamma, 1 + \alpha + \frac{\gamma}{2}$, respectively. The process is stopped when it reaches 1. By the comparison in (4.4), we see that we can couple $|V|$ and $\tilde{V}$ such that $\tilde{V} \leq |V|$, at least until $\tilde{V}$ reaches 1. Since $\tilde{V}_t \xrightarrow{t \to \infty} \infty$ with positive probability, we find that also $\mathbf{P}(|V_t| \xrightarrow{t \to \infty} \infty) > 0$ holds. Still, we need to make sure that the cluster does not wander to $\pm \infty$. For this, consider both boundaries of the cluster if it has grown to a large extent. The right boundary is again bounded from below by a jump process of the form as in Corollary 4.13 with $\lambda_1(t) = 1 + \frac{\gamma}{2}, 0$ and $\mu(t) = 1 + \alpha, 1 + \alpha + \frac{\gamma}{2}$ for the cases A and B (note that the right boundary alone of case C is already captured by the right boundaries of the cases A and B). So, again, we see from Corollary 4.13 that the right border of the cluster goes to infinity with positive probability. The same holds for the left border of the cluster which tends to $-\infty$. Therefore, we have shown (4.5).

Now, we use (4.5) and the fact that a cluster of cooperators can merge with finitely many other clusters in finite time. To be more precise, if a cluster of cooperators goes extinct at time $T_0$ which happens with probability $\mathbf{P}(T_0 < \infty) < 1$ it fused only with finitely other clusters which then go extinct, too. Due to the translation invariance of the initial configuration there are infinitely many clusters of cooperators initially such that we can restart this argument arbitrarily often. Since $\mathbf{P}(T_0 = \infty) = \mathbf{P}(V_t \uparrow \mathbb{Z}) > 0$ one of these clusters almost surely succeeds and expands to infinity. Hence, we have $\mathbf{P}(\lim_{t \to \infty} X_t = 1) = 1$ and we are done.

For (iii), in order to prove clustering in the case $\alpha = \gamma > 0$, for $i_0, \ldots, i_k \in \{0, 1\}$ and $k = 0, 1, 2, \ldots$ we write $p_t(i_0 \cdots i_k) := \mathbf{P}(X_t(0) = i_0, \cdots X_t(k) = i_k)$. We have to show that

$$p_t(10) \xrightarrow{t \to \infty} 0, \quad p_t(01) \xrightarrow{t \to \infty} 0 \quad (4.6)$$

since then – by translation invariance – every configuration carrying both types has vanishing probability for $t \to \infty$.

We start with the dynamics of $p_t(1)$, which reads (recall that $\alpha = \gamma$)

$$\frac{\partial p_t(1)}{\partial t} = \frac{1}{2}(p_t(10) + p_t(01)) + \frac{\gamma}{2}(p_t(110) + p_t(011))$$

$$- \frac{1+\alpha}{2}(p_t(10) + p_t(01)) - \gamma p_t(101)$$

$$= - \frac{\alpha}{2}(p_t(10) + p_t(01)) + \frac{\gamma}{2}(p_t(10) + p_t(01) - 2p_t(010)) - \gamma p_t(101)$$

$$= - \gamma(p_t(101) + p_t(010)) \leq 0.$$

Since $p_t(1) \in [0, 1]$, this probability has to converge for $t \to \infty$, hence $\frac{\partial p_t(1)}{\partial t} \xrightarrow{t \to \infty} 0$, and therefore

$$p_t(101) \xrightarrow{t \to \infty} 0, \quad p_t(010) \xrightarrow{t \to \infty} 0. \quad (4.7)$$
Now, consider the dynamics of $p_t(11)$, which is
\[
\frac{\partial p_t(11)}{\partial t} = p_t(101) + \gamma \left( p_t(110) + p_t(011) \right) - \frac{1 + \alpha}{2} (p_t(110) + p_t(011)) - \gamma \left( p_t(101) + p_t(111) \right) = p_t(101) - \frac{1 + \alpha}{2} (p_t(110) + p_t(011)).
\]
Since we know that $p_t(101) \xrightarrow{t \to \infty} 0$ by (4.7), and because $p_t(11) \in [0, 1]$, we also have that $p_t(110) \xrightarrow{t \to \infty} 0$, $p_t(011) \xrightarrow{t \to \infty} 0$.

We now conclude with
\[
\begin{align*}
p_t(10) &= p_t(010) + p_t(110) \xrightarrow{t \to \infty} 0, \\
p_t(01) &= p_t(010) + p_t(011) \xrightarrow{t \to \infty} 0,
\end{align*}
\]
which shows (4.6).

4.3.2 Proof of Theorem 4.3

(i) We use the comparison with the biased voter model from Lemma 4.10. Therefore, we have that $\alpha > \gamma(d - 1)/d$ if and only if $\alpha + \gamma/(2d) > \gamma(2d - 1)/(2d)$. Since for this choice of parameters type 1 goes extinct in the biased voter model which dominates the aVMBC, we are done.

(ii) For $d = 1$ and the nearest neighbor random walk, the altruistic mechanism is such that a configuration 01 (or 10) turns into 00 at rate $\alpha/2 + \gamma/4$. The same holds for the cVMBC with bias $\alpha + \gamma/2$. In addition, 110 (or 011) turns to 111 at rate $\gamma/2$, which is the same for as for the cVMBC with cooperation parameter $\gamma$. This shows the claim. In particular, clustering follows from Theorem 4.2(iii).

4.3.3 Proof of Theorem 4.5

At time $t$, let $N_t$ be the number of finite clusters in $X_t$ with sizes $C^1_t, ..., C^{N_t}_t$. If the process starts with finitely many defectors (cooperators), $C^1_t, C^3_t, C^5_t, ...$ are sizes of clusters of defectors (cooperators), and $C^2_t, C^4_t, ...$ are sizes of clusters of cooperators (defectors). Note that $(N_t, C^1_t, ..., C^{N_t}_t)_{t \geq 0}$ is a Markov process. We will show the following:

1. Either, $N_t \xrightarrow{t \to \infty} 0$ or $N_t \xrightarrow{t \to \infty} 1$.

2. In cases (ii) and (iii), $N_t \xrightarrow{t \to \infty} 0$. 

41
3. If \( N_t \xrightarrow{t \to \infty} 1 \), then \( C_t^1 \xrightarrow{t \to \infty} \infty \).

Note that 1. and 3. together imply (i), i.e. X clusters in all cases. Of course, 2. implies (ii) and (iii).

1. The process \( N = (N_t)_{t \geq 0} \) is a non-increasing process, bounded from below by 0, so convergence of \( N \) is certain. We assume that \( N_t = n \geq 2 \) and prove that the hitting time \( \inf \{ s : N_s < n \} \) is finite almost surely. For this, it suffices to show that

\[
T := \inf \{ s : C_s^k = 1 \text{ for some } 1 \leq k \leq N_s \} < \infty
\]

almost surely, since by time \( T \), some cluster has size 1 and there is a positive chance that \( N \) decreases at the next transition. If \( N \) does not decrease, there is a next chance after another finite time and eventually, \( N \) will decrease.

If \( \alpha \geq \gamma \), consider the size \( C_t \) of a cluster of cooperators. Before time \( T \), all clusters have size at least 2, so \( C_t \) jumps

- from \( x \) to \( x + 1 \) at rate \( 1 + \gamma \),
- from \( x \) to \( x - 1 \) at rate \( 1 + \alpha \),

hence \( (C_{t,T})_{t \geq 0} \) is a non-negative super-martingale and converges almost surely, which implies that \( T < \infty \) almost surely. If \( \gamma \geq \alpha \), the same argument shows that \( T < \infty \) if the role of cooperators and defectors is exchanged. Hence we have proved (4.8) and 1. is shown.

2. If \( N_t \xrightarrow{t \to \infty} 1 \) and \( \alpha \geq \gamma \), the remaining finite cluster must contain defectors (since the argument above shows that a finite cluster of cooperators would die out). Therefore, in (ii), we must have that \( N_t \xrightarrow{t \to \infty} 0 \). If \( \gamma \geq \alpha \), the remaining finite cluster contains cooperators for the same reason. Hence, in (iii), we must have that \( N_t \xrightarrow{t \to \infty} 0 \). Thus, we have shown 2. Moreover, if \( \alpha = \gamma \), both cases would hold, which is impossible and therefore, \( N_t \xrightarrow{t \to \infty} 0 \) in this case.

3. As argued above, if \( N_t \xrightarrow{t \to \infty} 1 \) the remaining finite cluster must contain the stronger type, i.e. defectors for \( \alpha > \gamma \) and cooperators for \( \gamma > \alpha \). The size of the remaining finite cluster therefore is a biased random walk which goes to infinity on \( \{ N_t \xrightarrow{t \to \infty} 1 \} \) and the result follows. \( \square \)
So far we have obtained a limit result favoring cooperating individuals which is limited to one dimension. In order to extend this to higher dimensions we need to come up with different techniques as opposed to those used so far, i.e. we cannot rely on a cluster analysis in higher dimensions. Therefore, in this chapter we rescale the interaction parameters $\alpha$ and $\gamma$ such that the selection and cooperation mechanisms only have a weak impact on the dynamical behavior of the system. The advantage of this approach is that the transitions now resemble those of a voter model with perturbations like introduced by Cox, Durrett and Perkins in [CDP13]. We will see that this framework allows us to analyze the cVMBC in dimensions $d \geq 3$, i.e. the dimensions in which the voter model possesses non-trivial invariant laws.

However, biologically this rescaling is in general not applicable when dealing with microbes; see [DG12] for a review on this issue. Bacteria usually face strong selection pressure. Hence, the decisions they make have literally a ”live or die” impact. Thus, different phenotypes enforce notable distinct behavior of cells. The phenotypes are strongly separated from each other and therefore seldom continuously transferable to each other. For example, consider the situation of a sucrose rich environment. Either bacteria are able to break down sucrose or not. Here, bacteria have to make an irreversible decision. Hence, minor changes in their phenotype (the weak selection assumption) will still not make any difference, i.e. bacteria not able to break down sucrose will never be able to use the this sugar, resulting in the death of these cells.

Another relevant aspect are the short generation times of microbes which lead to fast evolution times. Due to the short time periods between reproduction events cells are likely to have only a few interactions making large differences between the phenotypes more relevant. This is e.g. in contrast to animals where the weak selection approximation is feasible. Here, individuals have much longer generation times such that
small differences in the phenotypes are still notable when the number of interactions is large. Mathematically spoken one can treat the total fitness of an individual as the sum of the benefits from the single interactions.

With this in mind we need to be very careful when applying weak interaction approximations in the context of microbial communities. But, as already mentioned it allows us to rigorously study the limit behavior of the introduced cVMBC which is the main reason why we consider this case here.

In the following, we start by reviewing the framework and results about voter models with perturbations from [CDP13] in Section 5.1. Thereafter, we rescale the cVMBC model and formulate the limit results (Theorem 5.8) of this model in Section 5.2. Finally, in Section 5.3 we give the proof of Theorem 5.8. The proof itself consists of an approximation argument which shows that the rescaled cVMBC behaves like a reaction-diffusion equation. Therefore, we first analyze the reaction-diffusion equation corresponding to the cVMBC dynamics and then argue by applying a Theorem proved by Cox, Durrett and Perkins that the voter model with perturbations satisfies the same limit dynamics.

5.1 Voter model with perturbations

We briefly repeat the terminology necessary for the definition of a voter model with perturbations like introduced in [CDP13]. Furthermore, we state some general results about voter models with perturbations which are necessary for the further analysis of the rescaled cVMBC. We can characterize the limit behavior of such a system by a reaction-diffusion equation and review some conditions for the stability and convergence to invariant laws by analyzing the perturbation term.

To begin with, let $Y$ be a stochastic process with state space $\{0, 1\}^{Z^d}$. Throughout this whole chapter we will assume that $d \geq 3$. The interactions are governed by an irreducible interaction kernel $p : Z^d \to [0, 1]$. We assume that $p$ is of finite range, is symmetric and has covariance matrix $\sigma^2 I$ for $I$ the $d$-dimensional unit matrix and $\sigma \in \mathbb{R}$. Furthermore, the kernel should assign no mass to the origin, i.e. $p(0) = 0$.

Later we will rescale the model by a parameter $\varepsilon \in (0, 1]$ such that the state space transforms to $\{0, 1\}^{\varepsilon Z^d}$. For $\varepsilon = 1$ this reduces to the original model and we will omit the parameter value in the notation in this case. Additionally, we also have to rescale the interaction kernel and write $p_\varepsilon(\varepsilon u) = p(u)$ for $u \in Z^d$.

