UNIVERSITY OF
OXFORD

DEPARTMENT OF
STATISTICS

# Long term behaviour of spatial population models with heterozygous or asymmetric homozygous selection

A thesis submitted in partial fulfilment of the requirements for the degree
Doctorate of Philosophy in Statistics

July 2018

Mitchell Gooding
St Peter's College

Department of Statistics
University of Oxford

# Abstract

We investigate the long term behaviour of two models for the spatial distribution of alleles in a diploid population, one with asymmetric selection in favour of the homozygotes, and one with selection in favour of the heterozygote.

We model the population with asymmetric homozygous selection using a version of the spatial $\Lambda$-Fleming-Viot process. We identify three regimes. For very small values of the asymmetry, the limiting behaviour of the process is the same as for the case with symmetric selection. This case was studied in [13], and they showed, under certain rescaling and initial conditions, that the hybrid zone, the interface between two homogeneous regions of each homozygote, evolves according to mean curvature flow. However, for larger, but not that much larger, values of asymmetry, we show a new behaviour, that the hybrid zone evolves according to a different type of flow, which we call constant curvature flow. Furthermore, there is a strength of asymmetry for which elements of both types of curvature flow are present in the limit. This suggests that the behaviour found in [13] is more sensitive to perturbations than first thought.

We then go on to investigate the fluctuations of this process about its limit. To do this, we specialise to the one-dimensional case. We show that, when time, space and the strength of the asymmetry are appropriately rescaled, the hybrid zone, which is a single point in one dimension, evolves according to a Brownian motion, with drift proportional to the asymmetry.

Finally, we turn to the model with heterozygous selection. We restrict ourselves to two dimensions, and investigate this process through its dual, the branching annihilating random walk. We show that, up to an arbitrary time, and with arbitrarily high probability, there exists a branching rate such that we may couple a branching annihilating random walk to a ternary branching Brownian motion. While interesting in its own right, this result lends support to a conjecture of [6], that a branching annihilating random walk in two dimensions has a positive probability of survival for all time for any positive branching rate.

# Contents

## Acknowledgements

The author would like to thank his supervisor, Professor Alison Etheridge, for her patience, assistance and detailed feedback. I regard myself very lucky to have been supervised by someone, not only of the academic calibre of Professor Etheridge, but also someone with a matching dedication to her responsibilities as a supervisor. Thanks are also due to Sarah Penington and Raphaël Forien for their many helpful conversations, without which would have made this research much more stressful.

Thanks also to Professor Jochen Blath, for travelling from Berlin in order to examine my thesis, and Professor James Martin, who not only was also an examiner for my viva, but for my confirmation and transfer of status also. Thank you both for your feedback.

I would also like to thank my partner, Talia Pettigrew. I am glad to have been able to share this adventure with you. This degree would have been very isolating without you.

Finally, thanks to my parents, Brenda and David Gooding. Thank you for providing me with the opportunities I needed in my youth in order to be accepted to undertake a doctorate at the University of Oxford. Thank you also for supporting me, and enduring my absence, while I travelled to the opposite side of the globe to do that doctorate.

# Notation

| | |
|---|---|
| $[\cdot]$ | The set of positive integers up to and including $\cdot$. |
| $\sim$ | If $f(n) \sim g(n)$, then $f$ equals $g$ asymptotically, i.e., $\lim\limits_{n\to\infty} \frac{f(n)}{g(n)} = 1$. If a $f$ and $g$ are functions of multiple variables, a subscript may be used to indicate the variable which is being taken to infinity, i.e., $f(n,\cdot) \underset{n\to\infty}{\sim} g(n,\cdot)$. |
| $A_{[t',t]}$ | Non-collision event during the interval $[t',t]$ for two instantaneous continuous time random walks, see page 145. |
| $A_t$ | Non-collision event during the interval $[0,t]$ for two instantaneous continuous time random walks, see page 145. |
| BARW | Branching annihilating random walk. |
| BBM | Branching Brownian motion. |
| $\mathbf{B}_t$ | One-dimensional historical branching Brownian motion, see Section 2.2. |
| $B_t^i$ | In Chapter 4, the $i^{th}$ continuous time random walk in a set of $n$ random walks, see page 145. |
| $B_t$ | One-dimensional Brownian motion. |
| $\mathbb{d}$ | Dimension. |
| $D_t^n$ | Non-collision event until time $t$ for $n$ instantaneous continuous time random walks, see page 145. |
| $E_t^n$ | Non-collision event until time $t$ for $n$ delayed continuous time random walks, see page 145. |
| $F_t$ | Non-collision event until time t for two delayed continuous time random walks, see page 145. |
| $H$ | Lifetime of a particle in a TBBM, see page 173. |
| $h$ | Lifetime of a vine, see page 172. |
| $\mathcal{H}$ | Lifetime of a branching vine, see page 177. |
| $m$ | In Chapter 4, the transition kernel for a BARW. In Chapters 2 and 3, $m_0$. |
| $N_i(t)$ | The number of particles at site $i \in \mathbb{Z}^{\dim}$ of a BARW, see page 20. |
| $N_t$ | The total number of particles in a BARW, see page 20. |
| $\mathcal{O}\left(\right)$ | If $f(n) = \mathcal{O}\left(g(n)\right)$, then $\lvert f\rvert$ is bounded above by $g$, up to constant factor, asymptotically, i.e., $\limsup\limits_{n\to\infty} \frac{\lvert f(n)\rvert}{g(n)} < \infty$. See also the entry for $\sim$. |
| $o\left(\right)$ | If $f(n) = o\left(g(n)\right)$, then $f$ dominates $g$ asymptotically, i.e., $\lim\limits_{n\to\infty} \frac{f(n)}{g(n)} = 0$. |
| $\mathcal{U}$ | The set of labels in Ulam-Harris notation, i.e., $\{\emptyset\} \cup \bigcup_{k=1}^{\infty} \{1,2,3\}^k$. |
| $\mathbf{W}_t$ | A multidimensional historical branching Brownian motion, see Section 2.2. |

# Chapter 1

# Introduction

The biological process of evolution is an often misunderstood phenomenon. The result of the process is easily grasped: that living things, from animals, to plants to bacteria, change slowly over time. However, the causes of these changes can be quite subtle. Although it may seem as though evolution acts toward a certain ‘destination’, as illustrated by the popular image, ‘The Road to Homo Sapiens’, modern theories of evolution do not support this idea. Rather, evolution occurs in a more stochastic fashion; mutations occur essentially at random, most of which are neutral or even detrimental to the survival of the organism experiencing that mutation. Determining the process by which some of these mutations become widespread through a population, resulting in an evolutionary step in the species, is thus as much a statistical question as a biological one. Hence, it behoves us to model evolution from a mathematical perspective.

In order for evolution to occur, there must be some mechanism for the traits of an individual to be inherited by its offspring. In most living organisms, this heritable material comes in the form of *genes*. Physically, genes manifest as particular regions of an extremely long molecule called DNA, copies of which are in each cell of an organism. Variations in the chemical composition of the DNA result in variations of a trait in a population. Each possible copy of a gene is called an *allele*.

Most sexually reproducing life forms have at least two copies of their DNA in each cell, one inherited from each parent. We call organisms with exactly two copies *diploid*. If both copies of an allele are the same, we call that organism a *homozygote*, and if they are different, a *heterozygote*. We say different organisms that have the same copies of a particular allele are the same *genotype*, and the corresponding trait expressed in the individual, controlled by that gene, is the corresponding *phenotype*. In this thesis, we study models for diploid populations which focus on a single gene.

Some of these phenotypes may be advantageous to the organism exhibiting them, such as immunity from a disease or the ability to digest a common food source. We say these advantaged organisms are *fitter* than their rivals. If these advantages are large enough, they could cause these organisms to be more successful at reproduction than those without these phenotypes. Hence, we would expect these phenotypes to become more prevalent in the population, over many generations. Eventually, they could replace all other phenotypes corresponding to that gene, resulting in a perceptible step in the evolution of a population. This is the process of natural selection, one of the key forces of evolution. Although there are many factors which determine the fitness of an individual, to model the strength of this selection, we can consider the relative fitness of each genotype, interpreted as the relative proportion of the offspring of an individual of that genotype that survive to their own reproductive stage.

Even if a particular allele confers an advantage, it is possible that it could be eliminated from the population, due to the randomness inherent in reproduction. This is not very likely for a common allele giving a strong advantage in a large population, but is an important consideration for a newly mutated allele, or an allele with a very low or non-existent advantage. We say that changes in the genetic composition of a population resulting from random chance are due to the evolutionary force of genetic drift. Because of this, even when there is no selection acting on a particular allele at all, we still require a statistical model to investigate the spread of that allele through a population.

Both of these genetic forces, selection and genetic drift, are further complicated by the presence of space. Typically, the individuals in a population do not travel the entire length of their environment within a reproductive cycle. The offspring of an individual with a selective advantage tend to live near their parent, and thus are more likely to be competing with each other rather than an arbitrary member of the population. Furthermore, if reproductive events are spatially correlated, the effects of genetic drift will be exaggerated, as many individuals of one genotype could be wiped out in a single event.

This thesis investigates two mathematical models for describing how a population evolves over time and space, one with asymmetric homozygotic selection, and one with heterozygotic selection. We perform a sensitive analysis of how changes in the strength of this selection affects the long term behaviour of the population.

## 1.1 Population Models

There are many models which attempt to describe the evolution of a population over time. Perhaps the most well-known is the Lotka-Volterra model. This model uses a pair of coupled ODEs to describe two interacting populations. These equations attempt to capture several important features that govern the growth of a population. Firstly, in the absence of any outside factors, the growth of the population should simply be proportional to the number of individuals in that population. Second, in any environment, there is usually a finite amount of the resources required by the populations to survive. We call this theoretical limit on the number of individuals which the environment can support the carrying capacity. Finally, each population interacts with the other, either directly, say, through predation, or indirectly, by competing for the same resources. Either way, we can model this intraspecies competition by decreasing the growth rate of each population in proportion to the other population. A simple example of the resulting ODEs is

$$\begin{aligned} \frac{dN_1}{dt} &= r_1 N_1 \left(1 - \tfrac{N_1}{K_1} - \alpha_{12}\tfrac{N_2}{K_1}\right) \\ \frac{dN_2}{dt} &= r_2 N_2 \left(1 - \tfrac{N_2}{K_2} - \alpha_{21}\tfrac{N_1}{K_2}\right), \end{aligned} \tag{1.1}$$

where, for $i \in \{1, 2\}$, $N_i$ is the population of species $i$, which has intrinsic growth rate $r_i$ and carrying capacity $K_i$, and $\alpha_{12}$ and $\alpha_{21}$ are the rates of intraspecies competition. There is no explicit parameter for selection in this model. However, as we shall see later, one can derive an implicit measure of selection, which depends on each of the parameters in the model.

The Lotka-Volterra model is very tractable - indeed, analytical solutions can be found quite easily for certain choices of these parameters - but it fails to shed light on two important questions in population modelling. Using ODEs cannot tell us how the spatial distribution of a population affects its dynamics, and as a deterministic model, it can only give us the ‘average’ trajectory of the population, rather than a distribution of possible outcomes. There are many models which extend simple dimensionless deterministic population models like (1.1) into models which can account for one of these shortcomings, but accounting for both is harder.

An obvious way to adapt the Lokta-Voltera model to include a spatial dimension is to replace the ODEs with PDEs. This is a well studied model, see, for example, Section 3 of [34]. To capture the random element of the evolution of a population, the ODEs could be replaced with stochastic differential equations (SDEs). Again, this is also well studied, see, for example [11]. However, when we try to capture both elements at once this way, and study population dynamics through stochastic partial differential equations (SPDEs), we encounter the problem that the corresponding nonlinear SPDEs are only well defined in one spatial dimension, a number of dimensions which is not the most useful for population modelling. This is because, in order to make sense of the coefficent of the stochastic term in higher dimensions, the solution to a linear higher-dimensional SPDE only has enough regularity to be interpreted as a distribution, rather than a function. Hence, there is no clear meaning for the nonlinear term. For more on the technical construction of solutions to SPDEs, see [42].

Fortunately, many other models have been proposed for studying populations evolving stochastically over time and space. Before discussing these more exotic models, let us discuss one-dimensional SPDEs. While they cannot easily be extended into higher dimensions, some populations can be modelled sufficiently well with a single spatial dimension, such as populations living along the bank of a river, or other boundaries between two geographical regions.

If $u : \mathbb{R} \to [0, 1]$ represents the proportion of a population of a particular type at each point in space, then a term proportional to $\sqrt{u(1-u)}\dot{W}$, where $\dot{W}$ is a two parameter white noise, is often used as the noise term in the SPDE. This can be understood heuristically; if individuals within location $x \in \mathbb{R}$ mate randomly, then only the matings between two different types will change the proportion at that location, which happens with probability $2u(1-u)$. Each reproductive event could cause the proportion $u$ of individuals to increase or decrease, so rather than count the total number of matings, we change the proportion of the population of that type by the standard deviation of each reproductive event, i.e. $\sqrt{u(1-u)}$, and let the 'direction' of that change be chosen randomly, hence the white noise. Of course, this argument can be made more rigorous, see for example Section 2.1 of [16].

This type of white noise term is called a Fisher-Wright noise, and it is the noise term used in the SPDEs studied in, for example, [32, 40]. These papers, among others, investigate the properties of the interface between two populations starting from a Heaviside initial condition, where the population is completely one type on the left half line, and entirely the other type on the right. Although the population models we study in this thesis are not defined in terms of SPDEs, each are closely related to stochastic, partial and ordinary DEs, so these DEs arise often. In particular, these SDEs and SPDEs will have Fisher-Wright noises.

Another population model closely related to SPDEs is a class of stochastic processes called superprocesses. These are finite $\mathbb{R}^d$-measure valued processes which can be thought of as a scaling limit of a Galton-Watson process, where each individual is born at the location of their parent, and performs an independent Lévy process in space during its lifetime. The equation which drives the evolution of the process is then a martingale equation, rather than an SPDE, which avoids the technical problem preventing the rigorous definition of solutions to higher-dimensional SPDEs. The textbook [16] gives a thorough introduction to superprocesses for those who wish for a more comprehensive discussion.

Superprocesses allow us to go a bit further in population modelling, but they have their own shortcomings. In particular, the superprocess described above does not account for any interaction between individuals, within the species or externally. This causes undesirable long term behaviour. For example, in one or two spatial dimensions, a particular superprocess, the Dawson-Watanabe superprocess, at large times, either goes extinct, or forms arbitrarily dense "clumps" of population [19]. It is possible to add inter- or intrapopulation interaction to the model, see for example Chapter 7 of [16] and the references therein, but this causes it to lose a significant part of this model's tractability.

Rather than use a measure to describe the absolute population size over space, one can assume a uniform population size over space, and use a measure to describe the proportion a particular subpopulation makes of the total population. This is the idea behind a process known as the Spatial Λ-Fleming-Viot process. This process has the advantages of the superprocess models described above, but intra-population interactions are much more simple to incorporate, as subpopulations are an explicit part of the model. This is one of the two population models we consider in this thesis. The second model is a stepping stone model, which overcomes the issues facing higher-dimensional SPDE's by coupling together a system of SDEs. We introduce these models in more detail in the next two subsections.

### 1.1.1 Spatial Λ-Fleming-Viot process

The spatial Λ-Fleming-Viot process (SΛFV) was introduced in [12, 4]. Unlike some of the models discussed in the previous subsection, rather than keep track of the locations of individuals in the population, the SΛFV keeps track of the events affecting the genealogies of each individual.

We assume a population has two alleles, $a$ and $A$, and hence individuals in our diploid population can have three possible genotypes: the homozygotes, $aa$ and $AA$, and the heterozygotes $aA$. In our model, we allow for each of these genotypes to have a different fitness; in particular, the heterozygotes have a lower fitness than the homozygotes, and we (arbitrarily) choose $aa$ to be fitter than $AA$.

This model could be modified slightly to account for different genotypic fitness hierarchies, but we expect this generalisation to merely extend to cases which have already been studied. For example, we could try to include the case where one homozygote is fitter than the heterozygote, which is in turn fitter than the other homozygote. However, in this case, selection and genetic drift would work in tandem to remove the less favoured allele from the population. We would expect this to result in behaviour similar to genic selection in a haploid population, which has been studied for the SΛFV in [15]. Furthermore, the model we introduce in the next subsection can be used to model heterozygotic selection.

On the other hand, homozygotic selection is quite interesting, as if we start with two geographically distinct populations of each homozygote, the force of selection acts in the opposite direction to the natural diffusion of alleles through the population. Symmetric selection against heterozygosity, in that both homozygotes are equally fit, was studied in [13]. They were able describe what happens when these two evolutionary forces act in opposite directions, and were able to describe how the boundary between two populations evolves under this model. We extend this work to the case described above, where the fitness of the two homozygotes are unequal.

Let us be a bit more mathematically precise, in order to introduce the model. We define a measure for each time, $M_t$, on $\mathbb{R}^d \times \{A, a\}$, where the marginal measure over $\mathbb{R}^d$ is Lebesgue. We denote the space of such measures by $\mathcal{M}_\lambda$. For any Lebesgue measurable subset $E$, and $\alpha \in \{A, a\}$, we interpret $M_t(E, \alpha)/\operatorname{Vol}(E)$ to be the proportion of allele $\alpha$ in region $E$.

It is more convenient to extend this definition to a function, $w_t$, describing the probability of sampling a type $A$ allele from the population at each location and time. Fortunately, [41] show that the measure $M_t$ has a density $w_t(x) : \mathbb{R}^{\mathsf{d}} \to [0,1]$, which arises in the natural way,

$$M_t(dx, d\alpha) = (w_t(x)\delta_A(d\alpha) + (1 - w_t(x)\delta_a(d\alpha))dx. \tag{1.2}$$

This defines $w_t$ up to a Lebesgue null set of $\mathbb{R}^{\mathsf{d}}$ (as two densities, $w_t$ and $\tilde{w}_t$ will be equivalent under this definition if $\mathrm{Vol}(\{x \in \mathbb{R}^{\mathsf{d}} : w_t(x) \neq \tilde{w}_t(x)\}) = 0$). Hence, to complete the definition, we arbitrarily take $w_t(x)$ to be zero in such Lebesgue null sets.

We slightly abuse notation, and also use $\mathcal{M}_\lambda$ to denote the state space of the process $(w_t)_{t\in\mathbb{R}}$. We can now introduce the process.

**Definition 1.1** (Spatial Λ-Fleming-Viot with asymmetric selection against heterozygosity)**.** *Fix* $u, \boldsymbol{s} \in (0,1]$, $\gamma \in (0,1)$ *and* $\mathcal{R} \in (0,\infty)$. *Let* $\mu$ *be a finite measure on* $(0,\mathcal{R}]$. *Further, let* $\Pi$ *be a Poisson point process on* $\mathbb{R}_+ \times \mathbb{R}^{\mathsf{d}} \times (0,\mathcal{R}]$ *with intensity measure* $dt \otimes dx \otimes \mu(dr)$. *The* spatial Λ-Fleming-Viot process with asymmetric homozygotic selection *(SΛFVA) driven by* $\Pi$ *is the* $\mathcal{M}_\lambda$*-valued process* $(w_t)_{t\geq 0}$ *with dynamics given as follows.*

*If* $(t,x,r) \in \Pi$, *a reproduction event occurs at time* $t$ *within the closed ball* $\mathcal{B}_r(x)$ *of radius* $r$ *centred on* $x$. *With probability* $1-\boldsymbol{s}$ *the event is* neutral, *in which case:*

1. *Choose a parental location* $z$ *uniformly at random within* $\mathcal{B}_r(x)$, *and a parental type,* $\alpha_0$, *according to* $w_{t-}(z)$, *that is* $\alpha_0 = A$ *with probability* $w_{t-}(z)$ *and* $\alpha_0 = a$ *with probability* $1 - w_{t-}(z)$.

2. *For every* $y \in \mathcal{B}_r(x)$, *set* $w_t(y) = (1-u)w_{t-}(y) + u\mathbb{1}_{\{\alpha_0 = A\}}$.

*With the complementary probability* $\boldsymbol{s}$ *the event is* selective, *in which case:*

1. *Choose three 'potential' parental locations* $z_1, z_2, z_3$ *independently and uniformly at random within* $\mathcal{B}_r(x)$, *and at each of these sites 'potential' parental types* $\alpha_1$, $\alpha_2$, $\alpha_3$ *according to* $w_{t-}(z_1), w_{t-}(z_2), w_{t-}(z_3)$ *respectively. Let* $\widehat{\alpha}$ *denote the most common allelic type in* $\alpha_1, \alpha_2, \alpha_3$. *However, if precisely one of the parents has the fitter allele, with probability* $\frac{2\gamma}{3+3\gamma}$, *replace* $\widehat{\alpha}$ *with the fitter allele.*

2. *For every* $y \in \mathcal{B}_r(x)$ *set* $w_t(y) = (1-u)w_{t-}(y) + u\mathbb{1}_{\{\widehat{\alpha} = A\}}$.

Taking $\gamma = 0$ in this definition gives the model studied in [13]. As the process is only a slight modification of the ones in [13] and [15], we appeal to these papers for a proof of the existence and uniqueness of this process.

The parameters $\mathbf{s}$ and $\gamma$ represent the strength of selection against the heterozygote, and the less fit homozygote, respectively. The selection mechanism against the less fit homozygote should be fairly obvious from the description of the model; when we have a selective event (against the homozygote), we choose the offspring allele in a way consistently biased towards one of the homozygotes. However, the selective mechanism against heterozygotes is perhaps not as obvious. It may also seem a bit strange that the mechanism involves three potential parents, as that is not usually the number of parents involved in a reproductive event. This mechanism is not trying to replicate the reproductive process mechanically, but rather merely trying to replicate the effect on the population of alleles.

To see this, recall that the proportion of type $A$ alleles at the site of a reproductive event occurring at time $t$ and location $x$ is $w_t(x)$. In this and the following paragraph, we will only consider one time and location, so we shall simply refer to this proportion as $w$. Furthermore, to clarify the explanation, we temporarily set $\gamma = 0$. Allowing this proportion to be modelled by continuous variable, we have implicitly assumed the population size at any point in space is large. We also may assume our selection parameter is small (indeed, after rescaling, the selection strength goes to 0). Hence, we may assume the population in Hardy-Weinberg equilibrium, i.e., the proportion of individuals of genotype $AA$ to be $w^2$, type $aa$ to be $(1-w)^2$ and type $aA$ to be $2w(1-w)$.

We model the reproductive event by selecting an allele at random from the population, but to model the fitness of each genotype, we penalise the less fit genotype by giving selecting it with a reduced probability, reduced according to the strength of selection $\mathbf{s}$. The probability that we sample an allele of type $A$ is then

$$\begin{aligned}\frac{w^2 + w(1-w)(1-\mathbf{s})}{1 - 2\mathbf{s}w(1-w)} &= (1-\mathbf{s})w + \mathbf{s}(3w^2 - 2w^3) + \mathcal{O}\left(\mathbf{s}^2\right) \\ &= (1-\mathbf{s})w + \mathbf{s}(w^3 + 3w^2(1-w)) + \mathcal{O}\left(\mathbf{s}^2\right). \qquad (1.3)\end{aligned}$$

The first term in (1.3) corresponds to the neutral events described in the Definition 1.1, as these events occur with probability $1-\mathbf{s}$, and we choose allele $A$ simply with probability given by the proportion of allele $A$ at that site, which is $w$. Conversely, the second term in (1.3) corresponds to the selective events in the definition, as they occur with probability $\mathbf{s}$, and $w^3+3w^2(1-w)$ is the probability of choosing at least two $AA$ alleles out of an independent sample of size three.

We can analyse (1.3) further. If we subtract off the initial proportion of allele $A$, we get the local change in $w$ per reproductive event. If reproductive events happen at a constant rate, and we allow for neighbouring alleles to migrate diffusively, then we might expect, in the long term, $w$ to evolve according to the PDE

$$\frac{\partial w}{\partial t} = \Delta w + \mathbf{s}w(1-w)(2w-1), \tag{1.4}$$

modulo multiplicative constants for each term. This argument is quite informal, but in fact, when $\gamma = 0$, [13] show that, when suitably rescaled, in at least two dimensions, the evolution of the SΛFV process with symmetric selection is indeed related to this equation. In Chapter 2, we generalise this result to the case with asymmetric selection.

The scaling used in the result of [13] speeds up time, compresses space, and correspondingly scales the selection parameter and $u$, the impact parameter. In particular, fix $\varrho \in (0, \frac{1}{4})$, and a sequence $\epsilon_n > 0$. For each $n \in \mathbb{N}$, let

$$u_n = \frac{u}{n^{1-2\varrho}}, \qquad \text{and} \qquad \mathbf{s}_n = \frac{1}{\epsilon_n^2}\frac{1}{n^{2\varrho}}.$$

Furthermore, we define the finite measure $\mu_n$ on $(0, \mathcal{R}_n]$, where $\mathcal{R}_n = n^{-\rho}\mathcal{R}$, by $\mu_n(A) = \mu(n^\rho A)$ for all Borel subsets $A$ of $(0,\infty)$. Our rescaled SΛFVA will be driven by the Poisson point process $\Pi_n$ on $\mathbb{R}_+ \times \mathbb{R}^{\mathbb{d}} \times (0,\infty]$ with intensity measure

$$ndt \otimes n^\rho dx \otimes \mu_n(dr). \tag{1.5}$$

Specifically, $n^\rho dx$ scales each linear dimension by $n^\rho$, so that upon integration, the volume of a region is scaled by $n^{\mathbb{d}\rho}$. We use this scaling for our results also, however, we will need to introduce the appropriate scaling the asymmetric heterozygotic selection parameter, $\gamma$. We delay introducing the scaling of $\gamma$ until Chapter 2.

We also comment on the perhaps unexpected choice to define the parameter $\gamma$, rather than parametrising the asymmetry by the quantity $\frac{2\gamma}{3+3\gamma}$ which appears in Definition 1.1. This is because when we do the same calculation going from (1.3) to (1.4) for nonzero $\gamma$, recalling that $w$ is the proportion of the less fit allele, we find

$$(1-\mathbf{s})w + \mathbf{s}(w^3 + \tfrac{3+\gamma}{1+\gamma}w^2(1-w)) - w = \tfrac{\mathbf{s}}{1+\gamma}w(1-w)(2w-(1+\gamma)). \tag{1.6}$$

Equation (1.4) is a special case of the Allen-Cahn equation, and it is central to the results in Chapter 2. Hence it makes our calculations much nicer if we define the asymmetry parameter $\gamma$ in terms of the parameter we find in the PDE, rather than the form found in the definition of the SΛFVA. We elaborate further on this equation in the beginning of Chapter 2. First, we introduce the other major population model investigated in this thesis.

### 1.1.2 Simple stepping stone model

As for the SΛFVA, this new model considers a pair of interacting populations. This stepping stone model was studied in Blath, Etheridge, and Meredith [6], itself a simplification of one studied by Bolker and Pacala [7] and Murrell and Law [33]. Because of this, we shall refer to the model as the simple stepping stone model. Before introducing it, we explain how it is related to the classical Lotka-Volterra model in (1.1).

First, rather than keep track of the absolute population sizes, we can keep track of the total population, $N(t) := X_1(t) + X_2(t)$, and the proportion of one of the populations, $p(t) = \frac{X_1(t)}{X_1(t)+X_2(t)}$. The chain rule allows us to write the first equation of (1.1) as,

$$dp = \sigma p(t)(1-p(t))(1-\mu p(t))dt,$$

where

$$\sigma = r_1 - r_2 + \tfrac{N(t)}{K_2}(r_2 - \alpha_{12}r_1) \text{ and } \mu = \frac{r_1K_2 + r_2K_1 - \alpha_{12}r_1K_2 - \alpha_{21}r_2K_1}{K_1K_2(r_1-r_2) + K_1N(t)(r_2-\alpha_{12}r_1)}.$$

As discussed in the beginning of this section, to model the stochastic element, we can add a Fisher-Wright noise. Furthermore, if we assume a constant population size, $N$, the parameters

$\sigma$ and $\mu$ become constants, and we can describe both populations completely with a single SDE, namely,

$$dp = \sigma p(t)(1 - p(t))(1 - \mu p(t))dt + \sqrt{p(t)(1 - p(t))}dW(t), \tag{1.7}$$

We have seen that it is possible to model a population through SDEs, but we cannot add space to such a model by replacing the SDEs with SPDEs, if we want to model populations in two or more dimensions. However, we can couple multiple SDEs together, one for each of a countable number of locations in space. We can consider the SDE (1.7) as describing what happens on a single 'island' of population, an island in which each individual of each population competes equally with each other. However, we can allow individuals to migrate from island to island at a certain rate, and these migrations are what couple the system of SDEs together in space. A coupling of a system of dimensionless models through a spatial lattice is known as a stepping stone model.

We take a lattice in $\mathbb{Z}^{\mathsf{d}}$, and let $m_{ij} : \mathbb{Z}^{\mathsf{d}} \times \mathbb{Z}^{\mathsf{d}} \to [0, \infty)$ be a non-negative function of $|i - j|$ alone, with finite range, i.e., there exists an $R$ such that $m_{ij} = 0$ for $|i - j| > R$.

**Definition 1.2** (Simple Stepping stone model). *Let* $\{p(t)\}_{t\geq 0} = \{p_i(t), i \in \mathbb{Z}^{\mathsf{d}}\}_{t\geq 0}$ *satisfy*

$$dp_i(t) = \sum_{j\in\mathbb{Z}^{\mathsf{d}}} m_{ij}(p_j(t)-p_i(t))dt+\sigma p_i(t)(1-p_i(t))(1-\mu p_i(t))dt+\sqrt{p_i(t)(1-p_i(t))}dW_i(t), \tag{1.8}$$

*for constants* $\sigma, \mu \in \mathbb{R}$ *and where* $\{W_i(t), i \in \mathbb{Z}^{\mathsf{d}}\}_{t\geq 0}$ *is a family of independent Brownian motions. Then we say* $p(t)$ *evolves according to the simple stepping stone model.*

This model is a special case of one studied by Shiga and Shimizu [39]. They were able to prove that if $p(0) \in [0, 1]^{\mathbb{Z}^{\mathsf{d}}}$, then the simple stepping stone model has a continuous, pathwise unique strong solution in $[0, 1]^{\mathbb{Z}^{\mathsf{d}}}$ for all $t \geq 0$.

The authors of [6] studied this model because they were interested in developing population models where populations interact, but where one population is not guaranteed to eventually go extinct, a property that many simpler models do not enjoy. Indeed, they were able to prove that, under certain parameters, with positive probability, both populations in this model will coexist for all time (Theorem 1.4 of [6]).

We will focus on the symmetric case, where $r_1 = r_2$, $\alpha_{12} = \alpha_{21}$ and $K_1 = K_2$. In particular, this reduces $s$ to $r_1N(1-\alpha_{12})$ and $\mu = 2$. We make this assumption because it allows us to exploit a duality of this process which only occurs in this symmetric case. We will elaborate on duality and this dual process in the next section, but before that, we mention another consequence of this assumption.

In this symmetric case, the non-linear term in (1.8) is the same as that of (1.4), but with a sign change. Repeating the heuristic calculation in (1.3) reveals that this non-linearity is a first order approximation of sampling our types as if they were alleles from a diploid population with selection of strength $\sigma$ *in favour* of the heterozygosity. Hence, although we constructed this model for two interacting populations, with no assumptions on their genetic make up, this model can also be used to describe the evolution of a population experiencing heterozygous selection.

## 1.2 Duality for Stochastic Processes

Often, when studying a mathematical problem, insight can be found by studying a sightly different problem related to the original problem. In this thesis, we will make repeated use of the mathematical concept of duality, which links different Markov processes together, and allows us to make conclusions about the one Markov process by studying the other. We make this concept formal with a definition, from [24].

