Coexistence in competing species models†

Yu-Ting Chen‡ and Matthias Hammer§

March 2, 2018

Abstract

We discuss some stochastic spatial generalizations of the Lotka-Volterra model for competing species. The topics range from spin systems on general discrete sets to interacting diffusions on integer lattices. Methods for proving coexistence in these generalizations and related open questions are discussed around some central duality equations, which relate coexistence of the models to extinction of their dual processes.

Keywords: Interacting particle system, interacting diffusion, duality.

Mathematics Subject Classification (2000): 82C22, 60J10, 60J60

Contents

1 Introduction .......................................................... 1

2 Death-birth events in the Lotka-Volterra model ................. 3

3 Spin systems on discrete sets ...................................... 8
   3.1 Constructions and parity duality ................................ 8
   3.2 Invariance of equiparity coexistence .......................... 14

4 Interacting diffusions on lattices .................................. 18
   4.1 Duality and coexistence ....................................... 19
   4.2 Proof of the moment duality .................................. 23

5 References .......................................................... 25

1 Introduction

Competition of species is a basic phenomenon studied in ecological dynamics. The mathematical analysis goes back to the foundational model by Lotka and Volterra in [17, 27], which results in a multi-dimensional extension of the logistic differential equation. The model is a density dependent

†This article is a lecture note for the learning session ‘Competing species models’ organized by the authors and held during July and August, 2017 in the Institute for Mathematical Sciences at the National University of Singapore.
‡Department of Mathematics, University of Tennessee, Knoxville, TN, United States of America.
model in the sense that only sizes of species are considered, and large well-mixed populations are implicitly assumed. Dynamics for these sizes follow ordinary differential equations subject to constant parameters. There are restricted resources and competition among species which suppress the exponential growths. See the monograph by Hofbauer and Sigmund [12] for an excellent introduction to this classical model.

In the biological literature, choosing new features to generalize the Lotka-Volterra model has been the subject of various works during the last few decades. One popular consideration is the use of spatial structure and stochasticity. Namely, populations are arranged according to Euclidean spaces, lattices or possibly more general geometric structures. In particular, the two-dimensional Euclidean space and integer lattice are biologically relevant, and the two-dimensional discrete tori often appear when supporting computer experimental results are provided. In the presence of spatial structure, ranges of individual interactions are limited and the well-mixed property in the Lotka-Volterra model that anyone can interact with anyone else is no longer valid. Also, by stochasticity, types of new individuals in the population are randomly determined, although species which are more abundant still have better chances to replicate. Spatial structure and stochasticity are closely related and can be combined to create a wide array of models.

Spatial structure naturally points to the use of partial differential equations, and furthermore, their analogues in the presence of stochasticity as stochastic partial differential equations. See [13, 18] for some examples. On the other hand, stochastic models on discrete spatial structures also arise as popular candidates. They allow for a detailed understanding of the implied biological phenomena. Connections to stochastic partial differential equations by taking rescaled limits of the models can be also attained on Euclidean-like spatial structures in several cases (e.g. [5, 7]).

This article is an introductory lecture note to stochastic spatial generalizations of the Lotka-Volterra model for competing species on discrete spatial structures. For definiteness, we mainly discuss two of the most well-known generalizations, one due to Neuhauser and Pacala [20] and the other by Blath, Etheridge and Meredith [1] and some related methods. These models are built on insights from biologists. Moreover, they fit well into the existing, well-developed probability methods, as their original forms go back to the voter model or the stepping stone model. Some motivations to link the models in [20, 1] to the Lotka-Volterra model are discussed in Section 2.

The model by Neuhauser and Pacala [20] is a spin system defined on a discrete space. The use of spins means that the model is at the microscopic level and every spatial point can only be occupied by one individual at most. Interactions are therefore simple, and it shall become clear to the reader that the model is a generalization of the Lotka-Volterra with minimal technical complications. Yet there are natural questions to relate the generalization and the classical differential equation model which remain open, as we will discuss in this article.

Blath et al. [1] introduced their model by keeping very closely the main features of some earlier models in the biological literature and taking into account appropriateness of mathematical investigations. In this case, the spatial scale is different from that in [20]. It is meant that population sizes, which are in the continuum, are discretized according to subdivisions of the underlying space $\mathbb{R}^d$ into boxes. Only interactions of species among these boxes are considered. Interactions within boxes are mean-field and no spatial structure is imposed. These features thereby lead to the use of countably many interacting diffusion processes attached to these boxes.

As in the Lotka-Volterra model, the major interests in these generalizations are centered around the question of characterizing large-time behavior of the models. In this lecture note, we choose to explain these properties in the spin system in [20] and the interacting diffusions in [1] in terms of duality. Hence, for the Neuhauser-Pacala model, the discussion will be to consider the associated spin system $(\eta_t)$ in terms of its dual process $(\check{\eta}_t)$, which are related by equations taking the following form.
for suitable bivariate dual functions $H$:

$$
\mathbb{E}[H(\eta_t, \hat{\eta}_0)] = \mathbb{E}[H(\eta_0, \hat{\eta}_t)]
$$

(1.1)

when $\eta_0$ and $\hat{\eta}_0$ are both constant. For example, the question of characterizing the large-time behavior of $(\eta_t)$ can be relegated to the same question for $(\hat{\eta}_t)$, as one would hope that the dual process is easier to study. In the case of [20], duality can even take a finer form of pathwise duality, namely, pathwise relation between $(\eta_t)$ and $(\hat{\eta}_t)$ on the same probability space. In this regard, we find the point of view as in Sturm and Swart in [25] useful to explain both forms of duality. It explicitly relates the two processes by random linear transformations. As for the case of interacting diffusions, we only discuss large-time behavior of the model by analytic duality as in (1.1). Finally, we refer the reader to the lecture note in the present series by Sturm, Swart and Völlinger [26] for further discussions of such analytic duality in general.

Organization of this article. This article is organized as follows. In Section 2, we first explain some features of the Lotka-Volterra model that motivate the generalizations in [20, 1]. Section 3 is specialized to a discussion of the Neuhauser-Pacala model where some symmetry is in force, and the duality is discussed in detail before its usefulness for finding the stationarity of the model is explained. Finally, Section 4 continues the spirit in Section 3 and places more emphasis on large-time probabilities in the Blath-Etheridge-Meredith model in terms of duality.

Acknowledgements. Partial supports from the Institute for Mathematical Sciences at the University of Singapore for both authors and a grant from the Simons Foundation for Y.-T. C. are gratefully acknowledged.

2 Death-birth events in the Lotka-Volterra model

In this section, we discuss the Lotka-Volterra model and its connections with the stochastic spatial generalizations in the main papers under consideration [20, 1]. Both generalizations are closely related to an equation for the density of a fixed type of individuals, which is (2.6) below and is in particular central to the generalization in [20]. It shows death-birth events in the Lotka-Volterra model in terms of the densities of species.

Recall that the competitive Lotka-Volterra ordinary differential equations for two species are based on the logistic equations. There are interactions within and between two given species. If we denote by $N_0 = N_0(t)$ and $N_1 = N_1(t)$ the total sizes of species 0 and 1 in the population, respectively, then the system of dynamical equations for $N_0$ and $N_1$ is given by

$$
\begin{cases}
\dot{N}_0 = r_0N_0 \left(1 - \frac{N_0 + \alpha_{01}N_1}{K_0}\right), \\
\dot{N}_1 = r_1N_1 \left(1 - \frac{N_1 + \alpha_{10}N_0}{K_1}\right).
\end{cases}
$$

(2.1)

Here, $r_i$ is the intrinsic growth rate of species $i$, $K_i$ is the carrying capacity of species $i$, and $\alpha_{ij}$ is called the strength of interspecific competition. These are all given positive constants. Without the term $(N_0 + \alpha_{01}N_1)/K_0$ to suppress the growth of 0-individuals, they grow exponentially with rate $r_0$. A similar interpretation applies to the equation for $N_1$. Note that $N_0 + N_1$ is not constant by (2.1).
Now we rewrite (2.1) in the following form:

\[
\begin{align*}
\dot{N}_0 &= N_0 \left( r_0 - \frac{r_0}{K_0} N_0 - \frac{r_0 \alpha_{01}}{K_0} N_1 \right) = N_0 (a - b N_0 - c N_1), \\
\dot{N}_1 &= N_1 \left( r_1 - \frac{r_1}{K_1} N_1 - \frac{r_1 \alpha_{10}}{K_1} N_0 \right) = N_1 (a' - b' N_1 - c' N_0),
\end{align*}
\]  

(2.2)

and then turn to the densities of 0-individuals and 1-individuals defined by

\[
p_0 = \frac{N_0}{N_0 + N_1} \quad \text{and} \quad p_1 = \frac{N_1}{N_0 + N_1}.
\]

By (2.2), we get

\[
\dot{p}_0 = \frac{N_0 N_1}{(N_0 + N_1)^2} \left[ (a - b N_0 - c N_1) - (a' - b' N_1 - c' N_0) \right] = -\frac{N_0 N_1}{(N_0 + N_1)^2} (a' + b N_0 + c N_1) + \frac{N_0 N_1}{(N_0 + N_1)^2} (a + b' N_1 + c' N_0).
\]

(2.3)

To simplify the last equation, we assume

\[
a = a'
\]

(2.4)

so that intrinsic growth rates \(r_0\) and \(r_1\) of the two species are identical. Then (2.3) reduces to the following equation:

\[
\dot{p}_0 = -\frac{N_0 N_1}{(N_0 + N_1)^2} (b N_0 + c N_1) + \frac{N_0 N_1}{(N_0 + N_1)^2} (b' N_1 + c' N_0).
\]

(2.5)

If we introduce suitable time changes which amount to the effect of dividing both sides by \((\lambda p_1 + p_0)(N_0 + N_1)\), then (2.5) can be transformed to a closed equation for \(p_0\):

\[
\dot{p}_0 = -p_0 \frac{\lambda p_1}{\lambda p_1 + p_0} (p_0 + \alpha_{01} p_1) + p_1 \frac{p_0}{\lambda p_1 + p_0} (p_1 + \alpha_{10} p_0).
\]

(2.6)

The coefficients of (2.5) follow since, under the present assumption in (2.4), the constants \(a, b, c, a', b', c'\) in (2.2) satisfy the following equations:

\[
\lambda \overset{\text{def}}{=} b/b' = K_1/K_0, \quad c/b = \alpha_{01} \quad \text{and} \quad c'/b' = \alpha_{10}.
\]

(2.7)

Equation (2.6) suggests natural interpretations similar to the Moran process in population genetics. The first term in (2.6), that is

\[
-p_0 \frac{\lambda p_1}{\lambda p_1 + p_0} (p_0 + \alpha_{01} p_1) = -p_0 (p_0 + \alpha_{01} p_1) \times \frac{\lambda p_1}{\lambda p_1 + p_0},
\]

has the interpretation as

\[-(\text{death rate of 0-individuals}) \times (\text{birth probability of 1-individuals}).\]

Here, taking

\[p_0 + \alpha_{01} p_1,
\]
as a death rate is consistent with the dynamical equation of \( N_0 \) in (2.2) now that \( r_0 = r_1 \) by our assumption in (2.4). Indeed, \( p_0 + \alpha_{01} p_1 \) is the same as \( N_0 + \alpha_{01} N_1 \) up to the multiplicative factor \( N_0 + N_1 \). Also, by the definition of \( \lambda \) in (2.7),

\[
\frac{\lambda p_1}{\lambda p_1 + p_0} = \frac{K_1 N_1}{K_1 N_1 + K_0 N_0}
\]

is taken as a birth probability of 1-individuals in view of the property that the higher the carrying capacity \( K_1 \) of 1-individuals is the more likely 1-individuals can replicate. This leads to the decrease of \( p_0 \), and so, the minus signs in (2.8). The second term in (2.1) can be interpreted similarly with the roles of 1 and 0 exchanged. But now this term is responsible for the increase of 0-individuals in the population.

