



# Motion by mean curvature in interacting particle systems

Xiangying Huang<sup>1</sup> · Rick Durrett<sup>2</sup>

*To Harry Kesten, prodigious problem solver, mentor, colleague, and friend*

Received: 5 June 2020 / Revised: 11 July 2021 / Accepted: 14 July 2021 /

Published online: 25 August 2021

© The Author(s), under exclusive licence to Springer-Verlag GmbH Germany, part of Springer Nature 2021

## Abstract

There are a number of situations in which rescaled interacting particle systems have been shown to converge to a reaction diffusion equation (RDE) with a bistable reaction term, see e.g., Cox et al. (Astérisque 349:1–127, 2013), Durrett (Ann Appl Probab 19:477–496, 2009, Electron J Probab 19:1–64, 2014), Durrett and Neuhauser (Ann Probab 22:289–333, 1994). These RDEs have traveling wave solutions. When the speed of the wave is nonzero, block constructions have been used to prove the existence or nonexistence of nontrivial stationary distributions. Here, we follow the approach in a paper by Etheridge et al. (Electron J Probab 22:1–40, 2017) to show that in a wide variety of examples when the RDE limit has a bistable reaction term and traveling waves have speed 0, one can run time faster and further rescale space to obtain convergence to motion by mean curvature. This opens up the possibility of proving that the sexual reproduction model with fast stirring has a discontinuous phase transition, and that in Region 2 of the phase diagram for the nonlinear voter model studied by Molofsky et al. (Theoret Pop Biol 55(1999):270–282, 1999) there were two nontrivial stationary distributions.

**Keywords** Voter model perturbation · Fast stirring · Sexual reproduction model · Nonlinear voter model

**Mathematics Subject Classification** Primary 60K35

## 1 Introduction

The literature on motion by mean curvature is extensive, so we will only cite the papers most relevant to our research. In 1992 Evans et al. [11] established that suitably rescaled

---

✉ Rick Durrett  
rtd@math.duke.edu

<sup>1</sup> University of British Columbia, Vancouver, Canada

<sup>2</sup> Duke University, Durham, NC, USA

versions of the Allen–Cahn equations converge to motion by mean curvature, in the sense that the solution converges to an indicator function of a region whose boundary evolves as the mean curvature flow. The big breakthrough made in this paper was that the limiting result was valid for all time despite the possible occurrence of geometric singularities. See the first four pages of [11] for the physical motivation and references to previous work.

In 1995 Katsoulakis and Souganidis [18] used the results developed in [11] to prove that stochastic Ising models with long range interactions, called Kac potentials, when rescaled converge to motion by mean curvature. The interaction kernel for their Ising model on  $\mathbb{Z}^d$  is

$$\begin{cases} K_\gamma(x, y) = \gamma^d J(\gamma|x - y|) \\ J : \mathbb{R}^d \rightarrow [0, \infty) \text{ has compact support and is symmetric, i.e., } J(x) = J(|x|). \end{cases} \quad (1.1)$$

The weighted sum of spins seen by  $x$  is

$$h_\gamma(x) = \sum_{y \neq x} K_\gamma(x, y) \sigma(y).$$

This formula is used to define the Gibbs measure with inverse temperature  $\beta$

$$\mu(\sigma) = \frac{1}{Z(\beta)} \exp \left( -\beta \sum_x h_\gamma(x) \sigma(x) \right),$$

where  $Z(\beta)$  is a normalization to make  $\mu$  a probability measure. For this formula to be meaningful we have to restrict to a finite box  $\Lambda = [-L, L]^d$  with boundary conditions imposed outside of  $\Lambda$  and then let  $L \rightarrow \infty$ . See Chapter 6 of Liggett [19] for more details.  $h_\gamma$  is also used to define the rates at which  $\sigma(x)$  flips to  $-\sigma(x)$ ,

$$c_\gamma(x, \sigma) = \frac{\exp(-\beta h_\gamma(x) \sigma(x))}{\exp(-\beta h_\gamma(x)) + \exp(\beta h_\gamma(x))}.$$

This is one in the large collection of flip rates for which Gibbs states are reversible stationary distributions. Again, see Chapter 6 of [19].

A very basic question is to understand the behavior of the process as  $\gamma \rightarrow 0$ . DeMasi et al. [4–6] studied the limits as  $\gamma \rightarrow 0$  of the averaged magnetization of the system

$$m_\gamma(x, t) = E_{\mu^\gamma}^\gamma \sigma_t(x), \quad (x, t) \in \mathbb{Z}^d \times \mathbb{R}^+,$$

where  $E_{\mu^\gamma}^\gamma$  is the expectation starting from the measure  $\mu^\gamma$ . To state the result in [4] we need the mean-field equation

$$\frac{\partial m}{\partial t} + m - \tanh(J * m) = 0 \quad \text{in } \mathbb{R}^d \times \mathbb{R}^+, \quad (1.2)$$

where  $J * m$  denotes the usual convolution in  $\mathbb{R}^d$ . Let  $\mathbb{Z}_n^d = \{\bar{x} = (x_1, \dots, x_n) \in (\mathbb{Z}^d)^n | x_1 \neq \dots \neq x_n\}$ .

**Theorem 1.1** [Theorem 2.1 in [18]]. *Assume that the initial measure is product measure  $\mu^\gamma$  with*

$$E_{\mu^\gamma}^\gamma \sigma(x) = m_0(\gamma x), \quad x \in \mathbb{Z}^d,$$

where  $m_0$  is Lipschitz continuous and (1.1) holds. Then for any fixed  $n$  and  $\bar{x} \in \mathbb{Z}_n^d$ ,

$$\lim_{\gamma \rightarrow 0} \left| E_{\mu^\gamma}^\gamma \left( \prod_{i=1}^n \sigma_t(x_i) \right) - \prod_{i=1}^n m(\gamma x_i, t) \right| = 0$$

where  $m$  is the unique solution of (1.2) with initial condition  $m_0$ .

In words, the distribution of the particle system at time  $t$  is almost a product measure in which the probabilities are given by  $m(\gamma x, t)$ . To prove convergence to motion by mean curvature [18] use a lengthy argument to examine the asymptotics of the mean-field equation (1.2) as  $t \rightarrow \infty$  and space and time are rescaled. Since the publication of [18] a number of similar results have been proved. [1, 15, 16, 22, 23] is a small sample of the papers that can be found in AMS subject classification 60.

### 1.1 A more probabilistic approach

Soon after the publication of [11], Chen [2] generalized much of this work and simplified the proofs. Etheridge et al. [10] use his paper as their primary source of information about motion by mean curvature, so we will as well. The object of study in [2] is the reaction diffusion equation (RDE)

$$\begin{cases} \frac{\partial u}{\partial t} = \Delta u - \frac{1}{\epsilon^2} f(u), & (x, t) \in \mathbb{R}^d \times \mathbb{R}^+, \\ u(x, 0) = p(x), & x \in \mathbb{R}^d, \end{cases} \quad (1.3)$$

where  $\epsilon$  is a small rescaling parameter,  $p$  is a bounded continuous function in  $\mathbb{R}^d$  and  $f$  is the derivative of a bistable potential. Chen gives general conditions on  $f$  in (1.3) of his paper [2] that guarantee motion by mean curvature will appear in the limit as  $\epsilon \rightarrow 0$ ,

$$\begin{aligned} f &\in C^2(\mathbb{R}), \quad f \text{ has exactly three zeros: } u_- < u_0 < u_+ \\ f(u) &< 0, \quad \text{for } u \in (-\infty, u_-) \cup (u_0, u_+) \\ f(u) &> 0, \quad \text{for } u \in (u_-, u_0) \cup (u_+, \infty) \\ f'(u_-) &> 0, \quad f'(u_+) > 0, \quad f'(u_0) < 0. \end{aligned}$$

We will restrict our attention to the case in which  $f$  is a third or fifth degree polynomial that is anti-symmetric around its central root  $u_0$ , i.e.,  $f(u_0 - x) = -f(u_0 + x)$ .

In the case of a cubic, the  $1/\epsilon^2$  in front of the reaction term suggests that when  $\epsilon$  is small the values of the solution will be close to one of the three fixed points ( $u_-$ ,  $u_0$  and  $u_+$ ) across most of the space. Chen's results prove this and give quantitative estimates when  $\epsilon$  is small.

To explain the phrase “motion by mean curvature”, we note that under some assumptions that we state later, he proved that the set of points  $\{x \in \mathbb{R}^d : u(x, t) = u_0\}$  can be written as a family of parameterized hyper-surfaces  $\Gamma_t : S^{d-1} \rightarrow \mathbb{R}^d$  where  $S^{d-1}$  is the unit sphere in  $\mathbb{R}^d$ , and  $\Gamma_t$  evolves by

$$\frac{\partial \Gamma_t(\theta)}{\partial t} = \kappa_t(\theta) n_t(\theta), \quad \theta \in S^{d-1}, \quad (1.4)$$

where  $n_t(\theta)$  is the vector normal to the hypersurface and  $\kappa_t(\theta)$  is the mean curvature, i.e., the sum of the principal curvatures. We refer to  $\mathbf{0} = \{\Gamma_t : t \geq 0\}$  as the mean curvature flow.

Etheridge et al. [10] used Chen's results to show that the spatial  $\Lambda$ -Fleming-Viot process with selection against heterozygosity when suitably rescaled in space and time converges to motion by mean curvature. We refer the reader to [10] for the description of the process. Their first step was to study the behavior of the PDE in  $d \geq 2$ ,

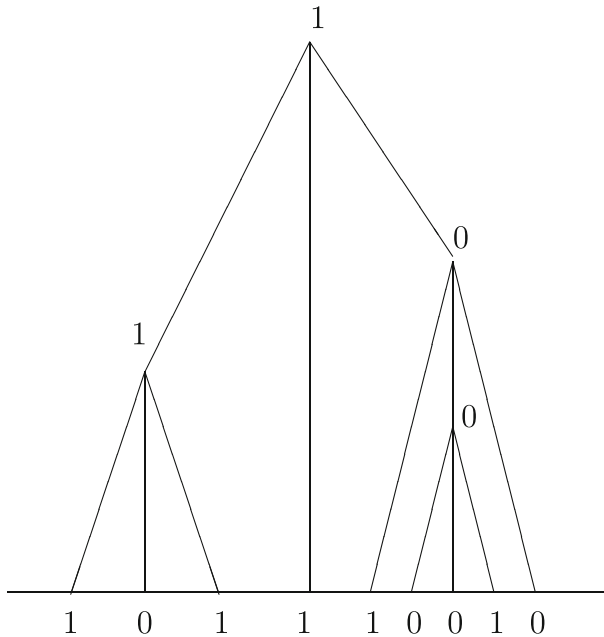
$$\frac{\partial v^\epsilon}{\partial t} = \Delta v^\epsilon + \frac{1}{\epsilon^2} v^\epsilon (1 - v^\epsilon) (2v^\epsilon - 1), \quad v^\epsilon(0, x) = p(x)$$

where  $p(x) : \mathbb{R}^d \rightarrow [0, 1]$  is the initial condition. To analyze the PDE [10] introduce a branching Brownian motion in which particles split into 3 at a fixed rate  $\epsilon^{-2}$ . As in the systems described in the next subsection, this is a dual process that can be used to compute solutions of the PDE. To find  $u(x, t)$  one starts with a particle at  $x$  at time  $t$  and runs the branching Brownian motion down to time 0. If a particle in the system ends up at  $y$  at time 0, its state is set to be 1 with probability  $p(y)$  and 0 with probability  $1 - p(y)$ . As we work upwards the branching tree, states of particles do not change until three lineages coalesce into one. At this point the one lineage that emerges after coalescence takes the value that is in the majority of the three coalescing particles.

A variety of particle systems have dual processes that are close to branching Brownian motions after rescaling. A similar treatment as in [10] can thus be taken to understand these systems. Instead of taking a majority vote at each branch point in the dual process, the specific interaction rule of the particle system considered would prescribe the value of the lineage that emerges after that.

In what follows we will discuss three examples. The sexual reproduction model which is a system with fast stirring and the Lotka–Volterra system and nonlinear voter models that are examples of voter model perturbations. In each case we will first consider a system  $\xi_t^\delta$  on  $\delta\mathbb{Z}^d$  run at rate  $\delta^{-2}$  that converges to a reaction diffusion equation. Then we will introduce a process  $\xi_t^\epsilon$  on  $\delta\epsilon\mathbb{Z}^d$  that is further sped up by a factor of  $\epsilon^{-2}$  that converges to motion by mean curvature. For reasons that we will explain later we will choose

$$\delta = \exp(-\epsilon^{-3}) \quad \text{or} \quad \epsilon = (\log(1/\delta))^{-1/3}. \quad (1.5)$$



**Fig. 1** Picture of the branching Brownian motion. We run from  $(x, t)$  down to time 0, and then work back up the structure to compute the state of  $x$  at time  $t$

Note that  $\epsilon$  is determined by  $\delta$  and vice versa so we can regard either as the small parameter in  $\xi_t^\epsilon$ , but given the notation for the processes it is more natural to choose  $\epsilon$ .

## 1.2 Systems with fast stirring

Particle systems with fast stirring were first introduced by Durrett and Neuhauser [9]. Let  $\delta > 0$  be a small rescaling parameter. They considered processes  $\xi_t^\delta : \delta\mathbb{Z}^d \rightarrow \{0, 1\}$  that evolve as follows:

- (i) There are translation invariant finite range flip rates  $c_\delta(x, \xi)$  that give the rate at which site  $x$  changes to the opposite state when the configuration is  $\xi$ .
- (ii) For each unordered pair  $x, y \in \delta\mathbb{Z}^d$  with  $\|x - y\|_1 = \delta$  we exchange the values at  $x$  and  $y$  at rate  $\delta^{-2}/2$ .

We will focus on the special case in which the particle system is the “sexual reproduction” model where state 1 means a site is occupied and state 0 means vacant. The flip rates is given by

$$c_\delta(x, \xi) = 1_{\{\xi(x)=1\}} + 1_{\{\xi(x)=0\}} \cdot \lambda n_1(x, \xi),$$

where  $\lambda > 0$  is the birth rate and  $n_1(x, \xi)$  is the number of pairs in the set

$$x + \mathcal{N}_b^\delta \equiv x + \delta \cdot \{ \{e_1, e_2\}, \{-e_1, e_2\}, \{-e_1, -e_2\}, \{e_1, -e_2\} \}$$

in which both sites are in state 1. To have a concrete example in mind we will restrict our attention to the case  $d = 2$ . As there are four possible pairs in  $\mathbb{Z}^2$ , we let  $\beta = 4\lambda$ .

Durrett and Neuhauser [9] showed that as  $\delta \rightarrow 0$  the density of 1's near  $x$  at time  $t$  converges to the solution of

$$\frac{\partial u}{\partial t} = \frac{1}{2} \Delta u + \phi(u) \quad \text{where} \quad \phi(u) = -u + \beta u^2(1 - u). \quad (1.6)$$

The term  $-u$  in  $\phi(u)$  accounts for deaths of individuals (i.e., the flips from 1 to 0 at rate 1), while the term  $\beta u^2(1 - u)$  accounts for the sexual reproduction. Writing  $\phi(u) = -u(1 - \beta u(1 - u))$  we see that when  $\beta < 4$  there is no positive solution of  $\phi(u) = 0$ . When  $\beta = 4$ ,  $1/2$  is a double root. When  $\beta > 4$  there are two positive roots  $\rho_1 < 1/2 < \rho_2 < 1$ . Based on this calculation one might guess that as  $\delta \rightarrow 0$ , the critical value for survival of the sexual reproduction model with fast stirring,  $\beta_c(\delta)$ , should converge to 4. However, the correct result, which is proved in [9], is  $\beta_c(\delta) \rightarrow 4.5$  as  $\delta \rightarrow 0$ .

To explain the intuition behind this, we look at the PDE (1.6) in  $d = 1$  for intuition. We note that if  $\beta > 4$  there are traveling wave solutions  $u(x, t) = w(x - ct)$  with  $w(-\infty) = \rho_2$  and  $w(\infty) = 0$ . A little calculus shows that  $w$  satisfies

$$-cw' = (1/2)w'' + \phi(w).$$

Multiplying by  $w'$  and integrating from  $-\infty$  to  $\infty$ , we find, see (1.6) in [9], that

$$c \int w'(x)^2 dx = \int_0^{\rho_2} \phi(y) dy.$$

We have no idea about the value of  $\int w'(x)^2 dx$ , but it is positive so this tells us that the sign of the wave speed  $c$  is the same as that of the integral on the right-hand side. When  $\beta = 4.5$ , the three roots are 0,  $1/3$ , and  $2/3$ , so symmetry around the central root  $1/3$  implies the integral is 0. Monotonicity (or calculus) tells us that  $c < 0$  when  $\beta < 4.5$ , and  $c > 0$  for  $\beta > 4.5$ . Convergence results for the PDE, see e.g. [13], and block constructions were used to show that

- When  $\beta > 4.5$  there is some  $\delta_0(\beta) > 0$  such that for  $\delta < \delta_0(\beta)$  there is a nontrivial stationary distribution with a density close to  $\rho_2$ . The second part of the conclusion about density is an improvement due to Cox et al. [3].
- When  $\beta < 4.5$  there is some  $\delta_0(\beta) > 0$  such that for  $\delta < \delta_0(\beta)$  the process  $\xi_t^\delta$  dies out.

Since  $\rho_2(\beta)$  approaches  $2/3$  as  $\beta \downarrow 4.5$ , it is conjectured in [9] the density of the upper invariant measure (which is obtained by starting with all 1's and letting  $t \rightarrow \infty$ ) has a positive density at  $\beta_c(\delta)$  when  $\delta$  is small.

Here we speed up the process  $\xi_t^\delta$  by an extra factor  $\epsilon^{-2}$  and rescale the space to  $\delta\epsilon\mathbb{Z}^d$  to obtain a new process

$$\xi_t^\epsilon : \delta\epsilon\mathbb{Z}^2 \rightarrow \{0, 1\}.$$

If  $\epsilon$  is kept fixed then the limiting differential equation as  $\delta \rightarrow 0$  is

$$\frac{\partial u^\epsilon}{\partial t} = \frac{1}{2} \Delta u^\epsilon + \frac{1}{\epsilon^2} \phi(u^\epsilon), \quad u^\epsilon(0, x) = p(x), \quad (1.7)$$

where  $p : \mathbb{R}^d \rightarrow [0, 1]$  is the initial condition and the reaction term remains the same

$$\phi(u) = -u + \beta u^2(1 - u).$$

(1.7) matches the form of an Allen–Cahn equation given in (1.3) except for a factor  $\frac{1}{2}$  in front of the Laplacian. This is because their underlying Brownian motions have different rates. The Brownian motion with generator  $\Delta u$  has rate 2, that is, at time 1 the Brownian motion has variance is 2, while the Brownian motion with generator  $\frac{1}{2} \Delta u$  runs at rate 1. We will adopt the convention in probability and assume that

$$\text{all Brownian motions have rate 1,} \quad (1.8)$$

which gives rise to PDEs with a factor  $\frac{1}{2}$  before the Laplacian like (1.7).

Fixing  $\epsilon$  and letting  $\delta \rightarrow 0$  shows us how the rescaled particle system is related to the Allen–Cahn equation. However, to prove our result we need to take both  $\epsilon$  and  $\delta$  to 0. In order to avoid collisions in the dual process (see Sect. 2 for a full discussion), we need to require that the branching rate  $\epsilon^{-2}$  is much slower than the stirring rate  $(\delta\epsilon)^{-2}/2$  so that newly born particles move away from each other before the next branching time. Choosing  $\delta = \exp(-\epsilon^{-3})$  guarantees this. Weaker conditions may suffice.

Let  $p(x) : \mathbb{R}^d \rightarrow [0, 1]$  be the initial density of the system that we consider. In the case of sexual reproduction  $d = 2$ . We will state our assumptions on  $p$  later in Sect. 1.5, see (C1)–(C3). We say the process  $\xi_t^\epsilon$  starts with initial condition  $p$  if the initial distribution is a product measure where  $P(\xi_0^\epsilon(x) = 1) = p(x)$  for  $x \in \delta\epsilon\mathbb{Z}^d$ .