**Definition 1.3.** *Let $X$ and $Y$ be Markov processes on polish state spaces $E$ and $F$, respectively, and let $H : E \times F \to \mathbb{R}$ be bounded and measurable. Then $X$ and $Y$ are dual with respect to $H$ if and only if for all $x \in E$, $y \in F$ and $t \geq 0$,*

$$\mathbb{E}_x H(X_t, y) = \mathbb{E}^y H(x, Y_t),$$

*where the expectation on the left is with respect to the process $X$, with initial value $X_0 = x$, and the similarly, the expectation on the right is with respect to the process $Y$, with initial value $Y_0 = y$. We say the process $Y$ is* dual *to the process $X$ (and vice versa), and call $H$ the* duality relation*.*

In this section, we introduce the processes which are dual to the process introduced in the last section, the SΛFVA and the simple stepping stone model.

### 1.2.1 A branching and coalescing dual for the spatial $\Lambda$-Fleming-Viot process

Forward in time, the S$\Lambda$FVA records the reproductive events occurring to lineages in the population. In the dual process, we essentially try to construct the lineages backwards in time. This is relatively straightforward for non-selective events, but selective events give three options for the possible ancestor, so in order to determine the genotype of the lineage, we need to follow each possible parent back in time to the initial time 0. At this time, we can determine the types of the ancestors by sampling according to the initial distribution of types, $w_0$, and we can use this information to resolve the types of the ancestors' descendants.

The selective events cause branches in the graph of potential ancestors for a given individual, but, as reproductive events can result in multiple offspring, the lineages of these potential parents can merge. Hence, this backwards in time lineage reconstruction results in a branching and coalescing process.

We are about to make this dual process more precise by stating the definition. Before doing so, we point out that we will state the definition in the natural direction of time for the dual process, which is opposite to the original process. Hence, the individuals who are potential parents in the original process will be referred to an 'offspring' in the dual process, as they will be the 'younger' particles when time is reversed. Although this naming convention may cause some confusion when linking the dual process back to the S$\Lambda$FVA, we will mostly work with the dual process, so using the language natural for this reversed direction of time will be ultimately more straightforward.

We define the dual to the scaled version of the S$\Lambda$FVA. The dual is driven by the same (but time reversed) scaled Poisson point process of 'events' that drives the S$\Lambda$FVA, which we denoted $\Pi_n$ in (1.5).

**Definition 1.4** (S$\Lambda$FVA dual)**.** *For $n \in \mathbb{N}$, the process $(\mathcal{P}_t^n)_{t\geq 0}$ is the $\bigcup_{l\geq 1}(\mathbb{R}^d)^l$-valued Markov process with dynamics defined as follows.*

*The process starts with a single individual, $\mathcal{P}_0^n = x$, and for any time $t \geq 0$, $\mathcal{P}_t^n = (\xi_1^n(t), \ldots, \xi_{N(t)}^n(t))$ for some $N(t) \in \mathbb{N}$. At each event $(t, x, r) \in \Pi^n$, independently of all else, the event is said to be neutral with probability $1 - \boldsymbol{s}_n$. In this case:*

1. *For each $\xi_i^n(t-) \in \mathcal{B}_r(x)$, independently mark the corresponding individual with probability $u_n$;*

2. *if at least one individual is marked, all marked individuals coalesce into a single offspring individual, whose location is drawn uniformly at random from within $\mathcal{B}_r(x)$.*

*With the complementary probability $\boldsymbol{s}_n$, the event is said to be selective, in which case:*

1. *For each $\xi_i^n(t-) \in \mathcal{B}_r(x)$, independently mark the corresponding individual with probability $u_n$;*

2. *if at least one individual is marked, all of the marked individuals are replaced by* three *offspring individuals, whose locations are drawn independently and uniformly from within $\mathcal{B}_r(x)$.*

*In both cases, if no individual is marked, then nothing happens.*

The duality is between the SΛFVA and the whole *historical process* of branching and coalescing lineages, which we define as

$$\Xi^n(t) := (\mathcal{P}_s^n)_{0 \le s \le t}.$$

For $\mathbf{i} \in \{1,2,3\}^{\mathbb{N}}$ we let $(\xi_{\mathbf{i}}^n(\cdot))_{0 \le s \le t} \subseteq \Xi(t)$ denote the $\mathbb{R}^{\mathsf{d}}$-valued jump process corresponding to the locations where reproductive events affect an individual in $\mathcal{P}_s^n$. We refer to $(\xi_{\mathbf{i}}^n(\cdot))_{0 \le s \le t}$ as an ancestral lineage.

The lack of a parameter corresponding to $\gamma$ in this definition may alarm some readers. However, this is because this parameter is accounted for by the duality relation, from Definition 1.3. We explicitly give this duality relation now. We refer to this relation as a voting procedure, and it is easiest to define algorithmically, tracing backward through time.

Let $p : \mathbb{R}^{\mathsf{d}} \to [0,1]$ be a fixed function, describing the proportion of the less fit individual at time 0. Initially (i.e., at time $t$), for each $j \le N(t)$, the individual $\xi_j^n(t)$ votes 1 with probability $p(\xi_j^n(t))$ and votes 0 otherwise. Each vote is independent. Tracing backward through time,

1. at each neutral event, all individuals that are marked in the event adopt the vote of the offspring individual of the event;

2. at each selective event in $\Pi^n$, all individuals that are marked in the event adopt the majority vote of the votes of the three offspring individuals of the event, unless precisely one vote is 0, in which case it votes 1 with probability $\frac{3+\gamma}{3+3\gamma}$.

We define the duality relation via this iterative procedure, running from the 'leaves' of $\Xi^n$ at time $t$ to the ancestral individual $\emptyset$ at time 0.

**Definition 1.5** ($\mathbb{V}_p^\gamma$)**.** *With the voting procedure described above, we define* $\mathbb{V}_p^\gamma(\Xi^n(t))$ *to be the vote associated to the root* $\emptyset$*.*

Recall that, by its definition via (1.2), the SΛFVA is a density, and hence can only be defined pointwise Lebesgue a.e., so we cannot necessarily define $w_t^n(x)$ for any fixed point $x \in \mathbb{R}^{\mathsf{d}}$. Hence, in the duality relation, we treat the SΛFVA as a density, and integrate it against $\psi \in C(\mathbb{R}^{\mathsf{d}}) \cap L^1(\mathbb{R}^{\mathsf{d}})$.

**Theorem 1.6.** *The spatial Λ-Fleming-Viot process with asymmetric selection driven by* $\Pi^n$*,* $(w_t^n(x), x \in \mathbb{R}^{\mathsf{d}})_{t\geq 0}$*, is dual to the historical process* $(\Xi^n(t))_{t\geq 0}$ *in the sense that for every* $\psi \in C(\mathbb{R}^{\mathsf{d}}) \cap L^1(\mathbb{R}^{\mathsf{d}})$*, we have*

$$\mathbb{E}_p\left[\int_{\mathbb{R}^{\mathsf{d}}} \psi(x) w_t^n(x)\, dx\right] = \int_{\mathbb{R}^{\mathsf{d}}} \psi(x) \mathbb{E}_x\left[\mathbb{V}_p^\gamma\left(\Xi^n(t)\right)\right] dx = \int_{\mathbb{R}^{\mathsf{d}}} \psi(x) \mathbb{P}_x\left[\mathbb{V}_p^\gamma\left(\Xi^n(t)\right) = 1\right] dx,$$

*where* $\mathbb{P}_x$ *is the law of* $\Xi^n$ *when* $\mathcal{P}_0^n$ *is the single point* $x$ *and* $\mathbb{E}_x$ *is the corresponding expectation. Similarly,* $\mathbb{E}_p$ *is the expectation for the law of the SΛFVA with initial distribution satisfying*

$$\int_{\mathbb{R}^{\mathsf{d}}} \psi(x) w_0^n(x) dx = \int_{\mathbb{R}^{\mathsf{d}}} \psi(x) p(x) dx,$$

*for all* $\psi \in C(\mathbb{R}^{\mathsf{d}}) \cap L^1(\mathbb{R}^{\mathsf{d}})$*.*

Again, we appeal to [15] for a formal proof of this theorem. They do not prove this relation exactly, but the proof of Theorem 1.6 is a simple extension of their proof, so we do not include it here. Furthermore, we illustrate the proof of a further extension of this result in Chapter 3, so another reason to omit this proof is to avoid repetition of that proof.

The advantage of working with this dual process, rather than the original SΛFVA process, is precisely because the duality relation separates the evolution of the process from the effect of the asymmetry. This means we can work with the relatively symmetric and well-behaved $\Xi^n$ to determine the locations of the ancestors, and we can exploit properties of the independent voting procedure to determine the genotypes of those ancestors. Hopefully this becomes clear as we illustrate the proof in Chapter 2.

### 1.2.2 The branching annihilating random walk

When $\mu = 2$, the simple stepping stone model also has a dual process, which we shall call the branching annihilating random walk.

**Definition 1.7.** *The Markov process* $(N_i(t), i \in \mathbb{Z}^{\mathsf{d}})_{t \geq 0}$ *with values* $N_i(t) \in \mathbb{N}_0$ *and dynamics described by*

$$\begin{aligned}
&\left.\begin{aligned} N_i &\to N_i - 1 \\ N_j &\to N_j + 1 \end{aligned}\right\} && \text{at rate} \quad N_i m_{ij} \\
&N_i \to N_i + 2 && \text{at rate} \quad \sigma N_i \\
&N_i \to N_i - 2 && \text{at rate} \quad \frac{1}{2} N_i (N_i - 1)
\end{aligned}$$

*and an initial configuration* $(N_i(0))_{i \in \mathbb{Z}}$ *such that* $N_t = \sum_{i \in \mathbb{Z}} N_i(0) < \infty$ *is a* delayed, double-branching annihilating random walk *with branching rate* $\sigma$*. As we don't consider instantaneous branching, nor single-branching annihilating walks, we will refer to this version simply as a* branching annihilating random walk (BARW)*.*

At first, it may seem a bit unusual that when the particles in the BARW branch, *two* new particles are created, rather than just one. However, this behaviour results in the pleasing property of parity preservation; particles are created and removed in pairs, so if the initial number of particles is even (or odd), it will remain even (or odd) throughout the duration of the model. Hence, extinction can only occur if the initial number of particles is even.

Before we state the duality relation, we linearly change the variable in (1.8) from $p_i(t)$ to $x_i(t) = 1 - 2p_i(t)$, so that the solution remains between $-1$ and 1, rather than 0 and 1. Under this transformation, (1.8) becomes

$$dx_i(t) = \sum_{j \in \mathbb{Z}^{\mathsf{d}}} m_{ij}(x_j(t) - x_i(t))dt + \frac{\sigma}{2}(x_i^3(t) - x_i(t))dt + \sqrt{1 - x_i^2(t)}dW_i(t), \ \forall i \in \mathbb{Z}^{\mathsf{d}}. \tag{1.9}$$

We can now state the duality relation.

**Lemma 1.8.** *The solution to the system of equations in* (1.9) *is dual to the branching annihi-*

*lating random walk with branching rate* $\frac{\sigma}{2}$, $\{N_i(t), i \in \mathbb{Z}^{\mathsf{d}}\}_{t\geq 0}$ *through the duality relation*

$$\mathbb{E}\Big[\underline{x}(t)^{\underline{N}(0)}\Big] = \mathbb{E}\Big[\underline{x}(0)^{\underline{N}(t)}\Big] \tag{1.10}$$

*where*

$$\underline{x}^{\underline{N}} := \prod_{i\in\mathbb{Z}^{\mathsf{d}}} x_i^{N_i}.$$

The proof of this duality relation is well know; it is a special case of the results of [36]. However, as it is not too involved, we briefly repeat the proof here.

*Proof.* The key to many duality proofs relating two stochastic processes is to show that the generators of the process and its dual, when applied to different arguments of a duality relation, are equal. In this case, the duality relation is $\mathcal{D}(\underline{N},\underline{x}) := \underline{x}^{\underline{N}}$. As a generator is essentially the derivative of the expectation of a stochastic process, (1.10) follows from the equivalence of the generators. For more of the technical details about this proof in particular, we appeal to classic texts, such as Corollary 4.4.13 of [17], and for more details about the existence and construction of dual process in general, we appeal to [24]. Here, we merely aim to show the equivalence of the generators.

We can represent the BARW as a measure-valued process, setting $\underline{N} := \sum_{i\in\mathbb{Z}^{\mathsf{d}}} N_i(t)\delta_i$. Hence, we can write the generator of the BARW as

$$\begin{aligned}\mathcal{G}f(\underline{N}) = &\sum_{i\in\mathbb{Z}^{\mathsf{d}}} N_i(t) \sum_{j\in\mathbb{Z}^{\mathsf{d}}} m_{ij} f(\underline{N} - \delta_j + \delta_i) - f(\underline{N}) + \sum_{i\in\mathbb{Z}^{\mathsf{d}}} \tfrac{\sigma}{2} N_i(f(\underline{N} + 2\delta_i) - f(\underline{N}) \\ &+ \sum_{i\in\mathbb{Z}^{\mathsf{d}}} \frac{N_i(t)(N_i(t)-1)}{2}(f(\underline{N} - 2\delta_i) - f(\underline{N})).\end{aligned}$$

Similarly, setting $\underline{x} := \{x_i(t) : i \in \mathbb{Z}^{\mathsf{d}}\}_{t\geq 0}$ as the solution to the system of differential equations in (1.9), we can calculate the generator of this process as

$$\mathcal{H}f(\underline{x}) = \sum_{i\in\mathbb{Z}^{\mathsf{d}}} \left[\sum_{j\in\mathbb{Z}^{\mathsf{d}}} m_{ij}(x_j(t) - x_i(t)) + \tfrac{\sigma}{2}(x_i^3(t) - x_i(t))\right] \frac{\partial f}{\partial x_i} - \tfrac{1}{2} \sum_{i\in\mathbb{Z}^{\mathsf{d}}} (1 - x_i^2(t)) \frac{\partial^2 f}{\partial x_i^2}$$

for a twice differentiable function $f$. Applying the generators $\mathcal{G}$ and $\mathcal{H}$ to the first and second arguments of the function $\mathcal{D}(\underline{N}, \underline{x})$ reveals that

$$\mathcal{G}\mathcal{D}(\cdot, \underline{x})(\underline{N}) = \mathcal{H}\mathcal{D}(\underline{N}, \cdot)(\underline{x}),$$

and hence the result follows. $\square$

This duality was used in [6] to extend their results about coexistence in the simple stepping stone model to a result about the BARW. A positive probability of coexistence for all time, starting from a configuration where each site has at least a proportion of each type uniformly bounded away from 0 means that $\underline{x}(t)$ never fixes to completely 1 or $-1$. Hence, the left side of (1.10) must be strictly less than one, starting from any non-empty configuration of $\underline{N}(0)$. If the BARW were to almost surely go extinct, then the right side of (1.10) would be equal to one, and by Lemma 1.8, this cannot happen. This gives the following corollary from their work.

**Corollary 1.9** (Corollary 1.6 from [6])**.** *There exists $s_0 > 0$ such that if $\sigma > \sigma_0$ the BARW, started from an even number of particles, has a positive probability of surviving for all time.*

Blath, Etheridge and Meredith prove their result by comparing their model to oriented percolation. Percolation is a good technique for proving the existence of critical model parameters like this, but is unfortunately notorious for excessively poor bounds on those values. This means if we were to attempt to calculate an explicit value for $\sigma_0$ using the methods in [6], it is likely that the value we would produce would be orders of magnitude larger than the infimum of values for $\sigma_0$ which would also satisfy Corollary 1.9. The explicit minimum value of $\sigma_0$ is of interest because of a conjecture made in [6].

**Conjecture 1.10** (Conjecture 2.2 of [6])**.** *For the one-dimensional simple stepping stone model, there is a critical value $\sigma_0 > 0$ such that the populations described by system* (1.9) *will both persist for all time with positive probability if and only if $\sigma > \sigma_0$. In two dimensions, the same result holds, but the critical value is $\sigma_0 = 0$. In three or more dimensions, this probability is positive if and only if $\sigma \geq 0$.*

Again, through the duality, we can reformulate these conjectures to be statements about the branching rate $\sigma$ required to ensure the persistence or extinction of the BARW.

In this thesis, we make a significant step towards verifying the second part of this conjecture. In particular, we find a coupling between the BARW and a branching Brownian motion, up to a finite, but arbitrary, time. This proof is the focus of Chapter 4. In that chapter, to aid the clarity of the poof, we specialise to the case where $m$ is the symmetric nearest-neighbour transition kernel. However, the same arguments should work for any symmetric finite range transition kernel.

## 1.3 Outline of results

As briefly mentioned earlier, Etheridge, Freeman and Penington studied a version of (1.4), the Allen-Cahn equation, using probabilistic methods, in [13]. They give a probabilistic proof of a result first observed by Allen and Cahn [1], and later made rigorous analytically by various authors (e.g. [8, 18]). The result is that, after suitable scaling, the solution to the Allen-Cahn equation converges to the indicator function of a set whose boundary evolves according to mean curvature flow. In Chapter 2, we investigate result further, and identify three possible regimes, with the results of [13] represented in only one of these regimes.

We perturb the Allen-Cahn equation studied in [13], by adding an asymmetry. Using similar probabilistic methods, we show that the solution to this asymmetric Allen-Cahn equation evolves according to a different boundary flow, which we call constant normal flow. Again, the analytical result is known by analysts, (for example, see [9]), but by giving a probabilistic proof of the result, we create a framework that will allow us to add noise, and ultimately allow us to study a stochastic version of the asymmetric Allen-Cahn equation, which we do in Chapter 3. Furthermore, we find that on the boundary of these two regimes, we find an additional boundary flow behaviour, which has aspects of both regimes. These results are stated precisely in Theorem 2.4.

At the end of Chapter 2, we connect this analytical result back to the S$\Lambda$FVA using the duality outlined in Theorem 1.6. This allows us to describe how the interface between homogeneous populations of each homozygote in the S$\Lambda$FVA evolves. We can consider this interface to be a hybrid zone of the population, as this will be the only place we find both homozygotes. Adapting Theorem 2.4 to this context, we see that if the asymmetry in the selection of the homozygotes is small, then the population evolves according to mean curvature flow. Similarly, if the asymmetry is large enough (but not necessarily that large), we show that the mean curvature flow is overwhelmed by a constant curvature flow in a direction toward the less fit homozygote. Furthermore, there is a strength of asymmetry which results in a constant curvature flow on the same scale as the mean curvature flow, resulting in a curvature flow with aspects of both flows. This result is summed up in Theorem 2.21, and suggests the limiting behaviour of described in [13] is perhaps not as robust as first though.

In Chapter 3, we investigate the fluctuations in the one-dimensional interface studied in Chapter 2. This was not studied in [13], so we study both the symmetric and asymmetric cases. The scaling used in this chapter does not match precisely to the scaling used in Chapter 2. However, it can be thought of as spanning the weaker asymmetry and intermediate regimes, but with a milder spatial scaling. This milder spatial rescaling means that, rather than the deterministic limits we saw in the last chapter, we see some fluctuations in the limit under this new rescaling.

We prove that the hybrid zone evolves according to a Brownian motion, with drift proportional to the asymmetry between the heterozygotes in the intermediate regime. This result is Theorem 3.2, and its proof builds on ideas from [20], which proves a similar result for a different stochastic version of the Allen-Cahn equation. This chapter is joint work with Raphaël Forien and Sarah Penington.

In Chapter 4, we turn our attention to the BARW. As mentioned in the previous section, we prove Theorem 4.1, which strongly supports Conjecture 1.10 in this chapter. Briefly, this result says we can fix a large time $T$ (or equivalently, a large number of particles $N$), and we can find branching rate $\sigma$ so that we can couple a BARW with this branching rate to a ternary branching Brownian motion (TBBM) until time $T$ (or until the branching Brownian motion has $N$ particles) with arbitrarily high probability.

This coupling clearly supports the conjecture, as the number of particles in a branching Brownian motion does not decrease, let alone go extinct. However, the coupling actually gets better, in that it will last longer or generate more particles, the *smaller* the branching rate is.

The idea behind the proof is that many of the annihilations in a BARW occur between sibling particles (i.e., the immediate offspring of a branching particle). We couple the BARW to the BBM by ignoring these counteracting branching and annihilation events. In the rare times that branching particles ‘escape’ their siblings, the lower the branching rate, the more time those particles have to continue separating, so that when they do branch again, their offspring have enough space to resolve their own branching events.

In order to prove this result, we need to adapt some arguments from [10] concerning the asymptotics of non-collision probabilities of random walks, in particular, Theorem 4.2. This is the aim of the first section of this chapter.

Both of these results demonstrate the complexity of models for population genetics. Perturbations of specific parameters, the strength of asymmetric selection between heterozygotes, $\gamma$, in the S$\Lambda$FVA, and the branching rate, $\sigma$, in the simple stepping stone model, both lead to qualitative, not just quantitative, differences in the long term behaviour of the populations in each model.

# Chapter 2

# Curvature flow and the Spatial $\Lambda$-Fleming-Viot Process with asymmetric selection

In this chapter we use duality to prove Theorem 2.21, which describes the long term behaviour of the scaled interface between regions of each homozygote in a population.

Rather than work with the S$\Lambda$FVA dual directly, we first prove a version of our main result for a simpler process known as historical ternary branching Brownian motion. Indeed, ignoring the possibility of coalescences, the S$\Lambda$FVA dual is a ternary branching process, and it should not be surprising that the mostly independent events driving the trajectory of the ancestral locations can be approximated by Brownian motions. In Section 2.6 we make the coupling between these processes precise.

## 2.1 Asymmetric Allen-Cahn equation and constant normal flow

Before introducing the perturbation we will study in this chapter, we briefly introduce the mean curvature flow from [13].

**Definition 2.1** (Mean curvature flow)**.** *Let* $\mathbb{d} \geq 2$ *and let* $S^{\mathbb{d}-1}$ *denote the unit* $\mathbb{d}-1$ *sphere in* $\mathbb{R}^{\mathbb{d}}$*. Let* $\mathbf{\Gamma} = (\mathbf{\Gamma}_t(\cdot))_t$ *be a family of smooth embeddings, indexed by* $t \in [0, \mathscr{T})$*, where for each* $t$*,* $\mathbf{\Gamma}_t : S^{\mathbb{d}-1} \to \mathbb{R}^{\mathbb{d}}$.

*For any $t$, take an orthonormal basis of the tangent space of the hypersurface $\mathbf{\Gamma}_t$ and form the matrix whose $(i,j)^{th}$ entry is the dot product of the unit normal to the hypersurface with the derivative of the $i^{th}$ vector in the basis in the direction of the $j^{th}$ (this matrix is know as the second fundamental form). The trace of this matrix is known as the* (scalar) mean curvature *of this hypersurface.*

*Let $\boldsymbol{n} = \boldsymbol{n}_t(\phi)$ denote the unit (inward) normal vector to $\mathbf{\Gamma}_t(\phi)$ at $\phi$ and let $\kappa_t(\phi)$ denote the mean curvature of $\mathbf{\Gamma}_t$ at $\phi$. We say that $\mathbf{\Gamma}$ is a* mean curvature flow *if, for all $t, \phi$,*

$$\frac{\partial \mathbf{\Gamma}_t(\phi)}{\partial t} = \kappa_t(\phi)\boldsymbol{n}_t(\phi).$$

Mean curvature flow is well understood, particularly in two dimensions, where mean curvature is simply curvature. In two dimensions, [21] shows that any smooth convex closed curve has a finite lifetime, $\mathscr{T} \geq 0$, and shrinks to a point at this time, when evolving according to curvature flow. Furthermore, [22] extended this to nonconvex curves, and showed that any smooth closed curve eventually becomes convex under curvature flow, at which point the results of [21] apply.

Mean curvature flow is more complicated in higher dimensions. In this case, mean curvature flow still has a finite lifetime, and [23] showed that starting from a compact convex set, the flow is defined up until surface shrinks to a single point, and in particular, $\mathscr{T} > 0$. However, these results do not extend to nonconvex surfaces, as singularities can form before the enclosed volume disappears.

Our interest is not in studying mean curvature flow itself, so like [9] and [13] before us, we will assume initial conditions sufficient to ensure that the solution exists for a positive time and stopping before we encounter any singularities. In particular, our surfaces will be sufficiently smooth to allow us to represent the mean curvature of the surface by the second derivative of the surface at that point.

We now introduce the perturbed version of the symmetric Allen-Cahn equation we will be studying in this chapter and its scaling. Following (1.6), we introduce the asymmetry through the non-linear term, yielding

$$\frac{\partial v}{\partial t} = \Delta v + \frac{\mathbf{s}}{1+\gamma} v(1-v)(2v-(1+\gamma)).$$

Each of the terms of the left hand side will be scaled, as well as the asymmetry, and the coefficient of the polynomial term, $\frac{\mathsf{s}}{1+\gamma}$ is absorbed into this rescaling. Let $\epsilon > 0$. We will denote the scaled asymmetry by $\gamma_\epsilon$, and we shall assume that we have $\gamma_\epsilon = \tilde{\nu}\epsilon^{\tilde{\alpha}}$, where $\tilde{\alpha} \geq 0$ and $\tilde{\nu} \geq 0$. It is necessary to also include the condition that $\tilde{\nu} < 1$ if $\tilde{\alpha} = 0$.

The scaling we will use depends on how large the asymmetry is; we can break up our scaling into three different regimes. If $\tilde{\alpha} > 2$, then the asymmetry decays away quickly, and we see mean curvature flow in this case. When $\tilde{\alpha} < 1$, then the asymmetry is quite large, and it is in this case we see constant normal flow. The intermediate case, $\tilde{\alpha} \in (1, 2]$, is more sensitive. We essentially show that there is a constant normal component to the flow in this case, however, this component decays with $\epsilon$, and hence we only see mean curvature flow in the limit. The boundary case, when $\tilde{\alpha} = 1$ is the case where the constant normal component does not decay, but does not overwhelm the mean curvature flow, which results in the mixed curvature flow.

Rather than treat these cases separately, we define notation which allows us to deal with all the cases simultaneously. We define $\alpha = \min(\tilde{\alpha}, 1)$, i.e., $\alpha$ is 1 except in the case corresponding to the constant normal flow, in which case it is $\tilde{\alpha}$. We also define

$$\nu_\epsilon = \begin{cases} \tilde{\nu} & \text{if } \tilde{\alpha} \leq 1 \\ \gamma_\epsilon \epsilon^{-1} & \text{if } \tilde{\alpha} \in (1, 2] \\ 0 & \text{if } \tilde{\alpha} > 2, \end{cases}$$

which can be thought of as a kind of pre-limiting normal flow speed, as it equals or tends to 0 in the cases which result in mean curvature flow, and gives the coefficient of the constant normal flow in the remaining cases.

Let $v^\epsilon_\gamma : \mathbb{R} \times \mathbb{R}^{\mathsf{d}} \to [0, 1]$. The particular scaled version of the Allen-Cahn equation that we will be studying is

$$\frac{\partial v^\epsilon_\gamma}{\partial t} = \epsilon^{1-\alpha} \Delta v^\epsilon_\gamma + \epsilon^{-1-\alpha} v^\epsilon_\gamma (1 - v^\epsilon_\gamma)(2v^\epsilon_\gamma - (1 + \gamma_\epsilon)), \qquad v^\epsilon_\gamma(0, x) = p(x), \tag{2.1}$$

for some initial condition, $p$, which is assumed to take values in $[0, 1]$. Note that the case $\gamma_\epsilon \equiv 0$, $\alpha = 1$ gives the case studied in [13]. We require that the initial condition $p$ satisfies the following assumptions. Set

$$\Gamma = \left\{ x \in \mathbb{R}^{\mathbb{d}} : p(x) = \frac{1+\gamma_\epsilon}{2} \right\}.$$

We assume

($\mathscr{C}0$) $\Gamma$ is the boundary of a bounded open set which is topologically equivalent to the sphere.

($\mathscr{C}1$) $\Gamma$ is $C^a$ for some $a > 3$.

($\mathscr{C}2$) For $x$ on one side of $\Gamma$ (either inside or outside), $p(x) < \frac{1+\gamma_\epsilon}{2}$. For $x$ on the other side of $\Gamma$, $p(x) > \frac{1+\gamma_\epsilon}{2}$.

($\mathscr{C}3$) There exist $r, \lambda > 0$ such that, for all $x \in \mathbb{R}^{\mathbb{d}}$, $|p(x) - \frac{1+\gamma_\epsilon}{2}| \geq \lambda \left( \operatorname{dist}(x, \Gamma) \wedge r \right)$.

The choice to define the boundary as the $\Gamma = \frac{1+\gamma_\epsilon}{2}$ level set, rather than, say, the $\Gamma = \frac{1}{2}$ level set may seem arbitrary, but it is chosen because $\frac{1+\gamma_\epsilon}{2}$, along with 0 and 1, are the only three constant solutions to (2.1). The solutions at 0 and 1 are stable, whereas $v_\gamma^\epsilon(x) = \frac{1+\gamma_\epsilon}{2}$ is unstable, so this level set is a sensible choice for the boundary between regions corresponding to the two other constant solutions. Condition ($\mathscr{C}3$) gives the existence of a minimum slope of $\Gamma$ near the interface.

The conditions we have imposed ensure that the constant normal flow, which we are about to define formally, starting from $\Gamma$, exists, at least up until some time $\mathscr{T} > 0$, as shown in, for example, [9]. Denote the flow by $(\mathbf{\Gamma}_t(\cdot))$ and let $d(x,t)$ be the signed distance from $x$ to $\mathbf{\Gamma}_t$, chosen to be negative on the side of $\mathbf{\Gamma}_t$ which is greater than $\frac{1+\gamma_\epsilon}{2}$ and positive on the other side. Note that, as sets,

$$\mathbf{\Gamma}_t = \{x \in \mathbb{R}^{\mathbb{d}} : d(x,t) = 0\}.$$

**Definition 2.2** (Constant normal flow). *Let* $\mathbb{d} \geq 2$ *and let* $S^{\mathbb{d}-1}$ *denote the unit* $\mathbb{d}-1$ *sphere in* $\mathbb{R}^{\mathbb{d}}$. *Let* $\mathbf{\Gamma} = (\mathbf{\Gamma}_t(\cdot))_t$ *be a family of smooth embeddings, indexed by* $t \in [0, \mathscr{T})$, *where for each* $t$, $\mathbf{\Gamma}_t : S^{\mathbb{d}-1} \to \mathbb{R}^{\mathbb{d}}$. *Let* $\boldsymbol{n} = \boldsymbol{n}_t(\phi)$ *denote the unit normal vector to* $\mathbf{\Gamma}_t(\phi)$, *oriented towards the region with negative signed distance from* $\mathbf{\Gamma}_t$. *We say that* $\mathbf{\Gamma}$ *is a* constant normal flow with speed $c \geq 0$ *if, for all* $t, \phi$,

$$\frac{\partial \mathbf{\Gamma}_t(\phi)}{\partial t} = c\boldsymbol{n}_t(\phi).$$

Unlike mean curvature flow, which in two dimensions exists until the curve shrinks to a single point [21], constant normal flow may cease to exist before reaching a single point. In particular, if the region of negative signed distance from the interface is compact, constant normal flow cannot exist beyond time $ck$, where $k$ is the maximum curvature of any point on the curve. However, if the region of negative signed distance is not compact, the flow may exist for all time; an outcome which is also not possible for mean curvature flow. This behaviour is trivially demonstrated by considering a circle with constant normal flow in the external direction.

We wish to include the possibility of a surface evolving according to a linear combination of mean curvature flow and constant normal flow. we call this flow mixed curvature flow.

**Definition 2.3** (Mixed curvature flow)**.** *Let* $\mathbf{\Gamma}$ *be as defined in Definition 2.1, and* $c \in \mathbb{R}$*. We say that* $\mathbf{\Gamma}$ *is a* mixed curvature flow *if, for all* $t, \phi$*,*

$$\frac{\partial \mathbf{\Gamma}_t(\phi)}{\partial t} = (c + \kappa_t(\phi))\boldsymbol{n}_t(\phi). \tag{2.2}$$

As $c$ can be 0, mean curvature flow is a specific kind of mixed curvature flow, but the same is not true for constant curvature flow. However, we shall abuse terminology slightly and used mixed curvature flow to refer to mean curvature, constant normal and mixed curvature flow, as defined above.