Equilibrium in the dynamical equation for \( p_0 \) in (2.6) is very easy to obtain. Since \( p_0 + p_1 = 1 \), we get

\[
\dot{p}_0 = \frac{p_0 (1 - p_0)}{\lambda (1 - p_0) + p_0} \left\{ (1 - \lambda \alpha_{01}) - p_0 \left[ (1 - \lambda \alpha_{01}) + (\lambda - \alpha_{10}) \right] \right\} = \frac{F(p_0)}{\lambda (1 - p_0) + p_0},
\]

where \( F \) is a polynomial in \( p_0 \). Setting the right-hand side to zero shows that the foregoing equation has a unique stable equilibrium given by

\[
p_0^* = \frac{(1 - \lambda \alpha_{01})}{(1 - \lambda \alpha_{01}) + (\lambda - \alpha_{10})}
\]

if \( 0 \leq \alpha_{10} < \lambda \) and \( 0 \leq \alpha_{01} < 1/\lambda \), where the restrictions on \( \alpha_{10} \) and \( \alpha_{01} \) ensure that \( p_0^* \) falls in \((0, 1)\). The stability of \( p_0^* \) is plain from the graph of \( F \) given in Figure 1.

A stochastic spatial generalization of the last differential equation to spin systems is introduced by Neuhauser and Pacala in [20]. Put in a general framework, the model can be described in the following way. Let \( q \) be an irreducible transition probability on a nonempty set \( E \) and suppose that \( q \) has a zero trace:

\[
\sum_{x \in E} q(x, x) = 0.
\]

For \( \sigma \in \{0, 1\}, x \in E \) and \( \eta \in \{0, 1\}^E \), we define the local frequencies of \( \sigma \)'s by

\[
f_\sigma(x, \eta) = \sum_{y \in E} q(x, y) \mathbb{1}_{\{\sigma\}}(\eta(y)).
\]

The canonical example here is that \( q \) is given by the transition probability of random walk on a connected graph. In this case, \( f_\sigma(x, \eta) \) reduces to the usual frequency of \( \sigma \)'s in the neighborhood of \( x \):

\[
f_\sigma(x, \eta) = \frac{1}{\deg(x)} \sum_{y: y \sim x} \mathbb{1}_{\{\sigma\}}(\eta(y)).
\]
The use of these kernels $(E, q)$’s are meant to allow perturbations of graph structures in the usual $\mathbb{Z}^2$ or two-dimensional discrete tori so that more realistic modeling may become possible, and thus, to view the spin system to be defined below from a more general perspective.

The Lotka-Volterra model is then generalized from a density-dependent model to a frequency dependent model and from a differential equation to a spin system. The flip rates of the sought-after generalization are set to be

$$
\begin{align*}
0 \rightarrow 1 & \text{ with rate } (f_0 + \alpha_{01} f_1) \left( \frac{\lambda f_1}{\lambda f_1 + f_0} \right), \\
1 \rightarrow 0 & \text{ with rate } (f_1 + \alpha_{10} f_0) \left( \frac{f_0}{\lambda f_1 + f_0} \right). 
\end{align*}
$$

(2.11)

In this spin system,

$$
\begin{align*}
f_0 + \alpha_{01} f_1 \quad \text{and} \quad f_1 + \alpha_{10} f_0
\end{align*}
$$

are the death rates and

$$
\begin{align*}
\frac{\lambda f_1}{\lambda f_1 + f_0} \quad \text{and} \quad \frac{f_0}{\lambda f_1 + f_0}
\end{align*}
$$

are the birth probabilities, with plain generalizations of the above interpretations for (2.9). Hence, in the sense of flip rates in [16, page 122–123], the $\{0, 1\}^E$-valued Markov process $(\eta_t)$ under consideration is characterized by the following limits:

$$
\begin{align*}
\mathbb{P}(\eta_t(x) = 1|\eta_0(x) = 0) = [f_0(x, \eta_0) + \alpha_{01} f_1(x, \eta_0)] \left( \frac{\lambda f_1(x, \eta_0)}{\lambda f_1(x, \eta_0) + f_0(x, \eta_0)} \right) + o(t), \quad t \searrow 0+, \\
\mathbb{P}(\eta_t(x) = 0|\eta_0(x) = 1) = [f_1(x, \eta_0) + \alpha_{10} f_0(x, \eta_0)] \left( \frac{f_0(x, \eta_0)}{\lambda f_1(x, \eta_0) + f_0(x, \eta_0)} \right) + o(t), \quad t \searrow 0+, \\
\mathbb{P}(\eta_t(x) \neq \eta_0(x), \eta_t(y) \neq \eta_0(y)) = o(t), \quad t \searrow 0+, \quad \forall x \neq y.
\end{align*}
$$

The flip rates introduced above in (2.11) are applicable for general $(E, q)$’s defined above. In particular, it can be shown by modern limit theorems of semimartingales (cf. [14]) that the density of 1’s in this particle system on a complete graph over $N$ vertices converges to the solution of (2.6) as $N \to \infty$ in the space of càdlàg functions equipped with Skorokhod’s $J_1$-topology.

Our discussions in the rest of this paper for this model will be on the constructions of this spin system in the symmetric case that $\alpha_{01} = \alpha_{10}$ and $\lambda = 1$ and on the analysis of its equilibrium. Note that if $\alpha_{01} = \alpha_{10} = 1$ and $\lambda = 1$, then the model reduces to the voter model and is henceforth excluded from the discussion below. See the second set of flip rates in (3.3).

A model introduced by Blath, Etheridge and Meredith [1] generalizes the Lotka-Volterra differential equations in (2.1) from a point of view very similar to those in the biology papers by Bolker and Pacala [2] and Murrell and Law [19]. Here, we only consider the case that the underlying spatial structure is an integer lattice $\mathbb{Z}^d$ for some $d \geq 1$. Population sizes at points in $\mathbb{Z}^d$ are in the continuum and subject to the new features of migration and stochastic growth defined by branching noises. Precisely, $(\eta_t) = (\eta_t(x); x \in \mathbb{Z}^d)$ and $(\eta'_t) = (\eta'_t(x); x \in \mathbb{Z}^d)$ to model population sizes of 0-individuals and 1-individuals at all sites $x \in \mathbb{Z}^d$, respectively, are given by the following system of stochastic differential equations: for all $x \in \mathbb{Z}^d$,

$$
\begin{align*}
d\eta_t(x) = & \sum_{y \in \mathbb{Z}^d} m_{xy}(\eta_t(y) - \eta_t(x)) \, dt \\
& + \eta_t(x) \left( \alpha - \sum_{y \in \mathbb{Z}^d} \beta_{xy} \eta_t(y) - \sum_{y \in \mathbb{Z}^d} \gamma_{xy} \eta'_t(y) \right) \, dt + \sqrt{\sigma \eta_t(x)} dB_t(x),
\end{align*}
$$

(2.13)
\[ d\eta'_t(x) = \sum_{y \in \mathbb{Z}^d} m'_{xy}(\eta'_t(y) - \eta'_t(x))\,dt \]
\[ + \eta'_t(x) \left( \alpha' - \sum_{y \in \mathbb{Z}^d} \beta'_{xy} \eta'_t(y) - \sum_{y \in \mathbb{Z}^d} \gamma'_{xy} \eta'_t(y) \right)\,dt + \sqrt{\sigma \eta'_t(x)} dB'_t(x), \]

where \{B(x), B'(x); x \in \mathbb{Z}^d\} are i.i.d. one-dimensional standard Brownian motions.

The equations in (2.13) and (2.14) now model spatial effects in a different way. For example, the sum
\[ \sum_{y \in \mathbb{Z}^d} m_{xy}(\eta_t(y) - \eta_t(x))\,dt \]
models migration of individuals by a nonnegative matrix \( m \) since
\[ \mathcal{L}f(x) = \sum_{y \in \mathbb{Z}^d} m_{xy}[f(y) - f(x)] \]
is the generator of a Markov chain moving along sites of \( \mathbb{Z}^d \) subject to the \( q \)-matrix \( \{m_{xy}; x \neq y\} \).

Also, the term
\[ \eta_t(x) \left( \alpha - \sum_{y \in \mathbb{Z}^d} \beta_{xy} \eta_t(y) - \sum_{y \in \mathbb{Z}^d} \gamma_{xy} \eta_t(y) \right)\,dt \]
has a natural correspondence to the logistic differential equations in (2.1). It is assumed in addition that these matrices \( m, m', \beta, \beta', \gamma, \gamma' \) are ‘homogeneous’ and have finite ranges in the sense that, for example, \( m_{xy} \) depends only on \( ||x-y||_\infty \) and \( m_{xy} = 0 \) for all \( ||x-y||_\infty \geq L \) for some \( L > 0 \) independent of \( x, y \).