We define the local frequencies for $i \in \{0, 1\}$ by

$$f_i^\varepsilon(u, Y) = \sum_{v \in \varepsilon Z^d} p_\varepsilon(v - \varepsilon u) 1\{Y(v) = i\}.$$  

Then, the transition rates of the unscaled voter model with perturbation are given by

$$c_i^\varepsilon(u, Y) = c_i^\varepsilon(u, Y) + \varepsilon^2 c_i^\varepsilon(u, Y) \geq 0,$$
for \( u \in \mathbb{Z}^d \) and \( Y \in \{0,1\}^\mathbb{Z}^d \). Here, \( c^\varepsilon_x \) are the transition rates of the classical voter model, i.e.

\[
c^\varepsilon_x(u, Y) = (1 - Y(u)) f_1^\varepsilon(u, Y) + Y(u) f_0^\varepsilon(u, Y),
\]

whereas the perturbation \( c^\varepsilon_p \) is assumed to be of the form

\[
c^\varepsilon_p(u, Y) = (1 - Y(u)) h^\varepsilon_1(u, Y) + Y(u) h^\varepsilon_0(u, Y),
\]

with \( h^\varepsilon_i \) being the perturbation functions. For the definition of these we assume that there is a vector \((\xi^1, \ldots, \xi^K)\) with \( \xi^i \in \mathbb{Z}^d \) and \( 0 < K < \infty \) which has law \( g \). Furthermore, let \( N_0 \in (0, \infty) \) and let \( g_i^\varepsilon \) be non-negative functions on \( \{0,1\}^K \) for \( i \in \{0,1\} \). Then, \( h^\varepsilon_i \) is given by

\[
h^\varepsilon_i(u, Y) = -N_0^{-2} f_i^\varepsilon(u, Y) + \mathbf{E}_g[g^\varepsilon_i(Y(u + \xi^1), \ldots, Y(u + \xi^K))],
\]

where \( \mathbf{E}_g \) is the expectation with respect to \( g \). This abstract form of the perturbation functions will later be simplified by letting \( N_0 \to \infty \). Hence, it suffices to consider the interaction functions \( g_i \) which will represent the dynamics induced by selection and cooperation from the cVMBC.

Having defined the standard voter model with perturbation \( Y \) we now study the rescaled process \( Y^\varepsilon \) given by \( Y_i^\varepsilon(\varepsilon u) := Y_{\varepsilon^{-2}i}(\varepsilon u) \in \{0,1\}^{\varepsilon \mathbb{Z}^d}, u \in \mathbb{Z}^d \) with transition rates

\[
c^\varepsilon(\varepsilon u, Y^\varepsilon) = \varepsilon^{-2} c^\varepsilon_x(\varepsilon u, Y^\varepsilon) + c^\varepsilon_p(\varepsilon u, Y^\varepsilon).
\]

In order to apply the results derived in [CDP13] we need to assume that there is a constant \( C < \infty \) and \( \kappa \in (0,1] \) such that

\[
P(\max_{i = 1, \ldots, K}(|\xi^i|) \geq u) \leq Ce^{-\kappa u}, \quad u > 0
\]

holds. Basically, this is an assumption on the interaction range to decrease exponentially with the distance of the particles. Additionally, we claim that there exist limiting maps \( g_i : \{0,1\}^K \to \mathbb{R}_+ \) such that

\[
\lim_{\varepsilon \to 0}\|g_i^\varepsilon - g_i\|_\infty = 0, \quad \text{for } i \in \{0,1\}.
\]

As Proposition 1.1 of [CDP13] shows, these conditions are satisfied whenever the interaction kernel \( p \) is of finite range. We cite their result.

**Proposition 5.1** (finite range interactions - Proposition 1.1 in [CDP13]) Let \( \zeta^1, \ldots, \zeta^K \in \mathbb{Z}^d \) be distinct points and \( \hat{g}_i^\varepsilon, \hat{g}_i : \{0,1\}^K \to \mathbb{R} \) such that

\[
h^\varepsilon_i(u, Y) = \hat{g}_i^\varepsilon(Y(u + \zeta^1), \ldots, Y(u + \zeta^K)), u \in \mathbb{Z}^d, Y \in \{0,1\}^{\mathbb{Z}^d},
\]

\[\{u : p(u) > 0\} \subset \{\zeta^1, \ldots, \zeta^K\}, \lim_{\varepsilon \to 0}\|\hat{g}_i^\varepsilon - \hat{g}_i\|_\infty = 0, \text{ for } i \in \{0,1\}.\]

Then conditions (5.1)-(5.3) hold for appropriate non-negative \( g_i^\varepsilon, g_i \) satisfying

\[
\|g_i^\varepsilon - g_i\|_\infty = \|\hat{g}_i^\varepsilon - \hat{g}_i\|_\infty \text{ and } \zeta^i = \zeta^i.
\]
5 Weak Interaction Limit

5.1.1 Hydrodynamic limit

In a next step we describe the limit behavior of such a perturbed voter model. Therefore, we identify the limit reaction-diffusion equation which then qualitatively describes the behavior of the rescaled voter model.

The general idea behind this procedure is that the system evolves on two time-scales. First, the system converges to the voter model equilibrium in \( d \geq 3 \). This happens due to the rescaling of the transition rates such that the voter model dynamics occur on the fast time-scale. As we know from Theorem 5.1.8 from [Lig85] the voter model \( \tilde{Y} \) with transition rates \( c(u, \tilde{Y}) = \sigma_1 u' \) possesses a one-parameter family of translation invariant extremal invariant distributions \( \{ \nu_\theta : \theta \in [0, 1] \} \) on \( \mathbb{Z}^d \) for \( d \geq 3 \) such that \( E^{\nu_\theta}[\tilde{Y}(u)] = \theta \) for all \( u \in \mathbb{Z}^d \). Hence it is clear that the system converges to an equilibrium. Once in this equilibrium the voter model with perturbation gets perturbed by the additional dynamics given in \( c^p \) which then induces a new equilibrium for the voter model. This process repeats itself until the perturbation process does not affect the system anymore. Hence, it is crucial to control the drift induced by the perturbation in order to make predictions about the limit behavior of the interacting particle system. In mathematical terms, when writing \( \langle g \rangle_\theta := E^{\nu_\theta}[g(Y)] \) we define the drift by

\[
\sigma \left( 1 - X(0) \right) h_1(0, X) - X(0) h_0.
\] (5.4)

Note, that without loss of generality we choose the origin as a reference site which is possible due to the spatial homogeneity of the transition rates and the translation invariance of the initial configuration.

Next, we formalize the convergence of the rescaled particle system to the partial differential equation given by

\[
\frac{\partial \theta(t, u)}{\partial t} = \frac{\sigma^2}{2} \Delta \theta(t, u) + f(\theta(t, u)), \quad \theta(0, \cdot) = l(\cdot),
\] (5.5)

where \( l \) is a continuous function representing the initial conditions. For its definition let \( R < \infty \) such that we can define the continuous function \( l : R^d \rightarrow [0, 1] \) with \( R^d \) being the \( d \)-dimensional cube with diameter \( 2R \). Furthermore, let \( \lambda_\varepsilon \) be a family of probability measures on \( \{0, 1\}^\varepsilon \mathbb{Z}^d \). Define \( a_\varepsilon = \lfloor \varepsilon r - 1 \rfloor \varepsilon \) for \( r \in (0, 1) \), the cube \( Q_\varepsilon = [0, a_\varepsilon)^d \cap \varepsilon \mathbb{Z}^d \) and let \( |Q_\varepsilon| \) be the cardinality of the cube \( Q_\varepsilon \). Then, for \( u \in a_\varepsilon \mathbb{Z}^d \) we set

\[
D(u, X) = \frac{1}{|Q_\varepsilon|} \sum_{v \in Q_\varepsilon} X(u + v),
\]

i.e. the local density of type 1 particles in the area \( Q_\varepsilon \) around \( u \). We say that for a continuous function \( l \) as above the family \( \lambda_\varepsilon \) has local density \( l \) if there is an \( r \in (0, 1) \) such that for all \( 0 < \delta, R < \infty \) we have

\[
\lim_{\varepsilon \rightarrow 0} \sup_{u \in a_\varepsilon \mathbb{Z}^d, |u| \leq R} \lambda_\varepsilon(|D(u, X) - l(u)| > \delta) = 0.
\]
Note, that in the following we consider \( \varepsilon \in (0, 1] \) such that \( a_\varepsilon = \varepsilon \) which means that we approximate the local densities by a continuous function on \( \mathbb{R} \) with compact support. Furthermore, it is obvious that a product of Bernoulli distributions satisfies these conditions. To see this, consider the Bernoulli product measure \( \lambda \) given by

\[
\lambda(X(u_i) = 1 : i = 1, \ldots, n) = \prod_{i=1}^{n} l(u_i) \text{ for all } n \in \mathbb{N} \text{ and } u_i \in \varepsilon \mathbb{Z}^d,
\]

which already yields the assertion.

Having introduced this notation we can now state a version of Theorem 1.2 in [CDP13] (see also Theorem 3.1 in [Dur14]) about the hydrodynamic limit of a voter model with perturbations.

**Theorem 5.2** (Hydrodynamic limit - Theorem 3.1 in [Dur14]) Let \( d \geq 3 \), \( l : \mathbb{R}^d \to [0, 1] \) be continuous and \( Y^\varepsilon \) be a voter model with perturbations satisfying conditions (5.1)-(5.3). Assume that the initial conditions \( Y^\varepsilon(0) \) have laws \( \lambda_\varepsilon \) with local density \( l \) and let

\[
\theta^\varepsilon(t, u_\varepsilon) = P(Y^\varepsilon_{\varepsilon - 2t}(u_\varepsilon) = 1).
\]

If \( u_\varepsilon \to u \), then

\[
\lim_{\varepsilon \to 0} \theta^\varepsilon(t, u_\varepsilon) = \theta(t, u), \text{ for all } t > 0 \text{ and } u \in \mathbb{R}^d.
\]

Here, \( \theta(t, u) \) is the solution of the partial differential equation given in equation (5.5) with initial condition \( \theta(0, u) = l(u) \) and \( f \) given in equation (5.4).

In [CDP13] it is furthermore proved that this measure-valued convergence also holds in \( L^2 \) on compact sets of \( \mathbb{R}^d \), see Theorem 1.3 in their paper.

The next step in the analysis is to characterize the limit behavior of the reaction-diffusion equation. Therefore, we will make use of a Proposition which was already proved in [CDP13] and originally developed by Aronson and Weinberger in [AW78].

**Proposition 5.3** (cf. Propositions 1.4, 1.5 in [CDP13]) Under the assumptions of Theorem 5.2 the following holds.

(i) Let \( f < 0 \) on \((0, 1)\) and \( f'(0) < 0 \). Then there exists a \( k > 0 \) such that if \( \delta > 0 \) there are positive constants \( L_\delta, c_\delta \) and \( C_\delta \). If \( L \geq L_\delta \) and \( l(u) \leq 1 - \delta \) for \( |u| \leq L \), then

\[
\theta(t, u) \leq C_\delta e^{-\delta t}, \text{ for } |u| \leq L + 2kt.
\]

(ii) Let \( f > 0 \) on \((0, 1)\), \( f(0) = f(1) = 0 \), \( f'(0) > 0 \) and \( f'(1) < 0 \). Furthermore, suppose that the initial value of \( \theta \) is non-trivial, i.e. \( \theta(0, u) \in (0, 1) \). Then there exists a \( k > 0 \) such that

\[
\liminf_{t \to \infty} \inf_{|u| \leq 2kt} \theta(t, u) = 1.
\]
This Proposition characterizes the limit behavior of the limiting partial differential equation from which we want to derive the limit behavior of the rescaled voter model with perturbation, at least for \( \varepsilon \) small enough. To complete this line of arguments we need to connect the asymptotic behavior of the reaction-diffusion equation to that of the particle system. Therefore, we need to specify the rate of convergence of the perturbation \( g^\varepsilon \), i.e.

\[
\sum_{i=0}^{1} \| g^\varepsilon_i - g_i \|_\infty \leq c \varepsilon^{r_0}, \tag{5.6}
\]

for some \( r_0 > 0 \) and a constant \( c \). Furthermore, we need an estimate for the reaction-diffusion equation. We formulate it in the following assumption.

**Assumption 5.4** There are constants \( 0 < v < 1, c_2, C_2, k > 0, L_0 \geq 3 \) such that for all \( L \geq L_0 \) the following holds. If \( \theta(0,u) \leq v \) for \( |u| \leq L \), then for all \( t \geq 0 \)

\[
\theta(t,u) \leq C_2 e^{-c_2 t}, \text{ for all } |u| \leq L + 2kt.
\]

Now, we can state the linking theorem between the particle system and the reaction-diffusion equation.