**Theorem 1.2** *Let  $\xi_t^\epsilon : \delta\epsilon\mathbb{Z}^2 \rightarrow \{0, 1\}$  denote the rescaled sexual reproduction model with fast stirring starting with an initial condition  $p(x)$  that satisfies (C1)–(C3). Choose  $\delta = \exp(-\epsilon^{-3})$ . If  $\beta = 4.5$  then as  $\epsilon \rightarrow 0$ ,  $P(\xi_t^\epsilon(x) = 1)$  converges to motion by mean curvature.*

Theorem 1.9 will explain explicitly what it means to converge to motion by mean curvature. Theorem 1.2 shows that the probabilities  $P(\xi_t^\epsilon(x) = 1)$  converge to a density  $u(x, t)$  that satisfies motion by mean curvature. As in Theorem 1.2 in [3] one can also prove that the rescaled particle system which takes values in  $\{0, 1\}$  on a fine grid also converges to  $u(x, t)$ . See the discussion before Theorem 1.2 in [3] for the necessary definition. This remark also applies to the next two examples.

In motion by mean curvature the interfaces become straight as time  $t \rightarrow \infty$ , so the regions in which the solution is close to one of the two stable fixed points get larger. This suggests that

**Conjecture 1.3** *If  $\beta = 4.5$  there exists some  $\epsilon_0(\beta) > 0$  so that when  $\epsilon < \epsilon_0(\beta)$  there is a translation invariant stationary distribution for the process  $\xi_t^\epsilon$  with density close to  $2/3$ .*

Theorem 1.2 suggests that there is a discontinuous phase transition but does not rule out the possibility that the phase transition could be continuous for any  $\epsilon > 0$ .

### 1.3 Voter model perturbations

Cox et al. [3] introduced a class of interacting particle systems called voter model perturbations. For simplicity we will restrict our attention to processes with two states. In this case the process is denoted by  $\xi_t : \mathbb{Z}^d \rightarrow \{0, 1\}$  and the rate at which  $\xi_t(x)$  flips to the opposite state given configuration  $\xi$  is

$$c^\delta(x, \xi) = c_v(x, \xi) + \delta^2 c_p(x, \xi)$$

where  $c_v(x, \xi)$  is the voter flip rate and  $c_p(x, \xi)$  is the perturbation flip rate. We rescale the system  $\xi_t$  by  $x \rightarrow \delta x$ ,  $t \rightarrow \delta^{-2}t$  and obtain the rescaled process  $\xi_t^\delta : \delta\mathbb{Z}^d \rightarrow \{0, 1\}$ . The perturbation  $c_p(x, \xi)$  is scaled down by  $\delta^2$  so that on the sped up time scale it is  $O(1)$  while the voter model runs at rate  $\delta^{-2}$ .

The voter model part of the process will depend on a symmetric (i.e.,  $K(x) = K(-x)$ ), irreducible probability kernel  $K : \mathbb{Z}^d \rightarrow [0, 1]$  with  $K(0) = 0$  and covariance matrix  $\sigma^2 I$ . Letting  $\mathcal{N}_v$  denote the neighborhood for voting (determined by  $K$ ), whenever there is a voter flip at  $x \in \mathbb{Z}^d$ , the voter at  $x$  chooses a site in  $x + \mathcal{N}_v$  randomly according to the probability kernel  $K$  and adopts its state. The voter flip rate can be formulated as

$$c_v(x, \xi) = (1 - \xi(x))f_1(x, \xi) + \xi(x)f_0(x, \xi),$$

where  $f_j(x, \xi) = \sum_{y \in \mathbb{Z}^d} K(y - x)1_{\{\xi(y)=j\}}$  is the local density.

Cox et al. [3] have shown (see their Theorem 1.2) that, under some mild assumptions on the perturbation  $c_p$ , if we run the system on  $\delta\mathbb{Z}^d$  with  $d \geq 3$  then the process converges to the solution of a reaction diffusion equation

$$\frac{\partial u}{\partial t} = \frac{\sigma^2}{2} \Delta u + \phi(u),$$

where  $\phi$  is the reaction term that depends on the particular perturbation. A general formula is given in Sect. 1.1 of [3]. See (1.30) Here,  $d \geq 3$  is needed so that the voter model has a one parameter family of stationary distribution. Four examples were studied in [3]. Two fall within the scope of this investigation.

*Lotka–Volterra systems* This model of the competition of two species were initially studied by Neuhauser and Pacala [21]. For more recent references see [3]. In this case the perturbation rate is given by

$$c_p(x, \xi) = \theta_0 f_1^2(1 - \xi(x)) + \theta_1 f_0^2 \xi(x)$$

where  $\theta_0$  and  $\theta_1$  are parameters in  $\mathbb{R}$ . In words we pick two nearest neighbors of  $x$  (with replacement, according to  $K$ ) and flip if both of the neighbors are of the opposite

type to  $x$ . Let  $\{e_1, e_2\}$  be i.i.d. with law  $K(\cdot)$  and let  $\langle \cdot \rangle_u$  denote the expectation on the product space where  $e_1, e_2$  and  $\xi$  are independent and  $\xi$  distributes as the voter equilibrium with density  $u$ . Then the limiting PDE has reaction term

$$\phi(u) = \theta_0 \langle (1 - \xi(0))\xi(e_1)\xi(e_2) \rangle_u - \theta_1 \langle \xi(0)(1 - \xi(e_1))(1 - \xi(e_2)) \rangle_u.$$

This term can be rewritten in the form

$$\phi(u) = u(1 - u)[\theta_0 p_2 - \theta_1(p_2 + p_3) + up_3(\theta_0 + \theta_1)], \quad (1.9)$$

where  $p_2 = p(0|e_1, e_2)$  is the probability that the rate 1 random walks with kernel  $K$  starting from  $e_1$  and  $e_2$  coalesce but they avoid the one starting at 0, and  $p_3 = p(0|e_1|e_2)$  is the probability that the random walks starting from 0,  $e_1, e_2$  never coalesce.

In [3] the phase diagram is described. There are five regions  $\{R_i, 1 \leq i \leq 5\}$ , see Fig. 1. At the boundary between  $R_4$  and  $R_5$ ,  $\theta_0 = \theta_1 = \theta > 0$  so (1.9) simplifies to

$$\phi(u) = \theta p_3 u(1 - u)(2u - 1).$$

In this case the reaction diffusion equation is bistable and the speed of traveling waves is 0. Next we further rescale the system  $\xi_t^\delta$  by  $x \rightarrow \epsilon x, t \rightarrow \epsilon^{-2}t$  to get the second rescaled process  $\xi_t^\epsilon$ . Following the same approach as our proof of Theorem 1.2, we have

**Theorem 1.4** *Let  $\xi_t^\epsilon : \delta\epsilon\mathbb{Z}^d \rightarrow \{0, 1\}$  denote the rescaled voter model perturbations where the perturbation is a Lotka–Volterra system, starting with an initial condition  $p(x)$  that satisfies (C1)–(C3). Choose  $\delta = \exp(-\epsilon^{-3})$ . In  $d \geq 3$  as  $\epsilon \rightarrow 0$ ,  $P(\xi_t^\epsilon(x) = 1)$  converges to motion by mean curvature.*

In the Lotka–Volterra system the stable fixed points are at 0 and 1, so reasoning as we did for the sexual reproduction model with fast stirring:

**Conjecture 1.5** *When  $\epsilon$  is sufficiently small there is clustering in the process  $\xi_t^\epsilon$ , i.e., for any finite box  $B$  the probability of seeing both types in the box tends to 0 as  $t \rightarrow \infty$ .*

**Nonlinear voter models** Molofsky et al. [20] used simulations and heuristic arguments to study a discrete time system with nearest neighbor interactions. We consider a continuous time version of the system with long range interactions. At times of the arrivals of a rate 1 Poisson process, a site  $x$  chooses four points  $x_1, \dots, x_4$  at random from  $x + [-L, L]^d$ . If there are exactly  $k$  one's at the sites  $x, x_1, \dots, x_4$  then  $x$  becomes 1 with probability  $a_k$  and 0 with probability  $1 - a_k$  where

$$a_0 = 0, \quad a_5 = 1, \quad a_1 = 1 - a_4 \quad a_2 = 1 - a_3.$$

This gives us a two-parameter family of models that are symmetric under interchange of 0 and 1.

It is complicated to compute the reaction term  $\phi_L(u)$  explicitly as the states of the chosen sites  $x, x_1, \dots, x_4$  might not be independent. However, when the neighborhood  $\mathcal{N}_b = [-L, L]^d$  is chosen to be large then coalescence in the dual process is rare and the states of these sites become nearly independent. A little calculation, see (1.67) in [3], shows that if they are independent then the reaction term is

$$\phi(u) = b_1 u(1-u)^4 + b_2 u^2(1-u)^3 - b_2 u^3(1-u)^2 - b_1(1-u)^4 u, \quad (1.10)$$

where  $b_1 = 4a_1 - a_4$  and  $b_2 = 6a_2 - 4a_3$ . For any  $L > 0$ ,  $\phi_L(u)$  has the same form as that in (1.10) with coefficients  $b_{1,L}, b_{2,L}$  instead of  $b_1, b_2$ . If  $L$  is large then the coefficients  $b_{1,L}, b_{2,L}$  are close to the coefficients  $b_1, b_2$  in the independent case.

The reaction term  $\phi(u)$  is a cubic in Region 1 and 3, but in Region 2 and 4 it is quintic. This leads to the following predictions about the behavior of the system.

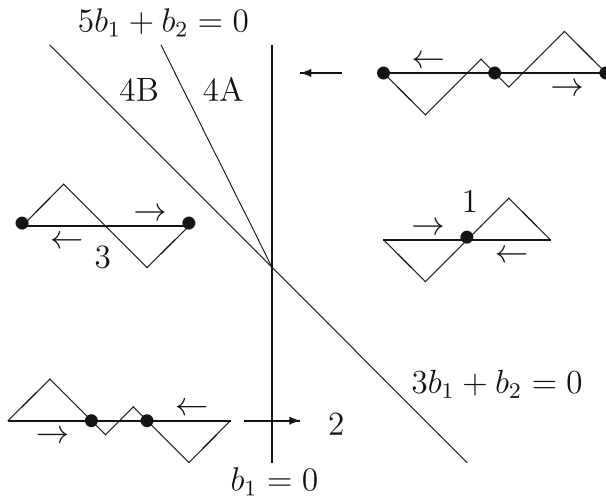
- In Region 1, the fixed point at  $1/2$  is attracting, so the system should exhibit coexistence.
- In Region 3, the fixed point at  $1/2$  is unstable, so when the process is sped up it should exhibit motion by mean curvature, and we expect clustering, i.e., for any finite box  $[-N, N]^d$  the probability that all sites in this box have the same state tends to 1.
- In Region 2, 0 and 1 are unstable fixed points, so if the fixed points are  $u^* < 1/2 < 1 - u^*$ , the values in  $[0, u^* - \varepsilon]$  and  $[1 - u^* + \varepsilon, 1]$  for any  $\varepsilon > 0$  should rapidly disappear from the solution. When the process is sped up then the system exhibits motion by mean curvature, resulting in large regions with 1's at density  $u^*$  separated by a thin boundary from large regions with density  $1 - u^*$ .
- In Region 4, there is a traveling wave solution  $w_1$  with  $w_1(-\infty) = 1$  and  $w_1(\infty) = 1/2$  with speed  $c_1$  and a traveling wave solution  $w_2$  with  $w_2(-\infty) = 1/2$  and  $w_2(\infty) = 0$  with speed  $c_2$ . By symmetry  $c_2 = -c_1$ . If  $c_1 < 0$  (Case 4A), the PDE converges to  $1/2$  and there is coexistence. If  $c_1 > 0$  (Case 4B) and  $L$  is sufficiently large, then there is a traveling wave solution  $w_0$  of the PDE in  $d = 1$  with  $w_0(-\infty) = 0$  and  $w_0(\infty) = 1$  with speed 0 (see p. 284 in [14]). When the process is sped up then it should exhibit motion by mean curvature, and we expect clustering.

In [3] the following result is proved, see their Theorem 1.13 (Fig. 2).

**Theorem 1.6** *Suppose  $(b_1, b_2)$  is in Region 1, 2 or 4A. If  $L$  is sufficiently large then (i) There is coexistence for sufficiently small  $\delta$  (depending on  $L$ ). (ii) Let  $\eta > 0$ . In Region 1 and 4A, there is a  $\delta_0(\eta)$  so that for  $\delta < \delta_0(\eta)$  and any stationary distributions  $\nu$  with  $\nu(\xi \equiv 0) = \nu(\xi \equiv 1) = 0$  have*

$$\sup_x |\nu(\xi(x) = 1) - 1/2| < \eta.$$

Again we need to further speed up the process to get convergence to motion by mean curvature. We rescale the process  $\xi_t^\delta$  a second time by speeding up time by an extra  $\epsilon^{-2}$  and rescaling space to  $\delta\epsilon\mathbb{Z}^d$  to define a process  $\xi_t^\epsilon$ .



**Fig. 2** Phase diagram for the continuous time nonlinear voter model with large range in  $d \geq 3$ . Piecewise linear curves show the shape of  $\phi$ . Black dots indicate the locations of stable fixed points

We define the voting and branching neighborhoods by

$$\mathcal{N}_v = \{\pm e_1, \dots, \pm e_d\} \quad \text{and} \quad \mathcal{N}_b = [-L, L]^d \cap \mathbb{Z}^d.$$

To prove our result we need several assumptions:

- (A1)  $b_1 > 0$  and  $3b_1 + b_2 < 0$ : the process is in Region 2.
- (A2)  $0 \leq a_1 \leq a_2 \leq 1/2$ : the process is attractive.
- (A3)  $6b_1 + b_2 > 0$ : the  $g$  function defined in (1.14) is concave on  $(1/2, 1 - u^*)$ .

**Theorem 1.7** Let  $\xi_t^\epsilon : \delta \mathbb{Z}^d \rightarrow \{0, 1\}$  denote the rescaled voter model perturbations where the perturbation is a nonlinear voter model, starting with an initial condition  $p(x)$  that satisfies (C1)–(C3). Choose  $\delta = \exp(-\epsilon^{-3})$ . Suppose the initial condition  $p(x)$  satisfies  $\epsilon \leq p(x) \leq 1 - \epsilon$  for some  $\epsilon > 0$ . In  $d \geq 3$  if (A1), (A2), and (A3) hold then as  $\epsilon \rightarrow 0$ ,  $P(\xi_t^\epsilon(x) = 1)$  converges to motion by mean curvature.

Using the reasoning from the two previous examples:

**Conjecture 1.8** If (A1), (A2), and (A3) hold then there exists some  $\epsilon_0(a_1, a_2) > 0$  so that when  $\epsilon < \epsilon_0(a_1, a_2)$  there is a translation invariant stationary distribution in which the density is close to  $u^*$ . By symmetry there is also one with density close to  $1 - u^*$ .

The statement in Conjecture 1.8 implies the existence of two translation invariant stationary distributions.

## 1.4 Overview of proofs

The key to the proof in [10] and in our three examples is understanding the dual process and using a special function  $g$  to compute the state after each branch point in the dual process.

### *Duality and the $g$ function*

The  $g$  function in [10]. [10] begins by constructing a dual process that produces the solution to

$$\frac{\partial u}{\partial t} = \Delta u + cu(1-u)(2u-1), \quad u(0, x) = p(x), x \in \mathbb{R}^d.$$

The initial condition  $p$  is assumed to take values in  $[0, 1]$  and satisfy some regularity conditions that we will state later.

The dual process in [10] is a branching Brownian motion in  $\mathbb{R}^d$  in which the Brownian motions are run at rate 2 and split into 3 particles at rate  $c$ . To compute the solution at  $x$  at time  $t$  they run the dual process backward in time down to time 0. A dual particle that lands at  $y$  at time 0 is set to be 1 with probability  $p(y)$  and to be 0 with probability  $1 - p(y)$ . The states for different particles at time 0 are independent. Then they work their way back up the tree performing majority vote whenever three lineages merge into one. In [10] an important role in the proof is played by the function

$$g_0(p) = p^3 + 3p^2(1-p) = 3p^2 - 2p^3, \quad (1.11)$$

which is the probability that the output of the majority vote operation is 1 when the inputs are independent Bernoulli( $p$ ) random variables.  $g_0$  has fixed points at 0,  $1/2$  and 1.

*The  $g$  function in sexual reproduction model with fast stirring* The sexual reproduction model with fast stirring has a dual process that was introduced by Durrett and Neuhauser [9]. The dual has particles that are moved by stirring, and have births when events in the sexual reproduction dynamics occur. In Sect. 2 we define this dual process rigorously and show that in the fast stirring limit it is almost a branching Brownian motion in  $\mathbb{R}^2$ . Since a birth event depends on the states of three particles (two particles in the chosen pair and the particle at the center), the dual branches into three particles at each branch point. However, we mark one lineage to indicate it came from the original particle while the other two are offspring. When  $\beta = 4.5$ , we have a birth event with probability  $9/11$  (i.e.,  $\beta/(1+\beta)$ ) and a death event with probability  $2/11$ . The analogous function to (1.11) for the sexual reproduction model with fast stirring is

$$g_1(p) = \frac{9}{11}[p^2(1-p) + p] = \frac{9}{11}[p + p^2 - p^3], \quad (1.12)$$

which has fixed points 0,  $1/3$  and  $2/3$ .

*The  $g$  function in voter model perturbations where the perturbation is a Lotka–Volterra system* Voter model perturbations also have duals that were defined by Cox et al. [3].

In the class of nonlinear voter models that we will study, the dual is a branching coalescing random walk. In the Lotka–Volterra system the system branches into three, and we mark one lineage to indicate it came from the original particle (call it  $x$ ) while the others are offspring (call them  $y$  and  $z$ ).

To simplify computation, in the dual process we only consider the “effective” perturbations and let them be branch points. When a perturbation occurs at  $x$  there are three possibilities: (i) If  $x$  coalesces with  $y$  or  $z$  (or with both) there is no change in the state of  $x$ . Hence we ignore this case. (ii) If  $y$  and  $z$  coalesce then  $y$  and  $z$  share the same state. This case is treated as a voter event since  $x$  would adopt the state of a randomly chosen neighbor ( $y$  or  $z$ ). Hence this case is not part of the perturbation. (iii) If there is no coalescence among the three particles,  $x$  changes state if  $y$  and  $z$  are both in the opposite state to itself. Case (iii) is the only effective perturbation and in this case

$$g_2(p) = (1-p)p^2 + p[1 - (1-p)^2] = 3p^2 - 2p^3 \quad (1.13)$$

which is the same as (1.11).

*The  $g$  function in voter model perturbations where the perturbation is a nonlinear voter model* In the nonlinear voter model the system branches into five, and we again mark one lineage to indicate it came from the original particle. Since the branching rate is 1  $\phi(p)$  has the form

$$\phi(p) = -ph_1(p) + (1-p)h_2(p),$$

where  $h_1(p)$  represents the probability of getting a 0 when the center is in state 1 and the states of the rest four sites are i.i.d. Bernoulli( $p$ ), while  $h_2(p)$  represents the probability of the center flipping from 0 to 1. It follows that

$$\begin{aligned} g_3(p) &= p(1 - h_1(p)) + (1-p)h_2(p) = \phi(p) + p \\ &= b_1p(1-p)^4 + b_2p^2(1-p)^3 - b_2p^3(1-p)^2 - b_1p^4(1-p) + p \end{aligned} \quad (1.14)$$

where  $b_1 = 4a_1 - a_4$  and  $b_2 = 6a_2 - 4a_3$ . In the collection of nonlinear voter models that we study  $g_3$  has fixed points at 0,  $1 - u^*$ ,  $1/2$ ,  $u^*$ , and 1. If the middle fixed point  $u_0$  is unstable then 0, 1 are stable if there are three zeros, and  $1 - u^*$ ,  $u^*$  are stable if there are five zeros. We collect these observations into an assumption

- (G0) There are fixed points  $0 \leq u_- < u_0 < u_+ \leq 1$  where  $u_0$  is unstable,  $u_+$ ,  $u_-$  are stable, and  $u_+ - u_0 = u_0 - u_-$ . There can be fixed points at 0 and 1 which must be unstable. To avoid absorption, the initial condition  $p(x)$  is uniformly bounded away from the fixed points at 0 and 1, i.e.,
- (i) if there is a fixed point at 0 we suppose the initial condition  $p(x) \geq \varepsilon$  for some  $\varepsilon > 0$ ,
  - (ii) if there is a fixed point at 1 we suppose the initial condition  $p(x) \leq 1 - \varepsilon$  for some  $\varepsilon > 0$ .