The objects corresponding to (2.6) are the density processes
\[ p_t(x) = \frac{\eta_t(x)}{\eta_t(x) + \eta'_t(x)} \]
of 0-individuals at all sites \( x \). The derivation in (2.3) can be generalized by Itô’s formula, but the change of time argument to obtain equation in (2.6) is not enough to close the equations of \( p_t(x) \) now since total population sizes at all sites are different in general. These differences should be evident from the use of branching noise terms \( \sqrt{\sigma \eta_t(x)} dB_t(x) \) and \( \sqrt{\sigma \eta'_t(x)} dB'_t(x) \) in (2.13) in particular. The consideration employed in [1] to address this technical feature is to introduce the model obtained by formally conditioning on \( \eta_t(x) + \eta'_t(x) \equiv N \) for all \( x \). Then, assuming also that \( m_{xy} = m'_{xy} \) and that \( \beta_{xy}, \beta'_{xy}, \gamma_{xy}, \gamma'_t \) are zero for \( x \neq y \) (purely local interactions), one is led to the following system:
\[ dp_t(x) = \sum_{y \in \mathbb{Z}^d} m_{xy}(p_t(y) - p_t(x))\,dt + sp_t(x)(1 - p_t(x))(1 - \mu p_t(x))\,dt \]
\[ + \sqrt{N^{-1}p_t(x)(1 - p_t(x))}dW_t(x), \quad x \in \mathbb{Z}^d, \]

where \( \{W(x); x \in \mathbb{Z}^d\} \) is a system of independent standard Brownian motions, and \( s \) and \( \mu \) are real parameters which can be expressed explicitly in terms of the parameters of \( \{\eta_t\} \) and \( \{\eta'_t\} \). We refer the reader to [1, pages 1482–1483] for details of the derivation. In particular, if we consider the symmetric case where \( \eta \) and \( \eta' \) follow the same parameters, then \( \mu = 2 \) and \( s = \beta_{xx} - \gamma_{xx} \). Note that existence and uniqueness (in the strong sense) of a \( [0,1]^{\mathbb{Z}^d} \)-valued solution of the system of SDEs in (2.15) can be obtained independently from classical results for infinite-dimensional SDEs due to Shiga and Shimizu [24, Theorem 2.1]. In Section 4, the discussion for these interacting diffusions in (2.15) is independent of the system defined by (2.13), and so we use general parameters \( s \) and \( \mu \) from now on.
3 Spin systems on discrete sets

Throughout this section, we focus on the symmetric Neuhauser-Pacala model, that is \( \lambda = 1 \) and 
\[
\alpha_{01} = \alpha_{10} = \alpha \in [0, 1),
\]
so that there is no bias in birth rates and death rates induced by these parameters, and the flip rates are now simplified to the following:
\[
\begin{align*}
0 \rightarrow 1 \text{ with rate } & (f_0 + \alpha f_1)f_1 = (1 - \alpha)f_0f_1 + \alpha f_1, \\
1 \rightarrow 0 \text{ with rate } & (f_1 + \alpha f_0)f_0 = (1 - \alpha)f_1f_0 + \alpha f_0.
\end{align*}
\]

The assumptions that \( \lambda = 1 \) and \( \alpha_{01} = \alpha_{10} \) mean that species 0 and 1 look like each other, which can be made precise by looking at the flip rates defined in (2.11). In addition to the obvious technical advantage of assuming \( \alpha \in [0, 1) \) from (3.2), this assumption also means that neither type can induce higher death rates for the other type.

Below we first introduce in Section 3.1 various constructions of this symmetric model. The constructions will progressively lead us to the duality for the model, which is called parity duality in this paper for reasons that shall become self-evident. Then when we discuss related basic results and open questions in Section 3.2, we will present a simple theorem to illustrate the idea that this parity duality is capable of making possible the study of large time equilibria of the model from different perspectives.

3.1 Constructions and parity duality

In the following, we first view the symmetric Neuhauser-Pacala model in terms of two sets of flip rates. These sets of flip rates define the symmetric model with a detailed coupling by independent Poisson processes. On a finite set, the construction leads to a matrix representation of the model and easily induce the parity duality. We will explain these steps in detail. By the end of this subsection, we give a sketch of how generalizations of the parity duality can be obtained on infinite sets.

First, the flip rates defined above in (3.2) are decomposed into two fundamental mechanisms given by
\[
\begin{align*}
\text{pairwise annihilation:} & \quad \begin{cases}
0 \rightarrow 1 \text{ with rate } (1 - \alpha)f_0f_1, \\
1 \rightarrow 0 \text{ with rate } (1 - \alpha)f_1f_0,
\end{cases} \\
\text{voting:} & \quad \begin{cases}
0 \rightarrow 1 \text{ with rate } \alpha f_1, \\
1 \rightarrow 0 \text{ with rate } \alpha f_0.
\end{cases}
\end{align*}
\]

As before, we suppress sites and configurations to be updated in this notation, and the corresponding \( \{0, 1\}^E \)-valued Markov processes can be characterized analogously as in (2.12).

The first set of flip rates in (3.3) corresponds to pairwise annihilation. The children of two neighbors randomly chosen according to the transition kernel \( q \) try to invade a focal site, namely the site chosen to be updated, subject to pairwise annihilation. To see how this interpretation arises, we write the flip rates for site \( x \) given population configuration \( \eta \) as follows:
\[
(1 - \alpha)f_0(x, \eta)f_1(x, \eta) = (1 - \alpha)f_1(x, \eta)f_0(x, \eta)
= (1 - \alpha) \left( \sum_{y \in E} q(x, y)\eta(y) \right) \left( \sum_{z \in E} q(x, z)[1 - \eta(z)] \right).
\]
Hence, a spin flip at \( x \) is triggered if and only if an 1-neighbor \( y \) and a 0-individual \( z \) are chosen.

If the underlying spatial structure is a graph and we let a filled vertex denote a vertex occupied by a 1-individual and an empty vertex denote one occupied by a 0-individual, then we can visualize the complete set of possible transitions as follows:

\[
\begin{align*}
\begin{array}{c}
    \bullet \ \bullet \ \circ \\
    y \ \ \ \ \ \ z
    \end{array} & \rightarrow \\
\begin{array}{c}
    \bullet \ \circ \ \circ \\
    y \ \ \ \ \ \ z
    \end{array} & \rightarrow \\
\begin{array}{c}
    \bullet \ \circ \ \circ \\
    y \ \ \ \ \ \ z
    \end{array} & \rightarrow
\end{align*}
\]

(3.4)

Consider the following two examples from (3.4):

\[
\begin{align*}
\begin{array}{c}
    \bullet \ \circ \\
    y \ \ \ \ \ \ z
    \end{array} & \rightarrow \\
\begin{array}{c}
    \bullet \ \circ \\
    y \ \ \ \ \ \ z
    \end{array} & \rightarrow \\
\begin{array}{c}
    \bullet \ \circ \\
    y \ \ \ \ \ \ z
    \end{array} & \rightarrow
\end{align*}
\] and

\[
\begin{align*}
\begin{array}{c}
    \bullet \ \circ \\
    y \ \ \ \ \ \ z
    \end{array} & \rightarrow \\
\begin{array}{c}
    \bullet \ \circ \\
    y \ \ \ \ \ \ z
    \end{array} & \rightarrow \\
\begin{array}{c}
    \bullet \ \circ \\
    y \ \ \ \ \ \ z
    \end{array} & \rightarrow
\end{align*}
\]

For the transition on the left-hand side, the focal site is \( x \), and the two neighbors at sites \( y \) and \( z \), randomly chosen according to the probability \( q(x, \cdot) \) without replacement, try to invade site \( x \) by their children. The resident 1-individual at \( x \) and the invading child of the 1-individual at \( z \) annihilate each other, and so the child of the 0-individual at \( y \) takes over the site \( x \) after the update. A similar interpretation applies to the focal site \( x \) for the transition on the right. But this time, two of the three 1-individuals (two from \( y \) and \( z \) plus the resident 1-individual) annihilate each other, so that there is only one 1-individual left at \( x \) (we do not care where this survivor comes from).

The second set of flip rates in (3.3) defines a voting mechanism. Now, the flip rates can be written as

\[
\begin{align*}
\alpha f_1(x, \eta) &= \alpha \sum_{y \in E} q(x, y) \eta(y) \\
\alpha f_0(x, \eta) &= \alpha \sum_{y \in E} q(x, y) [1 - \eta(y)]
\end{align*}
\]

so that there is a spin flip at \( x \) if and only if a neighbor of the opposite type is chosen:

\[
\begin{align*}
\begin{array}{c}
    \bullet \ \bullet \\
    y \ \ \ \ \ \ z
    \end{array} & \rightarrow \\
\begin{array}{c}
    \bullet \ \bullet \ \circ \\
    y \ \ \ \ \ \ z
    \end{array} \rightarrow \\
\begin{array}{c}
    \bullet \ \bullet \ \circ \\
    y \ \ \ \ \ \ z
    \end{array} \rightarrow \\
\begin{array}{c}
    \bullet \ \bullet \\
    y \ \ \ \ \ \ z
    \end{array} \rightarrow
\end{align*}
\]

(3.5)

In the following example of (3.5):

\[
\begin{align*}
\begin{array}{c}
    \bullet \ \circ \\
    y \ \ \ \ \ \ z
    \end{array} & \rightarrow \\
\begin{array}{c}
    \bullet \ \circ \\
    y \ \ \ \ \ \ z
    \end{array} & \rightarrow \\
\begin{array}{c}
    \bullet \ \circ \\
    y \ \ \ \ \ \ z
    \end{array} & \rightarrow
\end{align*}
\]

the focal site chosen for update is again \( x \), and the new type is chosen randomly from one of the neighbors, which is \( y \), so that a 0-individual replaces the resident 1-individual at \( x \).

The group structure of \( \mathbb{Z}_2 \) is compatible with the above two mechanisms and so we can ‘linearize’ the mechanisms by Poisson processes defined as follows. In essence, the argument relies heavily on the cancellative additivity \( 1 + 1 \equiv 0 \) in \( \mathbb{Z}_2 \). For any \( x, y, z \in E \), we first define an \( E \times E \) matrix with entries in \( \mathbb{Z}_2 \) as follows:

\[
\{x\} \times \{y, z\} = x \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}
\]
where the zeros denote block matrices of obvious dimensions with zero entries. Then for the pairwise
annihilation mechanism, we consider the standard superposition of Poisson processes and the following
Poisson updates by the linear transformations acting on $\eta \in \mathbb{Z}_2^E$ for all $x, y, z \in E$:
\begin{equation}
\eta \rightarrow \eta + \{x\} \times \{y, z\} \eta \quad \text{with rate } r(\{x\} \times \{y, z\}) = (1 - \alpha)q(x, y)q(x, z).
\end{equation}

Here, the configuration after update can be written as
\begin{equation*}
\eta + \{x\} \times \{y, z\} \eta = [\eta - \eta(x)\mathbf{1}_x] + [\eta(x) + \eta(y) + \eta(z)]\mathbf{1}_x.
\end{equation*}

The difference in the first pair of brackets shows that the types outside $x$ are kept fixed. The sum in
the second pair of brackets shows precisely pairwise annihilation due to the group structure of $\mathbb{Z}_2$ and
always gives us a value in $\{0, 1\}$.

For the voting mechanism, we introduce independent Poisson updates by the following linear
transformations:
\begin{equation}
\eta \rightarrow \eta + \{x\} \times \{x, y\} \eta \quad \text{with rate } r(\{x\} \times \{x, y\}) = \alpha q(x, y).
\end{equation}

Now the configuration after update can be written as
\begin{equation}
\{x\} \times \{x, y\} \eta = [\eta - \eta(x)\mathbf{1}_x] + [\eta(x) + \eta(x) + \eta(y)]\mathbf{1}_x = [\eta - \eta(x)\mathbf{1}_x] + \eta(y)\mathbf{1}_x,
\end{equation}

and so we have removal and adoption of types at $x$ by the first and second pairs of brackets, respectively.