**Theorem 5.5** (cf. Theorem 1.16 in [CDP13]) Suppose that Assumption 5.4 and equation (5.6) hold and that the configuration \( 0 \) is a trap, i.e.

\[
g^\varepsilon(0,\ldots,0) = 0, \text{ for } 0 < \varepsilon \leq \varepsilon_0.
\]

Furthermore, assume that the kernels \( p \) (interaction kernel) and \( q \) (law of \( (\xi_1,\ldots,\xi_K) \)) have finite support. Then for \( \varepsilon \) small \( 0 \)'s take over, i.e. for all \( L \) it holds

\[
P(Y_{\varepsilon^{-2}\varepsilon u} = 0 \text{ for all } u \in [-L,L]^d \text{ for } t \text{ large enough}) = 1, \tag{5.7}
\]

whenever the initial configuration has infinitely many sites in state \( 0 \).

Note, that this is a result about extinction of a species. The same kind of result can also be derived for the coexistence of the two species but since we do not need it, we refer the reader (once again) to the paper of Cox, Durrett and Perkins dealing with the general theory of voter models with perturbations, [CDP13].

### 5.2 The rescaled cVMBC

Finally, we have introduced all the necessary ingredients for the analysis of the cVMBC with weak interaction. Hence, this section is devoted to the definition of the cVMBC on \( \mathbb{Z}^d \) for \( d \geq 3 \) in terms of the presented notation above and its corresponding limit result.
To start with we review the transition rates of the cVMBC which are given by
\[
c(u, X) = (1 - X(u)) \left( \sum_v a(v, u)X(v) + \gamma \sum_{v,w} b(w, (v, u))X(v)X(w) \right) \\
+ X(u) \left( (1 + \alpha) \sum_v a(v, u)(1 - X(v)) + \gamma \sum_{v,w} b(w, (v, u))(1 - X(v))X(w) \right).
\]

The interaction kernels are given by
\[
a(u, v) = \frac{1}{2d} \mathbb{1}_{|u - v| = 1}
\]
and
\[
b(u, (v, w)) = \frac{1}{2d(2d - 1)} \mathbb{1}_{|u - v| = |v - w| = 1, u \neq w}.
\]

Rewriting the interaction parameters by \( \alpha^\varepsilon := \varepsilon^2 \alpha \) and \( \gamma^\varepsilon := \varepsilon^2 \gamma \) we obtain
\[
c^\varepsilon(u, X) = (1 - X(u)) \sum_v a(v, u)X(v) + X(u) \sum_v a(v, u)X(v) \\
+ (1 - X(u))\gamma^\varepsilon \sum_{v,w} b(w, (v, u))X(v)X(w) \\
+ X(u) \left( \alpha^\varepsilon \sum_v a(v, u)(1 - X(v)) + \gamma^\varepsilon \sum_{v,w} b(w, (v, u))(1 - X(v))X(w) \right) \\
= c^\varepsilon(u, X) + \varepsilon^2 c_p^\varepsilon(u, X).
\]

In order to apply the existing theory for voter models with perturbations we first need to check the conditions from equations (5.1)-(5.3).

**Proposition 5.6** (Rescaled cVMBC is a voter model with perturbations) Let \( X \) be the cVMBC on \( \mathbb{Z}^d \) with \( d \geq 3 \) and interaction kernels \( a(\ldots) \) as given in equation (5.8) and \( b(\ldots, \ldots) \) as in (5.9). Additionally, we rescale the selection and cooperation parameter such that the model is driven by \( \alpha^\varepsilon \) and \( \gamma^\varepsilon \). Then the transition rates are as in (5.10) and \( X \) satisfies conditions (5.1)-(5.3).

**Proof:** The form of the transition rates is immediate. For conditions (5.1)-(5.3) to be satisfied we apply Proposition 5.1. Therefore, with \( e_i \) being the unit vector in the \( i \)-th
Weak Interaction Limit

\[ \zeta^1 = e_1, \ldots, \zeta^d = e_d, \zeta^{d+1} = -e_1, \ldots, \zeta^{2d} = -e_d, \]
\[ \zeta^{2d+1} = 2e_1, \zeta^{2d+2} = e_1 + e_2, \ldots, \zeta^{3d} = e_1 + e_d, \]
\[ \ldots \]
\[ \zeta^{d^2+2d+1} = e_1 - e_2, \ldots, \zeta^{d^2+3d-1} = e_1 - e_d, \]
\[ \ldots \]
\[ \zeta^{2d^2+d+1} = -2e_1, \ldots, \zeta^{2d^2+2d} = -e_1 - e_d, \]
\[ \ldots \]
\[ \zeta^{3d^2+d+1} = -e_1 + e_2, \ldots, \zeta^{3d^2+2d-1} = -e_1 + e_d, \]
\[ \ldots \]
\[ \zeta^{4d^2-d+2} = -e_d + e_1, \ldots, \zeta^{4d^2} = -e_d + e_{d-1}. \]

Note, that the difference of the same unit vector is not included, e.g. \( \zeta^{d^2+2d+1} = e_1 - e_2 \) instead of \( e_1 - e_2 = 0 \). We immediately see that for \( u \in \mathbb{Z}^d \) we have that

\[ \{ v \in \mathbb{Z}^d : a(v, u) > 0 \} \subset \{ u + \zeta^1, \ldots, u + \zeta^{4d^2} \}. \]

Furthermore we define

\[ h_0^\varepsilon = \alpha f_0(u, X) + \frac{\gamma}{2d(2d-1)} \sum_{v \in \mathbb{N}_u} (1 - X(v)) \sum_{w \in \mathbb{N}_u \setminus \{u\}} X(w) \]
\[ =: \tilde{g}_0^\varepsilon(X(u + \zeta^1), \ldots, X(u + \zeta^{4d^2})), \]
\[ h_1^\varepsilon = \frac{\gamma}{2d(2d-1)} \sum_{v \in \mathbb{N}_u} X(v) \sum_{w \in \mathbb{N}_u \setminus \{u\}} X(w) \]
\[ =: \tilde{g}_1^\varepsilon(X(u + \zeta^1), \ldots, X(u + \zeta^{4d^2})). \]

Since both, \( \tilde{g}_0^\varepsilon \) and \( \tilde{g}_1^\varepsilon \) are independent of \( \varepsilon \) the condition

\[ \lim_{\varepsilon \to 0} \| \tilde{g}_i^\varepsilon - \tilde{g}_i \|_\infty = 0 \text{ for } i \in \{0, 1\} \]

is trivially satisfied by setting \( \tilde{g}_i = \tilde{g}_i^\varepsilon \). Hence, the statement follows with Proposition 5.1 and we are done.

Knowing that our model fits into the class of voter models with perturbations we can now proceed to the analysis of its limit behavior by the corresponding reaction-diffusion equation.
5.2 The rescaled cVMBC

5.2.1 Limit behavior of the rescaled cVMBC

We start by characterizing the limit behavior of the rescaled cVMBC $X^\varepsilon$.

**Corollary 5.7** (Hydrodynamic limit of the cVMBC) Let $X^\varepsilon$ be a cVMBC on $\mathbb{Z}^d$ with $d \geq 3$ having transition rates as in equation (5.10) and interaction kernels as in equations (5.8) and (5.9). Furthermore, let the initial configuration be a Bernoulli product measure with probability $p \in (0,1)$. Then, writing

$$\theta^\varepsilon(t,u_\varepsilon) = P(X^\varepsilon_{\varepsilon^{-2}t}(u_\varepsilon) = 1),$$

we have that if $u_\varepsilon \to u$, then

$$\lim_{\varepsilon \to 0} \theta^\varepsilon(t,u_\varepsilon) = \theta(t,u), \text{ for all } t > 0 \text{ and } u \in \mathbb{R}^d.$$

Here, $\theta(t,u) := P(X_t(u) = 1)$ satisfies the partial differential equation given by

$$\frac{\partial}{\partial t} P(X_t(u) = 1) = \frac{1}{2d} \Delta P(X_t(u) = 1) + f(P(X_t(u) = 1)).$$

**Proof:** Since $X^\varepsilon$ is a voter model with perturbations due to Proposition 5.6 this is an immediate consequence of Theorem 5.2.  

Next, we analyze the limit behavior of this system which in turn approximates the limit behavior of the rescaled cVMBC for $\varepsilon$ small enough. Therefore we need to calculate the drift term $f(\theta)$. Before we can do this explicitly we need some new notation.

We denote by $v_1$ a random vector chosen uniformly from the set $\{\pm e_i : i = 1, ..., d\}$ and $v_2$ a random vector chosen uniformly from the set $v_1 + \{\pm e_i : i = 1, ..., d\} \setminus \{-v_1\}$. Furthermore, let $B^u = (B^u_t)_{t \geq 0}$ be a random walk starting with a particle at $u \in \mathbb{Z}^d$. Then, we define for $u_1, u_2, u_3 \in \mathbb{Z}^d$ the probabilities

$$p(u_1|u_2) := P(B^u_{t_{1}} \neq B^u_{t_{2}} \text{ for all } t \geq 0),$$

$$p(u_1,u_2) := P(\exists t \geq 0 : B^u_{t_{1}} = B^u_{t_{2}}),$$

$$p(u_1,u_2|u_3) := P(\exists t \geq 0 : B^u_{t_{1}} = B^u_{t_{2}} \text{ and } \forall t \geq 0 : B^u_{t_{1}} \neq B^u_{t_{3}}, B^u_{t_{2}} \neq B^u_{t_{3}}).$$

In general we will separate vertices by a $|$ when their corresponding random walks never meet. Now, we can derive the drift term $(\tilde{v}_2 := v_1 + v_2)$:

$$f(\theta) = \langle (1 - X(0))\gamma X(v_1)X(v_1 + v_2)
- X(0)\alpha(1 - X(v_1)) - X(0)\gamma(1 - X(v_1))X(v_1 + v_2) \rangle_\theta
= \gamma(1 - \theta)\theta^2 p(0|v_1|\tilde{v}_2) + \gamma(1 - \theta)\theta p(0|v_1|\tilde{v}_2)
- \alpha \theta(1 - \theta)p(0|v_1) - \gamma(1 - \theta)\theta p(0|v_1|\tilde{v}_2) - \gamma \theta(1 - \theta)p(0|v_1|\tilde{v}_2)
= \theta(1 - \theta)(\gamma p(0|v_1|\tilde{v}_2) - p(v_1|0\tilde{v}_2) - \alpha p(0|v_1)).$$

(5.11)
Note, that the vertices $v_1$ and $\bar{v}_2$ are still random.

It is obvious that the only roots of $f$ are 0 and 1 which means that the probability of finding a site occupied by a particle is 0 or 1, respectively. Hence, these roots are equivalent to the trivial configurations $\emptyset$ and $\mathbb{1}$. It remains to derive conditions for $\gamma$ and $\alpha$ such that the cVMBC converges either to the "all-zero" or the "all-one" configuration.

In order to find such conditions we apply Proposition 5.3. Therefore, we need to control the derivative of $f$ at the fixed points. The derivative reads

$$f'(\theta) = (1 - 2\theta)(\gamma(p(0|v_1\bar{v}_2) - p(v_1|0\bar{v}_2)) - \alpha p(0|v_1)).$$

Trivially, we find $f'(0) < 0$ if $\alpha > \gamma$ which yields that $\emptyset$ is attracting. This is due to the following estimate

$$f'(0) = \gamma(p(0|v_1\bar{v}_2) - p(v_1|0\bar{v}_2)) - \alpha p(0|v_1) < \gamma p(0|v_1\bar{v}_2) - \alpha p(0|v_1) < (\gamma - \alpha)p(0|v_1) < 0.$$

The inequality follows by the strict positivity of the coalescence probabilities and the obvious equality

$$p(0|v_1) = p(0|v_1\bar{v}_2) + p(0\bar{v}_2|v_1) + p(0|v_1|\bar{v}_2).$$

Furthermore $f$ is strictly negative on $(0, 1)$ which allows us to apply the first statement of Proposition 5.3.

On the other hand cooperation is a winning strategy if $\alpha = 0$ and $\gamma > 0$. This is clear because in this case we have

$$f'(1) = -\gamma(p(0|v_1\bar{v}_2) - p(v_1|0\bar{v}_2)) < 0.$$

Since $f'(0) = -f'(1)$ and $f$ is obviously positive on $(0, 1)$, the stability of $\mathbb{1}$ follows with the second part of Proposition 5.3.

Of course these first estimates are not very satisfactory. In order to improve the conditions, we use the Markov property of the rescaled cVMBC and properties of the coalescence probabilities given in Lemma 5.9. A first step analysis together with Proposition 5.3 and Theorem 5.5 then yields the following Theorem.