We can observe the reaction term  $\phi(p)$  in the above three examples satisfies

$$\phi(p) = r(g(p) - p) \quad (1.15)$$

where  $r$  is the reaction rate.

### 1.5 Proof of convergence

The main result in this paper is the following result that contains Theorems 1.2, 1.4 and 1.7 as special cases. This result applies to any model with fast stirring or voter model perturbation where the  $g$  function satisfies (G0) and the following assumptions:

$$(G1) \quad g(u_+ - \delta) + g(u_- + \delta) = u_- + u_+ = 2u_0 \quad (1.16)$$

$$(G2) \quad g'(u_0) > 1 \quad \text{and} \quad g'(u_-) = g'(u_+) < 1. \quad (1.17)$$

$$(G3) \quad g''(p) > 0 \text{ if } p \in (u_-, u_0), g''(p) < 0 \text{ if } p \in (u_0, u_+). \quad (1.18)$$

(G4) There exists  $c_0 \in (0, 1 - g'(u_-))$  and  $\delta_* \equiv \inf\{x \geq 0 : g'(u_- + x) \geq 1 - c_0\} > 0$  so that for  $\delta \leq \delta_*$

$$u_+ - g(u_+ - \delta) = g(u_- + \delta) - u_- \leq (1 - c_0)\delta. \quad (1.19)$$

(G5)  $g$  is strictly increasing on  $[0, 1]$ .

In Sect. 4 we will show that the conditions hold in all our examples.

The initial condition  $p : \mathbb{R}^d \rightarrow [0, 1]$  is assumed to satisfy some regularity conditions given later. Given  $p$ , the initial interface is defined to be

$$\Gamma = \left\{ x \in \mathbb{R}^d : p(x) = u_0 \right\}.$$

Following [10] we suppose that  $\Gamma$  is a smooth hypersurface which is also the boundary of a bounded open set topologically equivalent to the sphere. Now we state the regularity conditions imposed on the initial condition  $p$ :

(C1)  $\Gamma$  is  $C^\alpha$  for some  $\alpha > 3$ .

(C2) For  $x$  inside  $\Gamma$ ,  $p(x) < u_0$ . For  $x$  outside  $\Gamma$ ,  $p(x) > u_0$ .

(C3) There exists  $r, \gamma > 0$  such that for all  $x \in \mathbb{R}^d$ ,  $|p(x) - u_0| \geq \gamma(\text{dist}(x, \Gamma) \wedge r)$ .

The conditions (C1)–(C3) guarantee that the mean curvature flow  $\mathbf{0} = \{\Gamma_t : t \geq 0\}$  started from  $\Gamma$  exists up to some finite time  $\mathcal{T}$ , see e.g., Evans and Spruck [12]. With  $\Gamma_t$  properly defined, the meaning of  $d(x, t)$  is now precise: it is the signed distance from  $x$  to  $\Gamma_t$ , positive outside  $\Gamma_t$  and negative inside. Note that  $\Gamma_t = \{x \in \mathbb{R}^d : d(x, t) = 0\}$ .

In the scope of this paper we consider only the evolution of a single interface. In more general scenarios there could be multiple interfaces evolving together, say nested interfaces.

**Theorem 1.9** *Let  $\xi_t^\epsilon$  denote a rescaled particle system on  $\delta\epsilon\mathbb{Z}^d$  within the two classes considered. Suppose  $\xi_t^\epsilon$  satisfies (G0)–(G5), and let  $u^\epsilon(t, x) = P(\xi_t^\epsilon(x) = 1)$  with  $u^\epsilon(0, x) = p(x)$ . Let  $T^* \in (0, \mathcal{T})$  and  $k \in \mathbb{N}$  be fixed. There exist  $\epsilon_d(k) > 0$  and  $b_d(k), c_d(k) \in (0, \infty)$  such that for all  $\epsilon \in (0, \epsilon_d)$  and  $t$  satisfying  $b_d(k)\epsilon^2 |\log \epsilon| \leq t \leq T^*$ ,*

1. for  $x$  such that  $d(x, t) \geq c_d(k)\epsilon |\log \epsilon|$ , we have  $u^\epsilon(t, x) \in (u_+ - \epsilon^k, u_+ + \epsilon^k)$ ,
2. for  $x$  such that  $d(x, t) \leq -c_d(k)\epsilon |\log \epsilon|$ , we have  $u^\epsilon(t, x) \in (u_- - \epsilon^k, u_- + \epsilon^k)$ .

Our proof follows [10] very closely. As we give our proof we will describe the correspondence between the two arguments. Here we give a brief outline of the proof to highlight the main steps. From now on we will let  $\mathbf{B}(t)$  denote the one-dimensional branching Brownian motion, while  $\mathbf{W}(t)$  denotes branching Brownian motion in  $d \geq 2$ .

*Step 1* Prove a result in one dimension. Let  $u_-$ ,  $u_+$  be the stable fixed points of  $g$ , and let  $\mathbb{V}(\mathbf{B}(t))$  be the result of applying the algorithm defined in Sect. 2.3 to compute the state when the initial density is  $p_0(x) = u_- \cdot 1_{\{x < 0\}} + u_+ \cdot 1_{\{x \geq 0\}}$ . In  $d = 1$  the interface is a single point and there is no curvature so it does not move. The one dimensional version of Theorem 1.9 is Theorem 3.6. These results are proved by combining facts about the iteration, with information on the structure of the tree and bounds on the movement of Brownian motion.

*Step 2* Generalize Theorem 3.6 to  $d \geq 2$  with  $x$  replaced by  $d(x, t)$ , the signed distance from the  $x$  to the interface  $\Gamma_t$ , see Proposition 3.10.

*Step 3* Proposition 3.10 takes care of the values away from the interface. The next step is to take care of the values near the interface by showing that the probability the dual gives a 1 (resp. 0) at  $x$  for a general initial condition  $p$  is almost the same as the probability the algorithm in Sect. 2.3 computes a 1 (resp. 0) at  $d(x, t) \pm K_1 e^{K_2 t} \epsilon |\log \epsilon|$  in the one dimensional system with the special initial condition  $p_0(x) = u_- \cdot 1_{\{x < 0\}} + u_+ \cdot 1_{\{x \geq 0\}}$ . See Proposition 3.11 for this result. The key step to proving Proposition 3.11 is Lemma 3.12, which compares the values computed by the algorithm in  $d = 1$  at

$$\begin{aligned} z_1^\pm &= d(\hat{X}_s^\epsilon, t - s) \pm \gamma(t - s)\epsilon |\log \epsilon|, \\ z_2^\pm &= B_s \pm \gamma(t)\epsilon |\log \epsilon|, \end{aligned}$$

where  $\hat{X}_s^\epsilon$  is an approximation of the rescaled dual process that will be defined in Sect. 2.1.2.

## 2 Dual process and branching random walk

### 2.1 The sexual reproduction model

#### 2.1.1 The graphical representation

We begin by constructing the process using a graphical representation that consists of a collection of independent Poisson processes. Here, we give only a brief description of the construction. More details can be found in Sect. 2a of [9]. Define

$$c^*(\epsilon) = \sum_i \sup_{\xi^\epsilon \in \{0, 1\}^{\delta\epsilon\mathbb{Z}^d}} c_i^\epsilon(0, \xi^\epsilon) = (\beta + 1)\epsilon^{-2}, \quad (2.1)$$

where  $c_i^\epsilon(0, \xi^\epsilon)$  is the rate that the origin changes to state  $i$  in the process  $\xi_t^\epsilon$  when the configuration is  $\xi^\epsilon$ .

- For every site  $x \in \delta\epsilon\mathbb{Z}^d$  we have a Poisson process  $\{T_n^{b,x}, n \geq 1\}$  with rate  $c^*(\epsilon)$  and a sequence of i.i.d. random variables  $\{U_n^x, n \geq 1\}$  uniform on  $(0, 1)$ . At time  $T_n^{b,x}$  we use  $U_n^x$  to determine the type of change that occurs:
  1. If  $U_n^x \in (0, \frac{\beta}{1+\beta})$ ,  $x$  gives birth to two particles on a randomly chosen pair from  $x + \mathcal{N}_b$ .
    - If the state of  $x$  is 1, then nothing occurs to the particle at  $x$ .
    - If the state of  $x$  is 0, then  $x$  flips to 1 if both of its children are 1's.
  2. If  $U_n^x \in (\frac{\beta}{1+\beta}, 1)$ :
    - If the state of  $x$  is 0, then nothing occurs to the particle at  $x$ .
    - If the state of  $x$  is 1, then  $x$  flips to 0.
- For every unordered pair  $x, y \in \delta\epsilon\mathbb{Z}^d$  with  $\|x - y\|_1 = \delta\epsilon$  we assign a Poisson process  $\{T_n^{x,y}, n \geq 1\}$  with rate  $(\delta\epsilon)^{-2}/2$ . At an arrival of this Poisson process, the states of  $x$  and  $y$  are exchanged.

### 2.1.2 The dual is almost a branching random walk

For a particle at site  $x \in \delta\epsilon\mathbb{Z}^d$  at time  $t$ , we denote by  $\{\mathbf{X}_s^\epsilon\}_{0 \leq s \leq t}$  its dual process. The dual process is naturally defined only for  $0 \leq s \leq t$  but it is convenient to assume that the Poisson processes and uniform random variables in the construction are defined for negative times and hence define  $\mathbf{X}_s^\epsilon$  for all  $s \geq 0$ . We will focus on the case where  $d = 2$  in later discussion of the sexual reproduction model, but the comparison to a branching random walk in this section is general in all dimensions  $d \geq 1$ .

Let  $R_0^\epsilon = 0$  and let  $R_m^\epsilon$  be the  $m$ -th time that a branching event occurs among the particles in  $\mathbf{X}_s^\epsilon$  and set  $X_0^\epsilon(0) = x$  to represent the initial location of the first particle.

- In between the branching time  $\{R_m^\epsilon : m \geq 1\}$  the particles move by stirring. If there is a particle at  $x$  or  $y$  at time  $s$  and there is an arrival in  $T_n^{x,y}$  at time  $t - s$  then the particle at  $x$  jumps to  $y$  and a particle at  $y$  jumps to  $x$ .
- At time  $R_1^\epsilon$  if the branching occurs at  $x_1$  we uniformly choose a pair of neighbors  $x_1, x_2$  from

$$x + \mathcal{N}_b^\epsilon = x + \delta\epsilon \cdot \{\{e_1, e_2\}, \{-e_1, e_2\}, \{-e_1, -e_2\}, \{e_1, -e_2\}\},$$

add  $x_1$  and  $x_2$  to the dual, and number them as 1 and 2.

- At later branching times  $R_m^\epsilon$  if the branching occurs at  $x_m$  then we add a randomly chosen pair from  $x_m + \mathcal{N}_b^\epsilon$ , and number the two new particles  $2(m - 1) + 1$  and  $2(m - 1) + 2$ . A *collision* is said to happen if a particle is born at the location already occupied by another particle. In this case the colliding particle is not added to  $\mathbf{X}^\epsilon$ . We also construct a (noncoalescing) branching random walk  $\hat{\mathbf{X}}^\epsilon$  in which two particles are always added, and if there is a collision an independent graphical representation is used to determine its movements.

Notice that in the sexual reproduction model once a particle flips to state 0 its future is then independent of all its past. When constructing the dual process, once a particle flips to state 0 we don't necessarily need to probe into its past anymore. We can either treat this particle as "dead" and do not let it branch again since we don't need information about its past, or we can let it branch despite its state so that the resulting dual process has the structure of a regular tree. Here we take the second treatment.

In order to have the probability of collisions in the dual process  $\mathbf{X}_t^\epsilon$  to be small we have chosen  $\delta = \exp(-\epsilon^{-3})$ , i.e.,  $\epsilon = (\log(1/\delta))^{-1/3}$ , so that  $\delta \ll \epsilon$ . Intuitively, if the stirring rate is large enough compared to the branching rate, then particles do not stay near each other for a sufficiently long time to have a birth that causes collisions. To simplify notation, we will write

$$\eta = \delta\epsilon$$

from now on. In this notation, the dual process  $\mathbf{X}^\epsilon$  on  $\eta\mathbb{Z}^d$  jumps at rate  $2d \cdot \eta^{-2}/2$  to a randomly chosen neighbor.

**Lemma 2.1** *Let  $T \in (0, \infty)$ ,  $k \in \mathbb{N}$  and  $x \in \mathbb{R}^d$  be fixed. Let  $\mathbf{X}^\epsilon$  and  $\hat{\mathbf{X}}^\epsilon$  be defined as above and both start at  $x$ . There exists  $\epsilon^*(k, T) > 0$  so that for  $\epsilon \in (0, \epsilon^*(k, T))$ ,*

$$P_x^\epsilon \left( \mathbf{X}^\epsilon(t) = \hat{\mathbf{X}}^\epsilon(t) \text{ for all } t \leq T \right) \geq 1 - \epsilon^k.$$

**Proof** This proof comes from Durrett and Neuhauser [9]. To be self-contained we will present most of the details. We say a particle  $X_k^\epsilon$  is *crowded* at time  $s$  if for some  $j \neq k$ ,  $\|X_k^\epsilon(s) - X_j^\epsilon(s)\|_1 \leq \eta$ . To bound the number of collisions, we need to estimate the amount of time  $X_k^\epsilon$  is crowded. Let  $j \neq k$ ,  $V_s^\epsilon = X_k^\epsilon(s) - X_j^\epsilon(s)$  and  $W_s^\epsilon$  be a random walk that jumps to a randomly chosen neighbor at rate  $2d\eta^{-2}$ . Let  $x, y \in \eta\{\pm e_1, \dots, \pm e_d\}$ . Then

jumps from $x$ to	rate in $V$	rate in $W$
$-x$	$\eta^{-2}/2$	0
0	0	$\eta^{-2}$
$x + y$	$\eta^{-2}$	$\eta^{-2}$

Since we are interested in  $\|V_s^\epsilon\|_1$  we can ignore the first line in the above table, which does not change the norm. Then we can couple the jumps so that  $|\{s \leq t : \|V_s^\epsilon\|_1 \leq \eta\}|$  is stochastically smaller than  $w_t^\eta = |\{s \leq t : \|W_s^\epsilon\|_1 \leq \eta\}|$ . Asymptotic results for random walks imply, see (2.1) at p. 301 of [9], that when  $t\eta^{-2} \geq 2$ ,

$$E w_t^\eta \leq \begin{cases} C\eta^2, & d \geq 3, \\ C\eta^2 \log(\eta^{-2}), & d = 2, \\ C\eta t^{1/2}, & d = 1. \end{cases}$$

Let  $\chi_\epsilon^k(t)$  be the amount of time  $X_k^\epsilon$  is crowded in  $[0, t]$  and  $\mathcal{K}_t$  be the total number of particles at time  $t$ . It is easy to see

$$\begin{aligned} E(\chi_\epsilon^k(t) | \mathcal{K}_t = K) &\leq K E w_t^\eta, \\ E \mathcal{K}_t &= \exp(\nu t), \text{ where } \nu = 3c^* \epsilon^{-2}, \\ E(\chi_\epsilon^k(t)) &\leq \exp(\nu t) E w_t^\eta. \end{aligned}$$

To see that with high probability no collisions occur, we note that the expected number of births from  $X_k^\epsilon$  while there is some other  $X_j^\epsilon$  in its neighborhood is (consider the worst case when  $d = 1$ )

$$\leq E(\chi_\epsilon^k(t)) c^* \epsilon^{-2} \leq C_0 \eta t^{1/2} \epsilon^{-2} \exp(\nu t)$$

Take  $K = \eta^{-0.2}$ . Then  $P(\mathcal{K}_t > K) \leq K^{-1} \exp(\nu t) = \eta^{0.2} \exp(\nu t)$ . When  $\mathcal{K}_t \leq K$ , the expected number of collisions is smaller than

$$K C_0 \eta t^{1/2} \epsilon^{-2} \exp(\nu t).$$

Combining the error probabilities we have the probability of a collision occurring before time  $T$  is smaller than

$$\eta^{0.2} \exp(\nu T) + K C_0 \eta T^{1/2} \epsilon^{-2} \exp(\nu T) \quad (2.2)$$

Since  $\eta = \delta \epsilon = \epsilon \exp(-\epsilon^{-3})$  the above term vanishes as  $\epsilon \rightarrow 0$  and decays faster than any polynomial of  $\epsilon$ . Then for any given  $k \in \mathbb{N}$ , there exists some  $\epsilon^*(k, T)$  so that when  $\epsilon < \epsilon^*(k, T)$  the probability of collision (2.2) is less than  $\epsilon^k$ . When there is no collision between  $[0, T]$ ,  $\mathbf{X}^\epsilon(t) = \hat{\mathbf{X}}^\epsilon(t)$  for all  $0 \leq t \leq T$ .

### 2.1.3 Our random walks are close to Brownian motion

Let  $\hat{X}_t^\epsilon$  represent a single lineage in the comparison process  $\hat{\mathbf{X}}^\epsilon(t)$ . At each branch point we will choose one lineage of the offspring particles to be  $\hat{X}_t^\epsilon$  uniformly at random. We start by showing that the trajectory of a single lineage  $\hat{X}_t^\epsilon$  of the dual process is close to a Brownian motion  $W_t$  in  $\mathbb{R}^d$  when  $\epsilon$  is small. Recall that  $\hat{X}_t^\epsilon$  is a random walk that jumps at rate  $d\eta^{-2}$  to a randomly chosen neighbor.

**Lemma 2.2** *Let  $\hat{X}_t^\epsilon$  be a single lineage started at  $x$  and let  $k \in \mathbb{N}$ . There exists some  $\epsilon_0(k)$  and a coupling between the Brownian motion  $W_t$  in  $\mathbb{R}^d$  and  $\hat{X}_t^\epsilon$  so that for  $\epsilon \in (0, \epsilon_0(k))$*

$$P\left(|W_t - \hat{X}_t^\epsilon| \geq \epsilon \text{ for some } t \leq k\epsilon^2 |\log \epsilon|\right) \leq \epsilon^{2k}.$$

**Proof** Write  $\hat{X}_t^\epsilon = (\hat{X}_t^{1,\epsilon}, \dots, \hat{X}_t^{d,\epsilon})$  where for each  $1 \leq i \leq d$   $\hat{X}_t^{i,\epsilon}$  is a random walk on  $\eta\mathbb{Z}$  with rate  $\eta^{-2}$ . Let  $\{N_i(t) : 1 \leq i \leq d\}$  be independent Poisson processes

with rate  $\eta^{-2}$  and let  $Y_1^{(i)}, Y_2^{(i)}, \dots$  be i.i.d. random variables uniform on  $\{-\eta, \eta\}$ . For  $1 \leq i \leq d$ , define discrete time random walks  $S_n^{(i)} := Y_1^{(i)} + Y_2^{(i)} + \dots + Y_n^{(i)}$ . Then we can observe that  $S_{N_i(t)}^{(i)}$  has the same distribution as  $\hat{X}_t^{i,\epsilon}$ . Furthermore, if we let  $N(t) = \sum_{i=1}^d N_i(t)$  then  $S_{N(t)} := (S_{N_1(t)}^{(1)}, \dots, S_{N_d(t)}^{(d)})$  has then same distribution as  $\hat{X}_t^\epsilon$ .