On a finite set $E$, the above linear construction of the symmetric Neuhauser-Pacala model by
Poisson events can be made more explicitly by a random-walk type construction, as we time-change a
discrete-time Markov chain by an independent Poisson process (see [21, Section 2.6]). In detail, with
initial condition $\mathbf{1}_A$, we define a $\mathbb{Z}_2^E$-valued process by
\begin{equation}
\eta^A_t \overset{\text{def}}{=} (\text{Id} + J_{N_t}) \cdots (\text{Id} + J_{1})\mathbf{1}_A,
\end{equation}

where $J_1, J_2, \cdots$ are i.i.d. random matrices with entries in $\{0, 1\}$ and law
\begin{equation*}
P(J_1 = J) = \frac{r(J)}{\sum_{J'} r(J')}
\end{equation*}

and $(N_t)$ is an independent Poisson process with rate
\begin{equation*}
\mathbb{E}N_t = \sum_{J'} r(J').
\end{equation*}

By the construction in (3.9), we can imitate the usual time-reversal duality of random walks and
easily get a duality of the symmetric Neuhauser-Pacala model:
\begin{equation}
\langle \mathbf{1}_B, \eta^A_t \rangle = \mathbf{1}^\top_B (\text{Id} + J_{N_t}) \cdots (\text{Id} + J_{1})\mathbf{1}_A
\end{equation}
\begin{equation*}
= [((\text{Id} + J_{N_t}) \cdots (\text{Id} + J_{1})\mathbf{1}_B)]^\top \mathbf{1}_A
\end{equation*}
\begin{equation}
= \langle \hat{\eta}^B_t, \mathbf{1}_A \rangle \overset{(d)}{=} \langle \hat{\eta}^B_t, \mathbf{1}_A \rangle.
\end{equation}

In other words, the inner product $(\xi, \eta) \mapsto \langle \xi, \eta \rangle = \langle \xi, \eta \rangle_{\mathbb{Z}_2^E}$ is used as the dual bivariate function and
\begin{equation}
\hat{\eta}^B_t \overset{\text{def}}{=} (\text{Id} + J_{N_t}) \cdots (\text{Id} + J_{1})\mathbf{1}_B
\end{equation}
is the dual Markov chain.
To specify the update rule of the dual chain defined in (3.12), it is now more convenient to say that there are only 1-individuals in the population and the sites not occupied by 1-individuals are left empty. Recall that for the symmetric Neuhauser-Pacala model, we use the interpretation that there are 0-individuals and 1-individuals in the population.

By transposing the linear transformation in (3.6), we obtain a mechanism featuring branching with pairwise annihilation:

$$\xi \rightarrow \xi + \{y, z\} \times \{x\} \xi \equiv [\xi - \xi(y)1_y - \xi(z)1_z] + [\xi(x) + \xi(y)]1_y + [\xi(x) + \xi(z)]1_z. \quad (3.13)$$

If $x$ is chosen as the focal site and is occupied by a 1-individual, it gives birth to two children and they try to invade two sites $y$ and $z$ subject to the pairwise annihilation by the last two terms as before due to the group structure of $\mathbb{Z}_2$. The set of all possible transitions can be visualized as follows:

Let us look at the following two examples in (3.14):

On the left-hand side, all of the three sites $x, y, z$ are occupied by 1-individuals, but after the update, the child of $x$ and the resident 1-individual at $y$ annihilate each other so that an empty site is created. The same holds at site $z$. In the example on the right-hand side, site $y$ becomes occupied by a 1-individual after the update, but pairwise annihilation still occurs at site $z$.

Now the transposed linear transformation in (3.8) defines a random walk with annihilation:

$$\xi \rightarrow \xi + \{x, y\} \times \{x\} \xi \equiv [\xi - \xi(x)1_x - \xi(y)1_y] + [\xi(x) + \xi(y)]1_y + [\xi(x) + \xi(y)]1_y \quad (3.15)$$

$$= [\xi - \xi(x)1_x - \xi(y)1_y] + [\xi(x) + \xi(y)]1_y, \quad (3.16)$$

so that the the individual at site $x$ moves to site $y$ subject to pairwise annihilation:

$$x \bullet y \quad \rightarrow \quad x \circ y \quad x \bullet y \quad \rightarrow \quad x \circ y \quad (3.17)$$

Now we look at the following two examples from (3.17):

On the left-hand side, the 1-individual moves to site $y$, but this individual and the resident 1-individual at $y$ annihilate each other. We get two empty sites in the end. On the right-hand side, site $x$ is empty and so has no effect on the type occupying $y$ from the update event.
In contrast to the Neuhauser-Pacala model where a death event is immediately followed by a birth event at each updating step, the dual process is defined by reversing the orders of birth events and death events, and thus, is an invasion process. Also, it is possible to have multiple sites where a change of types occurs in both mechanisms defining the dual process.

To fully define the dual process in (3.12), we also need to specify the rates of the associated Poisson processes. The Poisson process corresponding to (3.13) jumps with rate the same as that for (3.6):

\[ \hat{r}(\{y,z\} \times \{x\}) \overset{\text{def}}{=} r(\{x\} \times \{y,z\}), \] (3.18)

whereas the Poisson process corresponding to (3.15) jumps with rate

\[ \hat{r}(\{x,y\} \times \{x\}) \overset{\text{def}}{=} r(\{x\} \times \{x,y\}). \] (3.19)

Taking expectation of both sides of (3.11), we have the following theorem when \( E \) is a finite set. It still holds on infinite sets.

**Theorem 3.1 (Parity duality).** For any irreducible kernel \((E,q)\) with a zero trace, finite subsets \(A,B\) of \(E\), and \(t \geq 0\), we have

\[ P\left( (1_B, \eta^A_t) \equiv 1 \right) = P\left( (\hat{\eta}^B_t, 1_A) \equiv 1 \right), \]

where \(\equiv\) means an equality mod 2.

The validity of Theorem 3.1 on infinite sets can be obtained by various methods. For example, one can check the following Feynman-Kac duality equation by generator calculations. The plan is to consider the forward equation satisfied by

\[ \phi(B,A) \overset{\text{def}}{=} \phi_A(B) = \phi_B(A) = 1_{\{1_B,1_A\}=1} \]

under the symmetric Neuhauser-Pacala model (with semigroup \(P_t = \exp\{tL^{\text{NP}}\}\)) and then show by algebra that this forward equation is also a backward equation under the dual process (with semigroup \((Q_t = \exp\{tL^{\text{dual}}\})\):

\[ \frac{d}{dt}P_t \phi_B(A) = P_t L^{\text{NP}} \phi_B(A) = Q_t L^{\text{dual}} \phi_A(B) \] (3.20)

so that

\[ P_t \phi_B(A) = Q_t \phi_A(B). \]

In more detail, to do the algebra for the proof of the second equality in (3.20), one needs the possible transitions and rates defining \(L^{\text{NP}}\) and the possible transitions and rates in (3.13), (3.15), (3.18) and (3.19) defining \(L^{\text{dual}}\). We omit the details here, but in Section 4.2, we will illustrate this algebraic method by the interacting diffusions defined at the end of Section 2. See [20, Proposition 1] for this calculation and [16, Section III.4] or [26] for more general discussions.

The other method is to use Harris’s graphical representation by Poisson arrows [11]. This route can circumvent the possible technical issue from having infinite products of matrices in the case of infinite sets that the reader may have already worried about. More importantly, the graphical representation has the technical advantage of providing a simple way to visualize trajectories of the particle system.

In the present context, the representation can be explained as follows. Again let us start with finite sets and look at the definition (3.9) of \((\eta^A_t)\). For any matrices \(M_1, \cdots, M_n\) with entries in \(\{0,1\}\), the product

\[ 1_B^T M_n M_{n-1} \cdots M_1 1_A = \sum_{x_n, x_{n-1}, \cdots, x_0 \in E} 1_B(x_n) M_n(x_n, x_{n-1}) M_{n-1}(x_{n-1}, x_{n-2}) \cdots M_1(x_1, x_0) 1_A(x_0) \]
has a natural interpretation as the number of paths from $A$ to $B$. Indeed,

$$1_B(x_n)M_n(x_n, x_{n-1})M_{n-1}(x_{n-1}, x_{n-2}) \cdots M_1(x_1, x_0)1_A(x_0) = 1$$  \hspace{1cm} (3.21)

if and only if

$$1 = 1_A(x_0) = M_1(x_1, x_0) = \cdots = M_n(x_n, x_{n-1}) = 1_B(x_n).$$

So the nonzero summand in (3.21) corresponds to the following path:

$$x_0 \to x_1 \to \cdots \to x_n.$$  \hspace{1cm} (3.22)

In the above interpretation of the product $1_B^T M_n M_{n-1} \cdots M_1 1_A$, we stress that $M_m(x_m, x_{m-1}) = 1$ corresponds to the path $x_{m-1} \to x_m$ rather than to the one $x_m \to x_{m-1}$, since matrices are now multiplied from the right to the left in (3.9). To keep track of paths from $A$ to $B$ defined by a matrix with entries in $\{0, 1\}$, we call

$$\mathcal{R}(M) = \{ y : M(y, x) = 1 \text{ for some } x \in E \} \cup \{ x : M(y, x) = 1 \text{ for some } y \in E \}$$

the range of $M$.

To apply the interpretation of (3.21) by (3.22) to the coupling in (3.10), we can restrict our attention to the case that all of the matrices $M_m$ take the form

$$M \equiv \text{Id} + J,$$

and then the above path interpretation extends naturally on an infinite set. Indeed, we have an infinite product of matrices to define the entire population configuration at a fixed time, but since the sum of rates $r(J)$, for $J$ ranging over matrices such that $x \in \mathcal{R}(J)$, is finite for every $x$, the type at $x$ can only be changed finitely many times within each finite time interval. The factors $(\text{Id} + J)(x, y) = 1$ in a nonzero summand underlying (3.10) are contributed by the pairs of indices where $x = y$ except finitely many.

In summary, we can extend the definition of the process $(\eta^A_t)$ in (3.9) by the following recipe:

$$\eta^A_t(x) \equiv \# \{ \text{oriented paths from } A \times \{0\} \text{ to } \{(x, t)\} \}, \quad x \in E.$$  

Then we define $(\hat{\eta}^B_t, 1_A)$ as the number of reversely oriented paths between $A \times \{0\}$ and $B \times \{t\}$, which is equal to $(1_B, \eta^A_t)$ in $\mathbb{Z}_2$. The graphical representation of the dual process $(\hat{\eta}^B_t)$ can be done similarly.