**Theorem 5.8** (Asymptotic behavior) Let $d \geq 3$ and $X^\varepsilon$ be the rescaled cVMBC. Then for $X^\varepsilon(0)$ translation invariant, non-trivial and induced by a product Bernoulli measure with parameter $p \in (0, 1)$, we obtain the following statements:

(i) for $\alpha \geq \frac{4\varepsilon - 2d + 1}{3d(2d - 1)}$, defectors take over for $\varepsilon$ sufficiently small.

(ii) for $\alpha \leq \frac{\gamma}{(2d - 1)3d}$, cooperators take over for $\varepsilon$ sufficiently small.
5.3 Proof of Theorem 5.8

In the following proofs we need to calculate or approximate different coalescence probabilities. Therefore, we have to specify the locations of different vertices. In what follows, fixed sites will be labeled by \( u_k^{(i,j)} \) where \( k \) is the \( l_1 \)-distance (Manhattan distance) to the origin and \( (i,j) \) are the coordinates of the first two dimensions. If the distance \( k \) is larger than \( i + j \) the site can be located anywhere on the hyperplane \( (i,j,...) \) in \( \mathbb{Z}^d \) with distance \( k \) to the origin. Furthermore, we denote by \( v_k \) a random vertex with distance \( k \) from the origin.

Before proving Theorem 5.8 we state some basic observations about coalescence probabilities.

**Lemma 5.9 (Coalescence probabilities)** The following identities and estimates for coalescence probabilities hold:

1. \( p(0|v_1v_2) > p(0|v_i) \) for \( i \in \{1, 2\} \),
2. \( p(0|u_2^{(1,-)}) = \frac{2d}{2d-2} p(0|u_1^{(1,0)}) - \frac{1}{2d-2} p(0|u_2^{(2,0)}) \),
3. \( p(0|v_k) < p(0|v_1) \) for \( k, l \in \mathbb{N} \) and \( k < l \), (this is equivalent to \( p(0v_k) > p(0v_l) \)),
4. \( p(0|u_1^{(1,0)}u_1^{(0,1)}) < p(u_1^{(1,0)}|0u_1^{(0,1)}) \),
5. \( p(0|u_2^{(1,1)}u_2^{(1,-1)}) < p(u_2^{(1,1)}|0u_2^{(1,1)}) \),
6. for \( u, v, w \in \mathbb{Z}^d \) the probability \( p(u|vw) \) strictly increases if we decrease the distance between \( v \) and \( w \) and/or increase the distance between \( u \) and \( v \) or \( u \) and \( w \).

**Proof:** Statement (i) is clear from the definition of the probabilities. For statement (ii), we do a one step analysis which yields

\[
p(0|u_2^{(1,-)}) = \frac{2d}{2d-2} p(0|u_1^{(1,0)}) - \frac{1}{2d-2} p(0|u_2^{(2,0)}).
\]

Rearranging the terms gives the second assertion. Proving statement (iii), we analyze the coalescence probabilities \( p(0v_k) \). Therefore, we rewrite them in the following way:

\[
p(0v_k) = \mathbb{P}(\exists T \geq 0 : B_T^0 = B_T^{v_k}) = \sum_{n=k}^{\infty} \mathbb{P}(B_n^0 = B_n^{v_k} | B_m^0 \neq B_m^{v_k} \text{ for all } m < n).
\]
5 Weak Interaction Limit

Now, for \( k < l \) we obtain

\[
p(0v_k) = \sum_{n=k}^{\infty} P(B_n^0 = B_n^{v_k} | B_m^0 \neq B_m^{v_k} \text{ for all } m < n)
\]

\[
= \sum_{n=k}^{l-1} P(B_n^0 = B_n^{v_k} | B_m^0 \neq B_m^{v_k} \text{ for all } m < n)
\]

\[
+ \sum_{n=l}^{\infty} P(B_n^0 = B_n^{v_k} | B_m^0 \neq B_m^{v_k} \text{ for all } m < n)
\]

\[
> \sum_{n=l}^{\infty} P(B_n^0 = B_n^{v_k} | B_m^0 \neq B_m^{v_k} \text{ for all } m < n)
\]

\[
> \sum_{n=l}^{\infty} P(B_n^0 = B_n^{v_k} | B_m^0 \neq B_m^{v_k} \text{ for all } m < n) = p(0v_l),
\]

where the last inequality follows by comparing the number of possible paths with length \( n \geq l \) from \( v_k \) to 0 and from \( v_l \) to the origin. Obviously, this number is always higher for \( v_k \) and thus, the assertion follows.

For statement (iv) we define

\[
A_1 := \{ \exists T \geq 0 : B_T^{(1,0)} = B_T^{(0,1)} \},
\]

\[
A_2 := \{ \exists T \geq 0 : B_T^0 = B_T^{(0,1)} \},
\]

\[
B := \{ \forall t \geq 0 : B_t^0 \neq B_t^{(1,0)} \},
\]

\[
C_1 := \{ \forall t \geq 0 : B_t^{(1,0)} \neq B_t^0 \},
\]

\[
C_2 := \{ \forall t \geq 0 : B_t^{(1,0)} \neq B_t^{(0,1)} \}.
\]

Note that, \( P(A_1) = p(0u_2^{(1,1)}) \), \( P(A_2) = p(0v_1) \), \( P(C_1) = p(0|v_1) \) and \( P(C_2) = p(0|u_2^{(1,1)}) \). Therefore, with assertion (iii) we have

\[
p(0|u_2^{(1,0)}u_1^{(0,1)}) = P(A_1 \cap B \cap C_1) = P(A_1 \cap C_1 | B)P(B) = P(A_1 | B)P(B)
\]

\[
< P(A_2 | B)P(B) = P(A_2 \cap C_2 | B)P(B) = P(A_2 \cap B \cap C_2) = p(u_1^{(1,0)}|0u_1^{(0,1)})
\]

The proof of the fifth statement is similar to the line of arguments used for assertion (iv). Only, one has to note that \( p(u_2^{(1,1)}|u_2^{(1,-1)}) > p(u_2^{(1,1)}u_2^{(1,-1)}) \) which is clear by comparing the probabilities

\[
P(B_k^0 = B_k^{u_2^{(1,1)}}) \text{ and } P(B_k^{u_2^{(1,1)}} = B_k^{u_2^{(1,-1)}})
\]
for each $k \geq 2$. Again, we count the number of possible paths with length $k \geq 2$ which gives the result.

The last statement follows immediately by assertion (iii).

Now, we can start proving our main result about the rescaled cVMBC. Therefore, we first prove the limit behavior of the reaction-diffusion equation with the help of Proposition 5.3.

\textbf{Lemma 5.10} Assume the situation given in Theorem 5.8 and $\gamma \geq \frac{4d^2-2d+1}{3d(2d-1)}$. Then $f'(0) < 0$ and $f < 0$ on $(0,1)$.

\textit{Proof:} Once we have shown that $f'(0) < 0$, i.e.

$$f'(0) = \gamma(p(0|v_1\tilde{v}_2) - p(v_1|0\tilde{v}_2)) - \alpha p(0|v_1) < 0,$$

we also know that $f < 0$ on $(0,1)$ which then shows the result. To start with, we rewrite the random vertices by averaging over fixed vertices. This yields

$$p(0|v_1\tilde{v}_2) - p(v_1|0\tilde{v}_2) = \frac{2d-2}{2d-1} \left( p(0|u_1^{(1,0)}u_2^{(1,1)}) - p(u_1^{(1,0)}|0u_2^{(1,1)}) \right) + \frac{1}{2d-1} \left( p(0|u_1^{(1,0)}u_2^{(2,0)}) - p(u_1^{(1,0)}|0u_2^{(2,0)}) \right).$$

Here, we summed up terms which are transformationally equivalent, i.e.

$$p(0|u_1^{(1,0)}u_2^{(1,\pm1)}) = p(0|u_1^{(0,-1)}u_2^{(\pm1,-1)}) = p(0|u_1^{(-1,0)}u_2^{(-1,\pm1)}) = p(0|u_1^{(0,1)}u_2^{(\pm1,1)}).$$

These equations hold because of the spatial homogeneity of the transition rates of the voter model. In the following we will make use of this frequently.

We analyze terms $A$ and $B$ separately. For the approximation of these terms we need that all the coalescing random walks jump with the same rate which holds due to the dynamics of the coalescing random walks induced by the voter model. Furthermore, applying Lemma 5.9 we calculate (the underlined terms cancel out, the dashed underlines are summed up in the following lines, the dotted underlines are estimated from below by Lemma 5.9)

\begin{align*}
A &= p(0|u_1^{(1,0)}u_2^{(1,1)}) - p(u_1^{(1,0)}|0u_2^{(1,1)}) \\
&= P(B_t^0 \text{ jumps}) \left( \frac{1}{2d} p(0|u_1^{(0,1)}u_1^{(1,0)}u_2^{(1,1)}) + \frac{1}{2d} p(u_1^{(-1,0)}|u_1^{(1,0)}u_2^{(1,1)}) \right) \tag{5.41}
\end{align*}

\begin{align*}
\text{furthermore, applying Lemma 5.9 we calculate \ldots}
\end{align*}
5 Weak Interaction Limit

\[ \frac{1}{2d} p(0|u_2^{(1,0)}) + \frac{1}{2d} p(0|u_1^{(1,0)}u_1^{(0,1)}) + \frac{1}{2d} p(0|u_1^{(1,0)}u_3^{(1,1)}) \\
+ \frac{1}{2d} p(0|u_1^{(1,0)}u_3^{(2,1)}) + \frac{2d - 4}{2d} p(0|u_2^{(1,0)}u_2^{(1,1)}) \]

\[ + P(B_t^{n_2^{(1,1)}}) \left( \frac{1}{2d} p(0|u_2^{(1,0)}) + \frac{1}{2d} p(0|u_1^{(1,0)}u_1^{(0,1)}) + \frac{1}{2d} p(0|u_1^{(1,0)}u_3^{(1,1)}) \\
+ \frac{1}{2d} p(0|u_1^{(1,0)}u_3^{(2,1)}) + \frac{2d - 4}{2d} p(0|u_2^{(1,0)}u_2^{(1,1)}) \right) \]

\[ - P(B_t^{n_1^{(1,0)}}) \left( \frac{1}{2d} p(u_2^{(2,0)}|0u_2^{(1,1)}) + \frac{1}{2d} p(u_2^{(1,-1)}|0u_2^{(1,1)}) \\
+ \frac{2d - 4}{2d} p(u_2^{(1,0)}|0u_2^{(1,1)}) \right) \]

\[ - P(B_t^0) \left( \frac{1}{2d} p(u_1^{(1,0)}|u_2^{(0,0)}u_2^{(1,1)}) + \frac{1}{2d} p(u_1^{(1,0)}|u_1^{(-1,0)}u_2^{(1,1)}) \\
+ \frac{1}{2d} p(u_1^{(1,0)}|u_1^{(0,-1)}u_2^{(1,1)}) + \frac{2d - 4}{2d} p(u_1^{(1,0)}|u_1^{(0,0)}u_2^{(1,1)}) \right) \]

\[ - P(B_t^{n_2^{(1,1)}}) \left( \frac{1}{2d} p(u_1^{(1,0)}|0u_1^{(0,1)}) + \frac{1}{2d} p(u_1^{(1,0)}|0u_3^{(1,2)}) \\
+ \frac{1}{2d} p(u_1^{(1,0)}|0u_3^{(2,1)}) + \frac{2d - 4}{2d} p(u_1^{(1,0)}|0u_3^{(1,1)}) \right) \]

\[ = \frac{1}{6d} \left( p(u_1^{(-1,0)}|u_1^{(1,0)}u_2^{(1,1)}) + p(u_1^{(0,-1)}|u_1^{(1,0)}u_2^{(1,1)}) + (2d - 4)p(u_1^{(0,0)}|u_1^{(1,0)}u_2^{(1,1)}) \\
+ p(0|u_2^{(1,1)}) + p(0|u_2^{(1,-1)}u_2^{(1,1)}) + p(0|u_1^{(1,0)}u_1^{(0,1)}) \\
+ p(0|u_1^{(1,0)}u_3^{(1,2)}) + p(0|u_1^{(1,0)}u_3^{(2,1)}) + (2d - 4)p(0|u_1^{(1,0)}u_3^{(1,1)}) \\
- p(u_2^{(1,-1)}|0u_2^{(1,1)}) - p(u_1^{(1,0)}|u_1^{(-1,0)}u_2^{(1,1)}) - p(u_1^{(1,0)}|u_1^{(0,-1)}u_2^{(1,1)}) \\
- (2d - 4)p(u_1^{(1,0)}|0u_1^{(0,1)}) - p(u_1^{(1,0)}|0u_3^{(1,2)}) - p(u_1^{(1,0)}|0u_3^{(2,1)}) \\
- p(u_1^{(1,0)}|0u_3^{(1,1)}) - (2d - 4)p(u_1^{(1,0)}|0u_3^{(1,1)}) \right) \]