From now on we consider the first coordinate  $\hat{X}_t^{1,\epsilon}$  and  $S_{N_1(t)}^{(1)}$  of the two random walks. Write  $x = (x_1, \dots, x_d)$ . By Skorohod's embedding theorem (see [17] Theorem 12.1), there is a Brownian motion  $B_t$  in  $\mathbb{R}$  started at  $x_1$  and a sequence of stopping times  $0 = \tau_0 \leq \tau_1 \leq \dots$  such that  $B(\tau_i) = S_i^{(1)}$ . Moreover, the differences  $\tau_i - \tau_{i-1}$  are i.i.d. with  $E(\tau_i - \tau_{i-1}) = E|Y_1^{(1)}|^2 = \eta^2$  and  $E(\tau_i - \tau_{i-1})^2 \leq 4E|Y_1^{(1)}|^4 = 4\eta^4$ .

Note that

$$\tau_{N_1(t)} - t = (\tau_{N_1(t)} - N_1(t)E\tau_1) + (N_1(t)E\tau_1 - t)$$

is a martingale, so  $L_2$ -maximal inequality implies

$$E \left( \max_{0 \leq s \leq t} |\tau_{N_1(s)} - s|^2 \right) \leq 4E(\tau_{N_1(t)} - t)^2 \leq 4E[N_1(t)] \cdot \text{Var}(\tau_1) \leq 16t\eta^2.$$

By Chebyshev's inequality,

$$P \left( \max_{0 \leq s \leq t} |\tau_{N_1(s)} - s| \geq \eta^{1/2} \right) \leq \eta^{-1} E \left( \max_{0 \leq s \leq t} |\tau_{N_1(s)} - \sigma^2 s|^2 \right) \leq 16t\eta. \quad (2.3)$$

Write  $W_t = (W_t^{(1)}, \dots, W_t^{(d)})$ . Since  $W_t^{(1)}$  is itself a one dimensional Brownian motion, without loss of generality we can let  $W_t^{(1)} = B_t$ . Notice that  $|\hat{X}_t^{1,\epsilon} - W_t^{(1)}| = |S_{N_1(t)}^{(1)} - B(t)| = |B(\tau_{N_1(t)}) - B(t)|$ . Then applying (2.3)

$$\begin{aligned} & P \left( |W_t^{(1)} - \hat{X}_t^{1,\epsilon}| \geq \epsilon/\sqrt{d} \text{ for some } t \leq k\epsilon^2 |\log \epsilon| \right) \\ &= P \left( \max_{0 \leq t \leq k\epsilon^2 |\log \epsilon|} |B_{\tau_{N_1(t)}} - B_t| \geq \epsilon/\sqrt{d} \right) \\ &\leq P \left( \max_{0 \leq t \leq k\epsilon^2 |\log \epsilon|} |\tau_{N_1(t)} - t| \geq \eta^{1/2} \right) \\ &\quad + P \left( \max_{0 \leq t \leq k\epsilon^2 |\log \epsilon|} |\tau_{N_1(t)} - t| < \eta^{1/2}, \max_{0 \leq t \leq k\epsilon^2 |\log \epsilon|} |B_{\tau_{N_1(t)}} - B_t| \geq \epsilon/\sqrt{d} \right) \\ &\leq 16k\epsilon^2 |\log \epsilon| \eta + P \left( N_1(k\epsilon^2 |\log \epsilon|) > \eta^{-2.5} k\epsilon^2 |\log \epsilon| \right) \\ &\quad + \eta^{-2.5} k\epsilon^2 |\log \epsilon| \cdot P \left( \sup_{s \in [-\eta^{1/2}, \eta^{1/2}]} |B(s) - B(0)| \geq \eta^{1/6} \right). \end{aligned}$$

The second term is upper bounded by  $\eta^{1/2}$  due to Markov inequality. To estimate the third term, let  $Z$  be a standard Gaussian variable. By reflection principle,

$$\begin{aligned} P\left(\sup_{s \in [-\eta^{1/2}, \eta^{1/2}]} |B(s) - B(0)| \geq \eta^{1/6}\right) &\leq 2P\left(\sup_{s \in [0, \eta^{1/2}]} |B(s) - B(0)| \geq \eta^{1/6}\right) \\ &\leq 4P(|B(\eta^{1/2}) - B(0)| \geq \eta^{1/6}) = 4P(\eta^{1/4}Z \geq \eta^{1/6}) \\ &\leq 4e^{-\eta^{-1/6}/4}. \end{aligned}$$

Thus

$$\begin{aligned} P\left(|W_t^{(1)} - \hat{X}_t^{1,\epsilon}| \geq \epsilon/\sqrt{d} \text{ for some } t \leq k\epsilon^2|\log \epsilon|\right) \\ \leq 16k\epsilon^2|\log \epsilon|\eta + \eta^{1/2} + 4\eta^{-2.5}k\epsilon^2|\log \epsilon|e^{-\eta^{-1/6}/4} \leq C\eta^{1/2} \end{aligned} \quad (2.4)$$

for some  $C > 0$ . Finally, it follows from (2.4) that

$$\begin{aligned} P\left(|W_t - \hat{X}_t^\epsilon| \geq \epsilon \text{ for some } t \leq k\epsilon^2|\log \epsilon|\right) \\ \leq d \cdot P\left(|W_t^{(1)} - \hat{X}_t^{1,\epsilon}| \geq \epsilon/\sqrt{d} \text{ for some } t \leq k\epsilon^2|\log \epsilon|\right) \\ \leq dC\eta^{1/2} \leq \epsilon^{2k}. \end{aligned}$$

□

## 2.2 Voter model perturbations

### 2.2.1 The dual is close to a branching random walk

The dual process  $\mathbf{X}^\epsilon$  is a coalescing branching random walk. As a result of the coalescence, the dual process does not have the tree structure that leads to independence among subtrees. The situation is not too bad once we realize coalescence mostly happens between particles with the same parent in a short amount of time after their births. Hence we will construct a comparison process  $\hat{\mathbf{X}}^\epsilon$  that has the desired tree structure.

Recall that the voting and branching neighborhoods are

$$\mathcal{N}_v^\epsilon = \{-\eta, \eta\}^d \quad \text{and} \quad \mathcal{N}_b^\epsilon = [-\eta L, \eta L]^d \cap \eta\mathbb{Z}^d$$

for a fixed  $L$ . Let  $J(t)$  denote the set of particles in  $\mathbf{X}^\epsilon$  at time  $t$ . If two particles  $i$  and  $j$  coalesce at time  $s$ , then  $i \vee j$  is removed from  $J(s-)$  to form  $J(s)$ . Set  $R_0 = 0$  and let  $R_m$  be the  $m$ -th branching time in  $\mathbf{X}^\epsilon$ . Similarly, define  $\hat{J}(t)$  and  $\hat{R}_m$  for the process  $\hat{\mathbf{X}}^\epsilon$ .

The comparison process  $\hat{\mathbf{X}}^\epsilon$  is constructed as follows:

- At time  $\hat{R}_m$ , the parent particle at  $x$  gives birth to  $N_0 = 4$  particles at sites  $(Y_1, \dots, Y_4)$  chosen uniformly without replacement from  $x + \mathcal{N}_b^\epsilon$ .

- During  $[\hat{R}_m, \hat{R}_m + \eta^{1/2})$  we do not allow birth events. The particles move as coalescing random walks in  $\eta\mathbb{Z}^d$  and we allow the particles within the new family (i.e. the parent particle and its  $N_0$  children) to coalesce with each other.
- During  $[\hat{R}_m + \eta^{1/2}, \hat{R}_{m+1})$  the particles move as random walks without coalescing and give births at rate  $\epsilon^{-2}$ .

If we view the interval  $[\hat{R}_m, \hat{R}_m + \eta^{1/2})$  as one single point in time then the process  $\hat{\mathbf{X}}^\epsilon$  would have the desired tree structure where each vertex has a random number of offspring depending on the coalescence. Note that

$$\hat{R}_{m+1} - \hat{R}_m \stackrel{d}{=} \sqrt{\eta} + \text{Exponential}(\epsilon^{-2} \hat{J}(\hat{R}_m + \sqrt{\eta})).$$

The graphical representations of  $\mathbf{X}^\epsilon$  and  $\hat{\mathbf{X}}^\epsilon$  can be coupled until there is a coalescence in  $\mathbf{X}^\epsilon$  that is not in  $\hat{\mathbf{X}}^\epsilon$ . Whenever this happens we use an independent graphical representation to determine the movement of the non-coalesced particle in  $\hat{\mathbf{X}}^\epsilon$ . We hope to couple  $\mathbf{X}^\epsilon$  and  $\hat{\mathbf{X}}^\epsilon$  in a way that the former is dominated by the latter. The obstacles in doing so are (i)  $\mathbf{X}^\epsilon$  can have births during intervals  $\{[\hat{R}_m, \hat{R}_m + \eta^{1/2}) : m \geq 1\}$  (ii) if the scenario in (i) does not happen, since  $\hat{\mathbf{X}}^\epsilon$  has more particles ever since the first coalescence in  $\mathbf{X}^\epsilon$  that is not in  $\hat{\mathbf{X}}^\epsilon$ , the branching times  $\hat{R}_m$  could arrive faster than  $R_m$ . As we will prove soon, both (i) and (ii) will not be the case with high probability. Our goal is to establish the following coupling between  $\mathbf{X}^\epsilon$  and  $\hat{\mathbf{X}}^\epsilon$  in such a way that the former is dominated by the latter.

**Lemma 2.3** *Let  $T \in (0, \infty)$ ,  $k \in \mathbb{N}$  and  $x \in \mathbb{R}^d$  be fixed. Let  $\mathbf{X}^\epsilon$  and  $\hat{\mathbf{X}}^\epsilon$  be defined as above and both start at  $x$ . There exists  $\epsilon^*(k, T) > 0$  so that for  $\epsilon \in (0, \epsilon^*(k, T))$ ,*

$$P_x^\epsilon \left( \mathbf{X}^\epsilon(t) = \hat{\mathbf{X}}^\epsilon(t) \text{ for all } t \leq T \right) \geq 1 - \epsilon^k.$$

**Proof** Let  $N_T = \min\{m : R_m > T\}$  and define the good events

$$G_1 = \{R_m - R_{m-1} > \sqrt{\eta} \text{ for all } 1 \leq m \leq N_T\}$$

$$G_2 = \{R_m = \hat{R}_m \text{ for all } 1 \leq m \leq N_T\}$$

$$G_3 = \{J(s) = J(R_{m-1} + \sqrt{\eta}) \text{ for all } s \in [R_{m-1} + \sqrt{\eta}, R_m) \text{ and all } 1 \leq m \leq N_T\}.$$

Observe that on  $G \equiv G_1 \cap G_2 \cap G_3$  we can couple  $\mathbf{X}^\epsilon$  and  $\hat{\mathbf{X}}^\epsilon$  exactly. Hence it suffices to upper bound  $P(G^c)$ . The estimates have already been done in detail in [3] so we will cite the relevant results instead of repeating the arguments.

Lemma 2.4 in [3] gives

$$P(G_1^c) = P\left(\min_{1 \leq m \leq N_T} R_m - R_{m-1} \leq \sqrt{\eta}\right) \leq \eta^{1/8}.$$

Let  $X_s^{\epsilon,j}$  denote the location of particle  $j$  in  $\mathbf{X}^\epsilon$  at time  $s$ . Define

$$\tau_m = \inf \left\{ s \geq R_{m-1} + \sqrt{\eta} : \inf_{i \neq j \in J(s)} : |X_s^{\epsilon,i} - X_s^{\epsilon,j}| \leq \eta^{7/8} \right\}$$

Lemma 2.7 in [3] gives

$$P(G_3^c) = P(\tau_m < R_m \text{ for some } 1 \leq m \leq N_T) \leq \eta^{1/16}.$$

The memoryless property of exponential random variables implies that

$$(R_{m+1} - R_m | G_1 \cap G_3) \stackrel{d}{=} \sqrt{\eta} + \text{Exponential}(J(R_m + \sqrt{\eta})).$$

We will argue by induction that  $G_1 \cap G_3 \subseteq G_2$ . First note  $R_0 = \hat{R}_0 = 0$ . Suppose  $R_m = \hat{R}_m$  holds up to  $m = k$  on  $G_1 \cap G_3$ . Then we should have  $J(R_k + \sqrt{\eta}) = \hat{J}(\hat{R}_k + \sqrt{\eta})$  on  $G_1 \cap G_3$ . This means

$$\begin{aligned} (R_{k+1} - R_k | G_1 \cap G_3) &\stackrel{d}{=} \sqrt{\eta} + \text{Exponential}(\hat{J}(\hat{R}_k + \sqrt{\eta})) \\ &\stackrel{d}{=} \hat{R}_{k+1} - \hat{R}_k \end{aligned}$$

Therefore  $R_{k+1} = \hat{R}_{k+1}$  on  $G_1 \cap G_3$  and this concludes the proof of  $G_1 \cap G_3 \subseteq G_2$ . Finally,

$$P(G^c) \leq P(G_1^c) + P(G_3^c) \leq \eta^{1/8} + \eta^{1/16} \leq \epsilon^k$$

for any  $k \in \mathbb{N}$  when  $\epsilon$  is sufficiently small.

## 2.2.2 Our random walks are almost Brownian motions

We will show the trajectory of a single lineage  $\hat{X}_t^\epsilon$  of the dual process is close to a Brownian motion  $W_t$  in  $\mathbb{R}^d$ . Note that  $\hat{X}_t^\epsilon$  is a random walk in  $\eta\mathbb{Z}^d$  that jumps at rate  $\eta^{-2}$  to a site chosen uniformly random from its neighborhood of the form  $\mathcal{N}_v^\epsilon = \{-\eta, \eta\}^d$ . The following lemma is essentially the same as Lemma 2.2. Note that the random walk in Lemma 2.2 has jump rate  $2d \cdot \eta^{-2}/2$  while here the random walk has jump rate  $\eta^{-2}$ , implying that  $\hat{X}_t^\epsilon$  would converge to a time-changed Brownian motion. The proof is essentially the same as that of Lemma 2.2 and hence is omitted.

**Lemma 2.4** *Let  $\hat{X}_t^\epsilon$  be a single lineage started at  $x$  and let  $k \in \mathbb{N}$ . There exists  $\epsilon_0(k)$  and a coupling of the Brownian motion  $W_t$  in  $\mathbb{R}^d$  and  $\hat{X}_t^\epsilon$  so that for  $\epsilon \in (0, \epsilon_0(k))$*

$$P\left(|W_{\sigma^2 t} - \hat{X}_t^\epsilon| \geq \epsilon \text{ for some } t \leq k\epsilon^2 | \log \epsilon|\right) \leq \epsilon^{2k}$$

for  $\sigma = 1/\sqrt{d}$ .

### 2.3 Computing the state of $x$ at time $t$

To do this, we use the comparison process  $\hat{\mathbf{X}}^\epsilon$  constructed in Sects. 2.1.1 and 2.2.1 and work backwards in time.  $\hat{\mathbf{X}}^\epsilon$  has a tree structure so we can follow [10] to define a *time-labelled tree*  $\mathcal{T}(\hat{\mathbf{X}}^\epsilon(t))$  for  $\hat{\mathbf{X}}^\epsilon$ . Since  $\hat{\mathbf{X}}^\epsilon$  and  $\hat{\mathbf{X}}$  has the same tree structure except for the rescaling, to simplify notation we consider  $\mathcal{T}(\hat{\mathbf{X}}(t))$  from now on.

Each branch point in  $\{\hat{\mathbf{X}}\}_{0 \leq s \leq t}$  is a vertex in the tree  $\mathcal{T}(\hat{\mathbf{X}}(t))$  and is assigned a time label  $t_v$  that is the corresponding branching time in  $\hat{\mathbf{X}}$ . For the sexual reproduction model, at each branch point the parent gives birth to  $N_0 = 2$  children, so the tree  $\mathcal{T}(\hat{\mathbf{X}}(t))$  branches into 3 lineages. For the voter model perturbations, at each branch point the parent gives birth to  $N_0 = 4$  children some of whom will coalesce into one. There are two ways to look at  $\mathcal{T}(\hat{\mathbf{X}}(t))$ : we can either see it as a Galton-Watson tree where the offspring distribution is determined by the coalescence, or we can still see it as a regular tree where each vertex has  $N_0 + 1$  children and deal with the influence of coalescence in a computing process that will be introduced later as an *algorithm*. Here we take the second approach.

Now we will describe an *algorithm* that computes the state of  $x$  at time  $t$  given the graphical representation and the initial states of the particles in  $\hat{\mathbf{X}}(t)$ . Since we are considering the dual process without rescaling, with a little abuse of notation let  $p : \mathbb{Z}^d \rightarrow [0, 1]$  be the initial condition.

*Algorithm for sexual reproduction model with fast stirring*

1. Each particle  $i$  in  $\mathcal{T}(\hat{\mathbf{X}}(t))$  is independently assigned state 1 with probability  $p(\hat{X}_t^i)$  and state 0 with probability  $1 - p(\hat{X}_t^i)$ .
2. At each branch point  $v$  in  $\mathcal{T}(\hat{\mathbf{X}}(t))$ , we have an independent random variable  $U_v$  uniform on  $(0, 1)$  that determines the state of the parent particle according to rules specified in Sect. 2.1.1.

*Algorithm for nonlinear voter models* Let  $\{\pi_v\}$  be a collection of i.i.d. random partition of the set  $\{0, 1, \dots, N_0\}$ , where  $v$  represents a vertex in the  $N_0 + 1$  regular time-labelled tree  $\mathcal{T}(\hat{\mathbf{X}}(t))$ . The law of  $\pi_v$  is given by the coalescence of particles within the same family within time  $\sqrt{\eta}$  after birth.

1. Each particle  $i$  in  $\mathcal{T}(\hat{\mathbf{X}}(t))$  is independently assigned state 1 with probability  $p(\hat{X}_t^i)$  and state 0 with probability  $1 - p(\hat{X}_t^i)$ .
2. At each branch point  $v$  in  $\mathcal{T}(\hat{\mathbf{X}}(t))$ , we first sample a random partition  $\pi_v$ . For vertices in the same cell of  $\pi_v$ , we uniformly choose one of them and let its state be the state of every vertex in that cell. Let  $i_1$  denote the total number of 1's among these  $N_0 + 1$  particles. Then an independent random variable  $U_v$  uniform on  $(0, 1)$  is sampled. If  $U_v < a_{i_1}$  then set the output to be 1, otherwise set the output to be 0.

For Lotka–Volterra systems, since the effective perturbations only occur when there is no coalescence among the three children, see (1.13), we can consider only such branch points and effectively reduce the branching rate to  $\theta p_3 \epsilon^{-2}$ . At each branch point, the state of the chosen lineage only flips when it is opposite to both of the other lineages. This is essentially performing a majority vote, which is why (1.13) is the same as (1.11). Hence the proof for Lotka–Volterra systems is the same as that in [10].

### Algorithm for Lotka–Volterra systems

1. Each particle  $i$  in  $\mathcal{T}(\hat{\mathbf{X}}(t))$  is independently assigned state 1 with probability  $p(\hat{X}_t^i)$  and state 0 with probability  $1 - p(\hat{X}_t^i)$ .
2. Let the branching event occur at rate  $\theta p_3 \epsilon^{-2}$ . At each branch point  $v$  in  $\mathcal{T}(\hat{\mathbf{X}}(t))$ , we perform a majority vote.