**Example 3.2.** Let us visualize the graphical representation of $(\eta^A_t)$ by the following simple context. Write $\bullet = 1$ and $\bigcirc = 0$. Then for the symmetric Neuhauser-Pacala model, if $E = \{x, y, z\}$, the initial condition $1_y$ after the update $\{x\} \times \{y, z\}$ can be visualized by

```
<p>| | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>x</td>
<td>y</td>
<td>z</td>
</tr>
</tbody>
</table>
```

The two arrows are those defined by the following matrix which are not self-loops:

$$\begin{pmatrix} x & y & z \\ x & 0 & 1 & 1 \\ y & 0 & 0 & 0 \\ z & 0 & 0 & 0 \end{pmatrix} = \text{Id} + y \begin{pmatrix} x & y & z \\ 1 & 1 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$
Similarly, for $E = \{x, y\}$, the initial condition $\eta_0 = 1_y$ after the voting update of $\eta \mapsto \eta + \{x\} \times \{x, y\} \eta$ can be visualized by

![Diagram showing space-time visualization of the update](image)

One can check that the arrow is the one defined by the following matrix which is not a self-loop:

$$
\begin{pmatrix}
1 & 1 \\
0 & 0
\end{pmatrix}
\begin{pmatrix}
x \\
y
\end{pmatrix}
= 
\begin{pmatrix}
x \\
y
\end{pmatrix}
\begin{pmatrix}
0 & 1 \\
0 & 1
\end{pmatrix}.
$$

The usefulness of the parity dual equation in Theorem 3.1 follows from the fact that parity events $\{ \eta; |\eta \cap B| \equiv 1 \}, B \subseteq E$, uniquely determine a measure.

**Proposition 3.3.** For finite measures $\nu_1$ and $\nu_2$, $\nu_1 \{ \eta; |\eta \cap B| \equiv 1 \} = \nu_2 \{ \eta; |\eta \cap B| \equiv 1 \}$ for all finite subsets $B$ of $E$ implies that $\nu_1 = \nu_2$. Here, $\eta$ is identified as the subset of $E$ where $\eta(x) = 1$.

**Proof.** The proof follows upon rewriting the test functions $\prod_{x \in A} [2\eta(x) - 1]$ in two different ways:

$$
\prod_{x \in A} [2\eta(x) - 1] = (-1)^{|\eta \cap A| - 1} - 2 \mathbf{1}_{\{ \eta; |\eta \cap A| \equiv 1 \}},
$$

$$
\prod_{x \in A} [2\eta(x) - 1] = 2^{|A|} \prod_{x \in A} \eta(x) + \text{a lower-order polynomial in } \eta(x) \text{ for } x \in A.
$$

Then we use an induction on $|A|$ and the fact that the monomials $\prod_{x \in A} \eta(x)$ for $A$ ranging over all finite subsets of $E$ are measure-determining by the inclusion-exclusion principle.

### 3.2 Invariance of equiparity coexistence

If we start the symmetric Neuhauser-Pacala model with the Bernoulli product measure $\beta_{1/2}$ with density $1/2$, then the corresponding parity dual equation from Theorem 3.1 can be written as

$$
\mathbb{P}_{\beta_{1/2}}(\langle 1_B, \eta_t \rangle \equiv 1) = \mathbb{P} \left( \sum_{m=1}^{\hat{\eta}_t} X_m \equiv 1, \hat{\eta}_t^B \neq 0 \right).
$$

Here, $X_1, X_2, \cdots$ are i.i.d. $\{0, 1\}$-valued Bernoulli variables with mean $1/2$ and independent of $\hat{\eta}_t^B$. They come from the random spins at all sites under $\beta_{1/2}$. By conditioning on $\hat{\eta}_t^B$, one sees that

$$
\text{the right-hand side of } (3.23) = \frac{1}{2} \mathbb{P}(\hat{\eta}_t^B \neq 0).
$$

Hence, the problem of proving existence of $\mathbb{P}_{\beta_{1/2}}(\langle 1_B, \eta_t \rangle \equiv 1)$ is equivalent to the problem of calculating the absorbing probability in $(\hat{\eta}_t^B)$.
In the rest of Section 3, we give a brief discussion of calculating the limiting distribution of \((\eta_t^A)\), and the method is to refine what we just observe under the Bernoulli initial condition \(\beta_{1/2}\). In this regard, the reader may wish to recall the discussion for (2.10). Since \(\lambda = 1\) and \(\alpha = \alpha_{01} = \alpha_{10}\) for the symmetric Neuhauser-Pacala model, it seems reasonable to expect that equilibria of the model should universally be involved with the value \(p^*_0\), which is 1/2, in the case of the Lotka-Volterra model. This number 1/2 is ubiquitous in the rest of this section.

Recall that \(\alpha\) is the parameter defined in (3.1). It enters the weights of the two mechanisms in \((\eta^A_t)\) (see (3.6) and (3.7)) and the dual process \((\hat{\eta}^B_t)\) (see (3.18) and (3.19)). If \(\alpha = 0\), then there is no random walk with annihilation in \((\hat{\eta}^B_t)\) and so \((\hat{\eta}^B_t)\) cannot die out unless \(B = \emptyset\). The latter follows because any empty site has no effect on any other sites. Algebraically, this is attributable to the fact that 0 is the additive identity in \(\mathbb{Z}_2\). On the other hand, any 1-site \(x\) can only change its neighbors in three ways:

Here, the integer below each figure is the difference of the numbers of filled sites between the right-hand graph and the left-hand graph. Hence, by (3.23) and (3.24), we have the first theorem about 1/2 in equilibria of the model.

**Theorem 3.4.** If \(\alpha = 0\), then \(\beta_{1/2}\) is invariant for the symmetric Neuhauser-Pacala model.

There are two natural questions stemming from this simple theorem.

**Question 3.5.** Is \(\beta_{1/2}\) also the unique limiting distribution whenever we start with mixing initial conditions?

**Question 3.6.** What can be said about the set of invariant distributions of the symmetric Neuhauser-Pacala model if \(\alpha \in (0, 1)\)?

These questions are closely related to the following conjecture due to Neuhauser and Pacala [20, Conjecture 1].

**Conjecture 3.7.** There is coexistence in the symmetric Neuhauser-Pacala model on \(\mathbb{Z}^d\) for any \(d \geq 2\).

If \(\alpha > 0\), \(q(x, y)\) defines the nearest-neighbor random walk on \(\mathbb{Z}\) and \(\alpha_{01} = \alpha_{10} > 0\), coexistence in the sense that one can see both species occupy a fixed region after large times fails with probability one, and hence, clustering holds. See [20, Theorem 2 (b)] for the precise statement. On the other hand, in view the solution (2.10) under the classical Lotka-Volterra model and the convergence of the symmetric Neuhauser-Pacala model on large complete graphs (recall the discussion below (2.12)), it seems reasonable to expect that a mean-field phenomenon would be valid under these stochastic spatial generalizations in general. Coexistence in suitable notions should hold except in very few, if not pathological, cases. Some resolutions of Conjecture 3.7 are achieved in [5, 7, 6, 25], since [20, Theorem 1] takes the first few steps to do so. Yet to the knowledge of the authors, the full resolution of Conjecture 3.7 remains open.

The role of 1/2 in equilibria can be illustrated more generally by the following theorem.

**Theorem 3.8.** Let \(u \in (0, 1)\). If

\[
\langle \eta^B_{t_k}, 1 \rangle \xrightarrow{t_k \to \infty} Z_B \quad \text{in } \mathbb{Z}_+ \cup \{+\infty\}
\]

\[
(3.26)
\]
for some random variable $Z_B$, then

$$
P_{\beta_u}(\langle 1_B, \eta_{t_k} \rangle \equiv 1) \xrightarrow{t_k \to \infty} \frac{1}{2} E \left[ 1 - (1 - 2u)^{Z_B} \right]
$$

with the convention $0^0 = 1$ when $u = 1/2$.

This theorem is a generalization of what we discuss above for the case that $u = 1/2$. For example, on a finite set, $(\beta_u)_{0 < u < 1}$ can explicitly generate uniform distributions over sets of configurations with fixed sizes [4, pages 655–656] and provides links to distributions defined by exchangeable random variables over sites by de Finetti’s representation theorem [15, Theorem 11.10].

The following corollary is immediate from Theorem 3.8.

**Corollary 3.9.** Suppose that (3.26) holds and $u \neq 1/2$. Then $P_{\beta_u}(\langle 1_B, \eta_{t_k} \rangle \equiv 1) \to \frac{1}{2}$ if and only if the dual process survives:

$$
\lim_{t_k \to \infty} P(\langle \hat{\eta}^B_{t_k}, 1 \rangle \geq 1) = 1
$$

and there is **extinction versus unbounded growth** in the dual process:

$$
\lim_{t_k \to \infty} P(1 \leq \langle \hat{\eta}^B_{t_k}, 1 \rangle \leq L) = 0, \quad \forall L \in \mathbb{N}.
$$

(3.27)

The notion of extinction versus unbounded growth is introduced by Sturm and Swart in [25]. We will discuss this notion in more detail later on.

The proof of Theorem 3.8 is an application of the parity dual equation and the following lemma.

**Lemma 3.10 (Parity deviation).** Let $X_1, X_2, \ldots, X_N$ be independent $\mathbb{Z}_+\text{-valued}$ random variables with $P(X_m \equiv 1) = u_m$. Then it holds that

$$
P\left( \sum_{m=1}^{N} X_m \equiv 0 \right) - P\left( \sum_{m=1}^{N} X_m \equiv 1 \right) = \prod_{m=1}^{N} (1 - 2u_m).
$$

**Proof.** The identity follows upon writing the left-hand side as $E[(-1)^{\sum_{m=1}^{N} X_m}]$. ■

**Proof of Theorem 3.8.** With a slight abuse of notation, we write also $\beta_u$ as a random configuration independent of $(\hat{\eta}^B_{t_k})$. By the parity dual equation in Theorem 3.1, we have

$$
\lim_{k \to \infty} P_{\beta_u}(\langle 1_B, \eta_{t_k} \rangle \equiv 1) = \lim_{k \to \infty} P(\langle \hat{\eta}^B_{t_k}, \beta_u \rangle \equiv 1)
$$

$$
= \lim_{k \to \infty} \frac{1}{2} E \left[ 1 - (1 - 2u)^{\langle \hat{\eta}^B_{t_k}, 1 \rangle} \right]
$$

$$
= \frac{1}{2} E \left[ 1 - (1 - 2u)^{Z_B} \right]
$$

by (3.26), where the second equality follows from Lemma 3.10 since

$$
P\left( \sum_{m=1}^{N} X_m \equiv 1 \right) = \frac{1}{2} \left[ 1 - P\left( \sum_{m=1}^{N} X_m \equiv 0 \right) + P\left( \sum_{m=1}^{N} X_m \equiv 1 \right) \right] = \frac{1}{2} \left( 1 - \prod_{m=1}^{N} (1 - 2u_m) \right).
$$

■

More generally, the proof of Theorem 3.8 can be modified in an obvious way to show a similar formula if we use Bernoulli product measures with site-dependent densities, where the densities $\neq 0, 1$ are bounded away from 0 and 1.
Theorem 3.8 and Lemma 3.10 should provide the basic idea behind the related methods in [3, 20, 25]. There the method to resolve the main technical issue for not starting with Bernoulli measures is to identify ‘Bernoulli subsystems’ which have sizes growing to infinity as \( t \to \infty \). The frameworks along this direction in [3, 20, 25] can be roughly fit into the following principle. Let \( X_m(t), m \in \mathcal{I}_t \), be a collection of \( \{0, 1\} \)-valued random variables such that \( X_m(t), m \in \mathcal{B}_t \), are conditionally independent given \( \mathcal{F}_t \) and \( \sum_{m \in \mathcal{I}_t \setminus \mathcal{B}_t} X_m(t) \in \mathcal{F}_t \). The subset \( \{X_m(t); m \in \mathcal{B}_t\} \) defines a Bernoulli subsystem. If \( |\mathcal{B}_t| \) is large, then Lemma 3.10 shows that

\[
\mathbb{P}\left( \sum_{m \in \mathcal{I}_t} X_m(t) = 1 \mid \mathcal{F}_t \right) = \mathbb{P}\left( \sum_{m \in \mathcal{B}_t} X_m(t) = 1 - \sum_{m \in \mathcal{I}_t \setminus \mathcal{B}_t} X_m(t) \mid \mathcal{F}_t \right) \approx \frac{1}{2},
\]

where the error bound for the approximate equality depends only on \( |\mathcal{B}_t| \) and how

\[
u_m(t) = \mathbb{P}(X_m(t) = 1 \mid \mathcal{F}_t)
\]

are close to 0 and 1.