Suppose that $\mathbb{d} \geq 2$. The result we wish to show is that, as $\epsilon \to 0$, the solution of (2.1) with an initial condition satisfying the assumptions ($\mathscr{C}0$)-($\mathscr{C}3$) converges to the indicator function of a set whose boundary evolves according to mixed curvature flow.

**Theorem 2.4.** *Let* $v_\gamma^\epsilon$ *solve* (2.1) *with initial condition* $p$ *satisfying the conditions* ($\mathscr{C}0$)-($\mathscr{C}3$)*, and let* $\mathbf{\Gamma} = (\mathbf{\Gamma}_t(\cdot))_t$ *be a mixed curvature flow existing until time* $\mathscr{T}$*, with signed distance function* $d$*, starting from the set* $\{x \in \mathbb{R}^{\mathbb{d}} : p(x) = \frac{1+\gamma_\epsilon}{2}\}$ *and satisfying*

$$\frac{\partial \mathbf{\Gamma}_t(\phi)}{\partial t} = \nu_\epsilon \boldsymbol{n}_t(\phi) + \begin{cases} 0 & \text{if } \tilde{\alpha} \in [0,1) \\ \kappa_t(\phi)\boldsymbol{n}_t(\phi) & \text{if } \tilde{\alpha} \geq 1 \end{cases} \tag{2.3}$$

*for all* $t, \phi$*. Let* $k \in \mathbb{N}$*. There exists a* $T^* \in (0, \mathscr{T})$*, an* $\epsilon_{\mathbb{d}}(k) > 0$*, and* $a_{\mathbb{d}}(k), c_{\mathbb{d}}(k) \in (0, \infty)$ *such that for all* $\epsilon \in (0, \epsilon_{\mathbb{d}})$ *and* $t$ *satisfying* $a_{\mathbb{d}}\epsilon^{1+\alpha}|\log \epsilon| \leq t \leq T^*$*,*

1. *for* $x$ *such that* $d(x,t) \geq c_{\mathbb{d}}\epsilon|\log\epsilon|$*, we have* $v_\gamma^\epsilon(t,x) \geq 1 - \epsilon^k$*;*

*2. for $x$ such that $d(x,t) \le -c_{\mathsf{d}}\epsilon|\log\epsilon|$, we have $v^{\epsilon}_{\gamma}(t,x) \le \epsilon^k$.*

We prove this theorem using a similar method to that used in [13]. As constant normal flow happens on a faster time scale than the mean curvature flow observed in [13], one of the major tasks is adapting the proofs in [13] to a more general time scale. The other major difference is that Etheridge, Freeman and Penington probabilistically prove a one-dimensional version of Theorem 2.4. We avoid this by exploiting an explicit one-dimensional solution of (2.1).

## 2.2 A probabilistic dual to the asymmetric Allen-Cahn equation

In this section we illustrate the probabilistic dual, and how it helps prove Theorem 2.4. First, we note that, in order to maintain compatibility with the PDE literature, the authors of [13] adopt the convention that Brownian motions run at rate 2, as opposed to the probabilistic convention. However, our dual will run $\epsilon^{1-\alpha}$ times slower than the dual in their paper, so we will adopt the convention that

*all Brownian motions run at rate $2\epsilon^{1-\alpha}$.*

Otherwise, we adopt the same notation as used in [13], which we introduce now. We shall denote the set of Ulam-Harris labels by $\mathcal{U} = \{\emptyset\} \cup \bigcup_{k=1}^{\infty}\{1,2,3\}^k$. This set can be used to label the vertices of the infinite ternary tree, which is the graphical tree (connected graph with no cycles) which consists of a single vertex of degree 1, which we call the root, and every other vertex in the tree having degree 4. As the graph is a tree, there is a unique path from every vertex to the root, and this path can be used to distinguish exactly one of the four neighbours of any given vertex, $v$. This distinguished vertex can be considered the 'parent' of vertex $v$, and the remaining three vertices can be considered the 'offspring'. This number of offspring is the reason for the name. The root is given label $\emptyset$, and if vertex $v$ is given label $(i_1,\dots,i_n) \in \mathcal{U}$, its offspring are given labels $(i_1,\dots,i_n,1)$, $(i_1,\dots,i_n,2)$ and $(i_1,\dots,i_n,3)$.

**Definition 2.5.** *We say that $\mathcal{T}$ is a time-labelled ternary tree if $\mathcal{T}$ is a finite subtree of the infinite ternary tree, $\mathcal{U}$, and each internal vertex $v$ of the tree is labelled with a time $t_v > 0$, where $t_v$ is strictly greater than the label of the parent vertex of $v$.*

The death times of the particles in a ternary branching Brownian motion form a time-labelled ternary tree. Informally, a ternary branching Brownian motion is a set of particles performing independent Brownian motions, each for independent exponentially distributed lifetimes, and when one particle terminates, it is replaced by three new particles behaving in the same way. Making this definition more formal requires us to be more careful and explicit with our labels, which hopefully doesn't obscure this idea too much.

**Definition 2.6** (Ternary branching Brownian motion)**.** *For each $u \in \mathcal{U}$, let $h_u$ be an exponentially distributed random variable with parameter $\sigma > 0$, and inductively define $t_\emptyset = h_\emptyset$ and $t_u = t_v + h_u$ if $u = (v, i)$, for some $v \in \mathcal{U}$ and $i \in \{1, 2, 3\}$.*

*We define a* $\mathbb{d}$-dimensional ternary branching Brownian motion *with branching rate $\sigma$ starting at $x \in \mathbb{R}^{\mathbb{d}}$ to be a process $\overline{W}(t) = (W_u(t), t_u - h_u \leq t < t_u)_{u\in\mathcal{U}}$ which we define inductively. First, $(W_\emptyset(t))_{t=0}^{t_\emptyset}$ is a Brownian motion, running at rate $2\epsilon^{1-\alpha}$, starting at $x$. Now, assuming we have defined $t_v$ and $(W_v(t))_{t=t_v-h_v}^{t_v}$ for some $v \in \mathcal{U}$, we let $u = (v, i)$ for some $i \in \{1, 2, 3\}$ and define $(W_u(t))_{t=t_v}^{t_u}$ to be a Brownian motion, running at rate $2\epsilon^{1-\alpha}$, starting at $W_v(t_v-)$. Furthermore, let $\mathcal{U}(t) = \{u \in \mathcal{U} : t_u - h_u \leq t < t_u\}$ be the set of particles alive at time $t$.*

A historical ternary branching Brownian motion is the whole ternary branching Brownian motion process, as a collection of paths, up until a given time. In the previous definition, we only needed to introduce a new piece of Brownian motion between a parent particle and its offspring, however we consider historical branching Brownian motion to be a collection of $\mathcal{U}(t)$ paths, and for each particle alive at time $t$, we concatenate their path to the path traced out by all their ancestors.

**Definition 2.7** (Historical ternary branching Brownian motion)**.** *A $\mathbb{d}$-dimensional* historical *ternary branching Brownian motion with branching rate $\sigma$ starting at $x \in \mathbb{R}^{\mathbb{d}}$ is a ternary branching Brownian motion with all the genealogical information from the branching structure included. Formally, it is a collection of paths $\boldsymbol{W}(t) = (W_u(s), 0 \leq s < t)_{u\in\mathcal{U}(t)}$ for some branching Brownian motion $\overline{W}(t)$. The path $W_u(s)$, for any $u \in \mathcal{U}(t)$, is defined for any time $s < t_u - h_u$, (i.e., before the particle was born), as the path $W_v(s)$, for the unique ancestor of $u$ alive at time $s$, or more formally, the unique $v \in \mathcal{U}$ such that $u = (v, w)$ for some $w \in \mathcal{U}$ and $t_v - h_v \leq s < t_v$ .*

*We shall abuse notation slightly and write $(W_i(s))_{0\leq s\leq t}$ for the unique path that connects leaf $i$ to the root. We use $\mathcal{T}(\boldsymbol{W}(t))$ to denote the time-labelled ternary tree defined by the branching structure of $\boldsymbol{W}(t)$.*

We make particular note here that the branching Brownian motions we will be using differ from those used in [13] in two ways. First, as mentioned above, our Brownian motion is running $\epsilon^{1-\alpha}$ times slower. Second, our branching Brownian motion branches at a rate similarly scaled by $\epsilon^{1-\alpha}$ relative to the rate considered in [13]. However, we additionally increase the branching rate, to account for the new 'type' of event, which occurs as a consequence of the introduced asymmetry. We shall discuss this more below. As the branching rate of the branching Brownian motions in [13] is $\epsilon^{-2}$, our branching rate will be $(1+\gamma_\epsilon)\epsilon^{-1-\alpha}$.

We return to outlining notation. For $x \in \mathbb{R}^{\mathbb{d}}$, we write $\mathbb{P}_x^\epsilon$ for the probability measure under which $(\mathbf{W}(t), t \geq 0)$ has the law of the historical process of ternary branching Brownian motion in $\mathbb{R}^{\mathbb{d}}$ with Brownian motions running at rate $2\epsilon^{1-\alpha}$, branching rate $(1+\gamma_\epsilon)\epsilon^{-1-\alpha}$ and started from a single particle at location $x$ at time 0. We write $\mathbb{E}_x^\epsilon$ for the corresponding expectation. When we need to take expectations of a single Brownian motion, for clarity, we shall denote this expectation by $E$. It is similarly convenient to have a prominent distinction between one-dimensional and multidimensional Brownian motion in our notation. We therefore adopt the same convention as in [13], that $B$ will denote one-dimensional Brownian motion and $\mathbf{B}$ will represent the corresponding historical branching Brownian motion whereas $W$ and $\mathbf{W}$ will be used for Brownian motions and historical branching Brownian motions in dimensions $\mathbb{d} \geq 2$.

We now have the notation we need in order to set up our dual process. For a fixed function $p : \mathbb{R}^{\mathbb{d}} \to [0,1]$, we define a voting procedure on $\mathcal{T}(\mathbf{W}(t))$ as follows.

1. Each leaf $i$ of $\mathcal{T}(\mathbf{W}(t))$, independently, votes 1 with probability $p(W_i(t))$ and otherwise votes 0.

2. At each branch point in $\mathcal{T}(\mathbf{W}(t))$, the vote of the parent particle $j$ is the majority vote of the votes of its three children $(j,1)$, $(j,2)$ and $(j,3)$, unless precisely one vote is 0, in which case it votes 1 with probability $\frac{3+\gamma_\epsilon}{3+3\gamma_\epsilon}$.

This defines an iterative voting procedure, which runs inwards from the leaves of $\mathcal{T}(\mathbf{W}(t))$ to the root $\emptyset$.

**Definition 2.8** ($\mathbb{V}_p^\gamma$). *With the voting procedure described above, we define* $\mathbb{V}_p^\gamma(\boldsymbol{W}(t))$ *to be the vote associated to the root* $\emptyset$*. Note that we have dropped the subscript of* $\gamma_\epsilon$ *in this notation, to make it easier to read.*

In the symmetric case, the voting procedure described in [13] was purely majority voting. We can think of our voting system as a combination of two voting strategies. As in [13], 'symmetric' events still arrive at rate $\epsilon^{-1-\alpha}$. Under these events, at a single branching event, the root takes the value of the majority of the leaves. The asymmetry introduces a new kind of biased branching event, occurring at rate $\gamma_\epsilon \epsilon^{-1-\alpha} = \nu \epsilon^{\max(-1,\alpha-2)}$. Under these events, with probability $\frac{2}{3}$, a non-unanimous majority of votes of value 1 is overruled in favour of a vote of 0.

The following theorem gives the duality between the solution to our version of the Allen-Cahn equation (2.1) and the voting model we have just described.

**Theorem 2.9.** *Let* $p : \mathbb{R}^{\mathrm{d}} \to [0,1]$. *Then*

$$v_\gamma^\epsilon(t,x) = \mathbb{P}_x^\epsilon \left[ \mathbb{V}_p^\gamma(\boldsymbol{W}(t)) = 1 \right]$$

*is a solution to equation* (2.1) *with initial condition* $v_\gamma^\epsilon(0,x) = p(x)$.

*Proof.* (Sketch) Throughout this proof we shall neglect the superscripts $\epsilon$ and $\gamma$ and the subscript $p$ in $\mathbb{P}_x^\epsilon$, $\mathbb{E}_x^\epsilon$, $v_\gamma^\epsilon$ and $\mathbb{V}_p^\gamma$. This proof is only slightly different from the equivalent proof, of Theorem 2.2, in [13], itself only a sketch of the adaptation of the representation of solutions of the Fisher-KPP equation in terms of binary branching Brownian motion. Hence, we also only sketch it here.

As usual, the idea is to investigate what happens to $v_\gamma^\epsilon(t,x)$ during a small time interval by conditioning on what happens to the branching Brownian motion during that time. Using the same notation as [13], we write $S$ for the time of the first branching event in the branching Brownian motion and $W_S$ for the position of the ancestor at that time. Let $V_1, V_2, V_3$ denote the votes of the three offspring created at time $S$. These $V_i$ are conditionally independent given $(S, W_S)$, by the strong Markov property of the branching Brownian motion. Also by the Markov property, conditional on $S \leq \delta t$, the chance of another branching event in the time interval $\delta t$ is $\mathcal{O}(\delta t)$, so for $s \leq \delta t$,

$$\mathbb{E}_x[V_1 | (S, W_S) = (s,y)] = E_y[v(t, W_{\delta t - s})] + \mathcal{O}(\delta t).$$

The heat semigroup has enough regularity to allow us to write

$$\mathbb{E}_x[V_1|S \leq \delta t] = v(t,x) + \mathcal{O}(\delta t). \tag{2.4}$$

If the vote at the root is one, either all $V_1, V_2, V_3$ are one, or precisely two were 1 and their majority was successfully passed on, which occurs with probability $\frac{3+\gamma_\epsilon}{3+3\gamma_\epsilon}$. Thus, using (2.4) and conditional independence of the $V_i$ given $(S, W_S)$,

$$\mathbb{P}_x\left[\mathbb{V}(\mathbf{W}(t+\delta t)) = 1|S \leq \delta t\right] = v(t,x)^3 + 3 \cdot \frac{3+\gamma_\epsilon}{3+3\gamma_\epsilon} v(t,x)^2(1-v(t,x)) + \mathcal{O}(\delta t).$$

If $S > \delta t$, the ancestor of the branching Brownian motion is just a Brownian motion during the interval $[0, \delta t]$. Conditioning on whether our ancestor particle branches during the initial $\delta t$ of time gives

$$\begin{aligned}
v(t+\delta t, x) =& \mathbb{P}_x\left[\mathbb{V}(\mathbf{W}(t+\delta t)) = 1|S \leq \delta t\right] \mathbb{P}\left[S \leq \delta t\right] \\
&+ \mathbb{P}_x\left[\mathbb{V}(\mathbf{W}(t+\delta t)) = 1 \,|\, S > \delta t\right] (1 - \mathbb{P}\left[S \leq \delta t\right]) \\
=& \mathbb{P}_x\left[\mathbb{V}(\mathbf{W}(t+\delta t)) = 1 \,|\, S \leq \delta t\right] \mathbb{P}\left[S \leq \delta t\right] \\
&+ E_x\left[\mathbb{P}_{W_{\delta t}}\left[\mathbb{V}(\mathbf{W}(t)) = 1\right]\right] (1 - \mathbb{P}\left[S \leq \delta t\right]).
\end{aligned}$$

Now $\mathbb{P}[S \leq \delta t] = (1+\gamma_\epsilon)\epsilon^{-1-\alpha}\delta t + \mathcal{O}(\delta t^2)$ and so substituting and rearranging (and once again assuming enough regularity of $v(t,x)$ given by the heat semigroup) we obtain

$$\begin{aligned}
\lim_{\delta t \to 0} \frac{v(t+\delta t, x) - v(t,x)}{\delta t} =& (1+\gamma_\epsilon)\epsilon^{-1-\alpha}\left(v(t,x)^3 + \frac{3+\gamma_\epsilon}{1+\gamma_\epsilon} v(t,x)^2(1-v(t,x)) - v(t,x)\right) \\
&+ \lim_{\delta t \to 0} \frac{E_x\left[\mathbb{P}_{W_{\delta t}}\left[\mathbb{V}(\mathbf{W}(t)) = 1\right]\right] - v(t,x)}{\delta t} \\
=& \epsilon^{-1-\alpha}\left(-2v(t,x)^3 + (3+\gamma_\epsilon)v(t,x)^2 - (1+\gamma_\epsilon)v(t,x)\right) \\
&+ \lim_{\delta t \to 0} \frac{E_x\left[v(t, W_{\delta t})\right] - v(t,x)}{\delta t} \\
=& \epsilon^{1-\alpha}\Delta v(t,x) + \epsilon^{-1-\alpha} v(t,x)(1-v(t,x))(2v(t,x) - (1+\gamma_\epsilon)),
\end{aligned}$$

where the $\epsilon^{1-\alpha}$ coefficient, rather than the usual coefficient of $\frac{1}{2}$ for the Laplacian occurs as the Brownian motion is running at rate $2\epsilon^{1-\alpha}$. ☐

This duality reduces the the proof of Theorem 2.4 to the following theorem about the voting system.

**Theorem 2.10.** *Let* $\gamma_\epsilon$ *be defined as above. Suppose* $p : \mathbb{R}^{\mathbb{d}} \to [0,1]$ *is such that* ($\mathscr{C}0$)-($\mathscr{C}3$) *hold. Define* $\mathscr{T}$, $d(x,t)$ *as for Theorem 2.4; fix* $T^* \in (0,\mathscr{T})$ *and let* $k \in \mathbb{N}$. *There exist* $\epsilon_{\mathbb{d}}(k) > 0$, *and* $a_{\mathbb{d}}(k), c_{\mathbb{d}}(k) \in (0,\infty)$ *such that for all* $\epsilon \in (0,\epsilon_{\mathbb{d}})$ *and* $t$ *satisfying* $a_{\mathbb{d}}\epsilon^{1+\alpha}|\log\epsilon| \le t \le T^*$,

1. *for* $x$ *such that* $d(x,t) \ge c_{\mathbb{d}}\epsilon|\log\epsilon|$, *we have* $\mathbb{P}^\epsilon_x\left[\mathbb{V}^\gamma_p(\boldsymbol{W}(t)) = 1\right] \ge 1 - \epsilon^k$;

2. *for* $x$ *such that* $d(x,t) \le -c_{\mathbb{d}}\epsilon|\log\epsilon|$, *we have* $\mathbb{P}^\epsilon_x\left[\mathbb{V}^\gamma_p(\boldsymbol{W}(t)) = 1\right] \le \epsilon^k$.

We prove Theorem 2.10 by comparing the multidimensional case to a particular explicit solution to (2.1) in one dimension. We outline this solution and some of its particular properties that we need in Section 2.3. Although it follows along the same lines as [13], this subsection is the most significantly different from [13].

In Section 2.4, we use regularity properties that follow from the conditions ($\mathscr{C}0$)-($\mathscr{C}3$) to couple the distance between the (backwards in time) constant normal flow $(\boldsymbol{\Gamma}_{t-s})_{s\in[0,t]}$ and a (forwards in time) $\mathbb{d}$-dimensional Brownian motion $W$ with a (forwards in time) one-dimensional Brownian motion $B$ in such a way that $d(W_s, t-s)$ is well approximated by $B_s$ when $W_s$ is close to $\boldsymbol{\Gamma}_{t-s}$. The differences between this section and [13] are quite subtle, but significant.

Section 2.5 contains the proof of Theorem 2.10. After setting up the previous two sections, this proof proceeds in a similar way to [13]. However, it is in this section we see how the changes we have made in Sections 2.10 and 2.4 were what we required, in order to adapt the equivalent proof in [13].

Finally, in Section 2.6, we return to the SΛFVA, and outline how Theorem 2.10 can be extended to the SΛFVA dual process.

## 2.3 An explicit solution to the one-dimensional Asymmetric Allen Cahn Equation

Although we wish to show that solutions to the asymmetric Allen-Cahn equation evolve according to a different kind of geometric flow than solutions to the symmetric version, we may still use a similar strategy of proof to show this result. This is because the proof for the symmetric

Allen-Cahn equation involves coupling the evolution of the solution in higher dimensions to a particular one-dimensional solution. Thus, most of the work in adapting this proof is done by showing the one-dimensional solution to the asymmetric Allen-Cahn equation has the necessary properties, and adapting the coupling between higher and lower-dimensional solutions. In this section we do the former, and in the following section we do the latter.

The properties we need our one-dimensional solution to have are that it needs to form an interface sufficiently quickly, and that interface needs to propagate at a constant speed (for a given $\epsilon$). However, this speed can be 0. It is not too hard to check by direct calculation that

$$\tilde{u}_\epsilon(t,x) = \left(1 + \exp(-\epsilon^{-1}x + \gamma_\epsilon \epsilon^{-1-\alpha}t)\right)^{-1} \tag{2.5}$$

is the unique solution to (2.1) with initial condition

$$\tilde{p}(x) = \left(1 + \exp(-\epsilon^{-1}x)\right)^{-1}. \tag{2.6}$$

This solution is a travelling wave, which moves with speed $\gamma_\epsilon \epsilon^{-\alpha}$. Furthermore, the wave separates a region to the left of the real line, where the solution is near 0, from the right, where the solution is near 1. This interface is sufficient for our purposes, with the advantage that it essentially forms instantaneously. This makes (2.5) an obvious candidate for the one-dimensional solution we require.

We now note more rigorously a number of properties of this solution. First, it is symmetric about $x = \gamma_\epsilon \epsilon^{-\alpha}t$, in that $\tilde{u}_\epsilon(t,x) = 1 - \tilde{u}_\epsilon(t, 2\gamma_\epsilon \epsilon^{-\alpha}t - x)$. Through direct calculation, a one-dimensional version of Theorem 2.10 with our particular initial condition can be shown. Recall from the paragraph before Definition 2.8, that we denote the one dimensional version of the historical ternary branching Brownian motion by $(B(t), t \geq 0)$, and we reserve the notation $(W(t), t \geq 0)$ for the higher dimensional version.

**Theorem 2.11.** *Let $T^* \in (0,\infty)$. For all $k \in \mathbb{N}$, for all $t \in [0,T^*]$ and sufficiently small $\epsilon > 0$,*

1. *for $z \geq k\epsilon|\log \epsilon| + \nu_\epsilon t$, we have $\mathbb{P}^\epsilon_z \left[\mathbb{V}^\gamma_{\tilde{p}}(\boldsymbol{B}(t)) = 1\right] \geq 1 - \epsilon^k$*
2. *for $z \leq -k\epsilon|\log \epsilon| + \nu_\epsilon t$, we have $\mathbb{P}^\epsilon_z \left[\mathbb{V}^\gamma_{\tilde{p}}(\boldsymbol{B}(t)) = 1\right] \leq \epsilon^k$.*

This theorem is the equivalent of Theorem 2.5 in [13]. The proof of this theorem in [13] is probabilistic, but relies heavily on the symmetry about the origin of the one-dimensional solution to the symmetric Allen-Cahn equation. We do not have such symmetry, so we exploit the explicitness of (2.5).

*Proof.* By Theorem 2.9, $\mathbb{P}^{\epsilon}_{z}\left[\mathbb{V}^{\gamma}_{\tilde{p}}(\mathbf{B}(t))=1\right]=\tilde{u}_{\epsilon}(t,x)$. For $z\geq k\epsilon|\log\epsilon|+\nu_{\epsilon}t$,

$$\begin{aligned}\tilde{u}_{\epsilon}(t,z)&\geq\left(1+\exp(-k|\log\epsilon|-\epsilon^{-1}\nu_{\epsilon}t+\gamma_{\epsilon}\epsilon^{-1-\alpha}t)\right)^{-1}\\&=\left(1+\epsilon^{k}e^{\epsilon^{-1}t(\gamma_{\epsilon}\epsilon^{-\alpha}-\nu_{\epsilon})}\right)^{-1}.\\&=1-\epsilon^{k}\frac{e^{\epsilon^{-1}t(\gamma_{\epsilon}\epsilon^{-\alpha}-\nu_{\epsilon})}}{1+\epsilon^{k}e^{\epsilon^{-1}t(\gamma_{\epsilon}\epsilon^{-\alpha}-\nu_{\epsilon})}}\end{aligned}$$

By our assumptions on the asymptotics of $\gamma_{\epsilon}$ and the definition of $\nu_{\epsilon}$, the expression $\gamma_{\epsilon}\epsilon^{-\alpha}-\nu_{\epsilon}$ goes to 0 faster than $\epsilon$. Hence, the exponent of $e$ approaches 0, and so for sufficiently small $\epsilon$, the coefficient of $\epsilon^{k}$ is less than 1, and the conclusion follows. The second inequality follows similarly. ☐

It would be more satisfying if, as in [13], we could prove this theorem probabilistically, rather than by exploiting the explicit solution. This should be possible, but would take more work. By exploiting our explicit solution, we can also prove the following proposition, the equivalent of Corollary 2.12 in [13], which gives us a bound on the slope of the interface.

**Proposition 2.12.** *If* $\tilde{\alpha}=0$, *let* $\epsilon<\exp\left(\frac{-9}{(11+\tilde{\nu})(1-\tilde{\nu})}\right)$, *otherwise, let*

$$\epsilon<\min\left(\inf\{\epsilon:\gamma_{\epsilon}=1/2\},\exp\left(\frac{-9}{(11+1/2)(1-1/2)}\right)\right).$$

*Suppose that for some* $t\in[0,T^{*}]$ *and* $z\in\mathbb{R}$,

$$\left|\mathbb{P}^{\epsilon}_{z}\left[\mathbb{V}^{\gamma}_{\tilde{p}}(\boldsymbol{B}(t))=1\right]-\tfrac{1}{2}\right|\leq\tfrac{5+\gamma_{\epsilon}}{12},\tag{2.7}$$

*and let* $w\in\mathbb{R}$ *with* $|z-w|\leq\epsilon$. *Then*

$$\left|\mathbb{P}^{\epsilon}_{z}\left[\mathbb{V}^{\gamma}_{\tilde{p}}(\boldsymbol{B}(t))=1\right]-\mathbb{P}^{\epsilon}_{w}\left[\mathbb{V}^{\gamma}_{\tilde{p}}(\boldsymbol{B}(t))=1\right]\right|\geq\frac{|z-w|}{48\epsilon|\log\epsilon|}.$$

*Proof.* As we have an explicit form for the probability in (2.7), we may rewrite the equation to become,

$$\left|z-\gamma_\epsilon\epsilon^{-\alpha}t\right|\le \epsilon\log\left(\frac{11+\gamma_\epsilon}{1-\gamma_\epsilon}\right).$$

Fix such a $z$ and let $w\in\mathbb{R}$ satisfy $|z-w|\le\epsilon$.

Notice that $\tilde{u}_\epsilon(t,x)=\tilde{p}(x-\gamma_\epsilon\epsilon^{-\alpha}t)$. It is easy to calculate $\tilde{p}'(x)=\frac{\epsilon^{-1}\exp(-\epsilon^{-1}x)}{(1+\exp(-\epsilon^{-1}x))^2}$ and thus for $|x|\le A\epsilon$, $|\tilde{p}'(x)|\ge\frac{\epsilon^{-1}\exp(A)}{(1+\exp(A))^2}$. Let $A=1+\log\left(\frac{11+\gamma_\epsilon}{1-\gamma_\epsilon}\right)$. Then, for $|x-\gamma_\epsilon\epsilon^{-\alpha}t|\le A\epsilon$,

$$\begin{aligned}\left|\frac{d\tilde{u}(t,x)}{dx}\right|\ge&\frac{e\left(\frac{11+\gamma_\epsilon}{1-\gamma_\epsilon}\right)}{\epsilon\left(1+e\left(\frac{11+\gamma_\epsilon}{1-\gamma_\epsilon}\right)\right)^2}\\ \ge&\frac{1}{48\epsilon}\frac{(11+\gamma_\epsilon)(1-\gamma_\epsilon)}{9}\\ \ge&\frac{1}{48\epsilon|\log\epsilon|},\end{aligned}\tag{2.8}$$

where the second line simply uses the fact that $e<3$, and the final line follows by our bound on $\epsilon$. Taking $x=z$ in (2.8), and noticing that any point $c$ between $z$ and $w$ satisfies $|c-\gamma t|\le A\epsilon$, we may apply the mean value theorem to (2.8) to give

$$|\tilde{u}(t,z)-\tilde{u}(t,w)|\ge|z-w|\left|\frac{d\tilde{u}(t,c)}{dx}\right|\ge\frac{|z-w|}{48\epsilon|\log\epsilon|}.\qquad\square$$

## 2.4 A coupling argument

Next, we show how to couple $d(W_s,t-s)$ (the signed distance from a $\mathbb{d}$-dimensional Brownian motion $W_s$ to $\Gamma_{t-s}$) and a one-dimensional Brownian motion, when $W_s$ is near the interface $\Gamma_{t-s}$. To do this, we need some regularity properties which are a consequence of the assumptions ($\mathscr{C}0$)-($\mathscr{C}3$).

We write $\dot{d}$ for the time derivative of $d$. We may choose a $T^*\in(0,\mathscr{T})$ such that the following properties hold.

1. There exists $c_0>0$ such that for all $t\in[0,T^*]$ and $x\in\{y:|d(y,t)|\le c_0\}$, we have

$$|\nabla d(x,t)|=1.\tag{2.9}$$

   Moreover, $d$ is a $C^{a,\frac{a}{2}}$ function in $\{(x,t):|d(x,t)|\le c_0,t\le T^*\}$.

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 $-\dot{d}(x,t)$. Thus, (2.3) becomes

$$-\dot{d}(x,t) = \nu_\epsilon + \begin{cases} 0 & \text{if } \tilde{\alpha} \in [0,1) \\ \Delta d(x,t) & \text{if } \tilde{\alpha} \geq 1 \end{cases} \tag{2.10}$$

   for all $x, t$ such that $d(x,t) = 0$. We have replaced the arbitrary constant $c$ in Definition 2.3 with $\nu_\epsilon$, as this is the specific speed of the constant normal flow we wish to find.

3. There exists $C_0 > 0$ such that for all $t \in [0, T^*]$ and $x$ such that $|d(x,t)| \leq c_0$,

$$\left|\nabla\left(\dot{d}(x,t)\right)\right| \leq C_0 \quad \text{and} \quad |\Delta d(x,t)| \leq C_0. \tag{2.11}$$

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

$$|d(x,t) - d(x,s)| \leq V_0(s-t). \tag{2.12}$$

Properties 1 and 2 above come from [9], Theorem 2. Property 1 may seem particularly strong, but has quite an intuitive interpretation. Because our boundary is sufficiently smooth, if a point $x$ is within $c_0$ of a section of the boundary, that section of boundary will be the closest part of boundary to point $x$. As the shortest distance between two points is the straight line between them, we could parametrise space with this line as one of our co-ordinates, and then it is clear that $\nabla d(x,t)$ is just 1 in this direction, and 0 in the other co-ordinates. Alternative co-ordinate systems would not change the magnitude of this derivative.

Properties 3 and 4 follow easily from the fact that $\sup_{u \in S^1, t \leq T^*} |\Gamma_t(u)| < \infty$ and the regularity of $d$ provided by Property 1.

We can now state the necessary coupling argument.

**Proposition 2.13.** *Let $(W_s)_{s\geq 0}$ denote a $\mathbb{d}$-dimensional Brownian motion running at rate $2\epsilon^{1-\alpha}$ started at $x \in \mathbb{R}^{\mathsf{d}}$. Suppose that $\epsilon < 1$, $t \leq T^*$ and define a new parameter $\beta$ such that $\epsilon^{1+\alpha} \leq \beta \leq c_0$. Let*

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

*Then we can couple $(W_s)_{s\geq 0}$ with a one-dimensional Brownian motion $(B_s)_{s\geq 0}$, also running at* *rate* $2\epsilon^{1-\alpha}$*, started from* $z = d(x,t)$ *in such a way that for* $s \leq T_\beta$*,*

$$B_s + \nu_\epsilon s - C_1\beta s \leq d\left(W_s, t-s\right) \leq B_s + \nu_\epsilon s + C_1\beta s,$$

*where* $C_1 = 2C_0$*, i.e., twice the constant* $C_0$ *from* (2.11).