Starting from states of the leaves of  $\mathcal{T}(\hat{\mathbf{X}}(t))$ , the above algorithms compute the state of the root at  $x$ . From now on we use  $\mathbb{V}_p(\hat{\mathbf{X}}(t))$  to denote the output, i.e., the state of the root of  $\mathcal{T}(\hat{\mathbf{X}}(t))$ . Note that for a branching Brownian motion  $\mathbf{W}_t$  in  $\mathbb{R}^d$  we can define  $\mathbb{V}_p(\mathbf{W}_t)$  in the same way except that the initial condition  $p$  will be defined on  $\mathbb{R}^d$  instead of  $\mathbb{Z}^d$ .

## 3 Convergence to motion by mean curvature

Here we will prove the result assuming the  $g$  function has properties (G0)–(G5). In the next section we will check those conditions in our examples. A second consequence of concavity for  $p \in (u_0, u_+)$  is that if  $p \in [u_0 + \eta, u_+ - \eta]$

$$g(p + \eta) - 2g(p) + g(p - \eta) \leq 0. \quad (3.1)$$

To prove (3.1), we note that

$$\int_{p-\eta}^p \int_x^{x+\eta} g''(y) dy dx = g(p + \eta) - 2g(p) + g(p - \eta).$$

### 3.1 Branching Brownian motion in one dimension

Define the initial condition  $p_0 : \mathbb{R} \rightarrow [0, 1]$  to be  $p_0(x) = u_+ \cdot 1_{\{x \geq 0\}} + u_- \cdot 1_{\{x < 0\}}$  and write  $\mathbb{V} := \mathbb{V}_{p_0}$ . In this section we will consider one dimensional branching Brownian motion  $\mathbf{B}_t$ , beginning by listing the useful properties of  $\mathbb{V}(\mathbf{B}(t))$ .

**Monotonicity** When the interaction rule is attractive and the initial condition  $p_0$  is nondecreasing in  $x$  so for any  $x_1 \leq x_2 \in \mathbb{R}$ ,

$$P_{x_1}^\epsilon[\mathbb{V}(\mathbf{B}(t)) = 1] \leq P_{x_2}^\epsilon[\mathbb{V}(\mathbf{B}(t)) = 1].$$

**Antisymmetry** We use  $\mathcal{T}(\mathbf{B}(t))$  to denote the time-labelled tree for  $\mathbf{B}_t$  and write

$$P_x^t(\mathcal{T}) = P_x^\epsilon(\mathbb{V}(\mathbf{B}(t)) = 1 | \mathcal{T}(\mathbf{B}(t)) = \mathcal{T}).$$

Applying the reflection from  $z$  to  $-z$ , and using the symmetry of the Brownian motion conditioned on  $\{\mathcal{T}(\mathbf{B}(t)) = \mathcal{T}\}$ , we see that for any time-labelled tree  $\mathcal{T}$

$$P_z^t(\mathcal{T}) = 2u_0 - P_{-z}^t(\mathcal{T}).$$

The last property implies  $P_0^t(\mathcal{T}) = u_0$ . Using monotonicity we have

$$P_z^t(\mathcal{T}) \geq u_0 \quad \text{for } z \geq 0, \quad P_z^t(\mathcal{T}) \leq u_0 \quad \text{for } z \leq 0.$$

### 3.1.1 Useful inequalities

So far the function  $g : [0, 1] \rightarrow [0, 1]$  has a single variable. It is natural to extend  $g$  to be a function on  $[0, 1]^{N_0+1}$ . Let  $(p_1, \dots, p_{N_0+1}) \in [0, 1]^{N_0+1}$  and  $g(p_1, \dots, p_{N_0+1})$  is the probability that the output at the branch point is 1 when the inputs are independent Bernoulli random variables with rate  $p_1, \dots, p_{N_0+1}$  respectively. With a slight abuse of notation, we will use  $g(\cdot)$  to stand for both.

**Lemma 3.1** *For any time-labelled tree  $\mathcal{T}$ , and time  $t > 0$  and any  $z \geq 0$ ,*

$$P_z^t(\mathcal{T}) \geq u_+ P_z(B_t \geq 0) + u_- P_z(B_t < 0).$$

**Proof** The proof is by induction on the number of branching events in the tree  $\mathcal{T}$ . Suppose time  $\tau$  is the first branching event in  $\mathcal{T}$  and that the subtrees corresponding to the  $N_0 + 1$  offspring are  $\mathcal{T}_1, \dots, \mathcal{T}_{N_0+1}$ . Letting

$$P_z^t(\mathcal{T}*) = (P_z^t(\mathcal{T}_1), \dots, P_z^t(\mathcal{T}_{N_0+1})).$$

$$h(p_1, \dots, p_{N_0+1}) = g(p_1, \dots, p_{N_0+1}) - \frac{1}{N_0 + 1} (p_1 + \dots + p_{N_0+1}).$$

we can write

$$P_z^t(\mathcal{T}) = E_z(g(P_{B_\tau}^{t-\tau}(\mathcal{T}*)) = E_z(g(P_{B_\tau}^{t-\tau}(\mathcal{T}_1), \dots, P_{B_\tau}^{t-\tau}(\mathcal{T}_{N_0+1})))$$

$$= E_z(h(P_{B_\tau}^{t-\tau}(\mathcal{T}_1), \dots, P_{B_\tau}^{t-\tau}(\mathcal{T}_{N_0+1}))) + \frac{1}{N_0 + 1} \sum_{i=1}^{N_0+1} E_z(P_{B_\tau}^{t-\tau}(\mathcal{T}_i))$$

Write  $h(p) = h(p, \dots, p)$ . Observe that  $h(u_+ - p) = -h(u_- + p)$  due to (G1), which implies

$$h(P_{-z}^t(\mathcal{T}*)) = h(2u_0 - P_z^t(\mathcal{T}*)) = h(u_+ - (-u_- + P_z^t(\mathcal{T}**))) = -h(P_z^t(\mathcal{T}*)). \quad (3.2)$$

It follows that

$$E_z(h(P_{B_\tau}^{t-\tau}(\mathcal{T}*)) = E_z(h(P_{B_\tau}^{t-\tau}(\mathcal{T}*)) (1_{\{B_\tau \geq 0\}} + 1_{\{B_\tau < 0\}}))$$

$$= E_z(h(P_{B_\tau}^{t-\tau}(\mathcal{T}*)) 1_{\{B_\tau \geq 0\}}) - E_z(h(P_{-B_\tau}^{t-\tau}(\mathcal{T}*)) 1_{\{B_\tau < 0\}}) \quad (\text{by (3.2)})$$

$$= \int_0^\infty h(P_x^{t-\tau}(\mathcal{T}*)) (\phi_{z,\tau}(x) - \phi_{z,\tau}(-x)) dx$$

where  $\phi_{z,t}(x)$  is the probability density function of a Brownian motion starting at site  $z$  at time  $t$ . Since  $P_x^{t-\tau}(\mathcal{T}_i) \geq u_0$  for  $x \geq 0$  we have  $h(P_x^{t-\tau}(\mathcal{T}*)) \geq 0$ . Spatial

symmetry of Brownian motion and the fact that  $\phi_{z,t}(x)$  is decreasing on  $x \geq z$  implies  $\phi_{z,\tau}(x) - \phi_{z,\tau}(-x) \geq 0$  for all  $x \geq 0$ . That is,  $E_z(h(P_{B_\tau}^{t-\tau}(T_*))) \geq 0$ .

For  $i = 1, \dots, N_0 + 1$ , by the induction hypothesis

$$\begin{aligned} E_z(P_{B_\tau}^{t-\tau}(\mathcal{T}_i)) &\geq u_+ E_z(P_{B_\tau}(B_{t-\tau} \geq 0)) + u_- E_z(P_{B_\tau}(B_{t-\tau} < 0)) \\ &= u_+ P_z(B_t \geq 0) + u_- P_z(B_t < 0). \end{aligned}$$

It follows that

$$E_z(g(P_{B_\tau}^{t-\tau}(T_*))) \geq \frac{1}{N_0 + 1} \sum_{i=1}^{N_0+1} E_z(P_{B_\tau}^{t-\tau}(\mathcal{T}_i)) \geq u_+ P_z(B_t \geq 0) + u_- P_z(B_t < 0).$$

We define the iterates of  $g$ ,  $g^{(n)}(p)$ , by

$$g^{(n)}(p) = g(g^{(n-1)}(p)), \quad g^{(1)}(p) = g(p).$$

The fixed points at  $u_-$  and  $u_+$  of  $g$  are attracting and  $u_0$  is unstable. That is, if we start from  $u_0 + \epsilon$ , then iterating  $g$  will lead to  $u_+$  while if we start at  $u_0 - \epsilon$ , iterating  $g$  will take us down to  $u_-$ . Lemma 3.2 quantifies the rate of convergence.

**Lemma 3.2** *For all  $k \in \mathbb{N}$  there exists  $A(k) < \infty$  such that, for all  $\epsilon \in (0, u_+ - u_0 - \delta_*)$  where  $\delta_*$  is defined in (G4) and  $n \geq A(k) |\log \epsilon|$  we have*

$$g^{(n)}(u_0 + \epsilon) \geq u_+ - \epsilon^k \quad \text{and} \quad g^{(n)}(u_0 - \epsilon) \leq u_- - \epsilon^k.$$

**Proof** (G4) (i.e., (1.19)) implies that if  $\delta < \delta_*$  then  $u_+ - g(u_+ - \delta) \leq (1 - c_0)\delta$ . Iterating gives

$$u_+ - g^{(n)}(u_+ - \delta) \leq (1 - c_0)^n(u_+ - \delta).$$

That is, there is some constant  $C_k$  such that if  $\delta < \delta_*$  then for  $n \geq C_k |\log \epsilon|$  we have

$$g^{(n)}(u_+ - \delta) \geq u_+ - \epsilon^k.$$

It remains to find an  $M_\epsilon$ , which will depend on  $\epsilon$ , so that  $g^{(M_\epsilon)}(u_0 + \epsilon) \geq u_+ - \delta_*$ .

By (1.17) we know  $g'(u_0) > 1$ . Since  $u_0$  and  $u_+$  are two fixed points of  $g$  and  $g$  is strictly increasing, we have  $g(p) > p$  for  $p \in (u_0, u_+ - \delta_*]$ . It follows that

$$k_1 \equiv \inf_{x \in (0, u_+ - u_0 - \delta_*)} \frac{g(u_0 + x) - (u_0 + x)}{x} > 0$$

so for  $x \in [u_0 + \epsilon, u_+ - \delta_*]$  we have  $g(u_0 + x) - u_0 \geq (1 + k_1)x$ . Hence for  $m \in \mathbb{N}$  such that  $g^{(m)}(u_0 + \epsilon) < u_+ - \delta_*$  we have  $g^{(m)}(u_0 + \epsilon) \geq u_0 + (1 + k_1)^m \epsilon$ . This implies we can take  $M_\epsilon = B |\log \epsilon|$  where  $B = 1 / \log(1 + k_1)$ . Taking  $A(k) = B + C_k$  completes the proof.

Since the branching rate  $c^*\epsilon^{-2}$  is large when  $\epsilon$  is small, then even for a small  $t$  the tree  $\mathcal{T}(\mathbf{B}(t))$  should have a lot of vertices. For  $l \in \mathbb{R}$ , let  $\mathcal{T}_l^{reg}$  denote a ternary tree with depth  $\lceil l \rceil$ . For a time-labelled ternary tree  $\mathcal{T}$ , we write  $\mathcal{T} \supseteq \mathcal{T}_l^{reg}$  if  $\mathcal{T}_l^{reg}$  can be embedded in  $\mathcal{T}$  as a subtree. The next two results are Lemma 2.10 and 2.11 in [10]. The proofs are exactly the same so they are omitted.

**Lemma 3.3** *Let  $k \in \mathbb{N}$  and let  $A = A(k)$  be as in Lemma 3.2. Then there exists  $a_1 = a_1(k)$  and  $\epsilon_1 = \epsilon_1(k)$  such that, for all  $\epsilon \in (0, \epsilon_1)$  and  $t \geq a_1\epsilon^2|\log \epsilon|$ ,*

$$P^\epsilon \left[ \mathcal{T}(\mathbf{B}(t)) \supseteq \mathcal{T}_{A(k)|\log \epsilon|}^{reg} \right] \geq 1 - \epsilon^k.$$

**Lemma 3.4** *Let  $k \in \mathbb{N}$ , and let  $a_1(k)$  as in Lemma 3.3. Then there exists  $d_1(k), \epsilon_1(k)$  such that for all  $\epsilon \in (0, \epsilon_1(k))$  and all  $s \leq a_1\epsilon^2|\log \epsilon|$ ,*

$$P_x^\epsilon [\exists i \in N(s) : |B_i(s) - x| \geq d_1(k)\epsilon|\log \epsilon|] \leq \epsilon^k,$$

where  $N(s)$  is the set of indices of particles in  $\mathbf{B}$  up to time  $s$ .

While the proof of Lemma 3.2 is fresh on the reader's mind we will prove

**Lemma 3.5** *For a fixed  $k \in \mathbb{N}$ , there exists  $\sigma_1(k) > 0$  such that for  $t \geq \sigma_1(k)\epsilon^2|\log \epsilon|$  and  $x \in \mathbb{R}$*

$$P_x^\epsilon [\mathbb{V}_p(\mathbf{W}(t)) = 1] \leq u_+ + \epsilon^k$$

where  $p : \mathbb{R}^d \rightarrow [0, 1]$  is the initial condition satisfying (G0).

**Remark** The same conclusion also holds for  $P_x^\epsilon [\mathbb{V}_p(\hat{\mathbf{X}}^\epsilon(t)) = 1]$  following the same proof.

**Proof** First we consider the case where 1 is not a fixed point of  $g$ . Since  $u_+$  is a fixed point of  $g$  and  $g'(u_+) < 1$  by (G2), it is easy to see  $g(p) < p$  on  $(u_+, 1]$ . It follows that

$$k_2 \equiv \inf_{x \in (0, 1 - u_+]} \frac{(u_+ + x) - g(u_+ + x)}{x} \in (0, 1),$$

which implies that if  $\delta \in [0, 1 - u_+]$

$$g(u_+ + \delta) - u_+ \leq (1 - k_2)\delta.$$

Iterating as in the proof of Lemma 3.2

$$g^{(n)}(u_+ + \delta) - u_+ \leq (1 - k_2) \left( g^{(n-1)}(u_+ + \delta) - u_+ \right) \leq (1 - k_2)^n \delta.$$

By assumption (G0), since 1 is not a stable fixed point of  $g$  we get the largest value by setting  $p \equiv 1$ . In order to have  $g^{(n)}(1) \leq u_+ + \epsilon^k$  we need

$$g^{(n)}(u_+ + (1 - u_+)) - u_+ \leq (1 - k_2)^n(1 - u_+) \leq \epsilon^k.$$

It is easy to see that there exists  $C(k) > 0$  such that the above inequality holds for  $n \geq C(k)|\log \epsilon|$ . It follows from Lemma 3.3 that there exists  $\sigma_1(k) > 0$  such that for  $t \geq \sigma_1(k)\epsilon^2|\log \epsilon|$

$$P^\epsilon[T(\mathbf{W}(t)) \geq T_{C(k)|\log \epsilon|}^{\text{reg}}] \geq 1 - \epsilon^k.$$

Therefore, when  $t \geq \sigma_1\epsilon^2|\log \epsilon|$   $P_x^\epsilon[\mathbb{V}_p(\mathbf{W}(t)) = 1] \leq u_+ + \epsilon^k + \epsilon^k = u_+ + 2\epsilon^k$ .

The second case where 1 is a fixed point of  $g$  follows similarly. By assumption (G0) we can set  $p \equiv 1 - \varepsilon$  for some arbitrarily small  $\varepsilon > 0$ . Modify the definition of  $k_2$  to be

$$k_2(\varepsilon) \equiv \inf_{x \in (0, 1 - \varepsilon - u_+]} \frac{(u_+ + x) - g(u_+ + x)}{x} \in (0, 1).$$

In order to have  $g^{(n)}(1 - \varepsilon) \leq u_+ + \epsilon^k$  we need

$$g^{(n)}(u_+ + (1 - \varepsilon - u_+)) - u_+ \leq (1 - k_2)^n(1 - \varepsilon - u_+) \leq \epsilon^k.$$

The rest of the argument for the second case is the same.  $\square$

### 3.1.2 The main result in one dimension

We are now ready to prove

**Theorem 3.6** Fix any  $T^* \in (0, \infty)$ . For all  $k \in \mathbb{N}$  there exist  $c_1(k)$  and  $\epsilon_1(k) > 0$  such that, for all  $t \in [0, T^*]$  and all  $\epsilon \in (0, \epsilon_1)$ ,

1. for  $x \geq c_1(k)\epsilon|\log \epsilon|$ , we have  $P_x^\epsilon[\mathbb{V}(\mathbf{B}(t)) = 1] \geq u_+ - \epsilon^k$ ,
2. for  $x \leq -c_1(k)\epsilon|\log \epsilon|$ , we have  $P_x^\epsilon[\mathbb{V}(\mathbf{B}(t)) = 1] \leq u_- + \epsilon^k$ ,

**Proof of Theorem 3.6** For all  $\epsilon < 1/2$ , define  $z_\epsilon$  implicitly by the relation

$$P_0(B_{T^*} \geq -z_\epsilon) = 1/2 + (u_+ - u_-)^{-1}\epsilon \quad (3.3)$$

and note that  $z_\epsilon \sim (u_+ - u_-)^{-1}\epsilon\sqrt{2\pi T^*}$  as  $\epsilon \rightarrow 0$ . Let  $\epsilon_1(k) < 1/2$  be sufficiently small so that Lemmas 3.3 and 3.4 hold for  $\epsilon \in (0, \epsilon_1)$ . Let  $d_1(k)$  be given by Lemma 3.4 and let  $c_1(k) = 2d_1(k)$  so that, for  $\epsilon \in (0, \epsilon_1)$ ,

$$d_1(k)\epsilon|\log \epsilon| + z_\epsilon \leq c_1(k)\epsilon|\log \epsilon|.$$

Let  $a_1(k)$  be given by Lemma 3.3 and let  $\delta_1 = \delta_1(k, \epsilon) = a_1(k)\epsilon^2|\log \epsilon|$ .

Note that  $g(u_+) = u_+$ , which means if we start with initial condition  $p(x) \equiv u_+$  then

$$P_z^\epsilon(\mathbb{V}_p(\mathbf{B}(t)) = 1) = u_+ \quad \text{for all } t > 0, z \in \mathbb{R}. \quad (3.4)$$

If  $t \in (0, \delta_1)$  and  $z \geq c_1\epsilon|\log \epsilon|$ , then Lemma 3.4 and (3.4) gives

$$\begin{aligned} P_z^\epsilon(\mathbb{V}(\mathbf{B}(t)) = 0) &\leq P_z^\epsilon(\exists i \in N(t) \text{ such that } |B_i(t) - z| \geq d_1\epsilon|\log \epsilon|) + P_z^\epsilon(\mathbb{V}_p(\mathbf{B}(t)) = 0) \\ &\leq 1 - u_+ + \epsilon^k. \end{aligned}$$

We now suppose that  $t \in [\delta_1, T^*]$  and  $z \geq c_1\epsilon|\log \epsilon|$ , and define

$$p_{t-\delta_1}(z) = P_z^\epsilon(\mathbb{V}(\mathbf{B}(t - \delta_1)) = 1),$$

and let  $\psi^\epsilon \equiv p_{t-\delta_1}(z_\epsilon)$ . Write  $\{\mathbf{B}(\delta_1) > z_\epsilon\}$  for the event  $B_i(\delta_1) > z_\epsilon$  for all  $i \in N(\delta_1)$ . Then

$$\begin{aligned} P_z^\epsilon(\mathbb{V}(\mathbf{B}(t)) = 1) &= P_z^\epsilon(\mathbb{V}_{p_{t-\delta_1}}(\mathbf{B}(\delta_1)) = 1) \\ &\geq P_z^\epsilon(\{\mathbb{V}_{\psi^\epsilon}(\mathbf{B}(\delta_1)) = 1\} \cap \{\mathbf{B}(\delta_1) > z_\epsilon\}) \\ &\geq P_z^\epsilon(\mathbb{V}_{\psi^\epsilon}(\mathbf{B}(\delta_1)) = 1) - \epsilon^k \end{aligned}$$

By definition of  $z_\epsilon$  in (3.3) and  $t - \delta_1 < T^*$ ,

$$\begin{aligned} \psi^\epsilon &= P_{z_\epsilon}^\epsilon(\mathbb{V}(\mathbf{B}(t - \delta_1)) = 1) \geq u_+ P_{z_\epsilon}(B_{t-\delta_1} \geq 0) + u_- P_{z_\epsilon}(B_{t-\delta_1} < 0) \\ &= u_+(1/2 + (u_+ - u_-)^{-1}\epsilon) + u_-(1/2 - (u_+ - u_-)^{-1}\epsilon) = u_0 + \epsilon. \end{aligned}$$

It follows from Lemmas 3.2 and 3.3 that

$$\begin{aligned} P_z^\epsilon(\mathbb{V}_{\psi^\epsilon}(\mathbf{B}(\delta_1)) = 1) &\geq g^{(A(k)|\log \epsilon|)}(u_0 + \epsilon) P^\epsilon\left(\mathcal{T}(\mathbf{B}(t)) \supseteq \mathcal{T}_{A(k)|\log \epsilon|}^{reg}\right) \\ &\geq (u_+ - \epsilon^k)(1 - \epsilon^k) \geq u_+ - 2\epsilon^k \end{aligned}$$

Therefore,  $P_z^\epsilon(\mathbb{V}(\mathbf{B}(t)) = 1) \geq u_+ - 3\epsilon^k$ .  $\square$

### 3.1.3 Slope of the interface

To prepare the proof of Theorem 3.13, i.e., the extension of Theorem 3.6 to higher dimensions, we state the following result on the “slope” of the interface.