\[ < \frac{1}{6d} \left( p(u_1^{(-1,0)}|u_1^{(1,0)}u_2^{(1,1)}) + p(u_1^{(0,-1)}|u_1^{(1,0)}u_2^{(1,1)}) + (2d - 4)p(u_1^{(0,0)}|u_1^{(1,0)}u_2^{(1,1)}) \\
+ p(0|u_2^{(1,1)}) + p(0|u_2^{(1,-1)}u_2^{(1,1)}) + p(0|u_1^{(1,0)}u_1^{(0,1)}) \\
+ p(0|u_1^{(1,0)}u_3^{(1,2)}) + p(0|u_1^{(1,0)}u_3^{(2,1)}) + (2d - 4)p(0|u_1^{(1,0)}u_3^{(1,1)}) \\
- 2p(u_1^{(1,0)}|u_1^{(-1,0)}u_2^{(1,1)}) - 2p(u_1^{(1,0)}|u_1^{(0,-1)}u_2^{(1,1)}) - p(0|u_2^{(1,-1)}u_2^{(1,1)}) \right) \]
5.3 Proof of Theorem 5.8

\[-p(0|u_1^{(1,0)}u_1^{(0,1)}) - (4d - 8)p(u_1^{(1,0)}|u_1^{(0,0)}u_2^{(1,1)}) < 0\]

\[< \frac{1}{6d} \left( p(u_1^{(-1,0)}|u_1^{(1,0)}) + p(u_1^{(0,-1)}|u_1^{(1,0)}) + (2d - 4)p(u_1^{(0,0)}|u_1^{(1,0)}) \right.\]
\[\left. + p(0|u_2^{(1,1)}) + (2d - 1)p(0|u_1^{(1,0)}) \right)\]

\[= \frac{1}{6d} \left( p(u_1^{(-1,0)}|u_1^{(1,0)}) + (2d - 2)p(0|u_2^{(2,1)}) + (2d - 1)p(0|u_1^{(1,0)}) \right)\]

\[= \frac{1}{6d} \left( p(u_1^{(-1,0)}|u_1^{(1,0)}) + \frac{(2d - 2)2d}{2d - 2}p(0|u_1^{(1,0)}) - \frac{2d - 2}{2d - 2}p(0|u_2^{(2,0)}) \right.\]
\[\left. + (2d - 1)p(0|u_1^{(1,0)}) \right)\]

\[= \frac{4d - 1}{6d}p(0|u_1^{(1,0)}).\]

For term \(B\) we do the simple estimate

\[0 < p(0|u_1^{(1,0)}u_2^{(2,0)}) - p(u_1^{(1,0)}|0u_2^{(2,0)}) < p(0|u_1^{(1,0)}).\]

This yields

\[p(0|v_1\bar{v}_2) - p(v_1|0\bar{v}_2) = \frac{2d - 2}{2d - 1} \left( p(0|u_1^{(1,0)}u_2^{(2,1)}) - p(u_1^{(1,0)}|0u_2^{(2,1)}) \right)\]
\[+ \frac{1}{2d - 1} \left( p(0|u_1^{(1,0)}u_2^{(2,0)}) - p(u_1^{(1,0)}|0u_2^{(2,0)}) \right)\]
\[< \frac{(2d - 2)(4d - 1)}{(2d - 1)6d}p(0|u_1) + \frac{1}{2d - 1}p(0|u_1)\]
\[= \frac{4d^2 - 2d + 1}{3d(2d - 1)}p(0|u_1).\]

Now, inserting this estimate into \(f'(0)\) gives

\[f'(0) = \gamma(p(0|v_1\bar{v}_2) - p(v_1|0\bar{v}_2)) - \alpha p(0|v_1) < \left( \frac{4d^2 - 2d + 1}{3d(2d - 1)}\gamma - \alpha \right)p(0|v_1) \leq 0.\]

This finishes the proof.

\[\square\]

Lemma 5.11 Assume the situation of Theorem 5.8 and \(\alpha \leq \frac{\gamma}{(2d - 1)3d}\). Then \(f'(1) < 0\) and \(f > 0\) on \((0, 1)\).

Proof: Again, we only need to check the condition \(f'(1) < 0\), i.e.

\[f'(1) = -\gamma(p(0|v_1\bar{v}_2) - p(v_1|0\bar{v}_2)) + \alpha p(0|v_1) < 0\]
from which it follows that $f > 0$ on $(0, 1)$. In order to show this, we basically repeat the steps from the proof of the previous lemma, i.e. we perform a first step analysis of

$$p(0|v_1 \bar{v}_2) - p(v_1|0 \bar{v}_2) = \frac{2d - 2}{2d - 1} \left( p(0|u_1^{(1,0)} u_2^{(1,1)}) - p(u_1^{(1,0)}|0 u_2^{(1,1)}) \right)$$

$$= \frac{1}{2d - 1} \left( p(0|u_1^{(1,0)} u_2^{(2,0)}) - p(u_1^{(1,0)}|0 u_2^{(2,0)}) \right)$$

Again, estimating terms $A$ and $B$ separately we obtain the following. For $A$ by applying Lemma 5.9 we obtain the following simple estimate:

$$A = p(0|u_1^{(1,0)} u_2^{(1,1)}) - p(u_1^{(1,0)}|0 u_2^{(1,1)}) > 0.$$

For term $B$ we can calculate (solid, dashed and dotted underlined are summed up to the corresponding positive or negative term in the following line)

$$B = p(0|u_1^{(1,0)} u_2^{(2,0)}) - p(u_1^{(1,0)}|0 u_2^{(2,0)})$$

$$= \frac{1}{6d} \left( p(u_1^{(0,1)} u_2^{(1,0)} u_2^{(2,0)}) + p(u_1^{(-1,0)} u_1^{(1,0)} u_2^{(2,0)}) + p(u_2^{(-1,0)} u_1^{(1,0)} u_2^{(2,0)}) \right)$$

$$+ \frac{1}{d} \left( p(0|u_1^{(0,1)} u_1^{(1,0)} u_2^{(2,0)}) + p(0|u_2^{(1,0)} u_1^{(1,0)} u_2^{(2,0)}) \right)$$

$$- p(u_2^{(-1,0)}|0 u_2^{(2,0)}) - p(u_1^{(-1,0)}|0 u_2^{(2,0)}) - p(u_2^{(-1,0)}|0 u_2^{(2,0)})$$

$$= \frac{1}{6d} \left( (2d - 2)p(u_1^{(0,1)} u_1^{(1,0)} u_2^{(2,0)}) + p(u_1^{(-1,0)} u_1^{(1,0)} u_2^{(2,0)}) + p(0|u_2^{(2,0)}) \right)$$

$$+ \frac{1}{d} \left( p(0|u_1^{(0,1)} u_1^{(1,0)} u_2^{(2,0)}) + p(0|u_1^{(0,1)} u_2^{(1,1)} u_2^{(2,0)}) \right)$$

$$- p(u_1^{(-1,0)}|0 u_2^{(3,0)}) - p(u_1^{(-1,0)}|0 u_2^{(3,0)})$$

$$> \frac{1}{6d} \left( (2d - 2)p(u_1^{(0,1)} u_1^{(1,0)} u_2^{(2,0)}) + p(u_1^{(-1,0)} u_1^{(1,0)} u_2^{(2,0)}) + 2p(0|u_1^{(1,0)}) \right)$$

58
5.3 Proof of Theorem 5.8

\[ + (2d - 2) p(0|u_2^{(1,1)} u_2^{(2,0)}) + (2d - 2) p(0|u_1^{(1,0)} u_2^{(2,1)}) + p(0|u_1^{(1,0)} u_3^{(3,0)}) \]
\[ - (2d - 2) p(0|u_2^{(1,1)} u_2^{(2,0)}) - (2d - 2) p(u_1^{(0,1)}|u_1^{(1,0)} u_2^{(2,0)}) \]
\[ - p(u_1^{-1,0}|u_1^{(1,0)} u_2^{(2,0)}) - (2d - 2) p(0|u_1^{(1,0)} u_3^{(2,1)}) - p(0|u_1^{(1,0)} u_3^{(3,0)}) \]
\[ = \frac{1}{3d} p(0|u_1^{(1,0)}). \]

Hence, we obtain
\[ f'(1) = -\gamma (p(0|v_1 \bar{v}_2) - p(v_1|0 \bar{v}_2)) + \alpha p(0|v_1) < \left( \alpha - \frac{1}{(2d - 1)3d} \right) p(0|v_1) \leq 0, \]
if \( \alpha \leq \frac{1}{(2d - 1)3d} \gamma \) and we are done. \( \square \)

Finally, we can prove the statements in Theorem 5.8.

**Proof of Theorem 5.8:** We want to apply Theorem 5.5 which links the reaction-diffusion equation to the interacting particle system. Hence, once we can apply this Theorem, the statements follow. We start with the case \( \alpha \geq \frac{4d^2 - 2d + 1}{3d^2} \gamma \). We need to show the following assertions which then allow us to apply Theorem 5.5:

1. Decay condition from Assumption 5.4.
2. Rate of convergence condition from equation (5.6).
3. \( 0 \) is a trap for \( X^\varepsilon \).

Since the initial configuration is non-trivial and translation invariant Lemma 5.10 holds. Hence, the first statement of Proposition 5.3 is valid which gives the decay condition in Assumption 5.4, thus showing the first assertion. Assertion 2 is trivially satisfied since \( g_\varepsilon^i \) is independent of \( \varepsilon \) for the rescaled cVMBC. Finally, the claimed stationarity of \( 0 \), assertion 3, immediately follows from the dynamics of the rescaled cVMBC and equation (5.11). Hence, we can apply Theorem 5.5 which completes the proof of the first statement.

For the case \( \alpha \leq \frac{4d^2 - 2d - 1}{3d^2} \gamma \) we follow the same line of arguments. Instead of Lemma 5.10 we now apply Lemma 5.11. In order to apply Proposition 5.3 we reverse the roles of cooperators and defectors such that assertion 1 is satisfied. This follows from having \( f'(1) < 0 \) and \( f > 0 \) on \((0,1)\) by Lemma 5.11 which now translates to \( f'(0) < 0 \) and \( f < 0 \) on \((0,1)\) by the switched roles of cooperators and defectors. Again, the rate of convergence condition is trivially satisfied. The same holds for the stationarity of \( 0 \). Thus, once again we can apply Theorem 5.5 which shows statement (ii) and we are done. \( \square \)
CHAPTER 6

Limit behavior for large cooperation rates

Until now, we have analyzed the survival of cells producing a public good in a heterogeneous population in dimensions $d = 1$ and $d \geq 3$. We proved that under suitable parameter configurations cooperators are able to survive, in fact if they survive they take over the colony. In one dimension we have a fairly strong result yielding that for $\gamma > \alpha$ this holds almost surely. On the other hand we could prove that in dimensions $d \geq 3$ and under an adequate scaling of the parameters cooperation is the winning strategy when $\gamma \geq 3d(2d - 1)\alpha$. But still, the biologically most interesting case dimension $d = 2$ remains to be studied. Unfortunately, all the techniques applied so far fail in this dimension. Hence, we need to come up with a different idea.

In the following we will use tools provided by percolation theory on oriented graphs for the proof of a limit result in dimensions $d \geq 1$ (and thus also $d = 2$). By applying percolation theory we loose accuracy in the conditions on the parameter, i.e. our result holds for unspecified large cooperation rates $\gamma$. This is due to the approximation of probabilities of so called good events which need to be arbitrarily close to one.

Additionally, we have to alter the dynamics of the so far studied cVMBC in order to apply these methods. As it turns out it is necessary to decrease the interaction rate between cooperators and defectors.

The chapter is structured as follows. We will first make the necessary changes to the cVMBC in Section 6.1 and state the result, see Theorem 6.3. Thereafter, in Section 6.2 we review some definitions and results needed for the proof of this Theorem which will be given in Section 6.3. The proof itself consists of two steps:

(i) the comparison with a pure growth-process given in Definition 6.4 and

(ii) a percolation argument which then shows the extinction of defectors for $\gamma$ large enough.
Thereafter, in Section 6.4 we take a look at the reason for changing the dynamics of the cVMBC, review other potential interaction kernels and discuss simulation results.

### 6.1 Definition and Result for the sVMBC

Throughout this whole chapter we consider the case $V = \mathbb{Z}^d$ with $d \geq 1$.