*Proof.* By Itô's formula, we have that for $s \leq t$

$$d\left(W_s, t-s\right) = \int_0^s A_u \, du + B_s,$$

where

$$A_u = -\dot{d}\left(W_u, t-u\right) + \epsilon^{1-\alpha}\Delta d\left(W_u, t-u\right)$$

and

$$B_s = \sum_{i=1}^{\mathbb{d}} \int_0^s \frac{\partial}{\partial x_i} d(W_u, t-u) dW_u^{(i)}.$$

We will handle $A_u$ and $B_s$ in turn. For each $u \in [0, T_\nu]$ there exists some $x_u \in \mathbb{R}^{\mathbb{d}}$ such that $|x_u - W_u| \leq \nu$, and $d(x_u, t-u) = 0$. By (2.10) we have

$$-\dot{d}(x_u, t-u) = \nu_\epsilon + \mathbb{1}_{\{\tilde{\alpha} \geq 1\}}\Delta d(x,t).$$

Since $\beta \leq c_0$, by (2.11) we have that, for $x$ on the line segment connecting $x_u$ to $W_u$, the gradient of $\dot{d}(x, t-u)$ is bounded by $C_0$. Also by (2.11), we can bound the Laplacian term, so we obtain

$$|A_u - \nu_\epsilon| \leq C_0(\beta + \epsilon^{1-\alpha}) \leq C_1\beta$$

for $\epsilon < 1$. Since $\beta \leq c_0$, it follows by (2.9) and Lévy's characterisation (recall that our Brownian motions run at rate $2\epsilon^{1-\alpha}$) that $(B_s)_{0 \leq s \leq T_\beta}$ is a (stopped) Brownian Motion. This completes the proof. □

The difference between this coupling, and the coupling used in [13] is subtle, but significant. In our coupling, the quadratic variation term is scaled by $\epsilon^{1-\alpha}$. If $\alpha < 1$, this term makes an insignificant impact on our coupling as we take $\epsilon$ to 0. However, the curvature of the interface $\mathbf{\Gamma}_t$ at $x$ is $\Delta d\left(x, t-u\right)$, so it is this term that gives the coupling in [13] its dependence on the curvature of the interface. Conversely, our coupling has an additional term, $\nu_\epsilon u$. This constant drift allows this coupling to keep up with the moving reference frame of our particular solution (2.5), which gives us the constant normal component of the curvature flow.

Now that we have our one-dimensional solution and a coupling for this solution to higher-dimensional solutions, the rest of the proof follows the same lines as the proof in [13].

## 2.5 Biased majority voting for branching Brownian motion

As in [13], we split the proof of Theorem 2.10 into two parts. In Subsection 2.5.1 we show that the interface is generated in a time $\delta_{\mathbb{d}} \sim \epsilon^{1+\alpha}|\log \epsilon|$. In Subsection 2.5.2, we show that the interface propagates according to mixed curvature flow. It is in this second subsection that we use the coupling we constructed in Proposition 2.13, coupling the multidimensional BBM to the explicit one-dimensional solution. The propagation of the interface in higher dimensions is then just a consequence of Theorem 2.11. The proof of a lemma crucial to the proof of Theorem 2.10 is delayed to the last subsection, Subsection 2.5.3.

As we will be comparing the one-dimensional process to the multidimensional one, we will use the convention that in one dimension we always take $\mathbb{V}^\gamma = \mathbb{V}^\gamma_{\tilde{p}}$ with $\tilde{p}(x)$ defined in (2.6). We reserve the subscript $p$ for the multidimensional $\mathbb{V}^\gamma_p(\mathbf{W}(t))$, where we assume that $p$ satisfies ($\mathscr{C}0$)-($\mathscr{C}3$).

### 2.5.1 Generation of the interface

In this section we prove that in dimension $\mathbb{d} \geq 2$ an interface of width $\mathcal{O}(\epsilon|\log \epsilon|)$ is generated in time $\mathcal{O}\left(\epsilon^{1+\alpha}|\log \epsilon|\right)$. This proof involves finding a sufficiently large tree in the branching process, whose particles have not travelled too far from the interface, and iterating the voting procedure enough times so that small biases in the initial condition are magnified as the voting procedure progresses up to the root. This strategy is essentially unchanged from the proof of Theorem 2.5 of [13]. However, some of the details need significant modification, as [13] benefits heavily from the symmetry of the voting system.

Fortunately, as the asymmetry only appears in the voting system, rather than the dual process itself, the most significant modifications are restricted to the lemmas regarding the voting procedure. We deal with these lemmas first, Lemmas 2.14 and 2.15, deferring the Lemmas 2.16

and 2.17, which deal with the branching process, for later.

We introduce notation for the majority voting procedure. Let $g : [0,1]^3 \to [0,1]$ be given by

$$g(p_1, p_2, p_3) = p_1p_2p_3 + \tfrac{3+\gamma_\epsilon}{3+3\gamma_\epsilon}\left(p_1p_2(1-p_3) + p_2p_3(1-p_1) + p_3p_1(1-p_2)\right).$$

This is the probability the root of a single ternary branch votes 1, in the special case where the three voters are independent and have probabilities $p_1, p_2$ and $p_3$, respectively, of voting 1. With a slight abuse of notation, we let

$$g(p) = g(p,p,p) = p^3 + \tfrac{3+\gamma_\epsilon}{1+\gamma_\epsilon}p^2(1-p)$$

for $p \in [0,1]$, and define $g^{(n)}(p)$, inductively, by

$$g^{(1)}(p) = g(p), \qquad g^{(n+1)}(p) = g^{(n)}(g(p)).$$

Thus, $g^{(n)}(p)$ describes the probability of voting 1 at the root of an $n$-level regular ternary tree if the votes of the leaves are i.i.d. Bernoulli($p$). Note also that

$$g(p) - p = \tfrac{1}{1+\gamma_\epsilon}p(1-p)(2p - (1+\gamma_\epsilon))$$

has three roots, 0, 1 and $\frac{1+\gamma_\epsilon}{2}$, so the same three values are fixed points of $g$. As $g(p) - p$ is positive between the latter two roots, recursive applications of $g$ to a $p$ starting in this interval produce an increasing sequence. The next lemma gives some control over how quickly this sequence grows.

**Lemma 2.14.** *For all $k \in \mathbb{N}$ there exists $A(k) < \infty$ such that the following holds:*

1. *for all $\epsilon \in (0, \frac{1-\gamma_\epsilon}{2}]$ and $n \geq A(k)|\log \epsilon|$ we have*

$$g^{(n)}(\tfrac{1+\gamma_\epsilon}{2} + \epsilon) \geq 1 - \epsilon^k.$$

2. *for all $\epsilon \in (0, \frac{1+\gamma_\epsilon}{2}]$ and $n \geq A(k)|\log \epsilon|$ we have*

$$g^{(n)}(\tfrac{1+\gamma_\epsilon}{2} - \epsilon) \leq \epsilon^k.$$

*Proof.* We begin with the first statement. We shall carry out two phases of iteration of $g$. First, we will show that it takes $\mathcal{O}\left(|\log \epsilon|\right)$ iterations to obtain

$$g^{(n)}(\tfrac{1+\gamma_\epsilon}{2}+\epsilon) \geq \tfrac{1}{2}+\sqrt{\tfrac{1+\gamma_\epsilon^2}{8}}. \tag{2.13}$$

Note that $\frac{1+\gamma}{2} < \frac{1}{2} + \sqrt{\frac{1+\gamma 2}{8}} < 1$ for $\gamma \in (0,1)$. After establishing this, we see that $\mathcal{O}\left(k|\log \epsilon|\right)$ iterations are required to obtain

$$g^{(n)}\left(\tfrac{1}{2}+\sqrt{\tfrac{1+\gamma_\epsilon^2}{8}}\right) \geq 1-\epsilon^k. \tag{2.14}$$

Since $g$ is monotone, combining the two phases completes the proof of the first statement.

For the first phase, if $\delta \in (0, \sqrt{(1+\gamma_\epsilon^2)/8} - \gamma_\epsilon/2)$ then a simple calculation shows that

$$\begin{aligned} g(\tfrac{1+\gamma_\epsilon}{2}+\delta) =& \tfrac{1+\gamma_\epsilon}{2} + \tfrac{\delta}{2(1+\gamma_\epsilon)}(3+2\gamma_\epsilon - (\gamma_\epsilon+2\delta)^2) \\ \geq& \tfrac{1+\gamma_\epsilon}{2} + \tfrac{5-\gamma_\epsilon}{4}\delta. \end{aligned}$$

Thus if $g^{(n)}(\frac{1+\gamma_\epsilon}{2}+\epsilon) - \frac{1}{2} < \sqrt{\frac{1+\gamma_\epsilon^2}{8}}$, we have

$$g^{(n+1)}(\tfrac{1+\gamma_\epsilon}{2}+\epsilon) - \tfrac{1+\gamma_\epsilon}{2} \geq \tfrac{5-\gamma_\epsilon}{4}\left(g^{(n)}(\tfrac{1+\gamma_\epsilon}{2}+\epsilon) - \tfrac{1+\gamma_\epsilon}{2}\right) \geq (\tfrac{5-\gamma_\epsilon}{4})^n\epsilon.$$

It follows immediately that $\mathcal{O}\left(|\log \epsilon|\right)$ iterations are required to achieve (2.13).

For the second phase, as $g$ is monotone increasing on $[0,1]$, it is easy to see that

$$\begin{aligned} 1-g(1-\delta) =& \tfrac{\delta}{1+\gamma_\epsilon}\left((3-\gamma_\epsilon)\delta + 2\gamma_\epsilon - 2\delta^2\right) \\ \leq& \tfrac{\delta}{1+\gamma_\epsilon}\left((3-\gamma_\epsilon)\left(\tfrac{1}{2}-\sqrt{\tfrac{1+\gamma_\epsilon^2}{8}}\right) + 2\gamma_\epsilon\right) \\ :=& a_{\gamma_\epsilon+}\delta, \end{aligned}$$

where the inequality holds for $0 \leq \delta \leq \frac{1}{2} - \sqrt{\frac{1+\gamma_\epsilon^2}{8}}$. Another simple calculation shows that $0 < a_{\gamma_\epsilon+} < 1$, which means that the sequence created by iterating $1-g(1-\delta)$ starting with

$0 \leq \delta \leq \frac{1}{2} - \sqrt{\frac{1+\gamma_\epsilon^2}{8}}$ remains in the interval $\left[0, \frac{1}{2} - \sqrt{\frac{1+\gamma_\epsilon^2}{8}}\right]$. Thus,

$$1 - g^{(n+1)}\left(\tfrac{1}{2} + \sqrt{\tfrac{1+\gamma_\epsilon^2}{8}}\right) \leq a_{\gamma_\epsilon +}\left(1 - g^{(n)}\left(\tfrac{1}{2} + \sqrt{\tfrac{1+\gamma_\epsilon^2}{8}}\right)\right) \leq a_{\gamma_\epsilon +}^{n+1}\left(\tfrac{1}{2} - \sqrt{\tfrac{1+\gamma_\epsilon^2}{8}}\right).$$

Noting again that $0 < \frac{1}{2} - \sqrt{\frac{1+\gamma_\epsilon^2}{8}} < \frac{1}{2}$, it follows easily that the number of iterations required to obtain (2.14) is $\mathcal{O}\left(k|\log\epsilon|\right)$.

We now turn to the second statement. The proof of this statement is very similar to the first phase, but it does not simply follow by symmetry, so we include it here. Again, we split the proof into two phases. For the first phase, if $\delta \in (0, \sqrt{(1+\gamma_\epsilon^2)/8} + \gamma_\epsilon/2)$ then almost the same calculation as above shows that

$$\begin{aligned} g(\tfrac{1+\gamma_\epsilon}{2} - \delta) =& \tfrac{1+\gamma_\epsilon}{2} - \tfrac{\delta}{2(1+\gamma_\epsilon)}(3 + 2\gamma_\epsilon - (\gamma_\epsilon - 2\delta)^2) \\ \leq& \tfrac{1+\gamma_\epsilon}{2} - \tfrac{5-\gamma_\epsilon}{4}\delta. \end{aligned}$$

Thus if $\frac{1}{2} - g^{(n)}(\frac{1+\gamma_\epsilon}{2} - \epsilon) < \sqrt{\frac{1+\gamma_\epsilon^2}{8}}$, we have

$$\tfrac{1+\gamma_\epsilon}{2} - g^{(n+1)}(\tfrac{1+\gamma_\epsilon}{2} - \epsilon) \geq \tfrac{5-\gamma_\epsilon}{4}\left(\tfrac{1+\gamma_\epsilon}{2} - g^{(n)}(\tfrac{1+\gamma_\epsilon}{2} - \epsilon)\right) \geq (\tfrac{5-\gamma_\epsilon}{4})^n\epsilon.$$

It follows immediately that $\mathcal{O}\left(|\log\epsilon|\right)$ iterations are required to achieve

$$g^{(n)}(\tfrac{1+\gamma_\epsilon}{2} - \epsilon) \leq \tfrac{1}{2} - \sqrt{\tfrac{1+\gamma_\epsilon^2}{8}}. \tag{2.15}$$

To show the second phase for the second statement, observe that for $0 \leq \delta \leq \frac{1}{2} - \sqrt{\frac{1+\gamma_\epsilon^2}{8}}$, we have

$$g(\delta) \leq \frac{3+\gamma_\epsilon}{1+\gamma_\epsilon}\delta^2 \leq \frac{3+\gamma_\epsilon}{1+\gamma_\epsilon}\left(\tfrac{1}{2} - \sqrt{\tfrac{1+\gamma_\epsilon^2}{8}}\right)\delta := a_{\gamma_\epsilon -}\delta.$$

Another simple calculation shows that $0 < a_{\gamma-} < 1$, which means that $g^{(n)}(\delta) \in [0, \frac{1}{2} - \sqrt{\frac{1+\gamma_\epsilon^2}{8}}]$ if $0 \leq \delta \leq \frac{1}{2} - \sqrt{\frac{1+\gamma_\epsilon^2}{8}}$. Thus,

$$g^{(n+1)}\left(\tfrac{1}{2} - \sqrt{\tfrac{1+\gamma_\epsilon^2}{8}}\right) \leq a_{\gamma_\epsilon -}g^{(n)}\left(\tfrac{1}{2} - \sqrt{\tfrac{1+\gamma_\epsilon^2}{8}}\right) \leq a_{\gamma_\epsilon -}^{n+1}\left(\tfrac{1}{2} - \sqrt{\tfrac{1+\gamma_\epsilon^2}{8}}\right).$$

Noting again that $0 < \frac{1}{2} - \sqrt{\frac{1+\gamma_\epsilon^2}{8}} < \frac{1}{2}$, it follows easily that it takes $\mathcal{O}\left(k|\log \epsilon|\right)$ iterations to achieve

$$g^{(n)}\left(\tfrac{1}{2} - \sqrt{\tfrac{1+\gamma_\epsilon^2}{8}}\right) \leq \epsilon^k. \tag{2.16}$$

Combining (2.15) and (2.16) gives the second statement, which concludes the proof. □

We include here a simple lemma about $g$.

**Lemma 2.15.** *For* $\frac{1+\gamma_\epsilon}{2} \leq p_1, p_2, p_3 \leq 1$,

$$g(p_1, p_2, p_3) \geq \min(p_1, p_2, p_3).$$

*Similarly, for* $0 \leq p_1, p_2, p_3 \leq \frac{1+\gamma_\epsilon}{2}$,

$$g(p_1, p_2, p_3) \leq \max(p_1, p_2, p_3).$$

*Proof.* We define $h : [0,1]^3 \to \mathbb{R}$ by

$$h(p_1, p_2, p_3) = g(p_1, p_2, p_3) - \frac{1}{3}(p_1 + p_2 + p_3).$$

We can write $h$ in the form

$$h(p_1, p_2, p_3) = \frac{1}{3(1+\gamma_\epsilon)} \sum p_{i_1}\left((1-p_{i_2})(p_{i_3} - \tfrac{1+\gamma_\epsilon}{2}) + (1-p_{i_3})(p_{i_2} - \tfrac{1+\gamma_\epsilon}{2})\right)$$

where the sum is over $(i_1, i_2, i_3) \in \{(1,2,3), (2,3,1), (3,1,2)\}$. Hence, if $\frac{1+\gamma_\epsilon}{2} \leq p_1, p_2, p_3 \leq 1$, then $h(p_1, p_2, p_3) \geq 0$, which means

$$g(p_1, p_2, p_3) \geq \frac{1}{3}(p_1 + p_2 + p_3) \geq \min(p_1, p_2, p_3).$$

Similarly, if $0 \leq p_1, p_2, p_3 \leq \frac{1+\gamma_\epsilon}{2}$, then $h(p_1, p_2, p_3) \leq 0$, which means

$$g(p_1, p_2, p_3) \leq \frac{1}{3}(p_1 + p_2 + p_3) \leq \max(p_1, p_2, p_3).$$ □

We now want to find a large regular ternary tree inside $\mathcal{T}(\mathbf{W}(t))$, the ternary tree induced by the historical branching Brownian motion $\mathbf{W}$ up until time $t$. Let $\mathcal{T}_n^{reg} = \cup_{k\leq n}\{1,2,3\}^k \subset \mathcal{U}$ denote the $n$-level regular ternary tree and, for $l \in \mathbb{R}$, let $\mathcal{T}_l^{reg} = \mathcal{T}_{\lceil l\rceil}^{reg}$. For $\mathcal{T}$ a time-labelled ternary tree, we use the relation $\mathcal{T} \supseteq \mathcal{T}_l^{reg}$ to mean that as subtrees of $\mathcal{U}$, $\mathcal{T}_l^{reg}$ is contained inside $\mathcal{T}$ (ignoring its time labels).

**Lemma 2.16.** *Let $k \in \mathbb{N}$ and let $A = A(k)$ be as in Lemma 2.14. Then there exist $a_{\mathbb{d}} = a_{\mathbb{d}}(k)$ and $\epsilon_{\mathbb{d}} = \epsilon_{\mathbb{d}}(k)$ such that, for all $\epsilon \in (0, \epsilon_{\mathbb{d}})$ and $t \geq a_{\mathbb{d}}\epsilon^{1+\alpha}|\log\epsilon|$,*

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

*Proof.* First we establish control over the tail distribution of the sum of $n$ independent exponentially distributed (branching) times. Suppose $(X_j)_{j\geq 1}$ are i.i.d. Exp(1) random variables and let $S_n = \sum_{j=1}^n X_j$. Then

$$M_X(\lambda) := \mathbb{E}\left[e^{\lambda X}\right] = \begin{cases} \frac{1}{1-\lambda} & \text{if } \lambda < 1 \\ \infty & \text{if } \lambda \geq 1 \end{cases}$$

and for $a \geq 1$,

$$\Psi^*(a) := \sup_{\lambda\geq 0}(\lambda a - \log M_X(\lambda)) = \sup_{0\leq\lambda<1}(\lambda a + \log(1-\lambda)) = a - 1 - \log a.$$

By Cramér's theorem, for $a \geq 1$,

$$\lim_{n\to\infty}\left(-\frac{1}{n}\log\mathbb{P}[S_n \geq na]\right) = \Psi^*(a) = a - 1 - \log a. \tag{2.17}$$

For each leaf of $\mathcal{T}_l^{reg}$ we use (2.17) to estimate the probability that it is not in $\mathcal{T}(\mathbf{W}(t))$ and combine with a union bound (summing over leaves). For $t \geq \frac{a\epsilon^{1+\alpha}}{1+\gamma_\epsilon}\lceil A|\log\epsilon|\rceil$ we have

$$\begin{aligned} &\mathbb{P}^\epsilon\left[\mathcal{T}(\mathbf{W}(t)) \not\supseteq \mathcal{T}_{A|\log\epsilon|}^{reg}\right] \\ &\quad\leq 3^{\lceil A|\log\epsilon|\rceil}\,\mathbb{P}\left[\tfrac{\epsilon^{1+\alpha}}{1+\gamma_\epsilon}S_{\lceil A|\log\epsilon|\rceil} \geq \tfrac{a\epsilon^{1+\alpha}}{1+\gamma_\epsilon}\lceil A|\log\epsilon|\rceil\right] \\ &\quad= \exp\left(\lceil A|\log\epsilon|\rceil\left(\log 3 + \frac{1}{\lceil A|\log\epsilon|\rceil}\log\mathbb{P}\left[S_{\lceil A|\log\epsilon|\rceil} \geq a\lceil A|\log\epsilon|\rceil\right]\right)\right). \end{aligned} \tag{2.18}$$

By (2.17) (with $n = \lceil A|\log\epsilon|\rceil$), we can choose $\epsilon_{\mathbb{d}}(k) < e^{-1}$ such that, for all $\epsilon \in (0, \epsilon_{\mathbb{d}})$,

$$\frac{1}{\lceil A|\log\epsilon|\rceil} \log \mathbb{P}\left[S_{\lceil A|\log\epsilon|\rceil} \geq a\lceil A|\log\epsilon|\rceil\right] \leq -a + 3/2 + \log a.$$

Choose $a \geq 1$ sufficiently large that $-a + 3/2 + \log a \leq -\log 3 - k/A$. Putting this into (2.18) we obtain

$$\mathbb{P}^{\epsilon}\left[\mathcal{T}(\mathbf{B}(t)) \not\supseteq \mathcal{T}^{reg}_{A|\log\epsilon|}\right] \leq \exp\left(-|\log\epsilon|k\right)$$

for $t \geq \frac{a\epsilon^{1+\alpha}}{1+\gamma_\epsilon}\lceil A|\log\epsilon|\rceil$. Letting $a_{\mathbb{d}} = a(A+1) \geq \frac{a(A+1)}{1+\gamma_\epsilon}$ completes the proof. ☐

We now need to control the maximal displacement of individuals in the ternary branching Brownian motion at small times. Let $N(t)$ denote the set of individuals alive in $\mathbf{W}(t)$.

**Lemma 2.17.** *Let $k \in \mathbb{N}$, and let $a_{\mathbb{d}}(k)$ be as in Lemma 2.16. Then there exist $d_{\mathbb{d}}(k)$, $\epsilon_{\mathbb{d}}(k)$ such that, for all $\epsilon \in (0, \epsilon_{\mathbb{d}}(k))$ and all $s \leq (a_{\mathbb{d}}(k) + k + 1)\epsilon^{1+\alpha}|\log\epsilon|$,*

$$\mathbb{P}^{\epsilon}_x\left[\exists i \in N(s) : |W_i(s) - x| \geq d_{\mathbb{d}}(k)\epsilon|\log\epsilon|\right] \leq \epsilon^k.$$

*Proof.* Write $\delta_{\mathbb{d}} = (a_{\mathbb{d}}(k) + k + 1)\epsilon^{1+\alpha}|\log\epsilon|$ and let $Z = (Z_1, \ldots, Z_{\mathbb{d}})$ be a $N(0, \mathrm{Id})$ distributed $\mathbb{d}$-dimensional normal random variable. By Markov's inequality, for $s \leq \delta_{\mathbb{d}}$ we have

$$\begin{aligned}
&\mathbb{P}^{\epsilon}_x\left[\exists i \in N(s) : |W_i(s) - x| \geq d_{\mathbb{d}}\epsilon|\log\epsilon|\right] \\
&\quad \leq \mathbb{E}^{\epsilon}\left[|N(s)|\right] \mathbb{P}\left[\sqrt{2\epsilon^{1-\alpha}s}|Z| \geq d_{\mathbb{d}}\epsilon|\log\epsilon|\right] \\
&\quad \leq \mathbb{E}^{\epsilon}\left[|N(\delta_{\mathbb{d}})|\right] \mathbb{P}\left[\sqrt{2\epsilon^{1-\alpha}\delta_{\mathbb{d}}}|Z| \geq d_{\mathbb{d}}\epsilon|\log\epsilon|\right] \\
&\quad = e^{2\delta_{\mathbb{d}}\frac{1+\gamma_\epsilon}{\epsilon^{1+\alpha}}} \mathbb{P}\left[\sqrt{2(a_{\mathbb{d}}(k) + k + 1)}|Z| \geq d_{\mathbb{d}}|\log\epsilon|^{1/2}\right] \\
&\quad \leq 2\mathbb{d}e^{2(a_{\mathbb{d}}+k+1)|\log\epsilon|(1+\gamma_\epsilon)} \mathbb{P}\left[\sqrt{2(a_{\mathbb{d}}(k) + k + 1)}|Z_1| \geq d_{\mathbb{d}}|\log\epsilon|^{1/2}\right] \\
&\quad \leq 2\mathbb{d}\epsilon^{\frac{1}{4}\frac{d_{\mathbb{d}}^2}{a_{\mathbb{d}}(k)+k+1} - 2(a_{\mathbb{d}}(k)+k+1)(1+\gamma_\epsilon)}
\end{aligned}$$

where we bound $|Z|$ by the sum of the moduli of $\mathbb{d}$ scalar normal random variables in the fourth line, and in the last line, we use a tail bound for a standard normal variable, $\mathbb{P}[Z_1 \geq x] \leq e^{-x^2/2}$. The proof is completed by choosing $d_{\mathbb{d}} = d_{\mathbb{d}}(k)$ sufficiently large. ☐

We now put these elements together.

**Proposition 2.18.** *Let* $k \in \mathbb{N}$. *Then there exist* $\epsilon_{\mathbb{d}}(k), a_{\mathbb{d}}(k), b_{\mathbb{d}}(k) > 0$ *such that for all* $\epsilon \in (0, \epsilon_{\mathbb{d}})$, *if we set*

$$\delta_{\mathbb{d}}(k, \epsilon) := a_{\mathbb{d}}(k)\epsilon^{1+\alpha}|\log \epsilon| \quad \text{and} \quad \delta'_{\mathbb{d}}(k, \epsilon) := (a_{\mathbb{d}}(k) + k + 1)\epsilon^{1+\alpha}|\log \epsilon|, \tag{2.19}$$

*then for* $t \in [\delta_{\mathbb{d}}, \delta'_{\mathbb{d}}]$,

1. *for* $x$ *such that* $d(x, t) \geq b_{\mathbb{d}}\epsilon|\log \epsilon|$, *we have* $\mathbb{P}^\epsilon_x\left[\mathbb{V}^\gamma_p(\boldsymbol{W}(t)) = 1\right] \geq 1 - \epsilon^k$;
2. *for* $x$ *such that* $d(x, t) \leq -b_{\mathbb{d}}\epsilon|\log \epsilon|$, *we have* $\mathbb{P}^\epsilon_x\left[\mathbb{V}^\gamma_p(\boldsymbol{W}(t)) = 1\right] \leq \epsilon^k$.

*Proof.* By Lemma 2.16, given $k \in \mathbb{N}$, there exist $a_{\mathbb{d}}(k)$ and $\epsilon_{\mathbb{d}}(k) > 0$ such that, for all $\epsilon \in (0, \epsilon_{\mathbb{d}})$ and $t \geq a_{\mathbb{d}}\epsilon^{1+\alpha}|\log \epsilon|$,

$$\mathbb{P}^\epsilon\left[\mathcal{T}(\mathbf{W}(t)) \supseteq \mathcal{T}^{reg}_{A(k)|\log \epsilon|}\right] \geq 1 - \epsilon^k. \tag{2.20}$$

Similarly, by Lemma 2.17, given $k \in \mathbb{N}$, there exist $d_{\mathbb{d}}(k)$, $\epsilon_{\mathbb{d}}(k)$ such that for all $\epsilon \in (0, \epsilon_{\mathbb{d}})$, for $t \in [\delta_{\mathbb{d}}, \delta'_{\mathbb{d}}]$,

$$\mathbb{P}^\epsilon_x\left[\exists i \in N(t) : |W_i(t) - x| \geq d_{\mathbb{d}}\epsilon|\log \epsilon|\right] \leq \epsilon^k. \tag{2.21}$$

Set $b_{\mathbb{d}}(k) = 2d_{\mathbb{d}}(k) + V_0(a_{\mathbb{d}} + k + 1)$. By (2.12) there exist $v_0, V_0 > 0$ such that for $t \leq v_0$, and any $x \in \mathbb{R}^{\mathrm{d}}$, we have $|d(x, 0) - d(x, t)| \leq V_0 t$. Reducing $\epsilon_{\mathbb{d}}$ if necessary, for $\epsilon \in (0, \epsilon_{\mathbb{d}})$ we have $\delta'_{\mathbb{d}} \leq v_0$.

Thus, in the first case of the theorem, for $\epsilon \in (0, \epsilon_{\mathbb{d}})$, $t \in [\delta_{\mathbb{d}}, \delta'_{\mathbb{d}}]$, $x$ is such that $d(x, t) \geq b_{\mathbb{d}}\epsilon|\log \epsilon|$, and $|W_i(t) - x| \leq d_{\mathbb{d}}\epsilon|\log \epsilon|$, combining with the triangle inequality and (2.12) gives,

$$\begin{aligned} d\left(W_i(t), 0\right) &\geq d\left(x, t\right) - \left|d\left(x, t\right) - d\left(W_i(t), t\right)\right| - \left|d\left(W_i(t), t\right) - d\left(W_i(t), 0\right)\right| \\ &\geq b_{\mathbb{d}}\epsilon|\log \epsilon| - d_{\mathbb{d}}\epsilon|\log \epsilon| - V_0\delta'_{\mathbb{d}} \\ &= d_{\mathbb{d}}\epsilon|\log \epsilon| \end{aligned}$$

by our choice of $b_{\mathbb{d}}$. Applying ($\mathscr{C}2$) and ($\mathscr{C}3$),

$$\begin{aligned} p(W_i(t)) &\geq \tfrac{1+\gamma_\epsilon}{2} + \lambda\left(d_{\mathbb{d}}\epsilon|\log \epsilon| \wedge r\right) \\ &\geq \tfrac{1+\gamma_\epsilon}{2} + \epsilon, \end{aligned} \tag{2.22}$$

where we reduce $\epsilon_{\mathbb{d}} > 0$ (if necessary), to ensure that $\epsilon < \lambda r$, $\epsilon < \lambda d_{\mathbb{d}}\epsilon|\log\epsilon|$ for $\epsilon \in (0, \epsilon_{\mathbb{d}})$.

If $p_i \geq \frac{1+\gamma_\epsilon}{2}$ for $i = 1, 2, 3$ then the first case of Lemma 2.15 says that $g(p_1, p_2, p_3) \geq \min(p_1, p_2, p_3)$. Hence, if each leaf of $\mathcal{T}(\mathbf{W}(t))$ votes 1 independently with probability at least $\frac{1+\gamma_\epsilon}{2} + \epsilon$ and $\mathcal{T}(\mathbf{W}(t)) \supseteq \mathcal{T}^{reg}_{A|\log\epsilon|}$, then each of the leaves of $\mathcal{T}^{reg}_{A|\log\epsilon|}$ votes 1 independently with probability at least $\frac{1+\gamma_\epsilon}{2} + \epsilon$. Therefore,

$$\mathbb{P}^{\epsilon}_{z}\left[\mathbb{V}^{\gamma}(\mathbf{W}(t)) = 1\right] \geq g^{(\lceil A|\log\epsilon|\rceil)}\left(\tfrac{1+\gamma_\epsilon}{2} + \epsilon\right) - 2\epsilon^k \geq 1 - 3\epsilon^k$$

by combining (2.20), (2.21) and (2.22) and by Lemma 2.14.