**Proposition 3.7** *Suppose  $x \geq 0$  and  $\eta > 0$ . Then for any time-labelled regular tree  $\mathcal{T}$  with  $N_0 + 1$  offspring and any time  $t$ ,*

$$P_x^t(\mathcal{T}) - P_{x-\eta}^t(\mathcal{T}) \geq P_{x+\eta}^t(\mathcal{T}) - P_x^t(\mathcal{T}).$$

**Proof** The proof is essentially the same as that of Proposition 2.11 in [10]. We prove the result by induction on the number of branching events in  $\mathcal{T}$ . We begin by noting that for a time-labelled tree  $\mathcal{T}_0$  with a root and a single leaf, we easily get

$$P_x^t(\mathcal{T}_0) - P_{x-\eta}^t(\mathcal{T}_0) = \int_{x-\eta}^x \phi_{0,t}(u) du \geq \int_x^{x+\eta} \phi_{0,t}(u) du = P_{x+\eta}^t(\mathcal{T}_0) - P_x^t(\mathcal{T}_0)$$

where  $\phi_{\mu,\sigma^2}$  is the density function of a  $N(\mu, \sigma^2)$  random variable. To do the induction step let  $\tau$  be the first branching time and let  $\mathcal{T}_1, \dots, \mathcal{T}_{N_0+1}$  be the trees of the offspring of that branching. We have

$$\begin{aligned} & (P_x^t(\mathcal{T}) - P_{x-\eta}^t(\mathcal{T})) - (P_{x+\eta}^t(\mathcal{T}) - P_x^t(\mathcal{T})) \\ &= \left( E_x[g(P_{B_\tau}^{t-\tau}(\mathcal{T}^*))] - E_{x-\eta}[g(P_{B_\tau}^{t-\tau}(\mathcal{T}^*))] \right) \\ & \quad - \left( E_{x+\eta}[g(P_{B_\tau}^{t-\tau}(\mathcal{T}^*))] - E_x[g(P_{B_\tau}^{t-\tau}(\mathcal{T}^*))] \right). \end{aligned}$$

If we let  $\rho(x) = g(P_x^{t-\tau}(\mathcal{T}^*))$  then the above is

$$\begin{aligned} & - \int_{-\infty}^{\infty} g(\rho(y+\eta)) - 2g(\rho(y)) + g(\rho(y-\eta)) \phi_{x,\tau}(y) dy \\ &= - \int_0^{\infty} g(\rho(y+\eta)) - 2g(\rho(y)) + g(\rho(y-\eta)) (\phi_{x,\tau}(y) - \phi_{x,\tau}(-y)) dy \end{aligned}$$

Since  $x \geq 0$ , we have  $\phi_{x,\tau}(y) - \phi_{x,\tau}(-y) \geq 0$  for  $y \geq 0$  so it is enough to show (3.1), i.e.,

$$g(\rho(y+\eta)) - 2g(\rho(y)) + g(\rho(y-\eta)) \leq 0.$$

By the induction assumption  $\rho(y) - \rho(y-\eta) \geq \rho(y+\eta) - \rho(y) \equiv h$ . Let  $p = \rho(y)$ .

$$g(\rho(y+\eta)) - g(\rho(y)) = g(p+h) - g(p) \leq g(p) - g(p-h) \leq g(\rho(y)) - g(\rho(y-\eta))$$

by monotonicity of  $g$ , which completes the proof.

Exploiting the “concavity” in Proposition 3.7 gives a lower bound on the “slope” of the interface.

**Corollary 3.8** Fix any  $T^* \in (0, \infty)$ . Suppose that for some  $t \in [0, T^*]$  and  $z \in \mathbb{R}$ ,

$$|P_z^\epsilon[\mathbb{V}(\mathbf{B}(t)) = 1] - u_0| \leq (u_+ - u_0) - \delta_0, \quad (3.5)$$

Take  $\epsilon_1(1)$  and  $c_1(1)$  from Theorem 3.6 and  $\epsilon < \min(\epsilon_1(1), \delta_0/2)$ , and let  $w \in \mathbb{R}$  with  $|z - w| \leq c_1(1)\epsilon |\log \epsilon|$ . Then

$$|P_z^\epsilon[\mathbb{V}(\mathbf{B}(t)) = 1] - P_w^\epsilon[\mathbb{V}(\mathbf{B}(t)) = 1]| \geq \frac{\delta_0 |z - w|}{4c_1(1)\epsilon |\log \epsilon|}.$$

**Proof** Consider first the case  $0 \leq z < w$ . By Theorem 3.6 and (3.5) we have for small  $\epsilon$

$$P_{c_1(1)\epsilon|\log\epsilon|}^\epsilon[\mathbb{V}(\mathbf{B}(t)) = 1] - P_z^\epsilon[\mathbb{V}(\mathbf{B}(t)) = 1] \geq \frac{\delta_0}{2}.$$

Write  $\eta = w - z$ . Proposition 3.7 implies that  $P_{(j+1)\eta+z}^t - P_{j\eta+z}^t \leq P_w^t - P_z^t$  for  $j \in \mathbb{N}$ . Let  $n_0 = \lceil \eta^{-1}(c_1(1)\epsilon|\log\epsilon| - z) \rceil$ . Then

$$P_{c_1(1)\epsilon|\log\epsilon|}^t - P_z^t \leq \sum_{j=0}^{n_0-1} P_{(j+1)\eta+z}^t - P_{j\eta+z}^t \leq n_0(P_w^t - P_z^t).$$

That is,

$$P_w^t - P_z^t \geq \frac{P_{c_1(1)\epsilon|\log\epsilon|}^t - P_z^t}{n_0} \geq \frac{\delta_0|z - w|}{2(c_1(1)\epsilon|\log\epsilon| + |z - w|)} \geq \frac{\delta_0|z - w|}{4c_1(1)\epsilon|\log\epsilon|}.$$

## 3.2 BBM in higher dimensions

### 3.2.1 Properties of motion by mean curvature

A key fact in the proof in Etheridge et al. [10] is a coupling between a one dimensional Brownian motion  $B_s$  and  $d(W_s, t - s)$ , the signed distance from a  $d$ -dimensional Brownian motion  $W_s$  to the interface  $\Gamma_{t-s}$ . To prepare for the coupling we will state some regularity properties of the mean curvature flow, which are given in Sect. 2.3 of [10] and are derived based on assumptions (C1)–(C3). Recall that  $d(x, t)$  is the signed distance from  $x$  to the mean curvature flow  $\Gamma_t$ .

1. There exists  $\kappa_0 > 0$  such that for all  $t \in [0, T^*]$  and  $x \in \{t : |d(y, t)| \leq \kappa_0\}$  we have

$$|\nabla d(x, t)| = 1. \quad (3.6)$$

Moreover,  $d$  is a  $C^{\alpha, \alpha/2}$  function in  $\{(x, t) : |d(x, t)| \leq \kappa_0, t \leq T^*\}$ , where  $\alpha > 3$  as in (C1).

2. Viewing  $\mathbf{n} = \nabla d$  as the positive normal direction, for  $x \in \Gamma_t$ , the normal velocity of  $\Gamma_t$  at  $x$  is  $-\partial_t d(x, t)$ , and the curvature of  $\Gamma_t$  at  $x$  is  $-\Delta d(x, t)$ .
3. There exists  $\kappa_0 > 0$  such that for all  $t \in [0, T^*]$  and  $x$  such that  $|d(x, t)| \leq \kappa_0$ ,

$$\left| \nabla(\partial_t d(x, t) - \Delta d(x, t)) \right| \leq \kappa_0. \quad (3.7)$$

4. There exists  $v_0, V_0 > 0$  such that for all  $t \in [T^* - v_0]$  and all  $s \in [t, t + v_0]$ ,

$$|d(x, t) - d(x, s)| \leq V_0(s - t). \quad (3.8)$$

We state Proposition 2.13 in [10]:

**Proposition 3.9** Let  $(W_s)_{s \geq 0}$  denote a  $d$ -dimensional Brownian motion started at  $x \in \mathbb{R}^d$ . Suppose that  $t \leq T^*$ ,  $\beta \leq \kappa_0$  and let

$$T_\beta = \inf(\{s \in [0, t) : |d(W_s, t - s)| \geq \beta\} \cup \{t\}).$$

Then we can couple  $(W_s)_{s \geq 0}$  with a one-dimensional Brownian motion  $(B_s)_{s \geq 0}$  started from  $z = d(x, t)$  in such a way that for  $s \leq T_\beta$ ,

$$B_s - \kappa_0 \beta s \leq d(W_s, t - s) \leq B_s + \kappa_0 \beta s.$$

By Lemma 2.1 we can establish the results for  $\hat{\mathbf{X}}^\epsilon$ , which will also hold for  $\mathbf{X}^\epsilon$  with high probability. Let  $W_t$  denote a Brownian motion in  $\mathbb{R}^d$  while  $\hat{X}_t^\epsilon$  denote a random walk on  $\eta\mathbb{Z}^d$  with jump rate  $\eta^{-2}/2$  to each neighboring site.

### 3.2.2 Generation of the interface

The following proposition is very similar to Proposition 2.15 in [10]. The major difference is that we work with the rescaled dual process  $\mathbf{X}_t^\epsilon$  and its comparison process  $\hat{\mathbf{X}}_t^\epsilon$  instead of the branching Brownian motion  $\mathbf{W}_t$  in  $\mathbb{R}^d$ .

**Proposition 3.10** Let  $k \in \mathbb{N}$  and  $\sigma_1(k)$  be defined as in Lemma 3.5. Then there exist  $\epsilon_d(k)$ ,  $a_d(k)$ ,  $b_d(k) > 0$  such that for all  $\epsilon \in (0, \epsilon_d)$ , if we set

$$\begin{aligned}\delta_d(k, \epsilon) &:= \max\{a_d(k), \sigma_1(k)\}\epsilon^2 |\log \epsilon| \\ \delta'_d(k, \epsilon) &:= (\max\{a_d(k), \sigma_1(k)\} + k + 1)\epsilon^2 |\log \epsilon|,\end{aligned}$$

then for  $t \in [\delta_d, \delta'_d]$ ,

1. for  $x$  such that  $d(x, t) \geq b_d \epsilon |\log \epsilon|$ , we have  $P_x^\epsilon(\mathbb{V}_p(\hat{\mathbf{X}}^\epsilon(t)) = 1) \geq u_+ - \epsilon^k$ ;
2. for  $x$  such that  $d(x, t) \leq -b_d \epsilon |\log \epsilon|$ , we have  $P_x^\epsilon(\mathbb{V}_p(\hat{\mathbf{X}}^\epsilon(t)) = 1) \leq u_- + \epsilon^k$ .

**Proof** For fixed  $k \in \mathbb{N}$  and  $A(k)$  specified as in Lemma 3.2, it follows from Lemma 3.3 that there exists  $a_d(k)$ ,  $\epsilon_d(k) > 0$  such that for all  $\epsilon \in (0, \epsilon_d)$  and  $t \geq a_d \epsilon^2 |\log \epsilon|$ .

$$P^\epsilon[\mathcal{T}(\hat{\mathbf{X}}^\epsilon(t)) \supseteq \mathcal{T}_{A(k)|\log \epsilon}^{\text{reg}}] \geq 1 - \epsilon^k.$$

It follows from the same argument as in Lemma 3.4 that for  $t \in [\delta_d, \delta'_d]$  there exists  $b'_d(k)$ ,  $\epsilon_d(k)$  such that for all  $\epsilon \in (0, \epsilon_d)$ ,

$$P_x^\epsilon[\exists i \in N(t) : |W_i(t) - x| \geq b'_d(k)\epsilon |\log \epsilon|] \leq \epsilon^k.$$

By (2.34) in [10] there exists  $v_0$ ,  $V_0 > 0$  such that for  $t \leq v_0$ , and any  $x \in \mathbb{R}^d$  we have  $|d(x, 0) - d(x, t)| \leq V_0 t$ . We can choose  $\epsilon_d$  sufficiently small so that  $\delta'_d \leq v_0$ .

Thus if  $d(x, t) \geq 2b'_d\epsilon|\log \epsilon|$  and  $|W_i(t) - x| \leq b'_d\epsilon|\log \epsilon|$  then

$$\begin{aligned} d(W_i(t), 0) &\geq d(x, t) - |d(x, t) - d(W_i(t), t)| - |d(W_i(t), t) - d(W_i(t), 0)| \\ &\geq 2b'_d\epsilon|\log \epsilon| - b'_d\epsilon|\log \epsilon| - V_0\delta'_d \geq \frac{2}{3}b'_d\epsilon|\log \epsilon|. \end{aligned}$$

It follows from Lemma 2.2 that

$$P(|W_i(t) - \hat{X}_i^\epsilon(t)| \geq \epsilon \text{ for some } t \leq \delta'_d) \leq \epsilon^{2k}.$$

The triangle inequality then implies that with probability at least  $1 - \epsilon^{2k}$

$$d(\hat{X}_i^\epsilon(t), 0) \geq d(W_i(t), 0) - |\hat{X}_i^\epsilon(t) - W_i| \geq \frac{2}{3}b'_d\epsilon|\log \epsilon| - \epsilon \geq \frac{1}{2}b'_d\epsilon|\log \epsilon|.$$

Applying (C2) and (C3),

$$p(\hat{X}_i^\epsilon(t)) \geq u_0 + \gamma\left(\frac{1}{2}b'_d\epsilon|\log \epsilon| \wedge r\right) \geq u_0 + \epsilon.$$

For  $x$  such that  $d(x, t) \geq 2b'_d\epsilon|\log \epsilon|$  and  $t \in [\delta_d, \delta'_d]$  it follows exactly from the proof of Theorem 3.6 that

$$P_x^\epsilon[\mathbb{V}_p(\hat{\mathbf{X}}^\epsilon(t)) = 1] \geq u_+ - 3\epsilon^k.$$

Taking  $b_d = 2b'_d$  completes the proof.

### 3.2.3 Propagation of the interface

In the Sect. 3.2.2 we established the existence of an interface develops for a short time interval  $[\delta_d, \delta'_d]$ . In this section we will show that the interface continue to exist for much longer. The key to proving Theorem 3.13 is the following proposition, which is an analogue of Proposition 2.17 in [10]. To make things easier to write we define  $\gamma(t) = K_1e^{K_2t}$  and introduce

$$z_0^\pm = d(x, t) \pm K_1e^{K_2t}\epsilon|\log \epsilon|$$

which are two points in  $\mathbb{R}$ . They depend on  $x$  and  $t$  but we do not record the dependence in notation.

**Proposition 3.11** *Let  $l \in \mathbb{N}$  with  $l \geq 4$ . Define  $\delta_d(l)$  as in Proposition 3.10 and  $C_1$  as in Lemma 3.12. There exists  $K_1(l), K_2(l) > 0$  and  $\epsilon_d(l, K_1, K_2) > 0$  so that for all  $\epsilon \in (0, \epsilon_d)$  and  $t \in [\delta_d(l), T^*]$  we have*

$$\sup_{x \in \mathbb{R}^d} \left( P_x^\epsilon[\mathbb{V}_p(\hat{\mathbf{X}}^\epsilon(t)) = 1] - P_{z_0^+}^\epsilon[\mathbb{V}(\mathbf{B}(t)) = 1] \right) \leq C_1\epsilon^l \quad (3.9)$$

$$\sup_{x \in \mathbb{R}^d} \left( P_x^\epsilon[\mathbb{V}_p(\hat{\mathbf{X}}^\epsilon(t)) = 0] - P_{z_0^-}^\epsilon[\mathbb{V}(\mathbf{B}(t)) = 0] \right) \leq C_1\epsilon^l \quad (3.10)$$

The key ingredient for proving Proposition 3.11 is the following lemma, which is an analogue of Lemma 2.18 in [10]. The idea of the proof remains the same but the coefficients are slightly different due to the differences in the  $g$ 's. Let  $B_t \in \mathbb{R}$  be a one-dimensional Brownian motion that can be thought of as a single lineage in the branching Brownian motion  $\mathbf{B}(t)$ .

$$\begin{aligned} z_1^\pm &= d(\hat{X}_s^\epsilon, t-s) \pm \gamma(t-s)\epsilon |\log \epsilon| \\ z_2^\pm &= B_s \pm \gamma(t)\epsilon |\log \epsilon| \end{aligned}$$

**Lemma 3.12** *Let  $l \in \mathbb{N}$  with  $l \geq 4$  and  $\sigma_1(l)$  be as in Lemma 3.5. Let  $\delta_0$  and  $c_0$  be chosen as in (G4). Choose  $C_1$  sufficiently large so that  $C_1 > \max\{2(1-c_0)/c_0, 3/(2c_0)\}$ . Let  $C_2 = \max_{0 \leq p \leq 1} C_1 |g'(p)|$ . Let  $K_1 > 0$ . There exists  $K_2 = K_2(K_1, l) > 0$  and  $\epsilon_d(l, K_1, K_2) > 0$  such that for all  $\epsilon \in (0, \epsilon_d)$ ,  $x \in \mathbb{R}^d$ ,  $s \in [0, (l+1)\epsilon^2 |\log \epsilon|]$  and  $t \in [s, T^*]$ ,*

$$\begin{aligned} E_x[g(P_{z_1^+}^\epsilon[\mathbb{V}(\mathbf{B}(t-s)) = 1] + C_1\epsilon^l)] \\ \leq (1-c_0/3)C_1\epsilon^l + E_{d(x,t)}[g(P_{z_2^+}^\epsilon[\mathbb{V}(\mathbf{B}(t-s)) = 1])] + C_2\epsilon^l 1_{s \leq \epsilon^4} \end{aligned} \quad (3.11)$$

$$\begin{aligned} E_x[g(P_{z_1^-}^\epsilon[\mathbb{V}(\mathbf{B}(t-s)) = 0] + C_1\epsilon^l)] \\ \leq (1-c_0/3)C_1\epsilon^l + E_{d(x,t)}[g(P_{z_2^-}^\epsilon[\mathbb{V}(\mathbf{B}(t-s)) = 0])] + C_2\epsilon^l 1_{s \leq \epsilon^4} \end{aligned} \quad (3.12)$$

To keep our approach parallel to the one in [10] we defer the proof of Lemma 3.12 to the next subsection. The only property of  $g$  that is used in the proof below is its monotonicity.