First of all we state the altered model dynamics. The change affects the interaction rate of defectors using the cooperation mechanism. We change this rate from $\gamma$ to $f(\gamma)$ for some continuous function $f : \mathbb{R}_{\geq 0} \rightarrow \mathbb{R}_{\geq 0}$ such that the aid provided by cooperators to defectors is smaller than to cooperators, i.e. $f(\gamma) < \gamma$ whenever $\gamma > 0$. Biologically spoken this resembles some kind of kin-selection meaning that cooperators favor to support other cooperators. Another interpretation could be that defectors are not able to use the resources produced by cooperators as efficiently as cooperators do, again yielding a reduced cooperation rate for the defecting individuals. To be more precise the transition rates now read:

If $X(u) = 0$, then

$$c(u, X) = \sum_{v \in \mathbb{Z}^d} a(v, u)X(v) + \gamma \sum_{v, w \in \mathbb{Z}^d} X(v)X(w)b(w, (v, u)); \quad (6.1)$$

if $X(u) = 1$, then

$$c(u, X) = (1 + \alpha) \sum_{v \in \mathbb{Z}^d} a(v, u)(1 - X(v)) + f(\gamma) \sum_{v, w \in \mathbb{Z}^d} (1 - X(v))X(w)b(w, (v, u)), \quad (6.2)$$

where $a(., .)$ is a Markov kernel and $b(., (., .))$ is the interaction kernel given in equation (3.4) (the kernel used in the definition of the cVMBC).

**Definition 6.1** (Selective Voter Model with Bias and Cooperation) Let $a(., .)$ be a Markov kernel from $\mathbb{Z}^d$ to $\mathbb{Z}^d$ with $d \geq 1$ satisfying (3.1) and $b(., (., .))$ be a Markov kernel from $\mathbb{Z}^d$ to $\mathbb{Z}^d \times \mathbb{Z}^d$ satisfying equations (3.2) and (3.4). Then, if the transition rates are given by equations (6.1) and (6.2), the VMBC is called a selective Voter Model with Bias and Cooperation (sVMBC).

Note, that the conditions in equations (3.1) and (3.2) ensure the existence of a Markov process possessing the claimed interaction rates.

Additionally, we will need the following assumption on the transition rates.

**Assumption 6.2** For $f : \mathbb{R}_{\geq 0} \rightarrow \mathbb{R}_{\geq 0}$ continuous we assume that $f(\gamma) \in o(\gamma)$.

We can now state our limit result for large cooperation rates $\gamma$ which is valid in arbitrary dimensions.
Theorem 6.3 (sVMBC-limit) Let $V = \mathbb{Z}^d$ with $d \geq 1$ and $a(., .)$ be the uniform nearest neighbor random walk kernel, i.e. $a(u, v) = \frac{1}{2d} 1_{\{|u-v|=1\}}$ for $u, v \in \mathbb{Z}^d$ and $X$ be the sVMBC-process with $\alpha, f(\gamma) \geq 0$ such that Assumption 6.2 is satisfied. Additionally let the process start in some non-trivial translation invariant initial configuration. Then there exists a $C < \infty$ such that for all $\gamma > C$ the defector-type dies out, i.e.

$$P(\lim_{t \to \infty} X_t = 1) = 1.$$ 

The proof, as already mentioned in the introduction, consists of two steps. First, we show that for $\gamma$ large enough we can couple the sVMBC with a pure growth-process, namely the Richardson model, in finite space-time boxes. This coupling will be used for the definition of so called good events, necessary for the comparison with a percolation process. The Richardson model provides an estimate for the occurrence of these good events which is then used for the comparison of the sVMBC with a oriented percolation. This yields, that the sVMBC stochastically dominates a super-critical percolation process which implies the almost sure survival of cooperators. To order to prove extinction of defectors, we introduce a second percolation graph which is more adapted to the model dynamics. The lack of percolation of defectors on this graph then shows that cooperators create an expanding cluster free of defectors. In general, the proof follows the line of arguments given in [Dur92, Section 3] and [EL16, Section 4].

6.2 Preliminaries

In this section we define a pure growth-process as introduced by Richardson in [Ric73] which will be used for the stochastic domination of a super-critical percolation process. This percolation process and further results on its limit behavior will be reviewed afterwards.

6.2.1 Pure growth-process

We start with the definition of the Richardson model as introduced and studied in [Ric73].

Definition 6.4 (Richardson model) Let $(R_t)_{t \geq 0}$ be a stochastic spin-system on $\mathbb{Z}^d$, i.e. a stochastic process on $\{0, 1\}^{\mathbb{Z}^d}$, with transition rate

$$c(u, R) = \lambda (1 - R(u)) \frac{1}{2d} \sum_{v \in \mathbb{Z}^d : |v-u|=1} R(v),$$

where $u \in \mathbb{Z}^d$ and $\lambda > 0$. We call $R_t$ the Richardson model or equivalently a pure growth-process with parameter $\lambda$. 

63
Taking a closer look at the dynamics, this process is quite easy to understand. Basically, it is a spin-system where each site occupied by an individual (labeled by a 1) reproduces with rate \( \lambda \) and replaces one randomly chosen neighbor. Once a site is labeled by a 1 it remains occupied forever. This explains the notion of a pure growth-process.

In [Ric73] the expansion behavior of this interacting particle system was studied. We state an immediate consequence of Richardson’s result adapted to our setting. Therefore, we write \( B_N := [-N, N]^d \) for the \( N \)-ball on \( \mathbb{Z}^d \) and denote by \( \widehat{R}_t \) the set of sites labeled by a 1 in the Richardson model, i.e.

\[
\widehat{R}_t := \{ u \in \mathbb{Z}^d : R_t(u) = 1 \}.
\]

Furthermore, let \([t]\) be the integer part of \( t \in \mathbb{R} \). Then, we have the following Theorem.

**Theorem 6.5** (Expansion of Richardson model - Theorem 1 in [Ric73]) Let \( R_t \) be the Richardson model on \( \mathbb{Z}^d \) with parameter \( \lambda > 0 \) started with a single individual at the origin, i.e. \( \widehat{R}_0 = \{ 0 \} \). Then for all \( \varepsilon > 0 \) there exist constants \( k > 0 \) and \( 0 < T < \infty \) such that for all times \( t \geq T \) it holds

\[
P(B_{[t]} \subset \widehat{R}_{kt}) \geq 1 - \varepsilon.
\]

After showing that the sVMBC can be coupled to the Richardson process in finite space-time boxes, this theorem will be used for the appropriate definition of good events. These are necessary for the comparison and stochastic domination of a super-critical percolation process which then implies that the sVMBC also stochastically dominates this percolation process.

**6.2.2 Percolation**

Next, we give a brief introduction to percolation theory and state some well-known results. For a more detailed presentation of this topic see e.g. [Gri99].

We begin with the definition of the underlying graph structure. Therefore, let

\[
H^d := \{(z, n) \in \mathbb{Z}^d \times \mathbb{N} : z_1 + \ldots + z_d + n \text{ is even}\}
\]

be the set of vertices. Between these sites we define oriented edges by

\[
(z, n) \rightarrow (z', n') \text{ if and only if } n' = n + 1 \text{ and } z' = z \pm e_j \text{ for some } j \in \{1, \ldots, d\},
\]

where \( e_j \) is the unit vector of the \( j \)-th coordinate. This oriented graph will be denoted by \( \mathcal{H}^d \), see Figure 6.1 for a visualization in one dimension.

A vertex (or site) will be labeled independently of all other sites as open or closed with probability \( p \) or \( 1 - p \), respectively. Furthermore, we say that there is a path from
Figure 6.1: The oriented graph $\mathcal{H}^1$. Filled dots represent open sites, empty bullets represent closed sites.

$(z,n)$ to $(z',n')$ if there is a sequence of open vertices which are connected by oriented edges such that

$$(z,n) \rightarrow (z_1,n_1) \rightarrow \ldots \rightarrow (z_k,n_k) = (z',n')$$

with $(z_i,n_i) \in H^d$ for all $i = 1,\ldots,k$ and some $k \geq 1$. A composition of open sites in $\mathcal{H}^d$ which are connected to each other by paths is called a cluster. Here, $C(z,n)$ denotes the cluster containing $(z,n) \in H^d$, i.e.

$$C(z,n) := \{(z',n') \in \mathcal{H}^d : \exists \text{ path from } (z,n) \text{ to } (z',n') \text{ or from } (z',n') \text{ to } (z,n)\}.$$  

We assume that $(z,n)$ is connected to itself by a (trivial) path such that each open site is contained in its own cluster.

Now, the quantity of interest is the percolation probability $\theta(p,d)$, i.e. the probability for a site $(z,n) \in H^d$ to belong to an infinite cluster if sites are open with probability $p$. Of course, this percolation probability depends on the underlying structure of the graph. In mathematical terms we define

$$\theta(p,d) = P_p(|C(0,0)| = \infty),$$

with $P_p$ being the probability measure induced by the probability $p \in [0,1]$. Note, that due to the translation invariance of both, the graph $\mathcal{H}^d$ and the probability measure, without loss of generality we may consider the cluster containing the origin.

General results on percolation on $\mathcal{H}^d$ show that for $p$ large enough, the percolation probability of each site is positive. Often these kinds of results are given for bond percolations, i.e. instead of the sites being open or closed we define the edges (bonds) to be open or closed with probability $p$ and $1 - p$, respectively. As shown in Proposition 2.1 in [Dep08] and mentioned in [Lig99] these two kinds of percolations are transferable to each other by adapting the probabilities for sites or bonds to be open. The following result was proved for a bond percolation model.
Theorem 6.6 (Positive Percolation Probability - Theorem 4.1 in [Dur95]) Let $\mathcal{H}^1$ be the oriented graph defined above. Then for $1 - p < 6^{-36}$ we have

$$P_p(|C(0,0)| < \infty) \leq 1/20.$$ 

Note, that this already implies that for $\mathcal{H}^d$ with $d > 1$ the same holds. This is true because we can embed the graph $\mathcal{H}^1$ into the graph $\mathcal{H}^2$ by considering only the first component of the vertex. Hence, if a site belongs to the infinite cluster on $\mathcal{H}^1$ it also belongs to an infinite cluster on $\mathcal{H}^2$ which implies that the percolation probability

$$p_c(d) := \inf(p : \theta(p, d) > 0)$$

satisfies

$$p_c(d) \geq p_c(d + 1) \text{ for all } d \geq 1.$$

As already mentioned in the text, Theorem 6.6 implies the following.

Corollary 6.7 (Site Percolation - Proposition 2.1 in [Dep08]) Given a bond percolation model with positive percolation probability, i.e. $P_p(|C(0,0)| = \infty) > 0$, we can find a site percolation model with an adapted parameter $\tilde{p} \in (0, 1)$ such that the site percolation dominates the bond percolation.

Hence, all the results obtained for a bond percolation are transferable to site percolation models. In the following, we do not specify the concrete choice of model. Yet, in the proof of Theorem 6.3 in Section 6.3 it is more convenient to think of a site percolation model.

Another easy implication of Theorem 6.6 and Kolmogorov’s 0-1-law is the following theorem.

Theorem 6.8 (Infinite Cluster - Theorem 1.11 in [Gri99]) Given the oriented graph $\mathcal{H}^d$ with $d \geq 1$ and $p \geq p_c(d)$ there almost surely exists an infinite cluster of open sites.

If there is an infinite cluster of open sites, we say that percolation occurs. On the other hand we say that closed sites do not percolate if the probability for the size of a cluster of closed vertices to be infinity is zero, i.e.

$$P((0,0) \text{ is contained in a closed cluster with size } \infty) = 0.$$ 

Note however, that this case is not excluded by the theorems above. But, by increasing $p$ the probability for closed sites to percolate on the graph $\mathcal{H}^d$ can be made arbitrarily small. The following result deals with this situation.

Theorem 6.9 (Lack of Percolation - Theorem 5 in [Dur92]) Let $\mathcal{H}^d$ be the oriented graph from above. For $p$, the probability of a site to be open, close enough to one closed sites do not percolate.

This theorem will help in step (ii) of the proof of Theorem 6.3 to exclude that defectors invade the expanding cluster of cooperators. Particularly, the lack of percolation shows
that for some time $T > 0$ the density of closed sites vanishes, i.e. all sites above
the $[T]$-th level are open. Thus, connecting the definition of open and closed sites
appropriately to the sVMBC-dynamics, we can say that defectors die out.