The second case of the theorem is similar. In this case, for $\epsilon \in (0, \epsilon_{\mathbb{d}})$, $t \in [\delta_{\mathbb{d}}, \delta'_{\mathbb{d}}]$, $x$ is such that $d(x,t) \geq b_{\mathbb{d}}\epsilon|\log\epsilon|$ and $|W_i(t) - x| \leq d_{\mathbb{d}}\epsilon|\log\epsilon|$, combining with the triangle inequality and (2.12), we get

$$\begin{aligned} d\left(W_i(t), 0\right) &\geq d\left(x,t\right) + \left|d\left(x,t\right) - d\left(W_i(t),t\right)\right| + \left|d\left(W_i(t),t\right) - d\left(W_i(t),0\right)\right| \\ &\geq -b_{\mathbb{d}}\epsilon|\log\epsilon| + d_{\mathbb{d}}\epsilon|\log\epsilon| + V_0\delta'_{\mathbb{d}} \\ &= -d_{\mathbb{d}}\epsilon|\log\epsilon| \end{aligned}$$

again, by our choice of $b_{\mathbb{d}}$. Applying ($\mathscr{C}2$) and ($\mathscr{C}3$),

$$\begin{aligned} p(W_i(t)) &\leq \tfrac{1+\gamma_\epsilon}{2} - \lambda\left(d_{\mathbb{d}}\epsilon|\log\epsilon| \wedge r\right) \\ &\leq \tfrac{1+\gamma_\epsilon}{2} - \epsilon, \end{aligned} \tag{2.23}$$

noting that we have already reduced $\epsilon_{\mathbb{d}} > 0$ (if necessary), to ensure that $\epsilon < \lambda r$, $\epsilon < \lambda d_{\mathbb{d}}\epsilon|\log\epsilon|$ for $\epsilon \in (0, \epsilon_{\mathbb{d}})$.

If $p_i \leq \frac{1+\gamma_\epsilon}{2}$ for $i = 1, 2, 3$ then the other case of Lemma 2.15 says that $g(p_1, p_2, p_3) \leq \max(p_1, p_2, p_3)$. Hence, if each leaf of $\mathcal{T}(\mathbf{W}(t))$ votes 1 independently with probability no more than $\frac{1+\gamma_\epsilon}{2} - \epsilon$ and $\mathcal{T}(\mathbf{W}(t)) \supseteq \mathcal{T}^{reg}_{A|\log\epsilon|}$, then each of the leaves of $\mathcal{T}^{reg}_{A|\log\epsilon|}$ votes 1 independently with probability no more than $\frac{1+\gamma_\epsilon}{2} - \epsilon$. Therefore,

$$\mathbb{P}^{\epsilon}_{z}\left[\mathbb{V}^{\gamma}(\mathbf{W}(t)) = 1\right] \leq g^{(\lceil A|\log\epsilon|\rceil)}\left(\tfrac{1+\gamma_\epsilon}{2} - \epsilon\right) + 2\epsilon^k \leq 3\epsilon^k$$

by combining (2.20), (2.21) and (2.23) and by Lemma 2.14. $\square$

### 2.5.2 Propagation of the interface and proof of Theorem 2.10

Now that we have shown that the interface forms quickly enough, all that remains is to show that the interface region propagates as it should under mixed curvature flow. As in [13], we do this by coupling the motion of any point in the interface in the direction normal to the interface to the motion of the ‘interface’ of the one-dimensional solution (2.5) with the initial condition (2.6) that we studied in Subsection 2.3. More precisely, we aim to show that

$$\mathbb{P}_x^\epsilon\left[\mathbb{V}_p^\gamma(\mathbf{W}(t))=1\right] \approx \mathbb{P}^\epsilon_{d(x,t)+\nu_\epsilon t+\mathcal{O}(\epsilon|\log\epsilon|)}\left[\mathbb{V}^\gamma(\mathbf{B}(t))=1\right].$$

The key difference between this and what is shown in [13] is the additional $\nu_\epsilon t$ term in the subscript of the right-hand side. This term allows the coupling to keep up with the constant speed of the one-dimensional solution, giving the resulting interface required for a mixed curvature flow.

The following proposition establishes a precise version of this approximation.

**Proposition 2.19.** *Let $l\in\mathbb{N}$ with $l\geq 4$. Define $a_{\mathbb{d}}(l)$ and $\delta_{\mathbb{d}}(l,\epsilon)$ as in Proposition 2.18. There exist $K_1(l), K_2(l)>0$ and $\epsilon_{\mathbb{d}}(l,K_1,K_2)>0$ such that for all $\epsilon\in(0,\epsilon_{\mathbb{d}})$ and $t\in[\delta_{\mathbb{d}}(l,\epsilon),T^*]$ we have*

$$\sup_{x\in\mathbb{R}^{\mathbb{d}}}\left(\mathbb{P}_x^\epsilon\left[\mathbb{V}_p^\gamma(\boldsymbol{W}(t))=1\right]-\mathbb{P}^\epsilon_{d(x,t)+\nu_\epsilon t+K_1e^{K_2t}\epsilon|\log\epsilon|}\left[\mathbb{V}^\gamma(\boldsymbol{B}(t))=1\right]\right)\leq\epsilon^l \tag{2.24}$$

*and*

$$\sup_{x\in\mathbb{R}^{\mathbb{d}}}\left(\mathbb{P}_x^\epsilon\left[\mathbb{V}_p^\gamma(\boldsymbol{W}(t))=0\right]-\mathbb{P}^\epsilon_{d(x,t)+\nu_\epsilon t-K_1e^{K_2t}\epsilon|\log\epsilon|}\left[\mathbb{V}^\gamma(\boldsymbol{B}(t))=0\right]\right)\leq\epsilon^l. \tag{2.25}$$

We structure the proof of this proposition in the same way the equivalent proposition is proven in [13], in that Proposition 2.19 requires a key lemma, which we state here, but delay the proof of until the following section. Furthermore, it is this key lemma which contains most of the interesting differences from the symmetric case, hence, the proof of Proposition 2.19 is not that different from [13].

Recall that $g:[0,1]\to[0,1]$ is given by $g(p)=\frac{3+\gamma_\epsilon}{1+\gamma_\epsilon}p^2-\frac{2}{1+\gamma_\epsilon}p^3$. It is convenient to extend this definition to a continuous, monotone function $g : \mathbb{R} \to [0,1]$ as follows:

$$g(p) = \begin{cases} 0 & \text{if } p < 0 \\ \frac{3+\gamma_\epsilon}{1+\gamma_\epsilon} p^2 - \frac{2}{1+\gamma_\epsilon} p^3 & \text{if } p \in [0,1] \\ 1 & \text{if } p > 1. \end{cases} \tag{2.26}$$

**Lemma 2.20.** *Let $l \in \mathbb{N}$ with $l \geq 4$ and $K_1 > 0$. There exists $K_2 = K_2(K_1, l) > 0$ and $\epsilon_{\mathbb{d}}(l, K_1, K_2) > 0$ such that for all $\epsilon \in (0, \epsilon_{\mathbb{d}})$, $x \in \mathbb{R}^{\mathbb{d}}$, $s \in [0, (l+1)\epsilon^{1+\alpha}|\log \epsilon|]$ and $t \in [s, T^*]$,*

$$\begin{aligned} &E_x\left[g\left(\mathbb{P}^\epsilon_{d(W_s, t-s)+\nu_\epsilon(t-s)+K_1 e^{K_2(t-s)}\epsilon|\log\epsilon|}[\mathbb{V}^\gamma(\boldsymbol{B}(t-s)) = 1] + \epsilon^l\right)\right] \\ &\qquad \leq \tfrac{3+5\gamma_\epsilon}{4(1+\gamma_\epsilon)}\epsilon^l + E_{d(x,t)}\left[g\left(\mathbb{P}^\epsilon_{B_s+\nu_\epsilon t+K_1 e^{K_2 t}\epsilon|\log\epsilon|}[\mathbb{V}^\gamma(\boldsymbol{B}(t-s)) = 1]\right)\right] + \mathbb{1}_{s \leq \epsilon^3}\epsilon^l \end{aligned} \tag{2.27}$$

*and*

$$\begin{aligned} &E_x\left[g\left(\mathbb{P}^\epsilon_{d(W_s, t-s)+\nu_\epsilon(t-s)-K_1 e^{K_2(t-s)}\epsilon|\log\epsilon|}[\mathbb{V}^\gamma(\boldsymbol{B}(t-s)) = 0] + \epsilon^l\right)\right] \\ &\qquad \leq \tfrac{3+5\gamma_\epsilon}{4(1+\gamma_\epsilon)}\epsilon^l + E_{d(x,t)}\left[g\left(\mathbb{P}^\epsilon_{B_s+\nu_\epsilon t-K_1 e^{K_2 t}\epsilon|\log\epsilon|}[\mathbb{V}^\gamma(\boldsymbol{B}(t-s)) = 0]\right)\right] + \mathbb{1}_{s \leq \epsilon^3}\epsilon^l. \end{aligned} \tag{2.28}$$

*Proof of Proposition 2.19.* Take $K_1 = b_{\mathbb{d}}(l) + l$ with $b_{\mathbb{d}}$ as defined in Proposition 2.18. Let $K_2 = K_2(K_1, l)$, as defined in Lemma 2.20. Take $\epsilon_{\mathbb{d}} > 0$ sufficiently small that Theorem 2.11, Proposition 2.18 and Lemma 2.20 apply for $\epsilon \in (0, \epsilon_{\mathbb{d}})$.

Initially, we quickly show that (2.24) holds for $t \in [\delta_{\mathbb{d}}, \delta'_{\mathbb{d}}]$ (where $\delta'_{\mathbb{d}}$ is defined in (2.19)). For such $t$ and for any $\epsilon \in (0, \epsilon_{\mathbb{d}})$ and $x \in \mathbb{R}^{\mathbb{d}}$,

$$\mathbb{P}^\epsilon_x\left[\mathbb{V}^\gamma_p(\mathbf{W}(t)) = 1\right] \leq \mathbb{P}^\epsilon_{d(x,t)+\nu_\epsilon t+K_1 e^{K_2 t}\epsilon|\log\epsilon|}\left[\mathbb{V}^\gamma(\mathbf{B}(t)) = 1\right] + \epsilon^l. \tag{2.29}$$

If $d(x,t) \leq -b_{\mathbb{d}}(l)\epsilon|\log\epsilon|$, then by Proposition 2.18, $\mathbb{P}^\epsilon_x\left[\mathbb{V}^\gamma_p(\mathbf{W}(t)) = 1\right] \leq \epsilon^l$, which means (2.29) holds. If $d(x,t) \geq -b_{\mathbb{d}}(l)\epsilon|\log\epsilon|$, then $d(x,t) + K_1 e^{K_2 t}\epsilon|\log\epsilon| \geq l\epsilon|\log\epsilon|$, by our choice of $K_1$, and so, by Theorem 2.11, the right-hand side of (2.29) is $\geq 1$, which also means (2.29) holds.

We are left with the case $t \in [\delta'_{\mathbb{d}}, T^*]$, which shall take the remainder of the proof. We assume, aiming for a contradiction, that there exists $t \in [\delta'_{\mathbb{d}}, T^*]$ such that, for some $x \in \mathbb{R}^{\mathbb{d}}$,

$$\mathbb{P}^\epsilon_x\left[\mathbb{V}^\gamma_p(\mathbf{W}(t)) = 1\right] - \mathbb{P}^\epsilon_{d(x,t)+\nu_\epsilon t + K_1 e^{K_2 t}\epsilon|\log\epsilon|}\left[\mathbb{V}^\gamma(\mathbf{B}(t)) = 1\right] > \epsilon^l. \tag{2.30}$$

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

$$T \in [T', \min(T' + \epsilon^{l+3}, T^*)] \tag{2.31}$$

which is also in the set of such $t$. Hence, we are assuming there exists some $x = x(l,\epsilon) \in \mathbb{R}^{\mathbb{d}}$ such that

$$\mathbb{P}^\epsilon_x\left[\mathbb{V}^\gamma_p(\mathbf{W}(T)) = 1\right] - \mathbb{P}^\epsilon_{d(x,T)+\nu_\epsilon T + K_1 e^{K_2 T}\epsilon|\log\epsilon|}\left[\mathbb{V}^\gamma(\mathbf{B}(T)) = 1\right] > \epsilon^l. \tag{2.32}$$

We now seek to show that

$$\mathbb{P}^\epsilon_x\left[\mathbb{V}^\gamma_p(\mathbf{W}(T)) = 1\right] \le \tfrac{7+9\gamma_\epsilon}{8(1+\gamma_\epsilon)}\epsilon^l + \mathbb{P}^\epsilon_{d(x,T)+\nu_\epsilon T + K_1 e^{K_2 T}\epsilon|\log\epsilon|}\left[\mathbb{V}^\gamma(\mathbf{B}(T)) = 1\right]. \tag{2.33}$$

For sufficiently small $\epsilon$, $\gamma_\epsilon \in [0,1)$, so $\frac{7+9\gamma_\epsilon}{8(1+\gamma_\epsilon)}\epsilon^l < \epsilon^l$. Hence, once we obtain (2.33) we will have a contradiction to (2.32), which completes the proof.

Denote the time of the first branching event in $\mathbf{W}(T)$ by $S$, and denote the position of the initial 'ancestor' particle at that time by $W_S$. By the strong Markov property at time $S \wedge (T - \delta_{\mathbb{d}})$, we may condition on which of these times occur first, yielding

$$\begin{aligned}\mathbb{P}^\epsilon_x\left[\mathbb{V}^\gamma_p(\mathbf{W}(T)) = 1\right] = \mathbb{E}^\epsilon_x&\left[g\left(\mathbb{P}^\epsilon_{W_S}\left[\mathbb{V}^\gamma_p(\mathbf{W}(T-S)) = 1\right]\right)\mathbb{1}_{S \le T-\delta_{\mathbb{d}}}\right] \\ &+ \mathbb{E}^\epsilon_x\left[\mathbb{P}^\epsilon_{W_{T-\delta_{\mathbb{d}}}}\left[\mathbb{V}^\gamma_p(\mathbf{W}(\delta_{\mathbb{d}})) = 1\right]\mathbb{1}_{S \ge T-\delta_{\mathbb{d}}}\right].\end{aligned} \tag{2.34}$$

As $T - \delta_{\mathbb{d}} \ge \delta'_{\mathbb{d}} - \delta_{\mathbb{d}} = (l+1)\epsilon^{1+\alpha}|\log\epsilon|$ and $S \sim \mathrm{Exp}((1+\gamma_\epsilon)\epsilon^{-1-\alpha})$, the second term on the right of (2.34) can be bounded:

$$\mathbb{E}^\epsilon_x\left[\mathbb{P}^\epsilon_{W_{T-\delta_{\mathbb{d}}}}\left[\mathbb{V}^\gamma_p(\mathbf{W}(\delta_{\mathbb{d}})) = 1\right]\mathbb{1}_{S \ge T-\delta_{\mathbb{d}}}\right] \le \mathbb{P}^\epsilon\left[S \ge (l+1)\epsilon^{1+\alpha}|\log\epsilon|\right] \le \epsilon^{l+1}. \tag{2.35}$$

To bound the first term on the right of (2.34), we partition further on the event $\{S \le \epsilon^{l+3}\}$

(which has probability $\leq (1+\gamma_\epsilon)\epsilon^{l+2}$):

$$\begin{aligned}
&\mathbb{E}^\epsilon_x \left[ g\left( \mathbb{P}^\epsilon_{W_S}\left[ \mathbb{V}^\gamma_p(\mathbf{W}(T-S)) = 1 \right] \right) \mathbb{1}_{S \leq T-\delta_{\mathbb{d}}} \right] \\
&\quad \leq \mathbb{E}^\epsilon_x \left[ g\left( \mathbb{P}^\epsilon_{W_S}\left[ \mathbb{V}^\gamma_p(\mathbf{W}(T-S)) = 1 \right] \right) \mathbb{1}_{S \leq T-\delta_{\mathbb{d}}} \mathbb{1}_{S \geq \epsilon^{l+3}} \right] + \mathbb{P}^\epsilon\left[ S \leq \epsilon^{l+3} \right] \\
&\quad \leq \mathbb{E}^\epsilon_x \left[ g\left( \mathbb{P}^\epsilon_{d(W_S, T-S)+\nu_\epsilon(T-S)+K_1 e^{K_2(T-S)}\epsilon|\log\epsilon|}\left[ \mathbb{V}^\gamma(\mathbf{B}(T-S)) = 1 \right] + \epsilon^l \right) \mathbb{1}_{S \leq T-\delta_{\mathbb{d}}} \right] \\
&\qquad + (1+\gamma)\epsilon^{l+2}.
\end{aligned} \tag{2.36}$$

The last line follows from the minimality of $T'$ (note that if $\epsilon^{l+3} \leq S \leq T-\delta_{\mathbb{d}}$, then $T-S \in [\delta_{\mathbb{d}}, T')$ by (2.31)), which is a consequence of our assumption (2.30), and from the monotonicity of $g$.

As the path of the ancestor particle $(W_\cdot)$ is independent of the time it branched, $S$, we may condition on this time, giving

$$\begin{aligned}
&\mathbb{E}^\epsilon_x \left[ g\left( \mathbb{P}^\epsilon_{d(W_S, T-S)+\nu_\epsilon(T-S)+K_1 e^{K_2(T-S)}\epsilon|\log\epsilon|}\left[ \mathbb{V}^\gamma(\mathbf{B}(T-S)) = 1 \right] + \epsilon^l \right) \mathbb{1}_{S \leq T-\delta_{\mathbb{d}}} \right] \\
&\quad \leq \int_0^{(l+1)\epsilon^{1+\alpha}|\log\epsilon|} (1+\gamma_\epsilon)\epsilon^{-1-\alpha} e^{-(1+\gamma_\epsilon)\epsilon^{-1}s} \\
&\qquad\qquad E_x\left[ g\left( \mathbb{P}^\epsilon_{d(W_s, T-s)+\nu_\epsilon(T-s)+K_1 e^{K_2(T-s)}\epsilon|\log\epsilon|}\left[ \mathbb{V}^\gamma(\mathbf{B}(T-s)) = 1 \right] + \epsilon^l \right) \right] ds \\
&\qquad + \mathbb{P}^\epsilon\left[ S \geq (l+1)\epsilon^{1+\alpha}|\log\epsilon| \right] \\
&\quad \leq \int_0^{(l+1)\epsilon^{1+\alpha}|\log\epsilon|} (1+\gamma_\epsilon)\epsilon^{-1-\alpha} e^{-(1+\gamma_\epsilon)\epsilon^{-1}s} \\
&\qquad\qquad E_{d(x,T)}\left[ g\left( \mathbb{P}^\epsilon_{B_s+\nu_\epsilon T+K_1 e^{K_2 T}\epsilon|\log\epsilon|}\left[ \mathbb{V}^\gamma(\mathbf{B}(T-s)) = 1 \right] \right) \right] ds \\
&\qquad + \tfrac{3+5\gamma_\epsilon}{4(1+\gamma_\epsilon)}\epsilon^l + \mathbb{P}^\epsilon\left[ S \leq \epsilon^3 \right] \epsilon^l + \epsilon^{l+1} \\
&\quad \leq \mathbb{E}^\epsilon_{d(x,T)} \left[ g\left( \mathbb{P}^\epsilon_{B_{S'}+\nu_\epsilon T+K_1 e^{K_2 T}\epsilon|\log\epsilon|}\left[ \mathbb{V}^\gamma(\mathbf{B}(T-S')) = 1 \right] \right) \mathbb{1}_{S' \leq T-\delta_{\mathbb{d}}} \right] \\
&\qquad + \tfrac{3+5\gamma_\epsilon}{4(1+\gamma_\epsilon)}\epsilon^l + \epsilon^{l+1} + (1+\gamma_\epsilon)\epsilon^{l+2}.
\end{aligned} \tag{2.37}$$

Here, the second inequality follows by Lemma 2.20. For the final inequality, we write $S'$ for the time of the first branching event in $(\mathbf{B}(s))_{s\geq 0}$ and $B_{S'}$ for the position of the ancestor at that time, and note that $S'$ has the same distribution as $S$. The inequality follows since $T \geq \delta'_{\mathbb{d}}$ and so $T-\delta_{\mathbb{d}} \geq (l+1)\epsilon^{1+\alpha}|\log\epsilon|$.

Putting (2.36), (2.37) and (2.35) into (2.34) we obtain

$$\begin{aligned}
&\mathbb{P}^{\epsilon}_{x}\left[\mathbb{V}^{\gamma}_{p}(\mathbf{W}(T))=1\right]\\
&\qquad\leq \mathbb{E}^{\epsilon}_{d(x,T)}\left[g\left(\mathbb{P}^{\epsilon}_{B_{S'}+\nu_\epsilon T+K_1 e^{K_2 T}\epsilon|\log\epsilon|}[\mathbb{V}^{\gamma}(\mathbf{B}(T-S'))=1]\right)\mathbb{1}_{S'\leq T-\delta_{\mathbb{d}}}\right]\\
&\qquad\quad+2\epsilon^{l+1}+2(1+\gamma_\epsilon)\epsilon^{l+2}+\tfrac{3+5\gamma_\epsilon}{4(1+\gamma_\epsilon)}\epsilon^{l}\\
&\qquad\leq \mathbb{P}^{\epsilon}_{d(x,T)+\nu_\epsilon T+K_1 e^{K_2 T}\epsilon|\log\epsilon|}\left[\mathbb{V}^{\gamma}(\mathbf{B}(T))=1\right]+2\epsilon^{l+1}(1+\epsilon+\gamma_\epsilon\epsilon)+\tfrac{3+5\gamma_\epsilon}{4(1+\gamma_\epsilon)}\epsilon^{l},
\end{aligned}$$

where the second line follows by the strong Markov Property for $(\mathbf{B}(\cdot))$ at time $S'\wedge(T-\delta_{\mathbb{d}})$, in similar style to (2.34). Reducing $\epsilon_{\mathbb{d}}$, if necessary, to ensure that $\frac{3+5\gamma_\epsilon}{4(1+\gamma_\epsilon)}\epsilon^{l}+2\epsilon^{l+1}(1+\epsilon+\gamma_\epsilon\epsilon)\leq\frac{7+9\gamma_\epsilon}{8(1+\gamma_\epsilon)}\epsilon^{l}$ for all $\epsilon\in(0,\epsilon_{\mathbb{d}})$, we obtain (2.33), which completes the proof of (2.24).

By a similar argument, using (2.28) in place of (2.27), we can also deduce (2.25). $\square$

With this proposition, the proof of Theorem 2.10 follows easily.

*Proof of Theorem 2.10.* It suffices to prove the result for sufficiently large $k\in\mathbb{N}$, and in particular we will show it for $k\geq 4$.

We choose $c_{\mathbb{d}}(k)=k+K_1e^{K_2T^*}$. Thus, for any $t\in[\delta_{\mathbb{d}},T^*]$ and $x\in\mathbb{R}^{\mathbb{d}}$ such that $d(x,t)\leq -c_{\mathbb{d}}(k)\epsilon|\log\epsilon|$ we have

$$d(x,t)+K_1e^{K_2t}\epsilon|\log\epsilon|\leq -k\epsilon|\log\epsilon|.$$

It follows from Theorem 2.11 (reducing $\epsilon_{\mathbb{d}}$ if necessary so that $\epsilon<1$) and (2.24) that, for such $x$ and $t$, $\mathbb{P}^{\epsilon}_{x}\left[\mathbb{V}^{\gamma}_{p}(\mathbf{W}(t))=1\right]\leq 2\epsilon^{k}$. Similarly, for $x$ such that $d(x,t)\geq c_{\mathbb{d}}(k)\epsilon|\log\epsilon|$, by Theorem 2.11 and (2.25) we have $\mathbb{P}^{\epsilon}_{x}\left[\mathbb{V}^{\gamma}_{p}(\mathbf{W}(t))=0\right]\leq 2\epsilon^{k}$. $\square$

### 2.5.3 Proof of Lemma 2.20

All that remains to complete the proof of Theorem 2.10, is to prove Lemma 2.20. Again, the proof of this lemma will follow the proof of Lemma 2.17 in [13], but there are some key differences.

The proof in [13] essentially considers three cases based on the distance a point $x$ is from the interface. In the first two cases, when $x$ is significantly far in front of or behind the interface, we appeal to Theorem 2.11. These cases hold similarly to [13], except we have adjusted some constants due to the different scaling. These are Cases (i) and (ii) in the proof.

Case (iii) in the proof is where the modifications to our coupling become necessary. Proposition 2.13 introduces the constant drift term which allows the multidimensional interface to keep up with the single-dimensional case. This case is split into two events, based on the gradient of $g$ evaluated at, roughly, $\tilde{u}_\epsilon(t,x)$ at points $x$ near the interface. Proposition 2.12 is used when the gradient is greater than 1. As $\gamma_\epsilon$ affects the gradient of $g$, both these cases need careful modification. We also note that it is in this case that we require the assumption that $\nu_\epsilon < 1$ if $\alpha = 0$.

We go into more detail in the proof, which we give now.

*Proof of Lemma 2.20.* We begin by proving (2.27). To simplify our expressions slightly, for the duration of this proof, for $u \geq 0$ and $z \in \mathbb{R}$ we write

$$\mathbb{Q}_z^{\epsilon,u} = \mathbb{P}_z^{\epsilon}\left[\mathbb{V}^{\gamma}(\mathbf{B}(u)) = 1\right].$$

Recall $C_1$ from Proposition 2.13 and $V_0$ from (2.12). Let

$$R = l + 2(l+1)(2\mathbb{d} + V_0). \tag{2.38}$$

Fix $K_2$ such that

$$K_1(K_2 - C_1) - C_1 R = 1. \tag{2.39}$$

Initially, let $\epsilon_{\mathbb{d}} = 1$, but we shall reduce this throughout the proof, where necessary.

First we need an estimate for the probability that a $\mathbb{d}$-dimensional Brownian motion running at rate $2\epsilon^{1-\alpha}$ moves further than $\sim \epsilon|\log\epsilon|$ in time $s$ (recall that $s \leq (l+1)\epsilon^{1+\alpha}|\log\epsilon|$). Let

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

Let $Z \sim N(0,1)$. Bounding $|W_u|$ by the sum of the moduli of $\mathbb{d}$ one-dimensional Brownian motions gives,

$$\begin{aligned} P_x\left[A_x^{\complement}\right] &\leq 2\mathbb{d}P_0\left[\sup_{u\in[0,s]} B_u > 2(l+1)\epsilon|\log\epsilon|\right] \\ &\leq 4\mathbb{d}P_0\left[Z > (2(l+1)|\log\epsilon|)^{1/2}\right] \end{aligned}$$

$$\leq 4\mathbb{d}\epsilon^{l+1}. \tag{2.40}$$

Here, since $s \leq (l+1)\epsilon^{1+\alpha}|\log \epsilon|$, the second line follows by the reflection principle, remembering that the Brownian motion is running at speed $2\epsilon^{1-\alpha}$. The last line follows by the tail bound $\mathbb{P}[Z \geq x] \leq e^{-x^2/2}$.

We consider three cases, according to how close to the interface the point $x$ is at time $t$:

(i) $d(x,t) \leq -\left(l + (l+1)(2\mathbb{d} + V_0) + K_1 e^{K_2(t-s)}\right)\epsilon|\log \epsilon|$,

(ii) $d(x,t) \geq \left(l + (l+1)(2\mathbb{d} + V_0) + K_1 e^{K_2(t-s)}\right)\epsilon|\log \epsilon|$,

(iii) $|d(x,t)| \leq \left(l + (l+1)(2\mathbb{d} + V_0) + K_1 e^{K_2(t-s)}\right)\epsilon|\log \epsilon|$.

Case (i): Recall that by (2.12) there exist $v_0 > 0$, and we have chosen $V_0$, such that if $s \leq v_0$ and $x \in \mathbb{R}^{\mathbb{d}}$ then

$$|d(x,t) - d(x,t-s)| \leq V_0 s. \tag{2.41}$$

We reduce $\epsilon_{\mathbb{d}}$, if necessary, to ensure that for $\epsilon \in (0, \epsilon_{\mathbb{d}})$ we have $(l+1)\epsilon|\log \epsilon| \leq v_0$. Then if the event $A_x$ occurs,

$$\begin{aligned}
&d(W_s, t-s) + K_1 e^{K_2(t-s)}\epsilon|\log \epsilon| \\
&\qquad \leq -\left(l + (l+1)(2\mathbb{d} + V_0)\right)\epsilon|\log \epsilon| + |d(W_s, t-s) - d(x,t)| \\
&\qquad \leq -\left(l + (l+1)(2\mathbb{d} + V_0)\right)\epsilon|\log \epsilon| + |d(x,t) - d(x,t-s)| + |W_s - x| \\
&\qquad \leq -l\epsilon|\log \epsilon| - V_0(l+1)|\epsilon \log \epsilon|(1 - \epsilon^{\alpha}) \\
&\qquad \leq -l\epsilon|\log \epsilon|.
\end{aligned}$$

Here, the second line follows from being in Case (i) and the third follows from the triangle inequality. The fourth line then follows from (2.41) and that $s \leq (l+1)\epsilon^{1+\alpha}|\log \epsilon|$, and because we have assumed $A_x$ occurs. The final line follows as $\alpha \geq 0$ and $\epsilon < 1$, so the final term in the line above is non-positive. Therefore,

$$\begin{aligned}
E_x\left[g\left(\mathbb{Q}^{\epsilon,t-s}_{d(W_s,t-s)+\nu_\epsilon(t-s)+K_1 e^{K_2(t-s)}\epsilon|\log \epsilon|} + \epsilon^l\right)\right] &\leq E_x\left[g(\epsilon^l + \epsilon^l)\mathbb{1}_{A_x}\right] + P_x\left[A_x^{\complement}\right] \\
&\leq \tfrac{12+4\gamma_\epsilon}{1+\gamma_\epsilon}\epsilon^{2l} + 4\mathbb{d}\epsilon^{l+1}.
\end{aligned}$$

Here the first inequality follows by Theorem 2.11 and the second inequality by the definition of $g$ in (2.26), the fact that $\gamma_\epsilon \in (0,1)$ (for sufficiently large $\epsilon$) and by (2.40). Again, reducing $\epsilon_{\mathbb{d}}$ if necessary, for $\epsilon \in (0, \epsilon_{\mathbb{d}})$ we have

$$E_x[g(\mathbb{Q}^{\epsilon,t-s}_{d(W_s,t-s)+\nu_\epsilon(t-s)+K_1e^{K_2(t-s)}\epsilon|\log\epsilon|} + \epsilon^l)] \leq \tfrac{3+5\gamma_\epsilon}{4(1+\gamma_\epsilon)}\epsilon^l,$$

and so (2.27) holds in this case.

Case (ii): In this case, we have that $d(x,t) \geq (l + 2(l+1))\epsilon|\log\epsilon|$, as $K_1$ and $V_0$ are positive, and $\mathbb{d} \geq 2$. A similar argument to that used for (2.40) gives us that

$$P_{d(x,t)}\left[B_s \leq l\epsilon|\log\epsilon|\right] \leq \epsilon^{l+1}. \tag{2.42}$$

It follows that in this case

$$\begin{aligned} E_{d(x,t)}\left[g\left(\mathbb{Q}^{\epsilon,t-s}_{B_s+\nu_\epsilon t+K_1e^{K_2t}\epsilon|\log\epsilon|}\right)\right] &\geq E_{d(x,t)}\left[g\left(\mathbb{Q}^{\epsilon,t-s}_{B_s+\nu_\epsilon t+K_1e^{K_2t}\epsilon|\log\epsilon|}\right)\mathbb{1}\{B_s \geq l\epsilon|\log\epsilon|\}\right] \\ &\geq g(1-\epsilon^l) - \epsilon^{l+1} \\ &\geq 1 - \tfrac{2\gamma_\epsilon}{1+\gamma_\epsilon}\epsilon^l - \tfrac{3-\gamma_\epsilon}{1+\gamma_\epsilon}\epsilon^{2l} - \epsilon^{l+1}, \end{aligned}$$

where the second line follows by Theorem 2.11 and (2.42) and the last line by the definition of $g$ in (2.26). Noting that $\frac{2\gamma_\epsilon}{1+\gamma_\epsilon} < \frac{3+5\gamma_\epsilon}{4+4\gamma_\epsilon}$ for all $\gamma_\epsilon \in (0,1)$, we may again reduce $\epsilon_{\mathbb{d}}$ if necessary, so that for $\epsilon \in (0,\epsilon_{\mathbb{d}})$ we have

$$E_{d(x,t)}\left[g\left(\mathbb{Q}^{\epsilon,t-s}_{B_s+\nu_\epsilon t+K_1e^{K_2t}\epsilon|\log\epsilon|}\right)\right] \geq 1 - \tfrac{3}{4}\epsilon^l \geq 1 - \tfrac{3+5\gamma_\epsilon}{4(1+\gamma_\epsilon)}\epsilon^l$$

and so (2.27) holds in this case.