**Proof of Proposition 3.11** We begin by proving (3.9) for  $t \in [\delta_d, \delta'_d]$ . Take  $K_1 = b_d(l) + c_1(l)$  where  $b_d(l)$  is as defined in Proposition 3.10 and  $c_1$  is as defined in Theorem 3.6. Let  $K_2 = K_2(K_1, l)$ , as defined in Lemma 3.12. If  $d(x, t) \leq -b_d(l)\epsilon |\log \epsilon|$ , then by Proposition 3.10,  $P_x^\epsilon[\mathbb{V}_p(\hat{\mathbf{X}}^\epsilon(t)) = 1] \leq \epsilon^l$ . Then (3.9) holds.

On the other hand, if  $d(x, t) \geq -b_d(l)\epsilon |\log \epsilon|$ , then  $d(x, t) + \gamma(t)\epsilon |\log \epsilon| \geq c_1(l)\epsilon |\log \epsilon|$ , and by Theorem 3.6

$$P_{d(x,t)+\gamma(t)\epsilon |\log \epsilon|}^\epsilon[\mathbb{V}(\mathbf{B}(t)) = 1] \geq u_+ - \epsilon^l.$$

By definition of  $\delta_d$  in Proposition 3.10,  $t \geq \sigma_1(l)\epsilon |\log \epsilon|$ . It follows from the same argument as in Lemma 3.5 that

$$P_x^\epsilon[\mathbb{V}_p(\hat{\mathbf{X}}^\epsilon(t)) = 1] \leq u_+ + \epsilon^l.$$

Therefore when  $\epsilon$  is sufficiently small (3.9) holds.

We follow the proof in [10] and assume that there exists  $t \in [\delta'_d, T^*]$  such that for some  $x \in \mathbb{R}^d$  (3.9) does not hold, i.e.,

$$P_x^\epsilon[\mathbb{V}_p(\hat{\mathbf{X}}^\epsilon(t)) = 1] - P_{d(x,t)+\gamma(t)\epsilon |\log \epsilon|}^\epsilon[\mathbb{V}(\mathbf{B}(t)) = 1] > C_1\epsilon^l.$$

Let  $T'$  be the infimum of the set of such  $t$  and choose

$$T \in [T', \min(T' + \epsilon^{l+3}, T^*)]$$

which is in the set of such  $t$ . Hence there exists some  $x \in \mathbb{R}^d$  such that

$$P_x^\epsilon[\mathbb{V}_p(\hat{\mathbf{X}}^\epsilon(T)) = 1] - P_{d(x,T)+\gamma(T)\epsilon|\log\epsilon|}^\epsilon[\mathbb{V}(\mathbf{B}(T)) = 1] > C_1\epsilon^l. \quad (3.13)$$

Our goal is to contradict (3.13) by showing that

$$P_x^\epsilon[\mathbb{V}_p(\hat{\mathbf{X}}^\epsilon(T)) = 1] \leq P_{d(x,T)+\gamma(T)\epsilon|\log\epsilon|}^\epsilon[\mathbb{V}(\mathbf{B}(T)) = 1] + (1 - c_0/4)C_1\epsilon^l. \quad (3.14)$$

We write  $S$  for the time of the first branching event in  $\hat{\mathbf{X}}^\epsilon(T)$  and  $\hat{X}^\epsilon(S)$  for the position of the initial particle at that time. By the strong Markov property

$$\begin{aligned} P_x^\epsilon[\mathbb{V}_p(\hat{\mathbf{X}}^\epsilon(T)) = 1] &\leq E_x^\epsilon[g(P_{\hat{X}^\epsilon(S)}^\epsilon[\mathbb{V}_p(\hat{\mathbf{X}}^\epsilon(T - S)) = 1]1_{S \leq T - \delta_d}] \\ &\quad + E_x^\epsilon[P_{\hat{X}^\epsilon(T - \delta_d)}^\epsilon[\mathbb{V}_p(\hat{\mathbf{X}}^\epsilon(\delta_d)) = 1]1_{S \geq T - \delta_d}] \end{aligned} \quad (3.15)$$

Let  $c^*$  be a constant such that  $c^*\epsilon^{-2}$  is the reaction rate for the process that we consider. For sexual reproduction model with fast stirring,  $c^* = (1 + \beta)$  as defined in (2.1). For voter model perturbations,  $c^* = 1$ . Without loss of generality we can assume  $c^* \geq 1$  since otherwise we can rescale time to obtain  $c^* \geq 1$ . Since  $S = \text{Exponential}(c^*\epsilon^{-2})$  and  $T - \delta_d \geq \delta'_d - \delta_d = (l + 1)\epsilon^2|\log\epsilon|$ , we have

$$E_x^\epsilon \left[ P_{\hat{X}^\epsilon(T - \delta_d)}^\epsilon[\mathbb{V}_p(\hat{\mathbf{X}}^\epsilon(\delta_d)) = 1]1_{\{S \geq T - \delta_d\}} \right] \leq P[S \geq (l + 1)\epsilon^2|\log\epsilon|] \leq \epsilon^{c^*(l+1)} \leq \epsilon^{l+1}.$$

To bound the first term in (3.15), partition on the event  $\{S \leq \epsilon^{l+3}\}$ ,

$$\begin{aligned} E_x^\epsilon[g(P_{\hat{X}^\epsilon(S)}^\epsilon[\mathbb{V}_p(\hat{\mathbf{X}}^\epsilon(T - S)) = 1]1_{\{S \leq T - \delta_d\}}] \\ \leq P[S \leq \epsilon^{l+3}] + E_x^\epsilon[g(P_{\hat{X}^\epsilon(S)}^\epsilon[\mathbb{V}_p(\hat{\mathbf{X}}^\epsilon(T - S)) = 1]1_{\{\epsilon^{l+3} \leq S \leq T - \delta_d\}}] \\ \leq \epsilon^{l+1} + E_x^\epsilon[g(P_{d(\hat{X}^\epsilon_S, T - S) + \gamma(T - S)\epsilon|\log\epsilon|}^\epsilon[\mathbb{V}(\mathbf{B}(T - S)) = 1] + C_1\epsilon^l)1_{\{S \leq T - \delta_d\}}]. \end{aligned} \quad (3.16)$$

The last line follows from the minimality of  $T'$  and the fact that  $T - S \leq T'$  on the event  $\{S \geq \epsilon^{l+3}\}$ .

$$\begin{aligned} E_x^\epsilon[g(P_{d(\hat{X}^\epsilon_S, T - S) + \gamma(T - S)\epsilon|\log\epsilon|}^\epsilon[\mathbb{V}(\mathbf{B}(T - S)) = 1] + C_1\epsilon^l)1_{S \leq T - \delta_d}] \\ \leq \int_0^{(l+1)\epsilon^2|\log\epsilon|} c^*\epsilon^{-2}e^{-c^*\epsilon^{-2}s} E_x[g(P_{d(\hat{X}^\epsilon_S, T - s) + K_1e^{K_2(T - s)}\epsilon|\log\epsilon|}^\epsilon[\mathbb{V}(\mathbf{B}(T - s)) = 1] + C_1\epsilon^l)] ds \\ + P[S \geq (l + 1)\epsilon^2|\log\epsilon|]. \end{aligned}$$

Using Lemma 3.12 we get

$$\leq (1 - c_0/3)C_1\epsilon^l + \int_0^{(l+1)\epsilon^2|\log \epsilon|} c^* \epsilon^{-2} e^{-c^* \epsilon^{-2}s} E_{d(x,t)}[g(P_{B_s+\gamma(t)\epsilon|\log \epsilon}^\epsilon [\mathbb{V}(\mathbf{B}(t-s)) = 1])] ds \\ + C_2 \epsilon^l P[S \leq \epsilon^4] + \epsilon^{l+1}.$$

Let  $S'$  denotes the first branching time in  $(\mathbf{B}(s))_{s \geq 0}$  and  $B_{S'}$  the position of the ancestor at that time. Noting that  $S'$  has the same distribution as  $S$  we have

$$\leq (1 - c_0/3)C_1\epsilon^l + 2\epsilon^{l+1} + E_{d(x,t)}^\epsilon [g(P_{B_{S'}+K_1 e^{K_2 T} \epsilon|\log \epsilon}^\epsilon [\mathbb{V}(\mathbf{B}(T-S')) = 1]) 1_{S' \leq T-\delta'_d}]. \quad (3.17)$$

Combining (3.15), (3.16) and (3.17),

$$P_x^\epsilon [\mathbb{V}_p(\hat{X}^\epsilon(T)) = 1] \leq 4\epsilon^{l+1} + (1 - c_0/3)C_1\epsilon^l + E_{d(x,t)}^\epsilon [g(P_{B_{S'}+K_1 e^{K_2 T} \epsilon|\log \epsilon}^\epsilon [\mathbb{V}(\mathbf{B}(T-S')) = 1])] \\ \leq (1 - c_0/4)C_1\epsilon^l + P_{d(x,T)+K_1 e^{K_2 T} \epsilon|\log \epsilon}^\epsilon [\mathbb{V}(\mathbf{B}(T)) = 1],$$

which proves (3.14) and hence we have proved (3.9) by an argument of contradiction. The proof of (3.10) is similar.  $\square$

Before giving the proof of Lemma 3.12 we prove the main result.

**Theorem 3.13** *Let  $u^\epsilon(t, x) = P(\xi_t^\epsilon(x) = 1)$  with  $u^\epsilon(0, x) = p(x)$ . Let  $T^* \in (0, \mathcal{T})$  and  $k \in \mathbb{N}$  be fixed. Choose  $\sigma_1(k)$  as in Lemma 3.5. There exist  $\epsilon_d(k) > 0$  and  $a_d(k), c_d(k) \in (0, \infty)$  such that for all  $\epsilon \in (0, \epsilon_d)$  and  $t$  satisfying  $\max\{a_d, \sigma_1\}\epsilon^2|\log \epsilon| \leq t \leq T^*$ ,*

1. *for  $x$  such that  $d(x, t) \geq c_d(k)\epsilon|\log \epsilon|$ , we have  $u^\epsilon(t, x) \geq u_+ - \epsilon^k$ ,*
2. *for  $x$  such that  $d(x, t) \leq -c_d(k)\epsilon|\log \epsilon|$ , we have  $u^\epsilon(t, x) \leq u_- + \epsilon^k$ .*

**Proof** We first prove the result for  $\hat{\mathbf{X}}^\epsilon(t)$ . We choose  $c_d(k) = c_1(k) + K_1 e^{K_2 T^*}$ . Thus for  $t \in [\delta_d, T^*]$  and  $x \in \mathbb{R}^d$  such that  $d(x, t) \leq -c_d(k)\epsilon|\log \epsilon|$  we have

$$d(x, t) + K_1 e^{K_2 T^*} \leq -c_1(k)\epsilon|\log \epsilon|.$$

It follows from Proposition 3.11 and Theorem 3.6 that  $P_x^\epsilon[\mathbb{V}_p(\hat{\mathbf{X}}^\epsilon(t)) = 1] \leq u_- + (C_1 + 1)\epsilon^k$ . Similarly, if  $d(x, t) \geq c_d(k)\epsilon|\log \epsilon|$  then  $d(x, t) - K_1 e^{K_2 T^*} \geq c_1(k)\epsilon|\log \epsilon|$ . Hence

$$P_x^\epsilon[\mathbb{V}_p(\hat{\mathbf{X}}^\epsilon(t)) = 0] \leq P_{d(x,t)-\gamma(t)\epsilon|\log \epsilon}^\epsilon[\mathbb{V}(\mathbf{B}(t)) = 0] + C_1\epsilon^k \leq 1 - u_+ + (1 + C_1)\epsilon^k.$$

It remains to show  $u^\epsilon(t, x)$  is close to  $P_x^\epsilon[\mathbb{V}_p(\hat{\mathbf{X}}^\epsilon(t)) = 1]$ . Let  $G = \{\mathbf{X}^\epsilon(t) = \hat{\mathbf{X}}^\epsilon(t) \text{ for } t \leq T^*\}$ . Lemma 2.1 implies that  $P(G) \geq 1 - \epsilon^k$ .

Then

$$\begin{aligned} u^\epsilon(t, x) &= P_x^\epsilon[\mathbb{V}_p(\mathbf{X}(t)) = 1] \\ &= P_x^\epsilon[\{\mathbb{V}_p(\hat{\mathbf{X}}^\epsilon(t)) = 1\} \cap G] + P_x^\epsilon[\{\mathbb{V}_p(\mathbf{X}(t)) = 1\} \cap G^c] \\ &\leq P_x^\epsilon[\mathbb{V}_p(\hat{\mathbf{X}}^\epsilon(t)) = 1] + \epsilon^k \end{aligned}$$

On the other hand,

$$u^\epsilon(t, x) \geq P_x^\epsilon[\{\mathbb{V}_p(\hat{\mathbf{X}}^\epsilon(t)) = 1\} \cap G] \geq P_x^\epsilon[\mathbb{V}_p(\hat{\mathbf{X}}^\epsilon(t)) = 1] - P[G^c] \geq P_x^\epsilon[\mathbb{V}_p(\hat{\mathbf{X}}^\epsilon(t)) = 1] - \epsilon^k.$$

$$\text{Therefore, } |u^\epsilon(t, x) - P_x^\epsilon[\mathbb{V}_p(\hat{\mathbf{X}}^\epsilon(t)) = 1]| \leq \epsilon^k.$$

### 3.2.4 Proof of Lemma 3.12

**Proof** We continue to write  $\gamma(t) = K_1 e^{K_2 t}$ . Define a good event by

$$G = \{|d(W_s, t-s) - d(\hat{X}_s^\epsilon, t-s)| \leq \epsilon \text{ for } s \in [0, (l+1)\epsilon^2 \lfloor \log \epsilon \rfloor]\}.$$

The triangle inequality implies  $d(W_s, t-s) \leq d(\hat{X}_s^\epsilon, t-s) + |\hat{X}_s^\epsilon - W_s|$ . There is a similar result with  $W$  and  $X$  interchanged so

$$|d(W_s, t-s) - d(\hat{X}_s^\epsilon, t-s)| \leq |\hat{X}_s^\epsilon - W_s| \quad (3.18)$$

Lemma 2.2 implies that for sufficiently small  $\epsilon$

$$P(G) \geq 1 - \epsilon^{2l}. \quad (3.19)$$

We choose  $\kappa_0$  as in (3.7) and  $c_1(k)$  from Theorem 3.6. Let

$$R = 2c_1(l) + 4(l+1)d + 1 \quad (3.20)$$

and fix  $K_2$  such that

$$(K_1 + 1)(K_2 - \kappa_0) - \kappa_0 R = c_1(1). \quad (3.21)$$

Let  $s \in [0, (l+1)\epsilon^2 \lfloor \log \epsilon \rfloor]$  and

$$A_x = \left\{ \sup_{u \in [0, s]} |W_u - x| \leq 2(l+1)d\epsilon \lfloor \log \epsilon \rfloor \right\}.$$

Using the reflection principle

$$\begin{aligned} P(A_x^c) &\leq 2dP_0 \left( \sup_{u \in [0, s]} B_u > 2(l+1)\epsilon \lfloor \log \epsilon \rfloor \right) \\ &\leq 4dP_0(B_s > 2(l+1)\epsilon \lfloor \log \epsilon \rfloor) \leq 4d\epsilon^{l+1} \end{aligned} \quad (3.22)$$

where we have used the tail bound the tail bound

$$P(B_s \geq x\sqrt{s}) \leq \exp(-x^2/4)$$

with  $s = (l+1)\epsilon^2|\log \epsilon|$  and  $x = 2\sqrt{(\ell+1)|\log \epsilon|}$ .

Recall that in Lemma 3.12  $s \in [0, (\ell+1)\epsilon^2|\log \epsilon|]$  is fixed and  $t \in [s, T^*]$ . We consider three cases:

1.  $d(x, t) \leq -(2c_1(l) + 2(l+1)d + \gamma(t-s))\epsilon|\log \epsilon|$ ,
2.  $d(x, t) \geq (2c_1(l) + 2(l+1)d + \gamma(t-s))\epsilon|\log \epsilon|$ ,
3.  $|d(x, t)| \leq (2c_1(l) + 2(l+1)d + \gamma(t-s))\epsilon|\log \epsilon|$ .

The first two are easy since  $x$  is far from the interface so the probabilities of interest are either close to  $u_+$  or close to  $u_-$ .

Case 1 By (3.8) there exists  $v_0, V_0 > 0$  such that if  $s \leq v_0$  and  $x \in \mathbb{R}^d$  then

$$|d(x, t) - d(x, t-s)| \leq V_0 s. \quad (3.23)$$

We take  $\epsilon_d$  sufficiently small in Lemma 3.12 so that  $(l+1)\epsilon^2|\log \epsilon| \leq v_0$  for all  $\epsilon \in (0, \epsilon_d)$ . Rearranging the definition of Case 1 and adding  $d(W_s, t-s)$  to both sides

$$\begin{aligned} d(W_s, t-s) + \gamma(t-s)\epsilon|\log \epsilon| &\leq -(2c_1(l) + 2(l+1)d) \\ \epsilon|\log \epsilon| + d(W_s, t-s) - d(x, t) &\end{aligned}$$

The triangle inequality implies  $d(x, t-s) + |W_s - x| \geq d(W_s, t-s)$  so

$$\begin{aligned} d(W_s, t-s) + \gamma(t-s)\epsilon|\log \epsilon| &\leq -(2c_1(l) + 2(l+1)d) \\ \epsilon|\log \epsilon| + |W_s - x| + |d(x, t) - d(x, t-s)| &\end{aligned}$$

Using (3.23) with  $s \leq (l+1)\epsilon^2|\log \epsilon|$  we see that on  $A_x$

$$d(W_s, t-s) + \gamma(t-s)\epsilon|\log \epsilon| \leq -2c_1(l)\epsilon|\log \epsilon| + V_0(l+1)\epsilon^2|\log \epsilon|.$$

On event  $G \cap A_x$  when  $\epsilon$  is sufficiently small,

$$\begin{aligned} z_1^+ = d(\hat{X}_s^\epsilon, t-s) + \gamma(t-s)\epsilon|\log \epsilon| &\leq d(W_s, t-s) + \epsilon + \gamma(t-s)\epsilon|\log \epsilon| \\ &\leq -c_1(l)\epsilon|\log \epsilon|. \end{aligned}$$

Hence it follows from Theorem 3.6 that

$$E_x[g(P_{z_1^+}^\epsilon[\mathbb{V}(\mathbf{B}(t-s)) = 1] + C_1\epsilon^l)] \leq E_x[g(u_- + (1+C_1)\epsilon^l)] + P_x[A_x^c] + P[G^c].$$

Using (G4), (3.22), and (3.19) the above is

$$\leq u_- + (1-c_0) \cdot (1+C_1)\epsilon^l + 4d\epsilon^{l+1} + \epsilon^{2l} \leq u_- + (1-c_0/3)C_1\epsilon^l$$

when  $\epsilon$  is sufficiently small. As  $u_-$  is a fixed point of  $g$  and we start with initial condition  $p_0(x) = u_+ \cdot 1_{\{x \geq 0\}} + u_- \cdot 1_{\{x < 0\}}$  for the one dimensional BBM, the second term on the right hand side of (3.11) satisfies

$$E_{d(x,t)}[g(P_{z_2^+}^\epsilon[\mathbb{V}(\mathbf{B}(t-s)) = 1])] \geq u_-.$$

The third term on the right hand side of (3.11) is non-negative so the result follows.