We close the discussion of this section with a very brief account of two applications of (3.28) in [3, 20, 25]. There, the underlying spaces are integer lattices.

**Domination by oriented percolation.** The method in [20], extended from [3], uses the following parity dual equation:

\[
P(\langle 1_B, \eta^A_t \rangle = 1) = \mathbb{P}(\langle \tilde{\eta}^B_t, \eta^A_t \rangle = 1),
\]

where \( \langle \tilde{\eta}^B_t \rangle \) is independent of \( \langle \eta^A_t \rangle \). (3.29) is a simple consequence of the Markov property of \( \langle \eta^A_t \rangle \) and the parity dual equation in Theorem 3.1.) We identify configurations as sets of vertices occupied by 1-individuals. Then one writes \( \langle \tilde{\eta}^B_t, \eta^A_t \rangle = \tilde{\eta}^B_t \cap \eta^A_t \) as a sum of the \( \{0, 1\} \)-valued random variables

\[
X_m(t) = |\tilde{\eta}^B_t \cap \eta^A_t \cap E_m|, \quad m \in \mathcal{I}_t = \mathbb{Z},
\]

where \( \{E_m\} \) is a partition of the whole space.

To apply (3.28) to the probability on the right-hand side of (3.29), the proof proceeds by picking out the space-time boxes where both of the processes \( \langle \eta^A_t \rangle \) and \( \langle \tilde{\eta}^B_t \rangle \) are ‘isolated’. Namely, the Poisson clocks in the constructions of these processes, which govern interactions between sites inside the spatial regions and sites outside, do not ring within the corresponding time intervals. This results in the subset

\[
X_m(t), \quad m \in \mathcal{B}_t,
\]

of independent Bernoulli variables when suitably conditioned, where \( E_m \) and \( E_{m'} \) are far away from each for every \( m \neq m' \).

The remaining task is to show that \( \mathcal{B}_t \) can grow to infinity with probability one. This question is reduced to a question about domination by oriented percolation. See [8, 9].

**Extinction versus unbounded growth.** The method in [25, Section 3] takes a different route, and the discussion above is more similar to it. Now we consider the parity dual equation in the form: for fixed \( s > 0 \),

\[
P(\langle 1_B, \eta^A_{s+t} \rangle = 1) = \mathbb{P}(\langle \tilde{\eta}^B_t, \eta^A_s \rangle = 1) = \mathbb{E}[\mathbb{P}(\langle 1_C, \eta^A_s \rangle = 1) \mid 1_C = \tilde{\eta}^B_t],
\]

(3.30)
where \((\eta^A_t)\) and \((\tilde{\eta}^B_t)\) are independent as above. The first step is to show that if there are sufficiently many types of matrices \(J\) such that the corresponding Poisson processes with rates \(r(J)\) can trigger a change in \((1_C, \eta^A_t)\) for \(r \in [0, s]\) by one occurrence, then we have
\[
\mathbb{P}(1_C, \eta^A_t) = 1 \simeq \frac{1}{2}.
\]
These matrices \(J\)'s are chosen such that \((1_C, (\text{Id} + J)1_A) = 1\), and their ranges are sufficiently apart so that one can write \((1_C, \eta^A_t) \equiv 1\) as a sum of these inner products and other terms. Then the principle in (3.28) comes into play because the Bernoulli variables \(X_m(t), m \in B_t\), are defined from conditioning these Poisson processes with rates \(r(J)\) to ring at most once by time \(s\).

The second step then aims to find the good \(A\)'s by using appropriate initial conditions and the good \(C\)'s for (3.30) by using the extinction versus unbounded growth condition in Corollary 3.9. (The following initial conditions \(\mu\) are enough for the required convergence: \(\mu\) is translation invariant and is supported in the set of configurations \(1_A\) where \(\eta^A_t\) can be any configuration with positive probability when restricted to an arbitrary finite set for every \(t > 0\).) See [25, Section 3.2] for these two steps.

The work in [25] also shows conditions for the extinction versus unbounded growth condition in terms of the survival of \((\eta_t)\) and a certain recurrence property of \((\tilde{\eta}_t)\) when \(\alpha \in (0, 1)\) [25, Sections 3.3 and 3.5].

4 Interacting diffusions on lattices

In this section, we turn to coexistence in the system of interacting Wright-Fisher diffusions defined in (2.15), where from now on we set \(N = 1\).

In addition to the connection to the Lotka-Volterra model explained in Section 2, the model can be considered by itself as a model for the evolution of gene frequencies in a spatially structured two-type population subject to selection, where \(p_t(x)\) describes the proportion of 0-individuals at site \(x \in \mathbb{Z}^d\) and time \(t \geq 0\). In this case, at each site \(x\), for \(\mu < 1\) there is selection in favor of the 0-type if \(s > 0\) and in favor of the 1-type if \(s < 0\). For \(\mu > 1\), we have selection in favor of either heterozygosity or homozygosity according to whether \(s > 0\) or \(s < 0\). For the neutral case \(s = 0\), the system (2.15) reduces to the well-known stepping stone model. See [22].

It is natural to conjecture that long-term coexistence of the two types is possible in the heterozygosity (‘balancing’) selection case, at least if \(s\) is large enough. Indeed, based on comparison arguments with oriented percolation, in [1, Theorem 1.4] a coexistence result in the following sense is proven: For \(\mu > 1\) and fixed small \(\varepsilon > 0\), there exists \(s_0 \geq 0\) such that for all \(s > s_0\) and all initial conditions \(p_0\) with \(p_0(x) \in (\varepsilon, 1 - \varepsilon)\) for all \(x \in \mathbb{Z}^d\), then
\[
\liminf_{t \to \infty} \mathbb{P}(\epsilon < p_t(x) < 1 - \epsilon) > 0, \quad \forall x \in \mathbb{Z}^d.
\]
Note that [1, Theorem 1.4] does not claim that coexistence fails for \(s < s_0\), nor does it specify the value of \(s_0\). But in fact, one would conjecture the following (see [1, Conjectures 2.2 and 2.3]):

**Conjecture 4.1.** Suppose \(\mu > 1\). There exists a critical value \(s_0 \geq 0\) such that we have coexistence for \(s > s_0\) and non-coexistence for \(s < s_0\).

As in Conjecture 3.7, the full resolution of this conjecture is still open to the knowledge of the authors.

Our discussion below will continue the point of view from Section 3 and relate coexistence to the question of survival for a suitable dual process. This time, the duality we can use is a **moment duality** extended from the well-known stepping stone model [22]. It requires a bit of algebra to see
and, unfortunately, seems to be available for restricted values of the parameters \( s \) and \( \mu \) only. In the ‘symmetric’ case \( \mu = 2 \), the moment dual is a \textbf{branching annihilating random walk}, which is a non-monotone system and difficult to analyze mathematically. Based on non-rigorous results in the physics literature, [1] actually conjectures the value of \( s_0 \) in Conjecture 4.1 to be equal to 0 in dimensions \( d \geq 2 \) and to be strictly positive in \( d = 1 \).

4.1 Duality and coexistence

We start with the particular case \( s = 0 \), the stepping stone model, for which there is a well-known moment duality due to Shiga [23] with a system of coalescing random walks \((\xi_t) = (\xi_t(x); x \in \mathbb{Z}^d)\) on \( \mathbb{Z}^d_+ \) defined with the following rates:

\[
\begin{aligned}
\text{migration:} & \quad \xi(x) \rightarrow \xi(x) - 1, \\
& \quad \xi(y) \rightarrow \xi(y) + 1, \\
\text{coalescence:} & \quad \xi(x) \rightarrow \xi(x) - 1 \\
& \quad \text{with rate } \frac{\xi(x)(\xi(x) - 1)}{2},
\end{aligned}
\tag{4.1}
\]

and the dual equation reads as follows. For each \( p_0 \in [0,1]^{\mathbb{Z}^d} \) and initial condition \( \xi_0 \in \mathbb{Z}^d_+ \) such that \( \sum_x \xi_0(x) < \infty \), we have

\[
\mathbb{E} \left[ \prod_{x \in \mathbb{Z}^d} p_t(x)^{\xi_0(x)} \right] = \mathbb{E} \left[ \prod_{x \in \mathbb{Z}^d} p_0(x)^{\xi(x)} \right].
\tag{4.2}
\]

The migration mechanism in (4.1) shows that individuals at site \( x \) can migrate to site \( y \) with rate \( m_{xy} \), whereas the coalescence mechanism shows that only two of the individuals at site \( x \) coalesce at a time.

The moment duality (4.2) can be used to quickly settle the question whether typically the underlying two types of individuals can coexist when the migration matrix \( m \) is such that starting from finitely many particles, the process \((\xi_t)\) will almost surely end up with a single particle due to coalescence, i.e.

\[
|\xi_t| = \sum_{x \in \mathbb{Z}^d} \xi_t(x) \xrightarrow{t \to \infty} 1 \quad \text{a.s.}
\]

In this case, if we consider the homogenous initial condition \( p_0(x) = 1/2 \) for all \( x \in \mathbb{Z}^d \), then the duality (4.2) gives convergence of all mixed moments:

\[
\mathbb{E} \left[ \prod_{x \in \mathbb{Z}^d} p_t(x)^{\xi_0(x)} \right] \xrightarrow{t \to \infty} \frac{1}{2}, \quad \forall \xi_0 \in \mathbb{Z}^d_+ \text{ with } \sum_{x \in \mathbb{Z}^d} \xi_0(x) < \infty,
\]

which implies

\[
\left( \frac{1}{2} \right)^{(d)} \xrightarrow{t \to \infty} \frac{1}{2} \delta_1 + \frac{1}{2} \delta_0.
\]

The two species cannot coexist.