Case (iii): We now turn to the case in which $x$ is close to the interface. If the event $A_x$ occurs, for $u \in [0,s]$ we have

$$\begin{aligned} |d(W_u, t-u)| \leq& |W_u - x| + |d(x,t)| + |d(x,t) - d(x,t-u)| \\ \leq& (l + 2(l+1)(2\mathbb{d} + V_0) + K_1e^{K_2(t-s)})\epsilon|\log\epsilon| \\ =& (R + K_1e^{K_2(t-s)})\epsilon|\log\epsilon|, \end{aligned} \tag{2.43}$$

where the second line follows by (2.41), and the third by recalling how we defined $R$ in (2.38). We now apply Proposition 2.13 with

$$\beta = (R + K_1 e^{K_2(t-s)})\epsilon|\log\epsilon|. \tag{2.44}$$

By reducing $\epsilon_{\mathbb{d}}$ if necessary, we have for $\epsilon \in (0, \epsilon_{\mathbb{d}})$ that $\beta \leq c_0$. As in Proposition 2.13, we set

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

Then by Proposition 2.13, we can couple $(W_u)_{u\geq 0}$ with $(B_u)_{u\geq 0}$, a one-dimensional Brownian motion started from $d(x,t)$, in such a way that for $u \leq T_\beta$,

$$d(W_u, t-u) \leq B_u + \nu_\epsilon u + C_1\beta u. \tag{2.45}$$

As mentioned before the proof, the introduction of the $\nu_\epsilon u$ term in this coupling is the key difference between this proof and the proof of mean curvature flow in [13]. Hence, we can replace our distance from the interface for the multidimensional problem with what is essentially the distance to the interface in the single-dimensional problem, as follows:

$$\begin{aligned}
E_x &\left[g(\mathbb{Q}^{\epsilon,t-s}_{d(W_s,t-s)+\nu_\epsilon(t-s)+K_1e^{K_2(t-s)}\epsilon|\log\epsilon|} + \epsilon^l)\right] \\
&\leq E_{d(x,t)}\left[g(\mathbb{Q}^{\epsilon,t-s}_{B_s+\nu_\epsilon t+C_1\beta s+K_1e^{K_2(t-s)}\epsilon|\log\epsilon|} + \epsilon^l)\right] + P_x\left[T_\beta \leq s\right] \\
&\leq E_{d(x,t)}\left[g(\mathbb{Q}^{\epsilon,t-s}_{B_s+\nu_\epsilon t+C_1\beta s+K_1e^{K_2(t-s)}\epsilon|\log\epsilon|} + \epsilon^l)\right] + 4\mathbb{d}\epsilon^{l+1}.
\end{aligned} \tag{2.46}$$

Here, the first inequality follows by (2.45), the monotonicity of $g$ and the monotonicity of the stationary solution, (2.5). The second inequality then follows by (2.40) (note that by (2.43), if $A_x$ occurs then $T_\beta \geq s$).

Now let

$$E = \left\{\left|\mathbb{Q}^{\epsilon,t-s}_{B_s+\nu_\epsilon t+C_1\beta s+K_1e^{K_2(t-s)}\epsilon|\log\epsilon|} - \frac{1}{2}\right| \leq \frac{5+\gamma_\epsilon}{12}\right\}.$$

We shall consider the events $E$ and $E^{\complement}$ separately to bound the right hand side of (2.46). This is the other key difference in this proof; in contrast with [13], this event, $E$, depends on $\gamma_\epsilon$. In the symmetric case, the value of $\mathbf{\Gamma}_t$ chosen to define the interface, and the peak derivative of $g$

are both $\frac{1}{2}$. This is no longer true in the asymmetric case, where these values are $\frac{1}{2} + \frac{\gamma_\epsilon}{2}$ and $\frac{1}{2} + \frac{\gamma_\epsilon}{6}$, respectively.

The argument used in [13] when the event $E^{\complement}$ occurs requires the gradient of $g$ to be less than 1 at those points where $E^{\complement}$ occurs. Thus, we have widened our event $E$ so that it contains all the points were the gradient of $g$ is greater than 1. In particular, we have extended it in a way that leaves it centred around $\frac{1}{2}$, rather than $\frac{1+\gamma_\epsilon}{2}$, which means the event $E$ is no longer symmetric about the interface. This lack of symmetry about the interface is not important. The important thing is that $E$ definitely still spans the interface, i.e., the interval defined in $E$ contains $\frac{1+\gamma_\epsilon}{2}$.

Consider first when the event $E$ occurs. Note that by the definition of $\beta$ in (2.44),

$$
\begin{aligned}
K_1 e^{K_2 t}\epsilon|\log\epsilon| - \Big(C_1\beta s + K_1 e^{K_2(t-s)}\epsilon|\log\epsilon|\Big)& \\
&= \Big(K_1 e^{K_2(t-s)}(e^{K_2 s} - 1 - C_1 s) - C_1 R s\Big)\,\epsilon|\log\epsilon| \\
&\geq (K_1(K_2 - C_1) - C_0 R)\, s\epsilon|\log\epsilon| \\
&= s\epsilon|\log\epsilon|, \qquad (2.47)
\end{aligned}
$$

where the second line follows since $K_2 > 0$, and the last line follows by our choice of $K_2$ in (2.39). Reducing $\epsilon_{\mathsf{d}}$ further, if necessary, so that $\epsilon_{\mathsf{d}} < \exp\left(\frac{-9}{(1-\gamma_\epsilon)(1+\gamma_\epsilon)}\right)$, for $\epsilon \in (0, \epsilon_{\mathsf{d}})$ we can apply Corollary 2.12 with $z = B_s + \gamma_\epsilon t + C_1\beta s + K_1 e^{K_2(t-s)}\epsilon|\log\epsilon|$ and $w = z + s\epsilon|\log\epsilon| \leq B_s + \gamma_\epsilon t + K_1 e^{K_2 t}\epsilon|\log\epsilon|$. Note that $|z - w| = s\epsilon|\log\epsilon| \leq (l+1)\epsilon^{2+\alpha}|\log\epsilon|^2$, so we reduce $\epsilon_{\mathsf{d}}$, if necessary, to ensure that $|w - z| < \epsilon$, so that we may apply Corollary 2.12. This gives

$$
\mathbb{Q}^{\epsilon,t-s}_{B_s+\nu_\epsilon t+C_1\beta s+K_1 e^{K_2(t-s)}\epsilon|\log\epsilon|}\mathbb{1}_E \leq (\mathbb{Q}^{\epsilon,t-s}_{B_s+\nu_\epsilon t+K_1 e^{K_2 t}\epsilon|\log\epsilon|} - \tfrac{1}{48}s)\mathbb{1}_E. \qquad (2.48)
$$

Finally, we consider the case when the event $E^{\complement}$ occurs. Note that for $p \in (0,1)$, $g'(p) = \frac{6p}{1+\gamma_\epsilon}\left(1 + \frac{\gamma_\epsilon}{3} - p\right)$. Hence, it is not too hard to see that if $p, \delta \geq 0$ with either $p + \delta \leq \frac{1-\gamma_\epsilon}{9}$ or $p \geq \frac{8+\gamma_\epsilon}{9}$ then

$$
g(p+\delta) \leq g(p) + \tfrac{2(1+2\gamma_\epsilon)}{3(1+\gamma_\epsilon)}\delta. \qquad (2.49)
$$

Let $C_\gamma = \frac{2(1+2\gamma_\epsilon)}{3(1+\gamma_\epsilon)}$. Reducing $\epsilon_{\mathbb{d}}$ if necessary so that $\frac{1-\gamma_\epsilon}{12} + \epsilon^l < \frac{1-\gamma_\epsilon}{9}$ for $\epsilon \in (0, \epsilon_{\mathbb{d}})$, we have

$$\begin{aligned} g\left(\mathbb{Q}^{\epsilon,t-s}_{B_s+\nu_\epsilon t+C_1\beta s+K_1e^{K_2(t-s)}\epsilon|\log\epsilon|} + \epsilon^l\right)\mathbb{1}_{E^c} &\le \left(g\left(\mathbb{Q}^{\epsilon,t-s}_{B_s+\nu_\epsilon t+C_1\beta s+K_1e^{K_2(t-s)}\epsilon|\log\epsilon|}\right) + C_\gamma\epsilon^l\right)\mathbb{1}_{E^c} \\ &\le \left(g\left(\mathbb{Q}^{\epsilon,t-s}_{B_s+\nu_\epsilon t+K_1e^{K_2t}\epsilon|\log\epsilon|}\right) + C_\gamma\epsilon^l\right)\mathbb{1}_{E^c}, \end{aligned} \tag{2.50}$$

where the first line follows by (2.49) and the last line by (2.47) and monotonicity of $g$.

Putting (2.48) and (2.50) into (2.46),

$$\begin{aligned} &E_x\left[g(\mathbb{Q}^{\epsilon,t-s}_{d(W_s,t-s)+\nu_\epsilon(t-s)+K_1e^{K_2(t-s)}\epsilon|\log\epsilon|} + \epsilon^l)\right] \\ &\qquad\le E_{d(x,t)}\left[g\left(\mathbb{Q}^{\epsilon,t-s}_{B_s+\nu_\epsilon t+K_1e^{K_2t}\epsilon|\log\epsilon|} - \tfrac{1}{48}s + \epsilon^l\right)\mathbb{1}_E\right] \\ &\qquad\quad + E_{d(x,t)}\left[\left(g\left(\mathbb{Q}^{\epsilon,t-s}_{B_s+\nu_\epsilon t+K_1e^{K_2t}\epsilon|\log\epsilon|}\right) + C_\gamma\epsilon^l\right)\mathbb{1}_{E^c}\right] + 4\mathbb{d}\epsilon^{l+1} \\ &\qquad\le E_{d(x,t)}\left[g\left(\mathbb{Q}^{\epsilon,t-s}_{B_s+\nu_\epsilon t+K_1e^{K_2t}\epsilon|\log\epsilon|}\right)\right] + C_\gamma\epsilon^l + \epsilon^l\mathbb{1}_{s\le 48\epsilon^l} + 4\mathbb{d}\epsilon^{l+1}, \end{aligned}$$

where the last inequality follows in the case $s \le 48\epsilon^l$ since $|g'(p)| \le \frac{3}{2} \le C_\gamma + 1$ for all $p \in [0,1]$. Notice that $C_\gamma + \frac{1-\gamma_\epsilon}{12(1+\gamma_\epsilon)} = \frac{3+5\gamma_\epsilon}{4(1+\gamma_\epsilon)}$. Reducing $\epsilon_{\mathbb{d}}$, if necessary, so that $4\mathbb{d}\epsilon^{l+1} \le \frac{1-\gamma_\epsilon}{12(1+\gamma_\epsilon)}\epsilon^l$ and $48\epsilon^l \le \epsilon^3$ for $\epsilon \in (0, \epsilon_{\mathbb{d}})$ completes the proof of (2.27) (also noting that here is where we require $l > 3$).

The second statement of the lemma, equation (2.28), is proved by the same argument, considering $\{\mathbb{V}^\gamma(\mathbf{B}(u)) = 0\}$ instead of $\{\mathbb{V}^\gamma(\mathbf{B}(u)) = 1\}$ and using $d(W_u, t-u) \ge B_u + \nu_\epsilon u - C_1\beta u$ for $u \le T_\beta$ in place of (2.45). $\square$

## 2.6 Application to the spatial Λ-Fleming-Viot process with asymmetric selection

In this section we prove an analogue of Theorem 2.4 but for a scaled version of the Spatial Λ-Fleming-Viot process. Before stating the result, we will restate the scaling we will use.

Let $\varrho \in \left(0, \frac{1}{4}\right)$. For each $n \in \mathbb{N}$, we define the finite measure $\mu^n$ on $(0, \mathcal{R}_n]$, where $\mathcal{R}_n = n^{-\varrho}\mathcal{R}$, by $\mu^n(A) = \mu(n^\varrho A)$ for all Borel subsets $A$ of $(0, \infty)$. Furthermore, define $\epsilon_n$ as a sequence such that $\epsilon_n \to 0$ and $(\log n)^{1/2}\epsilon_n \to \infty$. We use this sequence of $\epsilon_n$ to describe the asymptotic behaviour of the perturbation, $\gamma_\epsilon$. We will continue to denote the perturbation this way, even though this parameter should now be considered a function of $n$.

Our rescaled SΛFVA will be driven by the Poisson point process $\Pi^n$ on $\mathbb{R}_+ \times \mathbb{R}^{\mathrm{d}} \times (0,\infty)$ with intensity measure $ndt \otimes n^{\varrho}dx \otimes \mu^n(dr)$. Here $n^{\varrho}dx$ denotes the scaling in which the linear dimension of the infinitesimal region $dx$ is scaled by $n^{\varrho}$ (so that when we integrate, the volume of a region is scaled by $n^{\mathrm{d}\varrho}$). Let

$$u_n = \frac{u}{n^{1-2\varrho}}, \qquad \text{and} \qquad \mathbf{s}_n = \frac{1+\gamma_\epsilon}{\epsilon_n^{1+\alpha}}\frac{1}{n^{2\varrho}}. \tag{2.51}$$

Under this scaling, we can determine the rate an ancestral lineage in $\Xi^n$ branches. This will just be the rate that selective events hit the lineage, which will just be $\mathbf{s}_n$ times the rate that any event hits the lineage. Each event causes the the lineage to change location, so first, let us consider the rate that at which the lineage jumps from $y$ to $y+z$, which we shall denote $m_n(dz)$.

By spatial homogeneity, we may assume the lineage is at the origin before the jump. In order for the lineage to jump to $z$, it must be affected by an event that covers both 0 and $z$. If the event has radius $r$, then centre of the event, $x$, must have occurred somewhere in $\mathcal{B}_r(0) \cap \mathcal{B}_r(z)$, which we shall denote $V_r(0,z)$. Such a centre is selected with intensity $n\, n^{\mathrm{d}\varrho} V_r(0,z)\mu^n(dr)$. The parental location is chosen uniformly from the ball $\mathcal{B}_r(x)$, the volume of which we shall denote $V_r$, so the probability that $z$ is chosen as the parental location is $dz/V_r$. Finally, the lineage is only affected by this event if the individual is chosen, which occurs with probability $u_n$. Putting all this together gives

$$m_n(dz) = n u_n n^{\mathrm{d}\varrho} \int_0^{\mathcal{R}_n} \frac{V_r(0,z)}{V_r} \mu^n(dr)\, dz. \tag{2.52}$$

The total rate of jumps is then simply

$$\begin{aligned}\int_{\mathbb{R}^{\mathrm{d}}} m_n(dz) &= \int_0^{\mathcal{R}_n} n u_n\, n^{\mathrm{d}\varrho} \frac{1}{V_r} \int_{\mathbb{R}^{\mathrm{d}}} \int_{\mathbb{R}^{\mathrm{d}}} \mathbb{1}_{|x|<r} \mathbb{1}_{|x-z|<r} dx\, dz\, \mu^n(dr) \\ &= \int_0^{\mathcal{R}_n} n u_n\, n^{\mathrm{d}\varrho} V_r \mu^n(dr) \\ &= n^{2\varrho} u V_1 \int_0^{\mathcal{R}} r^{\mathrm{d}} \mu(dr) \\ &=: n^{2\varrho}\eta, \end{aligned} \tag{2.53}$$

The size of each jump is $\sim n^{-\varrho}$, so under the diffusive rescaling we have chosen, in the limit, this lineage evolves according to a time-changed Brownian motion. We can calculate this time change, denoting it $\varsigma$,

$$\begin{aligned} \varsigma =& \frac{1}{2\mathbb{d}} \int_{\mathbb{R}^{\mathbb{d}}} |z|^2 m_n(dz) \\ =& \frac{1}{2\mathbb{d}} \int_{\mathbb{R}^{\mathbb{d}}} |z|^2 n u_n \int_0^{\mathcal{R}_n} n^{\mathbb{d}\varrho} \frac{V_r(0,z)}{V_r} \mu^n(dr) dz \\ =& \frac{u}{2\mathbb{d}} \int_0^{\mathcal{R}} \int_{\mathbb{R}^{\mathbb{d}}} |z|^2 \frac{V_r(0,z)}{V_r} dz \mu(dr), \end{aligned} \tag{2.54}$$

We can now state the main theorem of this section.

**Theorem 2.21.** *Suppose that* $\varrho \in (0,1/4)$ *and let* $\epsilon_n$ *be a sequence such that* $\epsilon_n \to 0$ *and* $(\log n)^{1/2}\epsilon_n \to \infty$ *as* $n \to \infty$. *Let* $(w_t^n)_{t\geq 0}$ *be the SΛFVA driven by* $\Pi^n$ *and with* $u_n$, $\boldsymbol{s}_n$ *given by* (2.51), *and initial condition* $w_0^n(x) = p(x)$. *Assume that* $p$ *satisfies* ($\mathscr{C}0$)-($\mathscr{C}3$), *and define* $\mathscr{T}$, $d(x,t)$ *as for Theorem 2.4; take* $T^* < \mathscr{T}$. *For* $k \in \mathbb{N}$ *there exist* $n_*(k) < \infty$, *and* $a_*(k), d_*(k) \in (0,\infty)$ *such that for all* $n \geq n_*$ *and all* $t$ *satisfying* $a_*\epsilon_n^{1+\alpha}|\log\epsilon_n| \leq t \leq T^*$,

1. *for almost every* $x$ *such that* $d(x,\varsigma^2 t) \geq d_*\epsilon_n|\log\epsilon_n|$, *we have* $\mathbb{E}\left[w_t^n(x)\right] \geq 1 - \epsilon_n^k$;
2. *for almost every* $x$ *such that* $d(x,\varsigma^2 t) \leq -d_*\epsilon_n|\log\epsilon_n|$, *we have* $\mathbb{E}\left[w_t^n(x)\right] \leq \epsilon_n^k$.

As discussed in Subsection 1.2.1, the duality between the SΛFVA process and its dual allows us to prove Theorem 2.21 by proving the following equivalent theorem about the dual process $\Xi^n(t)$.

**Theorem 2.22.** *Suppose that* $\varrho \in (0,1/4)$ *and let* $\epsilon_n$ *be a sequence such that* $\epsilon_n \to 0$ *and* $(\log n)^{1/2}\epsilon_n \to \infty$ *as* $n \to \infty$. *Assume* $p$ *satisfies* ($\mathscr{C}0$)-($\mathscr{C}3$) *and define* $\mathscr{T}$, $d(x,t)$ *as for Theorem 2.4; take* $T^* < \mathscr{T}$. *Let* $k \in \mathbb{N}$. *There exist* $n_*(k) \in \mathbb{N}$, *and* $a_*(k), d_*(k) \in (0,\infty)$ *such that for all* $n \geq n^*$ *and all* $t$ *satisfying* $a_*\epsilon_n^{1+\alpha}|\log\epsilon_n| \leq t \leq T^*$,

1. *for* $x$ *such that* $d(x,\varsigma^2 t) \geq d_*\epsilon_n|\log\epsilon_n|$, *we have* $\mathbb{P}_x\left[\mathbb{V}_p^{\gamma_\epsilon}(\Xi^n(t)) = 1\right] \geq 1 - \epsilon_n^k$;
2. *for* $x$ *such that* $d(x,\varsigma^2 t) \leq -d_*\epsilon_n|\log\epsilon_n|$, *we have* $\mathbb{P}_x\left[\mathbb{V}_p^{\gamma_\epsilon}(\Xi^n(t)) = 1\right] \leq \epsilon_n^k$.

To prove this theorem, we couple the SΛFVA dual to a process very similar to the historical branching Brownian motion from Section 2.2. Through this coupling, the proof of Theorem 2.22 follows in essentially the exact same way as the proof of Theorem 2.4.

### 2.6.1 Coupling to the spatial Λ-Fleming-Viot process dual

In this subsection we outline the coupling between the SΛFVA dual and a process very similar to historical branching Brownian motion. As this coupling is almost identical to that used in [13] - the only modification is to the branching rate - we only state the results we need, and point to [13] for the proofs. Furthermore, in Chapter 4 we need a similar coupling between a Brownian motion and a jump process. This coupling is in Lemma 4.31 so the details for a similar proof can be found within this thesis also.

Let us define the process which we wish to couple the SΛFVA dual to, which we shall denote by $\Psi^n(t)$. This process is driven by the same Poisson process as $\Xi^n(t)$, however, lineages evolve independently after branching (and hence cannot coalesce).

**Definition 2.23** (Branching jump process). *For given $n \in \mathbb{N}$ and starting point $x \in \mathbb{R}^{\mathrm{d}}$, $(\Psi^n(t), t \geq 0)$ is the historical process of the branching random walk which is described as follows.*

1. *Each individual has an independent lifetime, exponentially distributed with parameter $\eta(1+\gamma_\epsilon)\epsilon_n^{-(1+\alpha)}$.*
2. *During its lifetime, each individual, independently, evolves according to a pure jump process with jump intensity $(1-\boldsymbol{s}_n)m_n(dz)$.*
3. *At the end of its lifetime an individual branches into three offspring.*
4. *The locations of the offspring are determined as follows. For each branching event, independently, pick $r \in (0, \mathcal{R}_n]$ according to $r^{\mathrm{d}}\mu^n(dr)/\int_0^{\mathcal{R}_n} r^{\mathrm{d}}\mu^n(dr)$. If the parent is at the point $z \in \mathbb{R}^{\mathrm{d}}$, then each of the three offspring, independently, samples its location uniformly from $B_r(z)$.*

As the branching events we have suppressed in defining $\Psi^n(t)$ are rare, we can couple $\Xi^n(t)$ and $\Psi^n(t)$ with high probability.

**Lemma 2.24** (Lemma 3.12 of [13]). *Let $T^* \in (0,\infty)$, $k \in \mathbb{N}$ and $z \in \mathbb{R}^{\mathrm{d}}$. There exists $n_* \in \mathbb{N}$ such that for all $n \geq n_*$, there is a coupling of $\Xi^n$ started from $z$ and $\Psi^n$ started from $z$ such that with probability at least $1-\epsilon_n^k$ we have*

$$\Xi^n(T^*) = \Psi^n(T^*).$$

As we have suppressed coalescence events, the genealogical structure of $\Psi^n(t)$ is a ternary branching tree. Hence, we can define $\mathbb{V}_p(\Psi^n(t))$ as in Definition 2.8. Specifically, leaves vote independently, voting 1 with probability $p(\Psi_i(t))$ when at location $\Psi_i(t)$ and zero otherwise, and working towards the root, an internal vertex votes with the majority of its offspring, unless precisely one offspring votes 0, in which case the vertex votes against the majority with probability $\frac{2\gamma_\epsilon}{3+3\gamma_\epsilon}$. The vote of the root is denoted $\mathbb{V}_p(\Psi^n(t))$.

The final ingredient required in our coupling is to show that between branching events, the jump process performed by ancestors in the SΛFVA dual can be coupled to a Brownian motion.

**Lemma 2.25** (Corollary 3.9 of [13]). *Let $\tau$ be the first branch time of $\Xi^n$, and let $(W(t))_{t\geq 0}$ be a Brownian motion in $\mathbb{R}^d$ started at $\xi_i^n(0)$. There is a coupling of $\Xi^n$ and $W$ under which $\tau$ and $W$ are independent, $\tau \sim Exp\left(\eta(1+\gamma_\epsilon)\epsilon_n^{-(1+\alpha)}\right)$, where $\eta = uV_1\int_0^{\mathcal{R}} r^d\mu(dr)$, and for $i = 1, 2, 3$,*

$$\mathbb{P}\left[|\xi_i^n(\tau) - W(\varsigma^2\tau)| \geq 3n^{-\varrho/6}\right] = \mathcal{O}\left(n^{-\varrho}\right).$$

With these couplings, the proof of Theorem 2.22 follows essentially the same way as it did for Theorem 2.9. For the rest of this section, we highlight a few points in the argument that require slight modification.

### 2.6.2 Generation of the interface

In this subsection we show that, as in Proposition 2.18, the interface is generated in time of order $\epsilon_n^{1+\alpha}|\log\epsilon_n|$.

**Proposition 2.26.** *Let $k \in \mathbb{N}$. Then there exist $n_*(k), a_*(k), d_*(k) > 0$ such that, for all $n \geq n_*$, if we set*

$$\delta_*(k,n) := a_*(k)\epsilon_n^{1+\alpha}|\log\epsilon_n| \text{ and } \quad \delta_*'(k,n) := (a_*(k) + \eta^{-1}(k+1))\epsilon_n^{1+\alpha}|\log\epsilon_n|, \tag{2.55}$$

*then for $t \in [\delta_*, \delta_*']$,*

1. *for $x$ such that $d(x, \varsigma^2 t) \geq d_*\epsilon|\log\epsilon|$, we have $\mathbb{P}_x\left[\mathbb{V}_p(\Xi^n(t)) = 1\right] \geq 1 - \epsilon_n^k$;*
2. *for $x$ such that $d(x, \varsigma^2 t) \leq -d_*\epsilon|\log\epsilon|$, we have $\mathbb{P}_x\left[\mathbb{V}_p(\Xi^n(t)) = 1\right] \leq \epsilon_n^k$.*

*Moreover, both these statements hold with $\Psi^n(t)$ in place of $\Xi^n(t)$.*

To prove this, we need the following lemma.

**Lemma 2.27.** *Let $k \in \mathbb{N}$ and let $A(k)$ be chosen as in Lemma 2.14. There exist $a_*(k)$ and $B_*(k) \in (0, \infty)$, and $n_*(k) < \infty$ such that for all $n \geq n_*$ and $\delta_*$, $\delta'_*$ as defined in* (2.55),

$$\mathbb{P}\left[\mathcal{T}(\Psi^n(\delta_*)) \supseteq \mathcal{T}^{reg}_{A(k)|\log \epsilon_n|}\right] \geq 1 - \epsilon_n^k, \tag{2.56}$$

$$\text{and} \qquad \mathbb{P}\left[\mathcal{T}(\Psi^n(\delta'_*)) \subseteq \mathcal{T}^{reg}_{B_*(k)|\log \epsilon_n|}\right] \geq 1 - \epsilon_n^k. \tag{2.57}$$

*Proof.* By (2.53), a given ancestral lineage in $\Psi^n$ branches into three after an exponential time with rate $n^{2\varrho}\mathbf{s}_n\eta = \eta(1+\gamma_\epsilon)\epsilon_n^{-(1+\alpha)}$. Furthermore, (2.56) follows immediately for $a_*$ sufficiently large by the same proof as Lemma 2.16.

We now turn to the proof of (2.57). Let $L^n$ be a Poisson distributed random variable with mean $\delta'_*\eta(1+\gamma_\epsilon)\epsilon_n^{-(1+\alpha)} = (a_* + \eta^{-1}(k+1))(1+\gamma_\epsilon)\eta|\log \epsilon_n|$. Recall that if $Z'$ is Poisson with parameter $\chi$, then a Chernoff bound allows us to write, for $k > \chi$,

$$\mathbb{P}[Z' > k] \leq \frac{e^{-\chi}(e\chi)^k}{k^k}. \tag{2.58}$$

Take $B_* = B_*(k)$ sufficiently large that $B_* \geq (a_* + \eta^{-1}(k+1))(1+\gamma_\epsilon)\eta$ and

$$e(a_* + \eta^{-1}(k+1))(1+\gamma_\epsilon)\eta B_*^{-1} < \frac{1}{3}e^{-k/B_*}.$$

Then applying (2.58) gives

$$\begin{aligned}\mathbb{P}\left[L^n > B_*|\log \epsilon_n|\right] &\leq \left(e(a_* + \eta^{-1}(k+1))(1+\gamma_\epsilon)\eta B_*^{-1}\right)^{B_*|\log \epsilon_n|} \\ &\leq \epsilon^k 3^{-B_*|\log \epsilon_n|},\end{aligned}$$

and, taking a union bound over each root to leaf path of $\mathcal{T}^{\text{reg}}_{B_*|\log \epsilon_n|}$,

$$\mathbb{P}\left[\mathcal{T}(\Psi^n(\delta'_*)) \not\subseteq \mathcal{T}^{\text{reg}}_{B_*(k)|\log \epsilon_n|}\right] \leq 3^{B_*|\log \epsilon_n|}\mathbb{P}\left[L^n > B_*|\log \epsilon_n|\right] \leq \epsilon_n^k,$$

which completes the proof. □

*Proof of Proposition 2.26.* We prove this result for $\Psi^n$ rather than $\Xi^n$. The result for $\Xi^n$ follows immediately using the coupling from Lemma 2.24. This proof follows along the same lines as the proof of Proposition 2.18.

Take $a_*$ from Lemma 2.27, and $t \in [\delta_*, \delta'_*]$. Let $(\xi^n(t))_{t\geq 0}$ be a pure jump process with rate of jumps from $y$ to $y+z$ given by the intensity measure $m^n(dz)$. By Lemma 2.25, we can couple $(\xi^n(t))_{t\geq 0}$ with a $\mathbb{d}$-dimensional Brownian motion $(W(t))_{t\geq 0}$ such that $\xi^n(0) = W(0)$ and

$$\mathbb{P}\left[|\xi^n(t) - W(\varsigma^2 t)| \geq n^{-\varrho/6}\right] = \mathcal{O}\left(n^{-\varrho}\right).$$

As $\epsilon_n^{-(1+\alpha)} = o\left(\log n\right)$, for large enough $n$ and constant $d_*(k)$, we have $\frac{1}{2}d_*\epsilon_n|\log\epsilon_n| \geq 2n^{-\varrho/6}$. Hence, for such $n$,

$$\begin{aligned}\mathbb{P}\left[|\xi^n(t) - \xi^n(0)| \geq \tfrac{1}{2}d_*\epsilon_n|\log\epsilon_n|\right] &\leq \mathbb{P}\left[|\xi^n(t) - W(\varsigma^2 t)| \geq n^{-\varrho/6}\right] \\ &\quad + \mathbb{P}\left[|W(\varsigma^2\delta'_*(k,n))) - W(0)| \geq \tfrac{1}{4}d_*\epsilon_n|\log\epsilon_n|\right] \\ &\leq \mathcal{O}\left(n^{-\varrho}\right) + 2\mathbb{d}\exp\left(-\frac{1}{64}\frac{d_*^2}{\varsigma^2(a_* + \eta^{-1}(k+1))}|\log\epsilon_n|\right) \\ &\leq 3^{-B_*|\log\epsilon_n|}\epsilon_n^k,\end{aligned}$$

where the second inequality follows by bounding the modulus of a $\mathbb{d}$-dimensional Brownian motion by the sum of the moduli of $\mathbb{d}$ one-dimensional Brownian motions, and the last inequality follows for $d_*$ sufficiently large. By applying (2.57), and taking a union bound over all paths from the root to each leaf $\mathcal{T}_{B_*|\log\epsilon_n|}$, for $t \in [\delta_*, \delta'_*]$,

$$\begin{aligned}\mathbb{P}_x\left[\exists\xi^n_{\mathbf{i}} \subseteq \Psi^n(\delta'_*) \text{ s.t. } |\xi^n_{\mathbf{i}}(t) - x| \geq \tfrac{1}{2}d_*\epsilon_n|\log\epsilon_n|\right] &\leq \epsilon_n^k + 3^{B_*|\log\epsilon_n|}3^{-B_*|\log\epsilon_n|}\epsilon_n^k \\ &\leq 2\epsilon_n^k. \end{aligned} \tag{2.59}$$

Combining (2.59) with Lemma 2.27, we obtain that, with probability $\geq 1 - 3\epsilon_n^k$,

1. $\mathbb{V}_p(\Psi^n(t))$ is given by independent votes at each of the leaves of $\mathcal{T}(\Psi^n(t))$.

2. $\mathcal{T}(\Psi^n(t)) \supseteq \mathcal{T}^{\text{reg}}_{A|\log\epsilon_n|}$ and the positions of the individuals corresponding to the leaves of $\mathcal{T}(\Psi^n(t))$ are all within $\frac{1}{2}d_*\epsilon_n|\log\epsilon_n|$ of their starting position.

The remainder of the proof now follows exactly the proof of Proposition 2.18. □

### 2.6.3 Propagation of the interface

We finish this section off by again noting the slight modifications we need to the key lemmas needed to prove Theorem 2.9, Lemma 2.20 and Proposition 2.19.