*Case 2* In this case  $d(x, t) \geq (c_1(l) + 2(l+1))\epsilon |\log \epsilon|$ . Repeating the proof of (3.22) gives

$$P_{d(x,t)}[B_s \leq c_1(l)\epsilon |\log \epsilon|] \leq P_0[B_s \geq 2(l+1)\epsilon |\log \epsilon|] \leq \epsilon^{l+1} \quad (3.24)$$

Recall  $z_2^+ = B_s + \gamma(t)\epsilon |\log \epsilon|$ . Using Theorem 3.6 and (3.24) and (1.19) it follows that

$$\begin{aligned} E_{d(x,t)}[g(P_{z_2^+}^\epsilon[\mathbb{V}(\mathbf{B}(t-s)) = 1])] \\ \geq E_{d(x,t)}[g(P_{z_2^+}^\epsilon[\mathbb{V}(\mathbf{B}(t-s)) = 1])1_{\{B_s \geq c_1(l)\epsilon |\log \epsilon|\}}] \\ \geq g(u_+ - \epsilon^l) - \epsilon^{l+1} \geq u_+ - (1 - c_0)\epsilon^l - \epsilon^{l+1} \geq u_+ - \epsilon^l \end{aligned}$$

when  $\epsilon$  is small. Therefore, the right hand side of (3.11) for small  $\epsilon$  is at least

$$(1 - c_0/3)C_1\epsilon^l + u_+ - \epsilon^l.$$

Since the initial condition is  $p_0(x) = u_+ \cdot 1_{\{x \geq 0\}} + u_- \cdot 1_{\{x < 0\}}$ , by the monotonicity of  $g$  it is easy to see that for any  $x \in \mathbb{R}$  and  $t \geq 0$ ,

$$P_x^\epsilon[\mathbb{V}(\mathbf{B}(t)) = 1] \leq u_+.$$

Hence using (G4) the left hand side of (3.11) is

$$\begin{aligned} E_x[g(P_{z_1^+}^\epsilon[\mathbb{V}(\mathbf{B}(t-s)) = 1] + C_1\epsilon^l)] &\leq E_x[g(u_+ + C_1\epsilon^l)] \\ &\leq u_+ + (1 - c_0) \cdot C_1\epsilon^l \leq u_+ + ((1 - c_0/3)C_1 - 1)\epsilon^l, \end{aligned}$$

where the last line follows from the choice of  $C_1$ . So (3.11) holds in this case.

*Case 3* We now turn to the case with

$$|d(x, t)| \leq (2c_1(l) + 2(l+1)d + \gamma(t-s))\epsilon |\log \epsilon|.$$

Using (3.23) we see that on the event  $A_x$ , we have for  $u \in [0, s]$

$$\begin{aligned} |d(W_u, t-u)| &\leq |W_u - x| + |d(x, t)| + |d(x, t) - d(x, t-u)| \\ &\leq (2c_1(l) + 4(l+1)d + \gamma(t-s))\epsilon |\log \epsilon| + V_0(l+1)\epsilon^2 |\log \epsilon| \\ &\leq (R + \gamma(t-s))\epsilon |\log \epsilon|, \end{aligned}$$

where  $R = 2c_1(l) + 4(l+1)d + 1$ , see (3.20). Applying Proposition 3.9 with

$$\beta = (R + \gamma(t-s))\epsilon |\log \epsilon|$$

shows we can couple  $(W_u)_{u \geq 0}$  with  $(B_u)_{u \geq 0}$  (which starts from  $d(x, t)$ ) in such a way that for  $u \leq T_\beta = \inf\{s \in [0, t) : |d(W_s, t-s)| > \beta\} \wedge t$ ,

$$d(W_u, t-u) \leq B_u + \kappa_0 \beta u.$$

Note that  $A_x \subseteq \{T_\beta > s\}$ . Let  $\eta > 0$ . Recall  $z_1^+ = d(\hat{X}_s^\epsilon, t-s) + K_1 e^{K_2(t-s)} \epsilon |\log \epsilon|$  and let

$$\begin{aligned} z_3^+ &= d(W_s, t-s) + \epsilon + \gamma(t-s)\epsilon |\log \epsilon| \\ z_4^+ &= B_s + \kappa_0 \beta s + \epsilon + \gamma(t-s)\epsilon |\log \epsilon| \end{aligned}$$

By the coupling between  $d(W_t, t-s)$  and  $B_s$  we have  $z_3^+ \leq z_4^+$ . By the convergence of  $\hat{X}_s^\epsilon$  to  $W_s$  proved in Lemma 2.2 and the monotonicity of  $g$

$$\begin{aligned} E_x[g(P_{z_1^+}^\epsilon[\mathbb{V}(\mathbf{B}(t-s)) = 1] + C_1 \epsilon^l)] \\ \leq E_x[g(P_{z_3^+}^\epsilon[\mathbb{V}(\mathbf{B}(t-s)) = 1] + C_1 \epsilon^l)] + P_x(A_x^c) + P(G^c) \\ \leq E_{d(x,t)}[g(P_{z_4^+}^\epsilon[\mathbb{V}(\mathbf{B}(t-s)) = 1] + C_1 \epsilon^l)] + 4d\epsilon^{l+1} + \epsilon^{2l}. \end{aligned} \quad (3.25)$$

where in the last step we have used (3.22). Let

$$E = \{|P_{z_4^+}^\epsilon[\mathbb{V}(\mathbf{B}(t-s)) = 1] - u_0| \leq (u_+ - u_0) - \delta_0\}.$$

where  $\delta_0$  is the constant defined before (1.19).

Consider first when the event  $E$  occurs.

$$\begin{aligned} &\gamma(t)\epsilon |\log \epsilon| - (\epsilon + \kappa_0 \beta s + \gamma(t-s)\epsilon |\log \epsilon|) \\ &\geq \gamma(t)\epsilon |\log \epsilon| - (\kappa_0 \beta s + (K_1 + 1)e^{K_2(t-s)} \epsilon |\log \epsilon|) \\ &= \left( (K_1 + 1)e^{K_2(t-s)} (e^{K_2 s} - 1 - \kappa_0 s) - \kappa_0 R s \right) \epsilon |\log \epsilon| \\ &\geq ((K_1 + 1)(K_2 - \kappa_0) - \kappa_0 R) s \epsilon |\log \epsilon| = c_1(1) s \epsilon |\log \epsilon| \end{aligned} \quad (3.26)$$

where the last line follows from the choice of  $K_2$  in (3.21). Take  $\epsilon_d$  sufficiently small so that  $\epsilon_d < \min(\epsilon_1(1), \delta_0/2)$ . For  $\epsilon \in (0, \epsilon_d)$  we can apply Corollary 3.8 to  $z = z_4^+$  and  $w = z_2^+$ . Using (3.26) to conclude  $z_2^+ - z_4^+ \geq c_1(1) s \epsilon |\log \epsilon|$  it follows that on  $E$

$$P_{z_2^+}^\epsilon[\mathbb{V}(\mathbf{B}(t-s)) = 1] - P_{z_4^+}^\epsilon[\mathbb{V}(\mathbf{B}(t-s)) = 1] \geq \frac{\delta_0 s}{4} \quad (3.27)$$

so we have

$$g\left(P_{z_4^+}^\epsilon[\mathbb{V}(\mathbf{B}(t-s))=1]+C_1\epsilon^l\right)\leq g\left(P_{z_2^+}^\epsilon[\mathbb{V}(\mathbf{B}(t-s))=1]-\delta_0 s/4+C_1\epsilon^l\right)$$

Recalling  $s \leq (\ell+1)\epsilon^2|\log \epsilon|$  and using the monotonicity of  $g$  we can replace  $-\delta_0 s/4+C_1\epsilon^l$  by 0 when  $s > 4C_1\epsilon^l/\delta_0$ . If  $\ell \geq 4$  and  $s \leq 4C_1\epsilon^l/\delta_0$  the  $s \leq \epsilon^3$  for small  $\epsilon$ . Since  $g'(p) \leq C_2$

$$\begin{aligned} g\left(P_{z_4^+}^\epsilon[\mathbb{V}(\mathbf{B}(t-s))=1]+C_1\epsilon^l\right) &\leq g\left(P_{z_2^+}^\epsilon[\mathbb{V}(\mathbf{B}(t-s))=1]\right) + \max_{0 \leq p \leq 1} |g'(p)| \cdot C_1\epsilon^l 1_{s \leq \epsilon^3} \\ &\leq g\left(P_{z_2^+}^\epsilon[\mathbb{V}(\mathbf{B}(t-s))=1]\right) + C_2\epsilon^l 1_{s \leq \epsilon^3} \end{aligned} \quad (3.28)$$

(1.19) implies that If  $p \geq u_+ - \delta_0$ ,  $\delta \geq 0$  then

$$g(p+\delta) \leq g(p) + (1-c_0)\delta. \quad (3.29)$$

Taking  $\epsilon_d$  sufficiently small so that  $C_1\epsilon^l < \delta_0$  for all  $\epsilon \in (0, \epsilon_d)$ , and using (3.29) we have on  $E^c$  that

$$\begin{aligned} g\left(P_{z_4^+}^\epsilon[\mathbb{V}(\mathbf{B}(t-s))=1]+C_1\epsilon^l\right) &\leq g\left(P_{z_4^+}^\epsilon[\mathbb{V}(\mathbf{B}(t-s))=1]\right) + (1-c_0) \cdot C_1\epsilon^l \\ &\leq g\left(P_{z_2^+}^\epsilon[\mathbb{V}(\mathbf{B}(t-s))=1]\right) + (1-c_0) \cdot C_1\epsilon^l \end{aligned} \quad (3.30)$$

since  $z_4^+ \leq z_2^+$ . Using (3.28) and (3.30) in (3.25)

$$\begin{aligned} E_x \left[ g\left(P_{z_1^+}^\epsilon[\mathbb{V}(\mathbf{B}(t-s))=1]+C_1\epsilon^l\right) \right] \\ \leq E_{d(x,t)} \left[ g\left(P_{z_2^+}^\epsilon[\mathbb{V}(\mathbf{B}(t-s))=1]\right) \right] + (1-c_0)C_1\epsilon^l + 4d\epsilon^{l+1} + \epsilon^{2l} + C_2\epsilon^l 1_{s \leq \epsilon^3} \\ \leq E_{d(x,t)} \left[ g\left(P_{z_2^+}^\epsilon[\mathbb{V}(\mathbf{B}(t-s))=1]\right) \right] + (1-c_0/3)C_1\epsilon^l + C_2\epsilon^l 1_{s \leq \epsilon^3}, \end{aligned}$$

which completes the proof of Lemma 3.12 and hence of Proposition 3.11.

## 4 Checking the conditions

Since (G0) is based on an observation on all the particle systems considered, it is satisfied trivially. Recall that (G5)  $g$  is strictly increasing on  $[0, 1]$  holds in all our examples and (G4) is a consequence of (G1), (G2) and (G3). That is, it suffices to check (G1)–(G3).

#### 4.1 Cubic $g$

As discussed in Sect. 1, both the sexual reproduction model with rapid stirring and the Lotka–Volterra systems fall into this category. In this case, according to (1.15) we must have

$$g(p) = p - c[(p - u_-)(p - u_0)(p - u_+)]$$

for some  $c > 0$ . To check (G1) we note that if we let  $g_1(p) = (p - u_-)(p - u_0)(p - u_+)$  then  $g_1(u_+ - \delta) = -g_1(u_- + \delta)$ . So  $g(u_+ - \delta) + g(u_- + \delta) = u_+ + u_- = 2u_0$  by (G0).

$$g'(p) = 1 - c[(p - u_0)(p - u_+) + (p - u_-)(p - u_+) + (p - u_-)(p - u_0)]$$

From this we see that

$$\begin{aligned} g'(u_+) &= 1 - c(u_+ - u_-)(u_+ - u_0) < 1, \\ g'(u_-) &= 1 - c(u_- - u_0)(u_- - u_+) < 1, \\ g'(u_0) &= 1 - c(u_0 - u_-)(u_0 - u_+) > 1, \end{aligned}$$

which proves (G2). Taking the second derivative we obtain

$$g''(p) = -2c[(p - u_+) + (p - u_0) + (p - u_-)] = -6c(p - u_0)$$

since  $u_+ + u_- = 2u_0$ . This proves (G3).

#### 4.2 Nonlinear voter model

Recall that for the nonlinear voter model we suppose

- (A1)  $b_1 > 0$  and  $3b_1 + b_2 < 0$ ;
- (A2)  $0 \leq a_1 \leq a_2 \leq 1/2$ ;
- (A3)  $6b_1 + b_2 > 0$ .

In Region 2 there are two extra roots of  $\phi(p)$  denoted by  $1 - u^* < 1/2 < u^*$ , where

$$u^* = 1/2 + \beta_0 \quad \text{with } \beta_0 = \frac{\sqrt{-(b_1 - b_2)(3b_1 + b_2)}}{2(b_1 - b_2)}.$$

The roots come from the following calculation:

$$\begin{aligned}\phi(p) &= b_1 p(1-p)^4 + b_2 p^2(1-p)^3 - b_2 p^3(1-p)^2 - b_1 p^4(1-p) \\ &= b_1 p(1-p)(1+3p^2-3p-2p^3) + b_2(1-p)^2 p^2(1-2p) \\ &= p(1-p)(1-2p) \left( b_1(1-p+p^2) + b_2 p(1-p) \right) \\ &= p(1-p)(1-2p)(b_1-b_2) \left( p^2 - p + \frac{b_1}{b_1-b_2} \right).\end{aligned}$$

Solving  $p^2 - p + b_1/(b_1 - b_2) = 0$  gives the two extra roots  $\frac{1}{2} \pm \beta_0$ .

To check our conditions we note that  $g(p) = p + \phi(p)$  where  $\phi(p)$  is the reaction term, see (1.15). In our notation  $u_0 = 1/2$ ,  $u_- = 1 - u^*$  and  $u_+ = u^*$ .

*Checking (G1)*  $\phi(p)$  is antisymmetric about  $u_0$  so  $\phi(u_+ - \delta) = -\phi(u_- + \delta)$  and hence  $g(u_+ - \delta) + g(u_- + \delta) = u_+ + u_- = 2u_0$ , proving (G1).

*Checking (G2)*  $u_-$ ,  $u_+$  are stable fixed points so  $\phi'(u_-) < 0$ ,  $\phi'(u_+) < 0$ .  $u_0$  is unstable so  $\phi'(u_0) > 0$  and (G2) follows.

*Checking (G3)* Since  $g''(p) = \phi''(p)$  the next step is to calculate  $\phi''(p)$  for  $p \in (1/2, u^*)$ . By symmetry it is easy to see

$$\phi(0) = \phi(1/2) = \phi(1) = 0 \quad \text{and} \quad \phi(p) = -\phi(1-p). \quad (4.1)$$

It follows that  $\phi''(p) = -\phi''(1-p)$  and  $\phi''(1/2) = 0$ . Since  $\phi(p)$  is quintic it has at most three inflection points. To check (G3) it suffices to show  $\phi''(u^*) < 0$ .

Let  $\phi_1(p) = p(1-p)(1-2p)$  and  $\phi_2(p) = (b_1 - b_2) \left( p^2 - p + \frac{b_1}{b_1 - b_2} \right)$ . Since  $\phi(p) = \phi_1(p)\phi_2(p)$  we have

$$\phi''(p) = \phi_1''(p)\phi_2(p) + \phi_1(p)\phi_2''(p) + 2\phi_1'(p)\phi_2'(p).$$

Notice that  $\phi_2(u^*) = 0$  so our problem simplifies to

$$\phi''(u^*) = \phi_1(u^*)\phi_2''(u^*) + 2\phi_1'(u^*)\phi_2'(u^*)$$

The calculation simplifies if we write  $u^* = 1/2 + \beta_0$ , i.e.,

$$\begin{aligned}\phi''(1/2 + \beta_0) &= \phi_1(1/2 + \beta_0)\phi_2''(1/2 + \beta_0) + 2\phi_1'(1/2 + \beta_0)\phi_2'(1/2 + \beta_0) \\ &= -2\beta_0 \left( \frac{1}{4} - \beta_0^2 \right) \cdot 2(b_1 - b_2) + 2 \left( 6\beta_0^2 - \frac{1}{2} \right) \cdot 2\beta_0(b_1 - b_2) \\ &= 4\beta_0(b_1 - b_2) \left( 7\beta_0^2 - \frac{3}{4} \right) = -4\beta_0(6b_1 + b_2) < 0,\end{aligned}$$

hence proving (G3).

## References

1. Bonaventura, L.: Interface dynamics in an interacting particle system. *Nonlinear Anal.* **25**, 799–819 (1995)
2. Chen, X.: Generation and propagation of interfaces for reaction–diffusion equations. *J. Differ. Equ.* **96**, 116–141 (1992)
3. Cox, J.T., Durrett, R., Perkins, E.: Voter model perturbations and reaction diffusion equations. *Astérisque* **349**, 1–127 (2013)
4. De Masi, A., Orlandi, E., Presutti, E., Trioli, L.: Glauber evolution with Kac potentials: I. Mesoscopic, macroscopic limits, and interface dynamics. *Nonlinearity* **7**, 633–696 (1994)
5. De Masi, A., Orlandi, E., Presutti, E., Trioli, L.: Glauber evolution with Kac potentials: II. Fluctuations. *Nonlinearity* **9**, 27–51 (1996)
6. De Masi, A., Orlandi, E., Presutti, E., Trioli, L.: Glauber evolution with Kac potentials: III. Spinoidal decomposition. *Nonlinearity* **9**, 27–51 (1996)
7. Durrett, R.: Coexistence in stochastic spatial models. (Wald lecture paper). *Ann. Appl. Probab.* **19**, 477–496 (2009)
8. Durrett, R.: Spatial evolutionary games with small selection coefficients. *Electron. J. Probab.* **19**, 1–64 (2014)
9. Durrett, R., Neuhauser, C.: Particle systems and reaction–diffusion equations. *Ann. Probab.* **22**, 289–333 (1994)
10. Etheridge, A., Freeman, N., Penington, S.: Branching Brownian motion, mean curvature flow and the motion of hybrid zones. *Electron. J. Probab.* **22**, 1–40 (2017)
11. Evans, L.C., Soner, H.M., Souganidis, P.E.: Phase transitions and generalized motion by mean curvature. *Commun. Pure. Appl. Math.* **45**, 1097–1123 (1992)
12. Evans, L.C., Spruck, J.: Motion of level sets by mean curvature I. Preprint (1992)
13. Fife, P.C., McLeod, J.B.: The approach of solutions of nonlinear diffusion equations to travelling front solutions. *Arch. Ration. Mech. Anal.* **65**, 335–361 (1977)
14. Fife, P.C., McLeod, J.B.: A phase plane discussion of convergence to traveling fronts for nonlinear diffusion. *Arch. Ration. Mech. Anal.* **75**, 281–314 (1981)
15. Funaki, T., Spohn, H.: Motion by mean curvature from the Ginzburg–Landau  $\nabla\phi$  interface model. *Commun. Math. Phys.* **185**, 1–36 (1995)
16. Funaki, T., Tsunoda, K.: Motion by mean curvature from Glauber–Kawasaki dynamics. *J. Stat. Phys.* **177**, 183–208 (2019)
17. Kallenberg, O.: *Foundations of Modern Probability*, vol. 2. Springer, New York (1997)
18. Katsoulakis, M.A., Souganidis, P.E.: Generalized motion by mean curvature as a macroscopic limit of stochastic Ising models with long range interactions and Glauber dynamics. *Commun. Math. Phys.* **169**, 61–97 (1995)
19. Liggett, T.M.: *Interacting Particle Systems*. Springer, New York (1985)
20. Molofsky, J., Durrett, R., Dushoff, J., Griffeth, D., Levin, S.: Local frequency dependence and global coexistence. *Theoret. Pop. Biol.* **55**(1999), 270–282 (1999)
21. Neuhauser, C., Pacala, S.: An explicitly spatial version of the Lotka–Volterra model with interspecific competition. *Ann. Appl. Probab.* **9**, 1226–1259 (1999)
22. Sowers, R.B.: Hydrodynamic limits and geometric measure theory: mean curvature limits from threshold voter models. *J. Funct. Anal.* **169**, 121–155 (1999)
23. Yip, N.K.: Stochastic motion by mean curvature. *Arch. Ration. Mech. Anal.* **144**, 313–355 (1998)

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.