Finally, we need to formally link the sVMBC to the percolation process. This can be
done by a comparison result formulated by Durrett in [Dur95]. In order to state this
theorem we need some assumptions on the processes of interest. Therefore, consider a
translation invariant nearest-neighbor interacting particle system $X_t : Z^d \rightarrow \{0, 1\}$
that can be constructed by a graphical representation as done exemplarily for the eVMBC
in Remark 3.6 in Chapter 3. Furthermore, we assume that there exist events $\Omega(z, n)$
which are determined by the configuration of $X_t$ on $\{Nz + (-N, N]^d \times \lfloor nT, (n + 1)T \rfloor$
for some $N, T > 0$. Then, adapting the comparison assumptions given in [Dur95,
Chapter 4] to our case we obtain the following:

**Assumption 6.10 (Comparison Assumptions)** For $(z, n) \in H^d$ there exist events
$G(z, n)$ such that

1. $G(z, n)$ is measurable with respect to the graphical representation in
   \[ \{Nz + (-3N, 3N]^d \times \lfloor nT, (n + 1)T \rfloor \text{ for all } (z, n) \in H^d. \]

2. $P(G(z, n)) \geq p$ for some $p \in (0, 1)$.

3. For all $i \in \{1, ..., d\}$ we have that
   \[ G(z, n) \cap \Omega(z, n) \subset \Omega(z \pm e_i, n + 1). \]

Later on we will call the events $\Omega(z, n)$ good events and the corresponding sites $(z, n)$
good sites or open sites. The event $G(z, n)$ on the other hand is the propagation event
of such a good site, i.e. if site $(z, n)$ is good and $G(z, n)$ happens then all sites of the
set $\{(z \pm e_i, n + 1) \in H^d : i \in \{1, ..., d\}\}$ are good as well (see the third assumption).
Furthermore, we require that $G(z, n)$ is measurable to a finite space-time box as given
in the first claim above. This implies that we can deduce the estimate in the second
statement based on states of sites in this finite box, i.e. the estimate $P(G(z, n)) \geq p$
holds regardless of the configuration outside of $\{Nz + (-3N, 3N]^d \times \lfloor nT, (n + 1)T \rfloor$.

Last but not least we need to know how to compare our process to the oriented
percolation process. Therefore, we cite a result of Durrett from [Dur95].

**Theorem 6.11 (Comparison result - Theorem 4.3 in [Dur95])** Assuming that the
comparison assumptions in 6.10 hold, we can define random variables $\omega(z, n)$ such
that $X_n$ dominates the oriented percolation process $(W_n)_{n \geq 0}$ with initial configuration
$W_0 = X_0$ and density at least $p$ for $p \in (0, 1)$, i.e. $W_n \leq X_n$ componentwisely for all
$n \geq 0$.

This comparison yields the survival of good sites in the process $X_n$ if the probability
$p$ is larger than the critical percolation probability $p_c(d)$. This is due to choosing
the random variables $\omega(z,n)$ as Bernoulli-distributed with probability $p$ such that $\mathbb{P}(G(z,n)) \geq \mathbb{P}(\omega(z,n))$. Then by the domination $X_n \geq W_n$ and Theorem 6.8 for $p \geq p_c(d)$ which says that the good sites in $W_n$ percolate the survival of good sites follows immediately. Such a percolation process $W_n$ with $p \geq p_c(d)$ is called a super-critical percolation process.

We are now able to prove Theorem 6.3.

### 6.3 Proof of Theorem 6.3

As already mentioned the proof is divided into two steps. First, by a comparison with the Richardson model, we show that the sVMBC stochastically dominates a percolation process where closed sites (referring to the presence of defectors) do not percolate. Then, altering the percolation structure we can show that this suffices to prove extinction of defectors or to be more precise to prove the expansion of a cluster of cooperators void of defectors. Summarizing this, we need to show that

(i) the sVMBC stochastically dominates a super-critical percolation process on $\mathcal{H}^d$

and

(ii) defectors die out for $t \to \infty$.

We start with the first assertion. We consider a block of cooperators, i.e. for some $N \geq 2$ the set $B_N$ consists of cooperators only. Due to the translation invariance condition on the initial configuration we almost surely find arbitrary many of such blocks for all values of $N \geq 2$. Note, that we do not assume that all the neighbors of such a block are defectors. The neighborhood of such a block $B_N$ may consist of defectors and cooperators.

In the following we estimate the transition rates for a defector and a cooperator to switch their type. Therefore, we denote by $p_D^i$, $i \in \{0, \ldots, 2d-1\}$, the transition rate of a defector at the edge of a cluster $B_N$ with $i$ cooperating neighbors other than the cooperator belonging to the cluster. E.g. if a defector is in the neighborhood of a cluster $B_N$ and has $i + 1$ cooperating neighbors (including the cooperator belonging to the cluster $B_N$), then its transition rate will be given by $p_D^i$. We estimate this rate from below by considering the minimal transition rate for such a defector, i.e.

$$p_D^i \geq \frac{1+i}{2d} + \gamma \frac{d}{2d(2d-1)} = \frac{1+i}{2d} + \gamma \frac{1}{2(2d-1)}. \quad (6.3)$$

The second term due to cooperation emerges from the cooperator contained in $B_N$ which has at least $d$ other cooperating neighbors. This worst case is given if the cooperator is at a corner of such a block.

Concerning the transition rate for a cooperator belonging to $B_N$ we denote it by $p_C$. We can neglect the index $i$ since our estimate is independent of the number...
of cooperating particles outside of $B_N$. Considering a function $f$ as assumed in Theorem 6.3 we obtain the following estimate:

$$p_C \leq (1 + \alpha) \frac{d}{2d} + f(\gamma) \frac{(i_1 \lor i_2 \lor \ldots \lor i_d)d}{2d(2d - 1)} \leq \frac{1 + \alpha}{2} + f(\gamma) \frac{1}{2},$$  

(6.4)

where $i_1, \ldots, i_d$ are the number of cooperating neighbors of the potentially $d$ defecting neighbors of a cooperator in $B_N$ and $i_1 \lor i_2 \lor \ldots \lor i_d$ denotes their maximum. Again, the worst case for a cooperator is the situation where the cooperator is surrounded by as many defectors as possible which have as many cooperating neighbors as possible. Hence, the first term scales with $d$ and the maximum of $i_1 \lor i_2 \lor \ldots \lor i_d$ can be estimated from above by $(2d - 1)$.

As mentioned, these estimates are the best, respectively the worst case scenarios for a defector, respectively a cooperator to maintain its type. These cases are given by analyzing particles at the corners of the cooperating cluster $B_N$. The reasoning now is as follows: given a cluster of cooperators $B_N$ this cluster expands slowly by transforming neighboring defectors to cooperators which then stay cooperators. Hence, the last estimate we need is the transition rate $p_C^n$ of a cooperator next to the cluster $B_N$, i.e.

$$p_C^n \leq (1 + \alpha) \frac{2d - 1}{2d} + f(\gamma) \frac{(2d - 1)^2}{2d(2d - 1)} = (1 + \alpha + f(\gamma)) \frac{2d - 1}{2d}. 

(6.5)$$

Next, rescaling time by $(2d - 1)/d\gamma$, fixing $\alpha \geq 0$ and $f$ such that it satisfies Assumption 6.2, the transition rates transform to

$$p_D' \geq \frac{(2d - 1)(1 + i)}{2d^2 \gamma} + \frac{1}{2d} \rightarrow \frac{1}{2d},

p_C \leq \frac{(2d - 1)(1 + \alpha)}{2d \gamma} + \frac{f(\gamma) 2d - 1}{2d} \rightarrow 0, 

p_C^n \leq \frac{1 + \alpha + f(\gamma) (2d - 1)^2}{2d^2} \rightarrow 0, 

(6.6)$$

for $\gamma \rightarrow \infty$ and for all $0 \leq i \leq 2d - 1$. Hence, whenever there is a block of cooperators of size $N \geq 2$ the rate for a defector next to the cluster to switch its type tends to $1/2d$ for $\gamma \rightarrow \infty$ independently of $i$. Additionally, a cooperator located in the neighborhood of another cooperator has a vanishing rate to change its type when $\gamma$ is sufficiently large. This means, that such a cooperator will not change its type in a finite time interval with high probability.

Due to this time-scaling and limit behavior for $\gamma \rightarrow \infty$ we can couple (at least in a finite space-time regime) the time-rescaled sVMBC $X_t$ to the Richardson model $R_t$ with rate 1 such that for all $\varepsilon > 0$ there exist constants $k = k(\varepsilon) > 0$ and $N = N(\varepsilon) > 0$ such that

$$P(B_{2N} \cap \tilde{R}_t \subset B_{2N} \cap \tilde{X}_t \text{ for all } t \in (kN, 2kN)] | \tilde{X}_0 = \tilde{R}_0 = B_N) > 1 - \frac{\varepsilon}{2},$$

(6.7)
for \( \gamma \) large enough. Here, \( \tilde{X}_t \) and \( \tilde{R}_t \) are the set-valued processes corresponding to \( X_t \) and \( R_t \) indicating the sites labeled by a 1.

Applying the shape-theorem of Richardson, see Theorem 6.5 above and the connectivity property of \( R_t \), i.e. cooperators only appear right next to other cooperators and then persist for all times, we can estimate that for some fixed \( k > 0 \) and \( N \) large enough it holds

\[
P(B_{2N} \not\subset \tilde{R}_t \text{ for some } t \in (kN, 2kN) | B_N \subset \tilde{R}_0) \leq P(B_{2N} \not\subset \tilde{R}_{kN} | B_N \subset \tilde{R}_0) \leq P(B_{2N} \not\subset \tilde{R}_{kN} | \tilde{R}_0 = \{0\}) \leq \varepsilon \overset{(6.8)}{2}.
\]

Now, we fix \( k \) and \( N \) in dependence of \( \varepsilon \) such that inequalities (6.7) and (6.8) hold. Then, using the continuity of \( f \) and thus the continuity of the transition probabilities with respect to \( \gamma \), there exists a constant \( C < \infty \) such that for \( \gamma > C \) we obtain

\[
P(B_{2N} \not\subset \tilde{X}_t \text{ for some } t \in (kN, 2kN) | B_N \subset \tilde{X}_0) \leq P(B_{2N} \not\subset \tilde{R}_t \text{ for some } t \in (kN, 2kN) | B_N \subset \tilde{R}_0) + P(B_{2N} \cap \tilde{R}_t \not\subset B_{2N} \cap \tilde{X}_t \text{ for some } t \in (kN, 2kN) | \tilde{X}_0 = \tilde{R}_0 = B_N) \leq \varepsilon \overset{(6.9)}{2}.
\]

This yields that a block of cooperators expands in all directions with arbitrary high probability.

Next, we use the inequality derived in (6.9) for the definition of good events on a percolation cluster. Therefore, we use the oriented graph \( H^d \) as introduced in Section 6.2.2, i.e. the set of vertices is given by

\[
H^d := \{(z, n) \in \mathbb{Z}^d \times \mathbb{N} : z_1 + \ldots + z_d + n \text{ is even}\},
\]

and sites are connected by oriented edges satisfying

\[
(z, n) \to (z', n') \leftrightarrow n' = n + 1 \text{ and } z' = z \pm e_j \text{ for some } j \in \{1, \ldots, d\}.
\]

We now connect our block construction from above with the percolation graph. Therefore, we say that a vertex \((z, n) \in H^d \) is open whenever

\[
B_N(z) \subset \tilde{X}_t \text{ for all } t \in nkN + (0, kN),
\]

with \( B_N(z) := Nz + [-N, N]^d \). Hence, by equation (6.9) we know that the probability for all sites which are reachable by one oriented edge from a open site to be open is greater than \( 1 - \varepsilon \), e.g. if \((z, n) \) is open, then with probability greater than \( 1 - \varepsilon \) all sites \((z \pm e_j, n + 1) \) for \( j \in \{1, \ldots, d\} \) are open, too.

Now, once we have checked the comparison assumptions from Assumption 6.10, the stochastic domination follows from Theorem 6.11. Concerning the measurability
proof of Theorem 6.3

6.3 Proof of Theorem 6.3

Figure 6.2: A defector at the bottom level can influence not only the nearest neighboring vertices of the following level on the percolation graph but also sites which are further away. Filled vertices satisfy the event condition \( \Omega \), white vertices have at least one defector in their corresponding space-time box.

assumption, we can construct the sVMBC by a graphical construction in the same way as we did it for the cVMBC in Remark 3.6 in Chapter 3. Furthermore, defining

\[
\Omega(z,n) := \{ B_N(z) \subset \bar{X}_t \text{ for all } t \in nkN + (0,kN) \},
\]

as a good event we immediately see that

\[
P(G(z,n)) := \mathbb{P}( (z \pm e_i, n+1) \text{ is good for all } i = 1, ..., d | (z,n) \text{ is good}) \geq \mathbb{P}(B_{2N}(z) \subset \bar{X}_t \text{ for all } t \in (n+1)kN + (0,kN) | B_N(z) \subset \bar{X}_{nkN}) \geq 1 - \varepsilon
\]

holds for \( \varepsilon > 0 \) by equation (6.9). Hence, by the definition of \( \Omega(z,n) \) and \( G(z,n) \) the claims 2 and 3 in Assumption 6.10 are satisfied. Applying Theorem 6.11 yields that the sVMBC stochastically dominates the set of open sites in the percolation process on \( \mathcal{H}^d \).