**Lemma 2.28.** *Let* $l \in \mathbb{N}$ *with* $l \geq 4$ *and* $K_1 > 0$. *There exists* $K_2 = K_2(K_1, l) > 0$ *and* $n_*(l, K_1, K_2) > 0$ *such that for all* $n \geq n_*$, $x \in \mathbb{R}^{\mathbb{d}}$, $s \in [\varsigma^2 \epsilon_n^{l+3}, \varsigma^2 (l+1) \eta^{-1} \epsilon_n^{1+\alpha} |\log \epsilon_n|]$ *and* $t \in [s, \varsigma^2 T^*]$,

$$\begin{aligned} E_x &\left[ g \left( \mathbb{P}^{\epsilon_n}_{d(W_s, t-s) + \nu_\epsilon (t-s) + K_1 e^{K_2 (t-s)} \epsilon_n |\log \epsilon_n| + 3n^{-\varrho/6}} [\mathbb{V}^\gamma (\boldsymbol{B}(t-s)) = 1] + \epsilon_n^l \right) \right] \\ &\leq \tfrac{3+5\gamma_\epsilon}{4(1+\gamma_\epsilon)} \epsilon_n^l + E_{d(x,t)} \left[ g \left( \mathbb{P}^{\epsilon_n}_{B_s + \nu_\epsilon t + K_1 e^{K_2 t} \epsilon_n |\log \epsilon_n|} [\mathbb{V}^\gamma (\boldsymbol{B}(t-s)) = 1] \right) \right] + \mathbb{1}_{s \leq \epsilon_n^3} \epsilon_n^l, \end{aligned}$$

*and*

$$\begin{aligned} E_x &\left[ g \left( \mathbb{P}^{\epsilon_n}_{d(W_s, t-s) + \nu_\epsilon (t-s) - K_1 e^{K_2 (t-s)} \epsilon_n |\log \epsilon_n| - 3n^{-\varrho/6}} [\mathbb{V}^\gamma (\boldsymbol{B}(t-s)) = 0] + \epsilon_n^l \right) \right] \\ &\leq \tfrac{3+5\gamma_\epsilon}{4(1+\gamma_\epsilon)} \epsilon_n^l + E_{d(x,t)} \left[ g \left( \mathbb{P}^{\epsilon_n}_{B_s + \nu_\epsilon t - K_1 e^{K_2 t} \epsilon_n |\log \epsilon_n|} [\mathbb{V}^\gamma (\boldsymbol{B}(t-s)) = 0] \right) \right] + \mathbb{1}_{s \leq \epsilon_n^3} \epsilon_n^l. \end{aligned}$$

*Proof.* The proof is essentially the same as that of Lemma 2.20. Let $R = l + 2\varsigma^2 \eta^{-1} (l+1)(2\mathbb{d} + V_0)$ and fix $K_2$ such that $K_1 (K_2 - C_1) - C_1 R = 1$; let

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

The proof for the cases $|d(x,t)| \geq (l + 2(l+1)(2\mathbb{d} + V_0) + K_1 e^{K_2 (t-s)}) \epsilon_n |\log \epsilon_n|$ are then the same as in the proof of Lemma 2.20, as the additional term corresponding to the error introduced by comparing the SΛFVA dual to a branching Brownian motion, $3n^{-\varrho/6}$, is order $o(\epsilon_n |\log \epsilon_n|))$.

Again, $n^{-\varrho/6} = o(s \epsilon_n |\log \epsilon_n|)$, so we have for $\beta = (R + K_1 e^{K_2 (t-s)}) \epsilon |\log \epsilon|$ as in (2.44), for $n$ sufficiently large

$$K_1 e^{K_2 t} \epsilon_n |\log \epsilon_n| - (C_1 \beta s + K_1 e^{K_2 (t-s)} \epsilon_n |\log \epsilon_n| + 3n^{-\varrho/6}) \geq s \epsilon_n |\log \epsilon_n|. \tag{2.60}$$

Using (2.60) in place of (2.47), the proof for for the case when

$$|d(x,t)| \leq (2c_1(l) + 2\varsigma^2 \eta^{-1} (l+1) \mathbb{d} + K_1 e^{K_2 (t-s)}) \epsilon_n |\log \epsilon_n|$$

is the same as for the corresponding case in the proof of Lemma 2.20. ☐

**Proposition 2.29.** *Let $l \in \mathbb{N}$ with $l \geq 4$. Define $a_*(l)$ and $\delta_*(l,n)$ as in Proposition 2.26. There exist $K_1(l), K_2(l) > 0$ and $n_*(l, K_1, K_2) > 0$ such that for all $n \geq n_*$ and $t \in [\delta_*(l,n), T^*]$ we have*

$$\sup_{x \in \mathbb{R}^d} \Big( \mathbb{P}_x \left[ \mathbb{V}_p^\gamma(\boldsymbol{\Psi}^n(t)) = 1 \right] - \mathbb{P}^{\epsilon_n}_{d(x,\varsigma^2 t) + \varsigma^2 \nu_\epsilon t + K_1 e^{K_2 \varsigma^2 t} \epsilon_n |\log \epsilon_n|} \left[ \mathbb{V}^\gamma(\boldsymbol{B}(\varsigma^2 t)) = 1 \right] \Big) \leq \epsilon_n^l$$

*and*

$$\sup_{x \in \mathbb{R}^d} \Big( \mathbb{P}_x \left[ \mathbb{V}_p^\gamma(\boldsymbol{\Psi}^n(t)) = 0 \right] - \mathbb{P}^{\epsilon_n}_{d(x,\varsigma^2 t) + \varsigma^2 \nu_\epsilon - K_1 e^{K_2 \varsigma^2 t} \epsilon_n |\log \epsilon_n|} \left[ \mathbb{V}^\gamma(\boldsymbol{B}(\varsigma^2 t)) = 0 \right] \Big) \leq \epsilon_n^l.$$

*Proof.* The proof exactly follows that of Proposition 2.19, with Corollary 2.25 and Lemma 2.28 in place of Lemma 2.20, and Proposition 2.26 in place of Proposition 2.18. ☐

Now have all we need in order to prove the main theorem of this section, Theorem 2.22.

*Proof of Theorem 2.22.* It suffices to prove the result for sufficiently large $k \in \mathbb{N}$, and in particular we will show it for $k \geq 5$. By Lemma 2.24, for $n$ sufficiently large and $t \in [0, T^*]$,

$$\left| \mathbb{P}_x \left[ \mathbb{V}_p(\boldsymbol{\Psi}^n(t)) = 1 \right] - \mathbb{P}_x \left[ \mathbb{V}_p(\boldsymbol{\Xi}^n(t)) = 1 \right] \right| \leq \epsilon_n^{k+1}.$$

The result now follows from Proposition 2.29 with $l = k+1$, in the same way as in the proof of Theorem 2.10. ☐

# Chapter 3

# Fluctuations in the position of the hybrid zone

In the last chapter, we considered the behaviour of the scaling limit of the hybrid zone between two populations evolving according to the SΛFVA. In this chapter, we wish to consider the fluctuations in this limit. To do so, we simplify to the one-dimensional case, which allows us to use a stochastic version of the Allen-Cahn equation to describe our model.

This work builds upon a paper by Funaki [20]. This paper derives a scaling limit for SPDEs of the form

$$\frac{\partial u}{\partial t} = \Delta u + f(u) + a\dot{W},$$

where $\dot{W}$ is a two parameter white noise, $a$ is only a function of space and has compact support, and $f$ is a generalisation of $f_0$ from the last chapter, in that it has three zeros with gradients of alternating signs, and the area under each interval between these zeros is equal. The results are proven by considering the manifold of solutions to this equation without noise, and showing that if we start with solutions to the stochastic equation that are close enough to this manifold, the solutions stay close to the manifold. Informally, the main result from [20] is that the interface between regions where the solution is near 0 and where the solution is near 1, when rescaled, evolve according to Brownian motion. We prove a similar result, but and we replace the compactly supported noise term with the classic Fisher-Wright noise, and we add an asymmetry in the non-linear term, resulting in a drift in the Brownian motion.

To adapt this work to our setting, we have two major tasks. The first is to extend these results to our explicit but non-compact noise. To do this, we exploit results proven by Mueller and Sowers in [32]. This is done in Sections 3.2 and 3.3. The second is to add an asymmetry to the non-linearity $f$. This requires quite a careful bound on the Fréchet derivative of the position of the limiting point on the manifold of solutions, which heavily exploits the particular form of the asymmetry we will use. We should note, however, that this form of the asymmetry is the natural choice for population modelling, as it is again of the form derived in (1.6). The adaptations required for the addition of the asymmetry are dealt with in Sections 3.4 and 3.5. We state our results explicitly in the next section.

## 3.1 The stochastic Allen-Cahn equation

Explicitly, the stochastic version of the Allen-Cahn equation, and the scaling we will be using in this chapter is

$$\begin{aligned} \frac{\partial u}{\partial t} &= \Delta u + \epsilon^{-1} f_\lambda(u) + \epsilon^\gamma \sqrt{u(1-u)}\dot{W}, && t > 0, x \in \mathbb{R}, \\ u(0,x) &= u_0(x), && x \in \mathbb{R}, \end{aligned} \tag{3.1}$$

where $\dot{W} = \dot{W}(t,x)$ is a white noise, $\gamma$ is some constant greater than $13/8$ and $u_0$ is continuous and bounded between 0 and 1. Furthermore, $f$ is similar to the last chapter, however, we scale the asymmetry differently, in that $f_\lambda(u) = u(1-u)(2u-(1+\lambda))$ for some $\lambda = \nu\epsilon^\varrho$ such that $\nu \in \mathbb{R}$, $\varrho \geq 1 + 2\gamma$. We postpone the proof of the existence and uniqueness of this SPDE to Subsection 3.1.2.

This scaling does not match exactly the scaling used in Chapter 2, as in that chapter we chose the scaling for the whole equation based on the asymmetry, where as in this chapter we fix the scaling of the equation, and then vary the strength of the asymmetry. However, ignoring the asymmetry, the scaling of the equation we use in this chapter corresponds to the weak asymmetry regime from Chapter 2, but with a milder spatial scaling. Hence, from what we have learned from Chapter 2, we would expect the scaling in (3.1) to result in an interface moving according to mean curvature flow. However, in one dimension, as mean curvature flow in one dimension is no flow at all, and we see no net motion in the interface. Fortunately, the milder spatial rescaling results in more interesting behaviour in this case, and we find the interface moves according to a Brownian motion.

We can extend this result to the intermediate regime from the last chapter, which, in that chapter, experiences mean curvature flow and a constant normal drift. In one dimension, this corresponds to an interface which performs a Brownian motion with a drift, proportional to the strength of the asymmetry.

We have not extended these results to the higher regime, where only constant normal flow is present in the limit. Under this scaling, if we try to apply a stronger asymmetry than this, the drift blows up. Extending our results to this regime could be possible, but would require generalising the scaling we use in this chapter, and would be unlikely to result in interesting behaviour, as we would expect the motion of the interface to consist only of the constant drift.

Our results are made precise in Theorem 3.2, the proof of which is the goal of this chapter. Briefly, it says that solutions of (3.1) with initial condition $u_0(x)$ equal to 0 for sufficiently small $x$, and 1 for sufficiently large $x$, among other conditions, when rescaled, converge to $\chi_{\xi_t}(x)$, where $\chi_y(x) = \mathbb{1}_{x \geq y}$ for $y \in \mathbb{R}$ is the indicator function, and $\xi_t$ is a Brownian motion with drift proportional to $\nu$. We prove this by identifying a manifold of solutions to the deterministic version of this equation, and showing that if the stochastic equation starts close enough to this manifold, it will stay close.

Before we can state our result precisely, we need to define some notation. Let $m : \mathbb{R} \to \mathbb{R}$ denote the standing wave solution of the PDE

$$\frac{\partial v}{\partial t} = \Delta v + f_0(v), \tag{3.2}$$

i.e. $m$ solves

$$\Delta m(y) + f_0(m(y)) = 0 \quad \forall y \in \mathbb{R} \quad \text{and } m(\infty) = 1,\, m(-\infty) = 0. \tag{3.3}$$

This function $m$ is the standing wave solution, but the travelling wave solution has the same shape, translated in time. Thus, we can define a manifold of solutions to (3.2). For $\eta \in \mathbb{R}$, we set $m_\eta(x) = m(x - \eta)$ and $M = \{m_\eta : \eta \in \mathbb{R}\}$. Then $M$ is a manifold of fixed points for the equation $\frac{\partial v}{\partial t} = \Delta v + f_0(v)$. This is the manifold to which our solutions will stay near.

For $n \in \{0, 1, 2, \ldots\}$, we write $H^n = H^n(\mathbb{R})$ for the Sobolev space equipped with the norm

$$\|v\|_{H^n}^2 = \sum_{k=0}^{n} \|\nabla^k v\|_{L^2}^2.$$

We also write $H^n + m := \{v : \mathbb{R} \to \mathbb{R} : v - m \in H^n\}$ and similarly $L^n + m := \{v : \mathbb{R} \to \mathbb{R} : v - m \in L^n\}$.

We now define Fermi coordinates which parametrise a neighbourhood of $M$. For $v \in L^2 + m$, let

$$\mathrm{dist}(v, M) := \min_{\eta \in \mathbb{R}} \|v - m_\eta\|_{L^2}.$$

Then there exists $\beta_0 > 0$ such that for $v \in L^2 + m$ with $\mathrm{dist}(v, M) \leq \beta_0$, there exists a unique $\eta = \eta(v)$ such that $\|v - m_{\eta(v)}\| = \min_{\eta' \in \mathbb{R}} \|v - m_{\eta(v)}\|_{L^2}$. Let $S(v) = v - m_{\eta(v)}$. Then $(\eta(v), S(v)) \in \mathbb{R} \times L^2$ are the Fermi coordinates of $v$. The following result from [20] makes precise the idea that solutions to the deterministic Allen-Cahn equation that start close to the manifold $M$ stay close to the manifold.

**Theorem 3.1** (Theorem 7.1 of [20]). *There exists* $\beta_2 > 0$ *such that the following holds. Suppose* $v_0 \in L^2 + m$ *with dist*$(v_0, M) \leq \beta_2$. *Let* $v$ *denote the solution of* (3.2) *with initial condition* $v(0, x) = v_0(x)$ *for all* $x \in \mathbb{R}$. *Then* $v(t, \cdot)$ *converges in* $H^1$ *as* $t \to \infty$ *and* $\lim_{t\to\infty} v(t, \cdot) \in M$. *We define* $\zeta(v_0)$ *such that* $\lim_{t\to\infty} v(t, \cdot) = m_{\zeta(v_0)}$.

Next, we define some stopping times. These stopping times correspond to ways in which we can lose control of our solution to (3.1), and a key part of the proof of 3.2 is to show that these stopping times are arbitrarily large in the limit. It is easier to define these stopping times in terms of a differently scaled version of our solution, rather than in terms of our solution directly, so we give this rescaling now.

Let $v^{\epsilon,\lambda}(t, x) = u(\epsilon^{-1/2-2\gamma} t, \epsilon^{1/2} x)$. Then $v^{\epsilon,\lambda}$ satisfies

$$\frac{\partial v^{\epsilon,\lambda}}{\partial t} = \epsilon^{-3/2-2\gamma} \left( \Delta v^{\epsilon,\lambda} + f_\lambda(v^{\epsilon,\lambda}) \right) + \epsilon^{-1/2} \sqrt{v^{\epsilon,\lambda}(1 - v^{\epsilon,\lambda})} \dot{W},$$

where $\dot{W}$ is a white noise. We will also consider a smooth approximation of $v^{\epsilon,\lambda}$ as follows. Let $\rho \in C_0^\infty(\mathbb{R})$ be a non-negative and symmetric mollifier, i.e., $\rho(y) = 0$ for all $|y| \geq 1$ and

$\int_{-\infty}^{\infty} \rho(y)dy = 1$. Then define $\rho^r(\cdot) = \frac{1}{r}\rho(\frac{\cdot}{r})$ for any $r > 0$ and let

$$v^{\delta,\lambda}(t,z) := (v^{\epsilon,\lambda} * \rho^{\delta})(z),$$

for $\delta = \epsilon^{1/10+2\gamma/5}$. This choice of $\delta$ is not obvious, but it is sufficient for our needs. We may abbreviate these new scalings of our solution to $v_t^{\delta,\lambda} = v^{\delta,\lambda}(t,\cdot)$ and $v_t^{\epsilon,\lambda} = v^{\epsilon,\lambda}(t,\cdot)$.

We now define the stopping times. Take $\kappa > 0$. We shall treat $\kappa$ as very small, in particular, $\kappa < 1/5$. As $v^{\epsilon,\lambda}$ is bounded between 0 and 1, and should be close to 0 on one half of the real line, and close to 1 on the other, the following stopping time can be thought of as the first time the interface in our rescaled solution becomes too wide. We define

$$\sigma_1 = \inf\left\{t \geq 0 : \int_{-\infty}^{\infty} v_t^{\epsilon,\lambda}(x)(1 - v_t^{\epsilon,\lambda}(x))dx \geq \epsilon^{-1/2-2\kappa}\right\}. \tag{3.4}$$

Recall that $S(v)$ is the component of $v$ perpendicular to the manifold. The second stopping time gives the first time our rescaled solution gets too far away from the manifold. *A priori*, it is not clear how far is too far, but $\beta = \epsilon^{1/20+\gamma/5-\kappa}$ is sufficient for our needs. Hence, let

$$\sigma_2 = \inf\left\{t \geq 0 : \|S(v_t^{\delta,\lambda})\|_{H^1} \geq \beta\right\}. \tag{3.5}$$

The next stopping time indicates if the interface of the rescaled solution, appropriately mollified, has become too rough, in the sense of having many regions that are very steep. Again, the threshold for measuring this, in this case $\beta' = \epsilon^{4\gamma/5-3/10-5\kappa}$ is not so clear, but nonetheless what we need. Define

$$\sigma_3 = \inf\left\{t \geq 0 : \left\|\int_{-\infty}^{\infty} \rho^{\delta}(y)(v_t^{\epsilon,\lambda}(\cdot) - v_t^{\epsilon,\lambda}(\cdot - y))^2 dy\right\|_{L^1} \geq \beta'\right\}. \tag{3.6}$$

Finally, we want our rescaled solution to stick to the manifold, but we can't let it travel excessively far along the manifold either. Hence, we define

$$\sigma_4 = \inf\left\{t \geq 0 : |\eta(v_t^{\delta,\lambda})| \geq \epsilon^{-1/2-\kappa}\right\}.$$

Naturally, we mostly care about the first time one of these situations where we lose control of our solution occurs, so define $\sigma = \min(\sigma_1, \sigma_2, \sigma_3, \sigma_4)$.

The last thing we need to do before we state our theorem, is state our assumptions on the initial condition of (3.1). We take a constant $C < \infty$ and assume that $v_0^{\epsilon,\lambda}(\cdot) := u_0(\epsilon^{1/2}\cdot)$ satisfies $0 \leq v_0^{\epsilon,\lambda} \leq 1$ and also satisfies

$$\max\left( \|\Delta v_0^{\epsilon,\lambda}\|_{L^1}, \|\nabla v_0^{\epsilon,\lambda}\|_{L^1}, \|\nabla v_0^{\epsilon,\lambda}\|_{L^\infty}, \int_{-\infty}^{0} v_0^{\epsilon,\lambda}(y)dy, \int_0^\infty (1 - v_0^{\epsilon,\lambda}(y))dy \right) \leq C. \tag{3.7}$$

We make an assumption about how close we start to the manifold,

$$\|v_0^{\epsilon,\lambda} - m\|_{H^1}^2 \leq \delta. \tag{3.8}$$

Also we assume that initially, the interface is bounded, i.e., there exists $R < \infty$ (which may depend on $\epsilon$) such that

$$v_0^{\epsilon,\lambda}(x) = 0 \ \forall x \leq -R \ \text{ and } \ v_0^{\epsilon,\lambda}(x) = 1 \ \forall x \geq R. \tag{3.9}$$

Finally, we may assume $\eta(v_0^{\epsilon,\lambda}) = 0$, as if not, we simply translate space so that this is true. We are now ready to state our main theorem. Recall that for $y \in \mathbb{R}$, $\chi_y(x) = \mathbb{1}_{x \geq y}$ for all $x \in \mathbb{R}$.

**Theorem 3.2.** *Let* $a = \frac{1}{\|\nabla m\|_{L^2}^4} \langle |\nabla m|^2, m(1-m)\rangle = 6/5$, *let* $\bar{u}^\epsilon(t,x) = u^\epsilon(\epsilon^{-1/2-2\gamma}t, x)$ *and let* $\xi_t^\epsilon = \epsilon^{1/2}\zeta(v_{t\wedge\sigma}^\epsilon)$. *Take* $T < \infty$. *There exists a constant* $C < \infty$ *such that as* $\epsilon \to 0$

$$\mathbb{P}\left( \sup_{0 \leq t \leq T} \|\bar{u}^\epsilon(t,\cdot) - \chi_{\xi_t^\epsilon}\|_{L^2} > C\epsilon^{1/4} \right) \to 0.$$

*Moreover, as* $\epsilon \to 0$, $\xi^\epsilon$ *converges in distribution on* $C[0,T]$ *to* $\xi$ *where* $\xi_t = \xi_0 + B_{at} + \nu t \mathbb{1}_{\{\varrho = 1+2\gamma\}}$.

The strategy for the proof of Theorem 3.2 is as follows. As mentioned before, our strategy is similar to that used in [20]. As this paper assumes a compactly supported noise term, we first prove Lemma 3.3, which is our main tool in adapting the proof in [20] to our non-compactly supported noise.

**Lemma 3.3.** *[from Lemma 2.7 in [32], adapted] For* $T < \infty$ *fixed, there exist a constant* $K$ *and a random variable* $X_0$ *such that for any* $t \leq T$, $u(t,x) \leq e^{-Kx^2}$ *for* $x \leq -X_0$ *and* $1 - u(t,x) \leq e^{-Kx^2}$ *for* $x \geq X_0$. *Also there exists a constant* $c$ *such that* $\mathbb{P}(X_0 \geq x) \leq c^{-1}e^{-cx}$

*for all* $x > 0$.

We prove this in the next subsection. Next, we must ensure we maintain control over our solution. This step is most similar to [20], but also the most technical. However, the broad strategy is rather intuitive. We fix a time $t$, and for each of the problematic stopping times, we show that, as $\epsilon$ goes to 0, either $t$ or one of the other stopping times occurs before it. This culminates in the following proposition:

**Proposition 3.4.** *For* $t < \infty$,

$$\lim_{\epsilon \downarrow 0} \mathbb{P}\left(\sigma \leq t\right) = 0.$$

This is not quite enough though. We further need that the solution has some minimum regularity throughout this time. We measure this regularity using Sobolev spaces in the following proposition.

**Proposition 3.5.** *There exists* $\alpha > 0$ *such that for* $T < \infty$,

$$\mathbb{P}\left(v_t^{\epsilon,\lambda} \in H^\alpha + m, \ \ \forall\, 0 \leq t \leq \sigma \wedge T\right) = 1.$$

In Section 3.2, we prove the three intermediate results regarding the first three stopping times. However, the corresponding result for $\sigma_4$, which is all that remains to show, and hence equivalent to the proof of, Proposition 3.4, is the penultimate proof of this chapter. We prove Proposition 3.5 in Section 3.3.

The final step in this proof is to incorporate the asymmetry. The asymmetry does not alter the proof of Theorem 3.2 greatly, but we need an additional estimate on the $L^1$ norm of Fréchet derivative of $\zeta(v)$, which is not used in [20]. We prove this estimate in Section 3.4.

With these steps completed, we are able to prove Theorem 3.2, which is done in Section 3.5. We end this subsection with another definition and theorem from [20]. Let $F_\lambda(u) := -\int_0^u f_\lambda(u')du'$ and define an energy functional $\mathcal{H}$ on $\mathrm{Dom}(\mathcal{H}) := H^1 + m$ by

$$\mathcal{H}(v) = \int_{-\infty}^{\infty} \left(\tfrac{1}{2}|\nabla v|^2(y) + F_0(v(y))\right) dy - C_* \tag{3.10}$$

for $v \in H^1 + m$; the constant $C_* > 0$ is chosen so that $\min_{v \in \mathrm{Dom}(\mathcal{H})} \mathcal{H}(v) = 0$. The minimum of $\mathcal{H}$ is attained on $M$ since every critical point of $\mathcal{H}$ satisfies $\Delta m + f_0(m) = 0$. For $v \in H^2 + m$, we denote the functional derivative of $\mathcal{H}$ by

$$D\mathcal{H}(y,v) = -\Delta v(y) - f(v(y)).$$

**Theorem 3.6** (Theorem 3.1 in [20]). *There exist* $c_1, c_2 > 0$ *and* $0 < \beta_1 \leq \beta_0$ *such that*

$$c_1 \|S(v)\|_{H^1}^2 \leq \mathcal{H}(v) \leq c_2 \|S(v)\|_{H^1}^2 \tag{3.11}$$

*for all* $v \in H^1 + m$ *with* $\|S(v)\|_{H^1} \leq \beta_1$ *and*

$$c_1 \|S(v)\|_{H^2}^2 \leq \|D\mathcal{H}(\cdot, v)\|_{L^2}^2 \leq c_2 \|S(v)\|_{H^2}^2$$

*for all* $v \in H^2 + m$ *with* $\|S(v)\|_{H^1} \leq \beta_1$.

### 3.1.1 Proof of Lemma 3.3

In this subsection, we prove Lemma 3.3, which is an adaptation of a result in [32]. Initially, define $\bar{u}$, satisfying

$$\begin{aligned} \frac{\partial \bar{u}}{\partial t} =& \Delta \bar{u} + \bar{u} + \sqrt{(\bar{u} \wedge 1)(1 - (\bar{u} \wedge 1)} \dot{W}, & t > 0, x \in \mathbb{R} \\ \bar{u}(0,x) =& u_0(x) & x \in \mathbb{R}. \end{aligned} \tag{3.12}$$

To prove an analogue of Lemma 3.3 for the SPDE they study, Mueller and Sowers simply prove this result for $\bar{u}$, and show that $\bar{u}$ bounds their SPDE. Consequently, to prove this result for $u$, we merely need to show the same bound holds for $u$, i.e., with probability 1, $u(t,x) \leq \bar{u}(t,x)$ for all $t \geq 0$ and $x \in \mathbb{R}$. Hence, the proof of this lemma will follow the proof of Lemma 2.4 of [32], which itself follows from Theorem 2.5 of [27] and Theorem 2.3 of [38].

*Proof.* Let $f_\lambda^{(n)}(u)$ and $a^{(n)}(u)$ be a sequence of Lipschitz functions converging uniformly on compact sets to $f_\lambda(u)$ and $\epsilon^\gamma \sqrt{(u \wedge 1)(1 - (u \wedge 1))}$, respectively, with the additional properties that $f_\lambda^{(n)}(0) = f_\lambda^{(n)}(1) = 0$ and $a^{(n)}(0) = a^{(n)}(1) = 0$. Then, we can define $u^{(n)}$ and $\bar{u}^{(n)}$ as solutions to, respectively,

$$\begin{aligned}\frac{\partial u^{(n)}}{\partial t} =&\Delta u^{(n)} + \epsilon^{-1} f_\lambda^{(n)}(u^{(n)}) + a^{(n)}(u^{(n)})\dot{W}, && t > 0, x \in \mathbb{R}\\ u^{(n)}(0,x) =&u_0(x) && x \in \mathbb{R}.\end{aligned}$$

and

$$\begin{aligned}\frac{\partial \bar{u}^{(n)}}{\partial t} =&\Delta \bar{u}^{(n)} + \bar{u}^{(n)} + a^{(n)}(\bar{u}^{(n)})\dot{W}, && t > 0, x \in \mathbb{R}\\ \bar{u}^{(n)}(0,x) =&u_0(x) && x \in \mathbb{R}.\end{aligned}$$

As $\epsilon^{-1} f_\lambda^{(n)}(x) \leq x$ for all $x \in \mathbb{R}$ and both functions are 0 when $x = 0$, we may apply Theorem 2.5 of [27], which implies that for each $n \geq 1$, there exists a pair of solutions $(u^{(n)}, \bar{u}^{(n)})$, such that, with probability 1, $0 \leq u^{(n)} \leq \bar{u}^{(n)}$.

If we define $\hat{u}^{(n)}(t,x) := 1 - u^{(n)}(t,-x)$, a simple calculation shows that $\hat{u}^{(n)}$ satisfies a similar SPDE,

$$\begin{aligned}\frac{\partial \hat{u}^{(n)}}{\partial t} =&\Delta \hat{u}^{(n)} - \epsilon^{-1} f_{-\lambda}^{(n)}(\hat{u}^{(n)}) + a^{(n)}(\hat{u}^{(n)})\dot{W}, && t > 0, x \in \mathbb{R}\\ \hat{u}^{(n)}(0,x) =&1 - u_0(-x) && x \in \mathbb{R}.\end{aligned}$$

In particular, this SPDE has the same properties as those of (3.14) which we used to show $u^{(n)}$ is non-negative. Hence, the same argument shows that $\hat{u}^{(n)}$ is non-negative, which implies $u^{(n)} \leq 1$.

Then by Corollary 2.4 of [38], we may take a subsequence of $u^{(n)}, \bar{u}^{(n)}$ which converges weakly to a pair of solutions $(u, \bar{u})$ to (3.12) and a version of (3.1) with the coefficient of the white noise replaced by $\sqrt{(u \wedge 1)(1 - (u \wedge 1)}$, which inherit the property that with probability 1, $0 \leq u \leq \bar{u}$ and $0 \leq u \leq 1$ for all $t \geq 0$ and $x \in \mathbb{R}$.

Since $0 \leq u \leq 1$ for all $t \geq 0$, then $\sqrt{(\bar{u} \wedge 1)(1 - (\bar{u} \wedge 1)} = \sqrt{\bar{u}(1 - \bar{u})}$ for all $t \geq 0$, so $u$ actually satisfies (3.1). Thus, the solution to (3.12) also bounds the solution to (3.1).

We are essentially done, as we can appeal to Lemma 2.7 in [32] for the remainder of this proof. The last bound in this lemma is not explicit in the statement of Lemma 2.7 in [32], however it is a direct consequence of their proof. We will very briefly sketch their proof in order to show this.

Firstly, Mueller and Sowers split the integral expression for $\bar{u}$ into a deterministic part, $\int_{-\infty}^{\infty} e^t G(t, x-y) u_0(y) dy$, and a stochastic part,

$$N^\epsilon(t,x) = \epsilon \int_0^t \int_{-\infty}^{\infty} e^{t-s} G(t-s, x-y) \sqrt{(\bar{u}(s,y) \wedge 1)(1-\bar{u}(s,y) \wedge 1)} W(dy, ds).$$

The deterministic part is bounded straightforwardly with deterministic estimates and the assumptions on the initial condition, contained in our (3.9). To show that $N^\epsilon(t,x)$ decays exponentially in $x$, Mueller and Sowers note that $N^\epsilon(0,x) \equiv 0$ trivially decays exponentially in $x$, and find a bound on the Hölder continuity of $N^\epsilon$ in both time and spatial dimensions strong enough to extend this trivial decay to non-zero times. Explicitly, they show for any fixed $0 \le \beta < 1/4$, some positive constant $C$ and a sufficiently large $n$,

$$\mathbb{E} \sup_{\substack{x,x' \in [n,n+1] \\ t,t' \in [0,T] \\ (t,x) \neq (t',x')}} \frac{|N^\epsilon(t,x) - N^\epsilon(t',x')|}{(|t-t'| + |x-x'|)^\beta} \le C\epsilon^2 \exp(-Cx^2)$$

Using this Hölder continuity, they can write,

$$\mathbb{P}\left( \sup_{0 \le t \le T, n \le x \le n+1} N^\epsilon(t,x) \ge e^{-Cn^2/2} \right) \le \epsilon^2 (T+1)^{-\beta} e^{-Cn^2/2}. \tag{3.13}$$

By Borel-Cantelli, almost surely there exists a random integer $n(\omega)$ such that $N^\epsilon(t,x) \le e^{-Cn^2/2} \le e^{-C(x-1)^2/2}$, for all $n > n(\omega)$ and $x \le n+1$. Taking $X_0$ as this $n(\omega)$ gives that $\bar{u}$ has the required exponential decay.