For the case \( s > 0 \) and \( \mu = 2 \), the key idea from [1] is to consider the transformed process:

\[
\sigma_t \stackrel{\text{def}}{=} 1 - 2p_t \in [-1,1].
\]

It satisfies the following system of SDEs:

\[
d\sigma_t(x) = \sum_{y \in \mathbb{Z}^d} m_{xy}(\sigma_t(y) - \sigma_t(x)) dt + \frac{s}{2} (\sigma_t(x)^3 - \sigma_t(x)) dt - \sqrt{1 - \sigma_t(x)^2} dW_t(x), \quad x \in \mathbb{Z}^d.
\tag{4.3}
\]
It is stated in [1, Lemma 2.1] that the moment duality in (4.2) still applies for \( \sigma_t \) in place of \( p_t \), but instead of a coalescing random walk, the dual process is now a branching annihilating random walk.

**Definition 4.2.** Fix \( s \geq 0 \). Define a Markov process \( (\xi_t) \) taking values \( \xi_t \in \mathbb{Z}^d_+ \) starting from finitely many particles at time zero (i.e. \( \sum_{x \in \mathbb{Z}^d} \xi_0(x) < \infty \)) and with the following transition rates:

\[
\begin{align*}
migration: \quad & \{ \begin{array}{l} 
\xi(x) \to \xi(x) - 1, \\
\xi(y) \to \xi(y) + 1,
\end{array} \quad \text{with rate } \xi(x)m_{xy}, \\
n\begin{array}{l}
double branching: \quad \xi(x) \to \xi(x) + 2, \\
an\text{nnihilation: } \xi(x) \to \xi(x) - 2,
\end{array} \quad \text{with rate } s\xi(x), \\
& \frac{\xi(x)(\xi(x) - 1)}{2},
\end{align*}
\]  

(4.4)

This process is called a double branching annihilating random walk (DBARW) with branching rate \( s \).

**Lemma 4.3 ([1]).** Fix \( \mu = 2 \) and \( s \geq 0 \). Then we have the following moment duality between the transformed process \( (\sigma_t) \) defined by (4.3) and the DBARW \( (\xi_t) \) defined in Definition 4.2 with branching rate \( s/2 \): For all \( \sigma_0 \in [-1, 1]^{\mathbb{Z}^d} \) and \( \xi_0 \in \mathbb{Z}^d_+ \) with \( \sum_{x \in \mathbb{Z}^d} \xi_0(x) < \infty \),

\[
E_{\sigma_0} \left[ \prod_{x \in \mathbb{Z}^d} \sigma_t(x)^{\xi_0(x)} \right] = E_{\xi_0} \left[ \prod_{x \in \mathbb{Z}^d} \sigma_0(x)^{\xi_t(x)} \right].
\]  

(4.5)

We will give the proof of Lemma 4.3 below in Section 4.2. As an application of the moment duality, we now show the equivalence of coexistence for the system (2.15) (for \( \mu = 2 \)) and survival of DBARW. See also [25, Lemma 1] for a related result in the context of spin systems.

**Proposition 4.4.** Fix \( s > 0 \). Let \( (p_t) \) denote the solution to the system of SDEs in (2.15) for \( \mu = 2 \), and let \( (\xi_t) \) denote the DBARW with branching rate \( s/2 \) defined above. Then the following four statements are equivalent:

(a) For all initial conditions \( p_0 \) such that \( p_0(x) \in (\varepsilon, 1 - \varepsilon) \) for all \( x \) for some small \( \varepsilon > 0 \), we have long-term coexistence of \( (p_t) \) with positive probability in the sense that there exists \( \kappa \in (0, 1) \) such that

\[
\liminf_{t \to \infty} \mathbb{P} (\kappa < p_t(0) < 1 - \kappa) > 0.
\]

(b) There exists some initial condition \( p_0 \) for which we have long-term coexistence of \( (p_t)_t \) with positive probability.

(c) The DBARW started with exactly two particles at the origin at time zero survives for all time with positive probability. That is, for the initial condition \( \xi_0 = 2\mathbf{1}_{\{0\}} \), we have

\[
\mathbb{P} (\xi_t \neq 0, \forall t \geq 0) > 0.
\]

(d) The DBARW started with any even number of particles at time zero survives for all time with positive probability.

**Proof.** The equivalence of (c) and (d) is clear from the double branching mechanism in (4.4), and it is plain that (a) implies (b). Below, we show that (b) implies (c) and (c) implies (a). We work with the transformed process \( \sigma_t = 1 - 2p_t \) taking values in \([-1, 1]\), which satisfies (4.3). Note that coexistence
for \((p_t)\) as in (a) and (b) is equivalent to the following condition on the process \((\sigma_t)\). For all initial conditions \(\sigma_0\) such that \(\sup_{x \in \mathbb{Z}^d} |\sigma_0(x)| < 1 - 2\varepsilon\) for some small \(\varepsilon > 0\), there exists \(\kappa > 0\) such that

\[
\liminf_{t \to \infty} \mathbb{P}(|\sigma_t(0)| < 1 - 2\kappa) > 0.
\]

The proof that (b) implies (c) is as follows. Suppose by way of contradiction that (c) is false. This means that if we start the DBARW with two particles at the origin at time zero, that is \(\xi_0 = 2\mathbf{i}_{\{0\}}\), then the process will die out with probability one, that is \(\xi_t \to 0\) almost surely. Consider any initial condition \(p_0\) for which the process \((p_t)\) coexists. Then by the moment duality and dominated convergence, we have

\[
\mathbb{E} \left[ \sigma_t(0)^2 \right] = \mathbb{E} \left[ \prod_{x \in \mathbb{Z}^d} \sigma_t(x)^{\xi_0(x)} \right] = \mathbb{E} \left[ \prod_{x \in \mathbb{Z}^d} \sigma_0(x)^{\xi_t(x)} \right] \xrightarrow{t \to \infty} 1. \quad (4.6)
\]

On the other hand, by the assumed long-term coexistence in (b), we know that there is some \(\kappa \in (0, 1)\) such that

\[
\delta \overset{\text{def}}{=} \liminf_{t \to \infty} \mathbb{P}(\kappa < p_t(0) < 1 - \kappa) = \liminf_{t \to \infty} \mathbb{P}(|\sigma_t(0)| < 1 - 2\kappa) > 0,
\]

which implies

\[
\mathbb{E} \left[ \sigma_t(0)^2 \right] \leq (1 - 2\kappa)^2 \mathbb{P}(|\sigma_t(0)| < 1 - 2\kappa) + \mathbb{P}(|\sigma_t(0)| \geq 1 - 2\kappa)
\]

and thus

\[
\liminf_{t \to \infty} \mathbb{E} \left[ \sigma_t(0)^2 \right] \leq (1 - 2\kappa)^2 + (1 - \delta) < 1,
\]

which is a contradiction to (4.6). We have proved that (b) implies (c).

Next, we prove that (c) implies (a). Assume that (c) holds. Consider \(\varepsilon > 0\) and any ‘permissible’ initial condition \(\sigma_0\) in (a), i.e. \(|\sigma_0(x)| \leq 1 - 2\varepsilon\) for all \(x \in \mathbb{Z}^d\). Suppose by way of contradiction that for all \(\kappa > 0\) we have

\[
\liminf_{t \to \infty} \mathbb{P}(|\sigma_t(0)| < 1 - 2\kappa) = 0.
\]

Then for any \(\kappa > 0\) we get

\[
\mathbb{E} \left[ \sigma_t(0)^2 \right] = \mathbb{E} \left[ \sigma_t(0)^2 \left( \mathbb{1}_{(|\sigma_t(0)| < 1 - 2\kappa)} + \mathbb{1}_{(|\sigma_t(0)| \geq 1 - 2\kappa)} \right) \right]
\]

\[
\geq 0 + (1 - 2\kappa)^2 \mathbb{P}(|\sigma_t(0)| \geq 1 - 2\kappa)
\]

\[
= (1 - 2\kappa)^2 \left( 1 - \mathbb{P}(|\sigma_t(0)| < 1 - 2\kappa) \right),
\]

from which we infer that

\[
\limsup_{t \to \infty} \mathbb{E} \left[ \sigma_t(0)^2 \right] \geq (1 - 2\kappa)^2 \left( 1 - \liminf_{t \to \infty} \mathbb{P}(|\sigma_t(0)| < 1 - 2\kappa) \right) = (1 - 2\kappa)^2.
\]

Since \(\kappa > 0\) was arbitrary, this implies

\[
\limsup_{t \to \infty} \mathbb{E} \left[ \sigma_t(0)^2 \right] = 1.
\]

On the other hand, since the DBARW survives by (c), we have \(\delta = \mathbb{P}(\xi_t \neq 0, \forall t \geq 0) > 0\). Then by the duality we have for all \(t > 0\)

\[
0 \leq \mathbb{E} \left[ \sigma_t(0)^2 \right] = \mathbb{E} \left[ \prod_{x \in \mathbb{Z}^d} \sigma_t(x)^{\xi_0(x)} \right] = \mathbb{E} \left[ \prod_{x \in \mathbb{Z}^d} \sigma_0(x)^{\xi_t(x)} \right]
\]

21
≤ E \left[ \prod_{x \in \mathbb{Z}^d} |\sigma_0(x)|^{\xi_t(x)} \right] \\
= E \left[ \prod_{x \in \mathbb{Z}^d} |\sigma_0(x)|^{\xi_t(x)} (\mathbbm{1}_{\{\xi_t \text{ survives}\}} + \mathbbm{1}_{\{\xi_t \text{ dies out}\}}) \right] \\
≤ (1 - 2\varepsilon) \delta + 1 - \delta < 1,

again giving a contradiction. We have proved that (c) implies (a). The proof is complete. ■

By Proposition 4.4, coexistence for just one initial condition \( p_0 \) implies coexistence for all initial conditions of the form considered in (a). Hence, to determine coexistence, it suffices to work with nice initial conditions such as \( p_0(x) = 1/2 \) for all \( x \in \mathbb{Z}^d \), which we considered when discussing the case \( s = 0 \) above. Recall that using a comparison with oriented percolation, [1] proves coexistence for sufficiently large values of \( s \) and so, in view of Proposition 4.4, they obtain as a corollary that the DBARW survives if the branching rate is large enough.

There is another range of parameters for which solutions to (2.15) admit nice duals and so whether coexistence occurs can be settled again under appropriate conditions.

**Lemma 4.5.** For \( \mu \in [-1, 0] \) and \( s < 0 \), the solution \( (p_t) \) to (2.15) is dual to a branching coalescing random walk (BCRW) \((\xi_t)\) taking values in \( \mathbb{Z}^2 \) with transition rates given by (4.1), and in addition, the following two:

\[
\begin{align*}
\xi(x) &\to \xi(x) + 1 \quad \text{with rate } (-s)(\mu + 1)\xi(x), \\
\xi(x) &\to \xi(x) + 2 \quad \text{with rate } (-s)(-\mu)\xi(x).
\end{align*}
\]

Clearly, BCRW survives for all time with probability one whenever the initial condition is not trivial. Note that Lemma 4.5 contains the coalescing random walk dual for the stepping stone model (that is the model with \( s = 0 \)) as a special case. We will also prove Lemma 4.5 in Section 4.2.