Since we can choose \( \varepsilon > 0 \) arbitrarily small by increasing \( \gamma \) we can make the probability from equation (6.10) arbitrarily close to one. If we choose \( \varepsilon > 0 \) such that \( 1 - \varepsilon \geq p_c(d) \) we can apply Theorem 6.8 for the percolation process. Using these parameters we have shown assertion (i).

Concerning statement (ii) we first note that by increasing \( \gamma \) we can push the density of open sites, i.e. the density of cooperators arbitrarily close to one. This is due to Theorem 6.9 above. Nevertheless, this is not sufficient for the defectors to die out because closed sites can emerge right next to other closed sites without being connected to this neighboring site through a path. For an explanatory picture of this situation see Figure 6.2.

In order to resolve this problem, we alter the percolation structure such that closed sites on an arbitrary level \( n \) will be connected by a path of closed edges to a closed site
on the initial level $n = 0$. Therefore, we consider the oriented graph $\tilde{H}^d$ given by the same vertex set as above, i.e. $H^d$ and the set of edges which are now given by

$$(z, n) \rightarrow (z', n')$$

if and only if

$$\{n' = n + 1 \text{ and } z' = z \pm e_j \text{ for some } j \in \{1, \ldots, d\}\}$$

or

$$\{n' = n \text{ and } z' = z \pm 2e_j \text{ for some } j \in \{1, \ldots, d\}\}.$$ 

For an illustration of this graph see Figure 6.3.

Comparing the two graphs $H^d$ and $\tilde{H}^d$ we immediately see the difference. Due to the system dynamics closed sites on $\tilde{H}^d$ need to be connected to the initial level, i.e. each $(z, n) \in H^d$ which is closed is contained in the closed cluster of a site $(z', 0) \in H^d$. In other words, in order for a site $(z, n) \in H^d$ to be closed there always exists a path of closed sites starting at a closed site $(z', 0) \in \tilde{H}^d$ on $\tilde{H}^d$ but not necessarily on $H^d$. Hence, the structure on $\tilde{H}^d$ allows tracing defectors back in time which invade clusters of cooperators. Thus, extinction of closed sites on $H^d$ is now equivalent to the lack of percolation of closed sites on $\tilde{H}^d$.

In order to see that closed sites do not percolate on $\tilde{H}^d$ we use Theorem 6.9. The cited version of the Theorem was originally proved by Durrett in Theorem 5 in [Dur92] for the graph $H^d$ and later adapted to the graph $\tilde{H}^d$ by Lanchier in Lemma 15 in [Lan13]. Hence, using the latter result from Lanchier and repeating the steps from the proof of statement (i) we can choose the probability for a site to be open arbitrarily large due to inequality (6.9). Thus, applying Theorem 6.9 we are able to prevent closed sites to percolate on $\tilde{H}^d$ which then yields that closed sites on $H^d$ go extinct in the limit. This again is equivalent for the defector type to die out which shows the result.  

Figure 6.3: The altered graph $\tilde{H}^d$. Closed sites are empty bullets, open sites are filled circles. The rightmost closed site in the levels $n, n + 1$ and $n + 2$ are now contained in the cluster of a closed site in level 0.
6.4 Discussion

In this section we point out the explicit reason for altering the transition rates of the cVMBC as considered in the previous chapters. Furthermore, we highlight that the proof given above allows to analyze a much wider class of models not necessarily restricted to the uniform nearest neighbor dynamics. We close this section with simulation results of the original cVMBC in dimension $d = 2$.

6.4.1 Justification of the sVMBC

Let us start by reconsidering the estimations given in equation (6.6). It is quite easy to see that plugging in the transition rates of the cVMBC would not allow a comparison with the Richardson model since the limit behavior for $\gamma \to \infty$ in this case yields the following time rescaled transition rate limits:

\[
\begin{align*}
    p_D' & \geq \frac{(2d - 1)(1 + i)}{2d^2 \gamma} + \frac{1}{2d} \to \frac{1}{2d'}, \\
    p_C & \leq \frac{(2d - 1)(1 + \alpha)}{2d\gamma} + \frac{2d - 1}{2d} \to 1 - \frac{1}{2d'}, \\
    p_C^n & \leq \frac{(1 + \alpha + \gamma)(2d - 1)^2}{\gamma 2d^2} \to \frac{(2d - 1)^2}{2d^2}.
\end{align*}
\]

Obviously these limits leave us in a very bad position when looking at clusters of cooperators in such a general way as we did it above. For this reason, we changed the interaction rate between cooperators and defectors such that the limits of $p_C$ and $p_C^n$ vanish or at least become arbitrarily small for $\gamma \to \infty$. This basically means that cooperators next to each other have a very low probability of switching their type in a finite time window which is enough for a coupling with a pure growth-process.

6.4.2 Alternative models

The method of proof we used in this chapter is not limited to the nearest-neighbor interaction dynamics. In fact, we can either increase the interaction range or use alternative interaction kernels or even apply both. The only calculations we need to adjust for the analysis of these kinds of models are those in equations (6.3), (6.4), (6.5) and (6.6). For example, if we increase the interaction range by changing the interaction kernels such that

\[
a(u, v) = \frac{1}{2Md} \mathbb{I}_{\{|u-v| \leq M\}}, \quad \text{for } M \in \mathbb{N},
\]

we need to make sure that the probabilities still satisfy the limit behavior in equation (6.6).
Another possibility is to consider non-uniform interaction kernels such as exponentially decaying interaction dynamics, i.e.

\[ a(u, v) = c \exp(-|u - v|), \]

where \( c \) is a normalizing constant, such that

\[ c \sum_{v \in \mathbb{Z}^d} \exp(-|u - v|) = 1. \]

Again, we need to verify equation (6.6) by calculating the estimates in equations (6.3)-(6.5) (and of course first think about the existence of such models).

For an overview on frequently used geometries (which induce the interaction kernels) in the general framework of interacting particle systems see for example Section 2.5 and Section 2.6 in [Ald13].

In the remaining parts of the proof after equation (6.6) we need to adapt the good events, the measurability condition in the comparison assumptions and the new percolation graph \( \tilde{H}^d \) such that these fit to the considered interaction kernel. These steps are non-trivial and need to be done separately for each case. Especially the alteration of the percolation graph \( H^d \) and the estimation of the transition rate limits can be challenging tasks in these kinds of models.

Summarizing these remarks, we note that the result in Theorem 6.3 should hold for a much wider class of biased voter models possessing cooperation mechanisms as considered in our model. Although, it seems that the proof needs to be done for each case separately, the structure of the proof should be similar to the steps as written out in Section 6.3.

### 6.4.3 Simulation results

Here, we present some simulation results on the cVMBC in dimension \( d = 2 \). As we have already seen in Figure 4.1 there seems to be a phase transition for the cVMBC not only in one dimension but also in dimensions \( d = 2 \) and \( d = 3 \). This leads to the conjecture that there exists a dimension-dependent relation between the cooperation parameter \( \gamma \) and the selection parameter \( \alpha \) such that either of the two types takes over the colony dependent on the concrete parameter configuration. However, we are not able to prove such a general result for the cVMBC model. In Figure 6.4 simulations of the cVMBC in dimension \( d = 2 \) are displayed. Even examining these graphics thoroughly, there does not appear to be a clear structure of how cooperators outcompete defectors. Clusters of cooperators coalesce, separate and vanish randomly which makes it hard to find a way of using features of clusters for a rigorous proof.
Figure 6.4: Simulation of the eVMBC on $\mathbb{Z}^2$. Yellow sites represent cooperators, red sites indicate the presence of defectors. The left figure shows the configuration after 5,000 transitions, the middle picture is taken after 10,000 transitions and the right graphic displays the configuration after 15,000 transitions. The system was started in a product Bernoulli-distribution with parameter 0.5. The reproduction rate was 1, the selection rate $\alpha$ was 0.5 and the cooperation rate $\gamma$ was 2.
In this work we modeled different aspects of non-genetic heterogeneity in bacterial colonies. We focused on two causes for this phenotypic diversity applying two different mathematical frameworks.

First, in Chapter 2 we modeled phenotypic heterogeneity due to intracellular variation in single cells leading to individually different gene expression profiles. We introduced a Markov chain model indexed by a tree and proved a weak limit result for the empirical measure process which turned out to be deterministic if the initial condition is fixed. Interestingly, this explains and unifies the observation of individuality on the single cell level versus deterministic colony behavior. Hence, the result resembles a law of large numbers. A thorough study of the variance could potentially yield a central limit theorem for this process but we did not tackle this issue in this dissertation.

Concerning phenotypic heterogeneity induced by cell-cell interactions we introduced the class of VMBC’s as interacting particle systems which we analyzed in Chapters 3-6. We aimed to model heterogeneity in a public good producing colony where cooperators are interacting with defecting individuals. We extended a two-type biased voter model by a cooperation mechanism and were able to describe the one-dimensional limit behavior of the two sub-systems, the aVMBC and cVMBC, completely in Chapter 4. For the more altruistic model, the aVMBC, the limit behavior in one dimension favors defectors as long as the selection parameter $\alpha$ is positive. In higher dimensions we could prove that for $\alpha > \gamma \frac{d-1}{d}$ defectors are outcompeting cooperators. In contrast to this, for the cVMBC we observe a phase transition at $\alpha = \gamma$ in one dimension, i.e. the limit configuration changes from defectors outcompeting cooperators for $\alpha > \gamma$ to cooperators taking over the colony for $\gamma > \alpha$. In the case $\alpha = \gamma$ we were able to prove that the process clusters which means that coexistence of the two types is not possible. Note, that this does not imply that either type dies out in finite time such
Conclusion and Outlook

that heterogeneity is not necessarily lost when observing this process for finite time.

Furthermore, we could prove a result which is related to the evolution of such cooperative mechanisms. We considered initial conditions with finitely many particles of cooperators (or defectors) and studied the limit behavior of these systems. Interestingly, for $\gamma > \alpha$ and one cooperating particle at the beginning the probability for the cooperative trait to survive is positive, at least in one dimension.

Analyzing the cVMBC on the lattice $\mathbb{Z}^d$ for dimensions $d \geq 3$ as done in Chapter 5 we were forced to consider a weak interaction limit such that the process dynamics resemble those of a voter model with perturbations. This model framework allowed us to approximate the limit behavior by a reaction-diffusion equation such that we were able to derive conditions on the parameters for the survival of one of the two species.

The last model we studied was the sVMBC in Chapter 6 where we decreased the cooperator-defector-interaction rate as compared to the cVMBC. This alteration enabled the coupling of the sVMBC to a pure growth-process, at least for large cooperation rates. This coupling then allowed us to compare the sVMBC with a super-critical oriented percolation process leading to a limit favoring cooperators in all dimensions $d \geq 1$, at least for $\gamma$ sufficiently large.

However, when comparing this last result to simulation studies performed with the cVMBC we see that cooperators in this model take over the colony for comparably small parameter values of $\gamma$ in the lower dimensions $d = 2$ and $d = 3$. These simulations suggest sharp, dimension dependent thresholds for the phase transitions in all dimensions. Unfortunately, it is still an open problem (at least to our knowledge) to find techniques for solving these kinds of questions in non-attractive models with no known duality features in dimensions $d \geq 2$. 
<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.1</td>
<td>Notions of a Markov chain indexed by a complete binary tree</td>
<td>7</td>
</tr>
<tr>
<td>3.1</td>
<td>Graphical representation of the cVMCB</td>
<td>27</td>
</tr>
<tr>
<td>4.1</td>
<td>Simulation results of the cVMBC in dimensions $d = 1, 2, 3$</td>
<td>33</td>
</tr>
<tr>
<td>6.1</td>
<td>Oriented graph $\mathcal{H}_1$</td>
<td>65</td>
</tr>
<tr>
<td>6.2</td>
<td>Trajectory of a defector</td>
<td>71</td>
</tr>
<tr>
<td>6.3</td>
<td>Altered graph $\tilde{\mathcal{H}}_1$</td>
<td>72</td>
</tr>
<tr>
<td>6.4</td>
<td>Time-course of a cVMBC simulation on $\mathbb{Z}^2$</td>
<td>75</td>
</tr>
</tbody>
</table>


Bibliography