**Proposition 4.6.** Let \( (p_t) \) denote the solution to (2.15) with parameters \( s < 0 \) and \( \mu \in [-1, 0] \), and let \( (\xi_t) \) denote the BCRW with branching rates defined in Lemma 4.5. Assume that for any initial condition \( \xi_0 \) with \( \sum_{x \in \mathbb{Z}^d} \xi_0(x) < \infty \),

\[
\sum_{x \in \mathbb{Z}^d} \xi_t(x) \xrightarrow{a.s. \ t \to \infty} +\infty.
\]

Then for each initial condition \( p_0 \) such that \( p_0(x) < 1 - \varepsilon \) for all \( x \in \mathbb{Z}^d \) and some \( \varepsilon > 0 \), it holds that

\[ p_t \xrightarrow{a.s. \ t \to \infty} 0. \]

**Proof.** By Lemma 4.5, we have

\[
E \left[ \prod_{x \in \mathbb{Z}^d} p_t(x)^{\xi_0(x)} \right] = E \left[ \prod_{x \in \mathbb{Z}^d} p_0(x)^{\xi_t(x)} \right] \leq E \left[ (1 - \varepsilon) \sum_{x \in \mathbb{Z}^d} \xi_t(x) \right] \xrightarrow{t \to \infty} 0
\]

by dominated convergence. This shows that all mixed moments of \( (p_t) \) converge to zero, which is enough for our assertion. ■
4.2 Proof of the moment duality

In this subsection, we give the proofs of the two duality results in Lemmas 4.3 and 4.5.

Proof of Lemma 4.3. Let \( L^\xi \) denote the generator of the DBARW \((\xi_t)\) from Definition 4.2 and \( L^\sigma \) denote the generator of \((\sigma_t)\) from (4.3). The bivariate duality function in use is now given by

\[
H(\sigma, \xi) = \sigma^\xi \overset{\text{def}}{=} \prod_{x \in \mathbb{Z}^d} \sigma(\delta^x),
\]

Then (4.3) implies that

\[
L^\sigma H(\cdot, \xi)(\sigma) = \sum_{x \in \mathbb{Z}^d} \left( \sum_{y \in \mathbb{Z}^d} m_{xy}(\sigma(y) - \sigma(x)) \right) \frac{\partial}{\partial \sigma(x)} \sigma^\xi
+ \frac{s}{2} \sum_{x \in \mathbb{Z}^d} (\sigma(x)^3 - \sigma(x)) \frac{\partial^2}{\partial \sigma(x)^2} \sigma^\xi
+ \frac{1}{2} \sum_{x \in \mathbb{Z}^d} (1 - \sigma(x)^2) \frac{\partial^2}{\partial \sigma(x)^2} \sigma^\xi
= \sum_{x \in \mathbb{Z}^d: \xi(x) \geq 1} \left( \sum_{y \in \mathbb{Z}^d} m_{xy}(\sigma(y) - \sigma(x)) \right) \xi(x) \sigma(x)^{\xi(x)-1} \prod_{y \neq x} \sigma(y)^{\xi(y)}
+ \frac{s}{2} \sum_{x \in \mathbb{Z}^d: \xi(x) \geq 1} (\sigma(x)^3 - \sigma(x)) \xi(x) \sigma(x)^{\xi(x)-1} \prod_{y \neq x} \sigma(y)^{\xi(y)}
+ \frac{1}{2} \sum_{x \in \mathbb{Z}^d: \xi(x) \geq 2} (1 - \sigma(x)^2) \xi(x)(\xi(x) - 1) \sigma(x)^{\xi(x)-2} \prod_{y \neq x} \sigma(y)^{\xi(y)}
= \sum_{x \in \mathbb{Z}^d: \xi(x) \geq 1} \xi(x) \left( \sum_{y \in \mathbb{Z}^d} m_{xy}(\sigma(y)\sigma(x)^{-1} - 1) \right) \sigma^\xi
+ \frac{s}{2} \sum_{x \in \mathbb{Z}^d: \xi(x) \geq 1} \xi(x) (\sigma(x)^2 - 1) \sigma^\xi
+ \frac{1}{2} \sum_{x \in \mathbb{Z}^d: \xi(x) \geq 2} \xi(x)(\xi(x) - 1) (\sigma(x)^2 - 1) \sigma^\xi.
\]

On the other hand, if we write \((\xi - \eta)^+\) for the configuration defined by \((\xi(x) - \eta(x))^+\) for any \(x \in \mathbb{Z}^d\) and \(\xi, \eta \in \mathbb{Z}^{2d}_+\), then the definition of \((\xi_t)\) shows that

\[
L^\xi H(\sigma, \cdot)(\xi) = \sum_{x \in \mathbb{Z}^d} \sum_{y \in \mathbb{Z}^d} \xi(x)m_{xy}(H(\sigma, (\xi - \delta_x + \delta_y)^+) - H(\sigma, \xi))
+ \frac{s}{2} \sum_{x \in \mathbb{Z}^d} \xi(x)(H(\sigma, \xi + 2\delta_x) - H(\sigma, \xi))
+ \frac{1}{2} \sum_{x \in \mathbb{Z}^d} \xi(x)(\xi(x) - 1)(H(\sigma, (\xi - 2\delta_x)^+) - H(\sigma, \xi))
= \sum_{x \in \mathbb{Z}^d: \xi(x) \geq 1} \sum_{y \in \mathbb{Z}^d} \xi(x)m_{xy}(\sigma(y)\sigma(x)^{-1} - 1) \sigma^\xi
\]
\[ + \frac{s}{2} \sum_{x \in \mathbb{Z}^d; \xi(x) \geq 1} \xi(x)(\sigma(x)^2 - 1)\sigma^\xi + \frac{1}{2} \sum_{x \in \mathbb{Z}^d; \xi(x) \geq 2} \xi(x)(\xi(x) - 1)(\sigma(x)^2 - 1)\sigma^\xi, \]

which is equal to the right-hand side of (4.8). Hence, by the arguments in [10, pages 189–190], the required moment duality equation (4.5) holds.

**Proof of Lemma 4.5.** In this proof, \( L^\xi \) denotes the generator of the BCRW \((\xi_t)\) with transition rates as defined in the statement of Lemma 4.5 and \( L^p \) denotes the generator of \((p_t)\) from (2.15). The duality function is now changed to

\[ H(p, \xi) = p^\xi \overset{\text{def}}{=} \prod_{x \in \mathbb{Z}^d} p(x)^{\xi(x)}. \]

Then (2.15) implies that

\[
\begin{align*}
L^p H(\cdot, \xi)(p) &= \sum_{x \in \mathbb{Z}^d} \left( \sum_{y \in \mathbb{Z}^d} m_{xy}(p(y) - p(x)) \right) \frac{\partial}{\partial p(x)} p^\xi \\
&\quad + s \sum_{x \in \mathbb{Z}^d} p(x)(1 - p(x))(1 - \mu p(x)) \frac{\partial}{\partial p(x)} p^\xi \\
&\quad + \frac{1}{2} s \sum_{x \in \mathbb{Z}^d} p(x)(1 - p(x)) \frac{\partial^2}{\partial p(x)^2} p^\xi \\
&= \sum_{x \in \mathbb{Z}^d; \xi(x) \geq 1} \left( \sum_{y \in \mathbb{Z}^d} m_{xy}(p(y) - p(x)) \right) \xi(x)p(x)^{\xi(x)-1} \prod_{y \neq x} p(y)^{\xi(y)} \\
&\quad + s \sum_{x \in \mathbb{Z}^d; \xi(x) \geq 1} (p(x) - (\mu + 1)p(x)^2 + \mu p(x)^3) \xi(x)p(x)^{\xi(x)-1} \prod_{y \neq x} p(y)^{\xi(y)} \\
&\quad + \frac{1}{2} s \sum_{x \in \mathbb{Z}^d; \xi(x) \geq 2} (p(x) - p(x)^2) \xi(x)(\xi(x) - 1)p(x)^{\xi(x)-2} \prod_{y \neq x} p(y)^{\xi(y)} \\
&= \sum_{x \in \mathbb{Z}^d; \xi(x) \geq 1} \sum_{y \in \mathbb{Z}^d} \xi(x)m_{xy}(p(y)p(x)^{-1} - 1) p^\xi \\
&\quad + s \sum_{x \in \mathbb{Z}^d; \xi(x) \geq 1} \xi(x)(1 - (\mu + 1)p(x) + \mu p(x)^2) p^\xi \\
&\quad + \frac{1}{2} s \sum_{x \in \mathbb{Z}^d; \xi(x) \geq 2} \xi(x)(\xi(x) - 1)(p(x)^{-1} - 1) p^\xi. 
\end{align*}
\]

On the other hand, the definition of \((\xi_t)\) shows that

\[
\begin{align*}
L^\xi H(p, \cdot)(\xi) &= \sum_{x \in \mathbb{Z}^d} \sum_{y \in \mathbb{Z}^d} \xi(x)m_{xy} H(p, (\xi - \delta_y + \delta_y)^+) - H(p, \xi) \\
&\quad + (s)(-\mu) \sum_{x \in \mathbb{Z}^d} \xi(x)(H(p, \xi + 2\delta_x) - H(p, \xi)) \\
&\quad + (s)(\mu + 1) \sum_{x \in \mathbb{Z}^d} \xi(x)(H(p, \xi + \delta_x) - H(p, \xi)) \\
&\quad + \frac{1}{2} \sum_{x \in \mathbb{Z}^d} \xi(x)(\xi(x) - 1)(H(p, (\xi - \delta_x)^+) - H(p, \xi)) \\
&= \sum_{x \in \mathbb{Z}^d; \xi(x) \geq 1} \sum_{y \in \mathbb{Z}^d} \xi(x)m_{xy}(p(y)p(x)^{-1} - 1) p^\xi \\
&\quad + s \sum_{x \in \mathbb{Z}^d; \xi(x) \geq 1} \xi(x)(1 - (\mu + 1)p(x) + \mu p(x)^2) p^\xi \\
&\quad + \frac{1}{2} s \sum_{x \in \mathbb{Z}^d; \xi(x) \geq 2} \xi(x)(\xi(x) - 1)(p(x)^{-1} - 1) p^\xi.
\end{align*}
\]
\[
\begin{align*}
&= \sum_{x \in \mathbb{Z}^d : \xi(x) \geq 1} \sum_{y \in \mathbb{Z}^d} \xi(x) m_{xy} (p(y)p(x)^{-1} - 1) p^\xi \\
&\quad + (-s)(-\mu) \sum_{x \in \mathbb{Z}^d} \xi(x) (p(x)^2 - 1) p^\xi + (-s)(\mu + 1) \sum_{x \in \mathbb{Z}^d} \xi(x) (p(x) - 1) p^\xi \\
&\quad + \frac{1}{2} \sum_{x \in \mathbb{Z}^d : \xi(x) \geq 2} \xi(x) (\xi(x) - 1) (p(x)^{-1} - 1) p^\xi,
\end{align*}
\]

which is again equal to the right-hand side of (4.9). \hfill \blacksquare

5 References


25