We now go a slight step further than the proof in Mueller and Sowers to get our additional claim that $\mathbb{P}(X_0 \ge x) \le c^{-1} e^{-cx}$ for all $x > 0$. From (3.13), we can bound the probability

$$\begin{aligned} \mathbb{P}(n(\omega) \ge x) \le & \sum_{n=\lfloor x \rfloor}^{\infty} \epsilon^2 (T+1)^{-\beta} e^{-Cn^2/2} \\ &= \epsilon^2 (T+1)^{-\beta} e^{-C\lfloor x \rfloor/2} \sum_{n=\lfloor x \rfloor}^{\infty} e^{-C(n^2 - \lfloor x \rfloor)/2} \\ &\le \epsilon^2 (T+1)^{-\beta} e^{-C\lfloor x \rfloor/2} \sum_{n=0}^{\infty} e^{-Cn^2/2} \end{aligned}$$

$$\begin{aligned}&\leq \epsilon^2(T+1)^{-\beta}e^{-C\lfloor x\rfloor/2}\left(1+\int_0^\infty e^{-Cz^2/2}dz\right)\\&=\epsilon^2(T+1)^{-\beta}e^{-C\lfloor x\rfloor/2}\left(1+\frac{\sqrt{\pi}}{\sqrt{2C}}\right)\end{aligned}$$

which gives the required bound for the appropriate choice of $C$. □

### 3.1.2 Existence and uniqueness

In this subsection, we make a small digression from the proof of Theorem 3.2 in order to outline the proof of the existence and uniqueness of solutions to (3.1).

**Lemma 3.7.** *The SPDE* (3.1)*,*

1. *has a solution satisfying* $0 \leq u(t,x) \leq 1$*, for all* $t \geq 0$*,* $x \in \mathbb{R}$*, almost surely, and*
2. *this solution is unique in law.*

*Proof.* As indicated by the enumeration in the lemma, we prove this result in two parts.

*Proof of 1.* This proof is adapted from Lemmas 2.3 and 2.4 of [32], and is similar to the beginning of the proof of Lemma 3.3. The first of these lemmas relies on Theorem 5.1, part (i), of [37], which shows that an SPDE of the same form as (3.1), but with Lipschitz non-linearity and coefficient of white noise, has a solution which is almost surely unique in law.

Let $f_\lambda^{(n)}(u)$ and $a^{(n)}(u)$ be a sequence of Lipschitz functions converging uniformly on compact sets to $f_\lambda(u)$ and $\epsilon^\gamma\sqrt{u(1-u)}$, respectively, with the additional properties that $f_\lambda^{(n)}(0) = f_\lambda^{(n)}(1) = 0$ and $a^{(n)}(0) = a^{(n)}(1) = 0$. Then, we can define

$$\begin{aligned}\frac{\partial u^{(n)}}{\partial t} =&\Delta u^{(n)} + \epsilon^{-1}f_\lambda^{(n)}(u^{(n)}) + a^{(n)}(u^{(n)})\dot{W}, && t>0, x\in\mathbb{R},\\ u^{(n)}(0,x) =&u_0(x), && x\in\mathbb{R}.\end{aligned} \tag{3.14}$$

As $u_0$ is continuous and bounded between 0 and 1, by Theorem 5.1, part (i), of [37], there is a unique pathwise solution to (3.14) for each $n$. Furthermore, Theorem 2.3 of [38] gives that each of these solutions is non-negative almost surely. Repeating the argument from Lemma 3.3, we can also bound $u^{(n)}$ above by 1. We may now take a subsequence of the solutions $u^{(n)}$, which converges weakly to a solution $u$ to (3.1), which is bounded almost surely between 0 and 1.

*Proof of 2.* All that remains is to show uniqueness of the solution in law. To do this, we will show that the solution to (3.1) is dual to a system of branching and coalescing random walks. The duality establishes the equivalence of the finite-dimensional distributions of any two solutions to (3.1), which in turn proves the uniqueness of solutions to (3.1) when interpreted as a martingale problem.

The duality we are about to show is based on similar dualities shown in Chapter 6 of [30] and Chapter 2 of [16]. Chapter 2 of [16] establishes a duality between a PDE and super Brownian motion, and [30] does so for coalescing Brownian motions. As the duality we wish to establish is not particularly new, and we will not be using it outside this lemma, what we demonstrate here will be more of a sketch than a rigorous proof; we refer to these two original sources, in particular [30], for those who want to see more detail.

We now introduce the dual process. The dual process will be very similar to the voting dual we used in Chapter 2, for the deterministic Allen-Cahn equation. However, we need to introduce coalescence to our branching Brownian motion to account for the stochastic term in (3.1). A further difference from the deterministic dual is that the dual we describe here will only work in one-dimension. This is all we need for this problem, as it is only in one dimension that (3.1) can be defined.

Our probability space will consist of countably many independent Brownian motions, all of which will be running at rate 2, and two sets of countably many independent exponential random variables, with parameters $(1+\lambda)\epsilon^{-1}$ and $\alpha$. We will construct the process progressively, so we shall assume the random variables within these two countable sets have an order, but we will make no attempt to explicitly keep track of the index of of the random variable we use in any particular event. We shall call the dual process, $\xi_t = (\xi_t^i, t \geq 0)_{i=1}^{N_t}$, and initially, let $\xi_0 = (\xi_0^0)$ consist of a single point $x \in \mathbb{R}$. We also start the first of the exponential clocks with parameter $(1+\lambda)\epsilon^{-1}$.

The particle $\xi^0$ performs a Brownian motion until this clock rings, at which point we replace $\xi^0$ with three particles, each of which performs an independent Brownian motion started at the point at which $\xi^0$ terminated. We shall call these offspring particles $\xi^1$, $\xi^2$ and $\xi^3$. We also initiate three more exponential clocks with parameter $(1+\lambda)\epsilon^{-1}$, calling them $E^1$, $E^2$ and $E^3$, and three exponential clocks with parameter $\alpha$, calling them $E^{12}$, $E^{13}$ and $E^{23}$. Finally, we set

$T^{ij} = \inf\{t \geq 0 : L_t^0(\xi^i - \xi^j) > E^{ij}\}$ for $1 \leq i < j \leq 3$, where

$$L_t^a(B_t) := \lim_{\varepsilon \to 0} \frac{1}{\epsilon} \int_0^t \mathbb{1}\{a - \varepsilon < B_s < a + \varepsilon\} ds$$

is the local time up to time $t$ of the Brownian motion $B$ (also running at rate 2) at $a \in \mathbb{R}$.

The particles $\xi^1$, $\xi^2$ and $\xi^3$ are Brownian motions until time $\min(E^1, E^2, E^3, T^{12}, T^{13}, T^{23})$ has elapsed. If $E^i$ is smallest, we replace $\xi^i$ with three new particles as we did for $\xi^0$, initiate the corresponding new exponential clocks; one with parameter $(1+\lambda)\epsilon^{-1}$ for each new particle, and one with parameter $\alpha$ for each new pair of particles in $\xi$ (and the corresponding $T^{ij}$); and discard any of the existing variables that were associated to $\xi^i$, i.e., the two $T^{ij}$ and the two $E^{ij}$.

If $T^{ij}$ is smallest, where $i < j$, then we remove particle $\xi^j$ from $\xi$, and remove any variable associated with $\xi^j$, i.e., $E^j$, both $T^{kj}$ and both $E^{kj}$, where $k \in [3] \setminus \{j\}$.

Once all the required variables have been removed and/or replaced, the remaining particles perform Brownian motion until the next exponential clock rings, at which point the process of replacing and removing particles repeats. We shall denote the number of particles in $\xi_t$ at time $t$ by $N_t$. Finally, we comment that we can start the process from finitely many particles, each starting at a possibly distinct point in $\mathbb{R}$, proceeding the same way as the above process immediately after the first time it reaches that number of particles. We shall abuse the notation from Chapter 2 and denote the graph structure of $\xi$ at time $t$ by $\mathcal{T}(\xi_t)$, which is not a ternary tree (nor even a graphical tree at all), due to the coalescence events. However, all the branching events are still ternary, and we can still designate the vertices which exist at the earliest time as roots (there may be more than one if we have started with more than one particle), and the remaining vertices of degree 1, which all should exist at the latest time, as the leaves.

Although the underlying process has changed, the voting procedure at the branch points remains the same. However, we need to specify what happens at the new coalescence points we have introduced. To be explicit, we define a voting procedure on $\mathcal{T}(\xi_t)$ as follows.

1. Each leaf of $\mathcal{T}(\xi_t)$, independently, votes 1 with probability $u_0(W_i(t))$ and otherwise votes 0.

2. At each branch point in $\mathcal{T}(\xi_t)$, the vote of the parent particle is the majority vote of the votes of its three children, unless precisely one vote is 0, in which case it votes 1 with

probability $\frac{3+\lambda}{3+3\lambda}$.

3. At a coalescence point in $\mathcal{T}(\xi_t)$, the two 'parents' each deterministically receive the vote of their common 'child'.

This defines an iterative voting procedure, which runs inwards from the leaves of $\mathcal{T}(W(t)$ to the roots. Denote the vote at the root starting at point $x \in \mathbb{R}$ by $\mathbb{V}_x(\xi_t)$.

Before introducing the duality relation, we will first describe the martingale problem which the SPDE (3.1) satisfies. Writing it in its integral form, after multiplying by a continuous, bounded, nonnegative test function $\phi : \mathbb{R} \to \mathbb{R}$ in the domain of the Laplacian, gives the equation

$$\langle u_t, \phi \rangle = \langle u_0, \phi \rangle + \int_0^t \langle u_s, \Delta\phi \rangle ds + \epsilon^{-1} \int_0^t \langle f_\lambda(u_s), \phi \rangle ds + \epsilon^\gamma M_t(\phi), \tag{3.15}$$

where $M_t(\phi)$ is a martingale with quadratic variation

$$[M(\phi)]_t = \int_0^t \langle u_s(1-u_s), \phi^2 \rangle ds.$$

The duality we expect between solutions to (3.1) and $\xi_t$ is expressed as

$$E_{u_0}\left[\prod_{i=1}^n \langle u_t, \phi_i \rangle\right] = \int_{\mathbb{R}^n} E_{x_1,\dots,x_n}\left[\prod_{i=1}^n \mathbb{V}_{x_i}(\xi_t)\right] \prod_{i=1}^n \phi_i(x_i) dx_1 \dots dx_n, \tag{3.16}$$

where the expectation on the RHS is taken with respect to the dual process $\xi$ starting from $n$ particles at locations $x_1$ to $x_n$.

As $\langle u_t, \phi \rangle$ is a semimartingale, its law is determined by its drift and quadratic variation. Thus, we really only need to establish (3.16) for $n = 1$ and $n = 2$, with $\phi_1 = \phi_2$ and $x_1 = x_2$. Furthermore, to make the calculations simpler, we will show this for an infinitesimally small time interval $\delta t$. The same arguments will work for longer time intervals, but care needs to be taken in order to properly bound the errors introduced. Again, we appeal to [30] for those interested in a more careful approach.

First, we note that

$$\frac{d}{dt} E_{u_0}\left[\langle u_t, \phi \rangle\right]\Big|_{t=0} = \langle u_0, \Delta\phi \rangle + \epsilon^{-1} \langle f_\lambda(u_0), \phi \rangle, \tag{3.17}$$

which should be clear from the semimartingale decomposition in (3.15). Next, we use Itô's formula to write

$$\begin{aligned} d\langle u_t, \phi\rangle^2 =& 2\langle u_t, \phi\rangle d\langle u_t, \phi\rangle + d[M(\phi)]_t \\ =& \langle u_t, \phi\rangle(2\langle u_t, \Delta\phi\rangle + 2\epsilon^{-1}\langle f_\lambda(u_t), \phi\rangle)dt + \langle u_s(1-u_s), \phi^2\rangle dt + \langle u_t, \phi\rangle dM_t(\phi), \end{aligned}$$

and we can calculate

$$\frac{d}{dt}E_{u_0}\left[\langle u_t, \phi\rangle^2\right]\bigg|_{t=0} = \langle u_0, \phi\rangle(2\langle u_0, \Delta\phi\rangle + 2\epsilon^{-1}\langle f_\lambda(u_0), \phi\rangle) + \langle u_0(1-u_0), \phi^2\rangle. \tag{3.18}$$

We now consider the expectation in the RHS of (3.16) for $n = 1$. When we start with a single particle, there is nothing for this particle to coalesce with, so the first event that can occur is a branching event. After this branching event, a coalescing event may occur, or perhaps another branching event. However, as these additional coalescing or branching events occur after exponentially distributed times, the probability of more than one occurring in the interval $\delta t$ is $o(\delta t)$. So we may write

$$\begin{aligned} E_x\left[\mathbb{V}_x(\xi_{\delta t})\right] =& (1-(1+\lambda)\epsilon^{-1}\delta t)E_x\left[u_0(\xi_0)\right] + o(\delta t) \\ &+ (1+\lambda)\epsilon^{-1}\delta t E_x\left[u_0(\xi^1_{\delta t})u_0(\xi^2_{\delta t})u_0(\xi^3_{\delta t}) + \frac{3+\lambda}{3+3\lambda}\sum_{\substack{(i,j,k)\in\\ \{(1,2,3),(2,3,1),(3,1,2)\}}}(u_0(\xi^i_{\delta t})u_0(\xi^j_{\delta t})(1-u_0(\xi^k_{\delta t})))\right]. \end{aligned}$$

As $\delta t$ is small, the heat semigroup is regular enough to write

$$E_x\left[u_0(\xi^1_{\delta t})u_0(\xi^2_{\delta t})u_0(\xi^3_{\delta t})\right] = u_0(x)^3 + \mathcal{O}(\delta t),$$

and as the $\xi^1_t$, $\xi^2_t$ and $\xi^3_t$ are independent once they have branched, we may similarly write

$$E_x\left[\sum_{\substack{(i,j,k)\in\\ \{(1,2,3),(2,3,1),(3,1,2)\}}}(u_0(\xi^i_{\delta t})u_0(\xi^j_{\delta t})(1-u_0(\xi^k_{\delta t})))\right] = 3u_0(x)^2(1-u_0(x)) + \mathcal{O}(\delta t).$$

Taking the derivative of RHS of (3.16) and evaluating it at $t = 0$ then gives

$$\begin{aligned}
&\lim_{\delta t\to 0}\frac{1}{\delta t}\left(\int_{\mathbb{R}} E_x\left[u_0(\xi^0_{\delta t})\right]\phi(x)dx - \int_{\mathbb{R}} E_x\left[u_0(\xi^0_0)\right]\phi(x)dx\right)\\
&\quad =(1+\lambda)\epsilon^{-1}\int_{\mathbb{R}}\left(u_0(x)^3 + \frac{3+\lambda}{1+\lambda}u_0(x)^2(1-u_0(x)) - u_0(x)\right)\phi(x)dx\\
&\qquad + \lim_{\delta t\to 0}\frac{1}{\delta t}\int_{\mathbb{R}}(E_x\left[\phi(\xi_{\delta t})\right] - \phi(x))u_0(x)dx\\
&\quad =\epsilon^{-1}\langle f_\lambda(u_0(\cdot)),\phi\rangle + \langle \Delta\phi, u_0\rangle
\end{aligned}$$

where the coefficient of $\frac{1}{2}$ from the generator of the Laplacian cancels with the factor of two from the increased diffusivity of the Brownian motions we are using, and we exchanged the order of the integral and expectation in the second line. This is the same expression as was calculated in (3.17), so all that remains is to check the second moment.

Consider now the expectation in the RHS of (3.16) for $n = 2$. This time, we have three possible outcomes which we expect to occur with probability greater than $o\left(\delta t\right)$: a coalescence occurs, one of the two particles branches, or no coalescence nor branching occurs. However, as coalescence occurs after an exponentially distributed local time, we take slightly more care with this variable.

First, let us condition on whether coalescence occurs in the interval $\delta t$. Denote the exponential random variable associated with the coalescence time between these two particles by $E$. Then we have

$$\begin{aligned}
E_{x,y}\left[\mathbb{V}_x(\xi^1_{\delta t})\,\mathbb{V}_y(\xi^2_{\delta t})\right] =&E_{x,y}\left[u_0(\xi^1_{\delta t})u_0(\xi^2_{\delta t})\mathbb{1}_{\{L^0_{\delta t}(\xi^1,\xi^2)<E\}}\right] + E_{x,y}\left[u_0(\xi^1_{\delta t})\mathbb{1}_{\{L^0_{\delta t}(\xi^1,\xi^2)\geq E\}}\right]\\
=&E_{x,y}\left[u_0(\xi^1_{\delta t})u_0(\xi^2_{t})e^{-L^0_{\delta t}(\xi^1,\xi^2)}\right] + E_{x,y}\left[u_0(\xi^1_{\delta t})\left(1-e^{-L^0_{\delta t}(\xi^1,\xi^2)}\right)\right]\\
=&E_{x,y}\left[u_0(\xi^1_{\delta t})u_0(\xi^2_{\delta t})\right] + E_{x,y}\left[u_0(\xi^1_{\delta t})(1-u_0(\xi^2_{\delta t}))\left(1-e^{-L^0_{\delta t}(\xi^1,\xi^2)}\right)\right]
\end{aligned}$$

We shall return this expectation to its integral in order to define,

$$\begin{aligned}
&\int_{\mathbb{R}^2} E_{x,y}\left[\mathbb{V}_x(\xi^1_{\delta t})\,\mathbb{V}_y(\xi^2_{\delta t})\right]\phi(x)\phi(y)dxdy\\
&\qquad\qquad = \int_{\mathbb{R}^2} E_{x,y}\left[u_0(\xi^1_{\delta t})u_0(\xi^2_{\delta t})\right]\phi(x)\phi(y)dxdy\\
&\qquad\qquad\quad + \int_{\mathbb{R}^2} E_{x,y}\left[u_0(\xi^1_{\delta t})(1-u_0(\xi^2_{\delta t}))\left(1-e^{\{L^0_{\delta t}(\xi^1,\xi^2)\geq E\}}\right)\right]\phi(x)\phi(y)dxdy
\end{aligned}$$

$$=: I_1 + I_2.$$

Next, we consider branching. The second integral, $I_2$, is $\mathcal{O}\left(L^0_{\delta t}(\xi^1, \xi^2)\right)$, and $L^0_{\delta t}(\xi^1, \xi^2) < \delta t$, so conditioning this term on whether branching occurs will only introduce terms of order $o\left(\delta t\right)$. Furthermore, we notice that in $I_1$, the motions of the two initial particles, and their possible descendants, are independent of each other, so we can take the product of their expectations. Applying calculations performed above in the first moment case, we see that

$$\begin{aligned}
I_1 = &\int_{\mathbb{R}^2} \Bigg( (1 - 2(1+\lambda)\epsilon^{-1} dt) E_x \left[u_0(\xi^1_{\delta t})\right] E_y \left[u_0(\xi^2_{\delta t})\right] + o\left(\delta t\right), \\
&\quad + (1+\lambda)\epsilon^{-1}\delta t u_0(x) \left( u_0(y)^3 + \frac{3+\lambda}{1+\lambda} u_0(y)^2 (1 - u_0(y)) \right) \\
&\quad + (1+\lambda)\epsilon^{-1}\delta t u_0(y) \left( u_0(x)^3 + \frac{3+\lambda}{1+\lambda} u_0(x)^2 (1 - u_0(x)) \right) \Bigg) \phi(x)\phi(y) dx dy \\
= &\left\langle E_\cdot \left[\phi(\xi^1_{\delta t})\right], u_0 \right\rangle^2 + 2\epsilon^{-1}\delta t \langle f_\lambda(u_0), \phi\rangle \langle u_0, \phi\rangle + o\left(\delta t\right),
\end{aligned}$$

where again, in the first term of the last line we switched the order of the expectation and integration, as well as the labels of the integrating variables. Next, consider $I_2$. We begin by translating the variables of integration, exchanging the order of the integral and expectation, and finally changing the variable from $(x, y)$ to $(x, y - x)$,

$$\begin{aligned}
I_2 = &E_{0,0} \left[ \int_{\mathbb{R}^2} u_0(\xi^1_{\delta t} + x)(1 - u_0(\xi^2_{\delta t} + y)) \left(1 - e^{-L^{y-x}_{\delta t}(\xi^1, \xi^2)}\right) \phi(x)\phi(y) dx dy \right] \\
= &E_{0,0} \left[ \int_{\mathbb{R}^2} u_0(u)(1 - u_0(u+v)) \left(1 - e^{-L^{v}_{\delta t}(\xi^1, \xi^2)}\right) \phi(u - \xi^1_{\delta t})\phi(u + v - \xi^2_{\delta t}) du dv \right]. \\
=: &\int_{\mathbb{R}^2} u_0(u)(1 - u_0(u+v)) E_{0,0}\left[Y_{\delta t}\right] du dv.
\end{aligned}$$

Applying Itô's formula to $Y_t$ gives,

$$\begin{aligned}
dY_t = &e^{-L^v_t(\xi^1, \xi^2)} \phi(u - \xi^1_t)\phi(v + u - \xi^2_t) dL^v_t(\xi^1, \xi^2) \\
&- \left(1 - e^{-L^v_t(\xi^1, \xi^2)}\right) \phi'(u - \xi^1_t)\phi(u + v - \xi^2_t) d\xi^1_t \\
&- \left(1 - e^{-L^v_t(\xi^1, \xi^2)}\right) \phi(u - \xi^1_t)\phi'(u + v - \xi^2_t) d\xi^2_t \\
&+ \tfrac{1}{2} \left(1 - e^{-L^v_t(\xi^1, \xi^2)}\right) \phi''(u - \xi^1_t)\phi(u + v - \xi^1_t) dt
\end{aligned}$$

$$+ \tfrac{1}{2}\left(1 - e^{-L_t^v(\xi^1,\xi^2)}\right)\phi(u-\xi_t^1)\phi''(u+v-\xi_t^2)dt.$$

The second and third terms in this expression are martingales, so they vanish when we take the expectation. Hence,

$$\begin{aligned}
E_{0,0}\left[Y_{\delta t}\right] =& E_{0,0}\left[\int_0^{\delta t} e^{-L_t^v(\xi^1,\xi^2)}\phi(u-\xi_t^1)\phi(u+v-\xi_t^2)dL_t^y(\xi^1,\xi^2)\right] \\
&+ \tfrac{1}{2}E_{0,0}\left[\int_0^{\delta t}\left(1 - e^{-L_s^v(\xi^1,\xi^2)}\right)\phi''(u-\xi_s^1)\phi(u+v-\xi_s^1)ds\right] \\
&+ \tfrac{1}{2}E_{0,0}\left[\int_0^{\delta t}\left(1 - e^{-L_s^v(\xi^1,\xi^2)}\right)\phi(u-\xi_s^1)\phi''(u+v-\xi_s^2)ds\right] \\
=& \delta t\phi(u)\phi(u+v) + o\left(\delta t\right),
\end{aligned}$$

where the last line follows because $L_0^y(\xi^1,\xi^2) = 0$, so the integrand and duration of the interval in the final two terms in this expression are both $\mathcal{O}\left(\delta t\right)$, and hence are combined of order $o\left(\delta t\right)$. We similarly take a first order approximation to the first integral. Hence

$$\begin{aligned}
I_2 =& \delta t\int_{\mathbb{R}^2} u_0(u)(1-u_0(u+v))\phi(u)\phi(u+v)dudv + o\left(\delta t\right) \\
=& \delta t\int_{\mathbb{R}^2} u_0(x)(1-u_0(y))\phi(x)\phi(y)dxdy + o\left(\delta t\right)
\end{aligned}$$

Putting $I_1$ and $I_2$ together gives,

$$\begin{aligned}
&\lim_{\delta t\to 0}\frac{1}{\delta t}\left(\int_{\mathbb{R}^2} E_{x,y}\left[\mathbb{V}_x(\xi_{\delta t}^1)\,\mathbb{V}_y(\xi_{\delta t}^2)\right]\phi(x)\phi(y)dxdy - \int_{\mathbb{R}^2} E_{x,y}\left[\mathbb{V}_x(\xi_0^1)\,\mathbb{V}_x(\xi_0^2)\right]\phi(x)\phi(y)dxdy\right) \\
&\quad= \lim_{\delta t\to 0}\frac{1}{\delta t}\left(I_1 + I_2 - \langle u_0,\phi\rangle^2\right) \\
&\quad= \lim_{\delta t\to 0}\frac{1}{\delta t}\left(\left\langle E_\cdot\left[\phi(\xi_{\delta t}^1)\right], u_0\right\rangle^2 - \langle u_0,\phi\rangle^2\right) + 2\epsilon^{-1}\langle f_\lambda(u_0(\cdot)),\phi\rangle\langle u_0,\phi\rangle + \langle u_0(1-u_0),\phi^2\rangle \\
&\quad= \frac{d}{dt}\left\langle E_\cdot\left[\phi(\xi_t^1)\right], u_0\right\rangle^2\Big|_{t=0} + 2\epsilon^{-1}\langle f_\lambda(u_0(\cdot)),\phi\rangle\langle u_0,\phi\rangle + \langle u_0(1-u_0),\phi^2\rangle \\
&\quad= 2\langle\phi,u_0\rangle\langle\Delta\phi,u_0\rangle + 2\epsilon^{-1}\langle f_\lambda(u_0(\cdot)),\phi\rangle\langle u_0,\phi\rangle + \langle u_0(1-u_0),\phi^2\rangle.
\end{aligned}$$

This agrees with (3.18), so we are finished. □

## 3.2 Control of stopping times

In this section, we provide necessary bounds on the stopping times $\sigma_1$, $\sigma_2$ and $\sigma_3$ of (3.4), (3.5) and (3.6). The strategy is essentially the same as in [20], however, we exploit Lemma 3.3 several times, in place of assuming the noise has compact support.

### 3.2.1 Control of $\sigma_1$

In this subsection we shall prove the following result.

**Proposition 3.8.** *For $t < \infty$,*

$$\lim_{\epsilon \downarrow 0} \mathbb{P}\left(\sigma_1 \leq \min(\sigma_2, \sigma_3, \sigma_4, t)\right) = 0.$$

Before proving Proposition 3.8, we need to prove the following lemma.

**Lemma 3.9.** *For $t < \infty$,*

$$\mathbb{P}\left(\sigma_1 \leq t, \int_{-\infty}^{\infty} v_{\sigma_1}^{\epsilon,\lambda}(x)(1 - v_{\sigma_1}^{\epsilon,\lambda}(x))dx < \epsilon^{-1/2-2\kappa}\right) = 0.$$

*Proof.* By Lemma 3.3, there exist $K < \infty$ and a random variable $X_0$ such that for all $s \leq t+1$ we have $v_s^{\epsilon,\lambda}(x) \leq e^{-Kx^2}$, for all $x \leq -X_0$, and $1 - v_s^{\epsilon,\lambda}(x) \leq e^{-Kx^2}$ for all $x \geq X_0$. By the definition of $\sigma_1$, on $\{\sigma_1 \leq t\}$, there exist times $t_1, t_2, \ldots$ such that $t_n \leq t+1$ for all $n$, $t_n \to \sigma_1$ as $n \to \infty$ and

$$\int_{-\infty}^{\infty} v_{t_n}^{\epsilon,\lambda}(x)(1 - v_{t_n}^{\epsilon,\lambda}(x))dx \geq \epsilon^{-1/2-2\kappa}.$$

On $\{X_0 < \infty\} \cap \{\sigma_1 \leq t\}$, since $v_{t_n}^{\epsilon,\lambda}(x)(1 - v_{t_n}^{\epsilon,\lambda}(x)) \leq \mathbb{1}_{x \in [-X_0, X_0]} + e^{-Kx^2}$, by dominated convergence,

$$\int_{-\infty}^{\infty} v_{t_n}^{\epsilon,\lambda}(x)(1 - v_{t_n}^{\epsilon,\lambda}(x))dx \to \int_{-\infty}^{\infty} v_{\sigma_1}^{\epsilon,\lambda}(x)(1 - v_{\sigma_1}^{\epsilon,\lambda}(x))dx$$

as $n \to \infty$. Hence on $\{X_0 < \infty\} \cap \{\sigma_1 \leq t\}$,

$$\int_{-\infty}^{\infty} v_{\sigma_1}^{\epsilon,\lambda}(x)(1 - v_{\sigma_1}^{\epsilon,\lambda}(x))dx \geq \epsilon^{-1/2-2\kappa}.$$

The result follows since $X_0 < \infty$ almost surely. □

*Proof of Proposition 3.8.* For $t \geq 0$, $x \in \mathbb{R}$, let $v_1(t,x) = v^{\epsilon,\lambda}(\epsilon^{3/2+2\gamma}t, x)$ and let $v_0(x) = u_0(\epsilon^{1/2}x)$. Then $v_1$ solves the SPDE

$$\frac{\partial v_1}{\partial t} = \Delta v_1 + f_\lambda(v_1) + \epsilon^{1/4+\gamma}\sqrt{v_1(1-v_1)}\dot{W}, \qquad v_1(0,\cdot) = v_0(\cdot), \tag{3.19}$$

where $\dot{W} = \dot{W}(t,x)$ is white noise. Writing (3.19) in integral form, we have

$$\begin{aligned} v_1(t,x) = &\int_{-\infty}^{\infty} e^{-t/4} G_t(x-y) v_0(y) dy \\ &+ \int_0^t \int_{-\infty}^{\infty} e^{-(t-s)/4} G_{t-s}(x-y) \left(\tfrac{1}{4} v_1(s,y) + f_\lambda(v_1(s,y))\right) dyds + \epsilon^{1/4+\gamma} M_t(x), \end{aligned} \tag{3.20}$$

where

$$M_t(x) = \int_0^t \int_{-\infty}^{\infty} e^{-(t-s)/4} G_{t-s}(x-y) \sqrt{v_1(s,y)(1-v_1(s,y))} W(ds\,dy). \tag{3.21}$$

Let $\tilde{\sigma} = \epsilon^{-3/2-2\gamma}\sigma$. Now

$$\begin{aligned} \int_0^t \int_{-\infty}^{\infty} & e^{-(t-s)/2} G_{t-s}(x-y)^2 v_1(s,y)(1-v_1(s,y)) dyds \\ &\leq \int_0^t \int_{-\infty}^{\infty} e^{-(t-s)/2} G_{t-s}(x-y)^2 dyds \\ &\leq \int_0^t e^{-(t-s)/2} \frac{1}{\sqrt{2\pi(t-s)}} \int_{-\infty}^{\infty} G_{t-s}(x-y) dyds \\ &< \int_0^{\infty} e^{-r/2} \frac{1}{\sqrt{2\pi r}} ds \\ &< \infty, \end{aligned}$$

where the second line holds a.s. by Lemma 3.7. It follows that $(M_t(x))_t$ is a martingale and therefore $\mathbb{E}M_{t\wedge\tilde{\sigma}}(x) = 0$. Hence

$$\begin{aligned} \mathbb{E}v_1(t\wedge\tilde{\sigma}, x) = &\mathbb{E}\int_{-\infty}^{\infty} e^{-(t\wedge\tilde{\sigma})/4} G_{t\wedge\tilde{\sigma}}(x-y) v_0(y) dy \\ &+ \mathbb{E}\int_0^{t\wedge\tilde{\sigma}} \int_{-\infty}^{\infty} e^{-((t\wedge\tilde{\sigma})-s)/4} G_{(t\wedge\tilde{\sigma})-s}(x-y) \left(\tfrac{1}{4} v_1(s,y) + f_\lambda(v_1(s,y))\right) dyds. \end{aligned}$$

Note that $f_\lambda(u) \leq -\frac{(3-\lambda)(1+\lambda)}{8} u$ if $0 \leq u \leq \frac{1+\lambda}{4}$ and $f(u) \leq 1$ for $0 \leq u \leq 1$. Therefore by

Lemma 3.7, a.s. for all $s \geq 0$, $y \in \mathbb{R}$,

$$\tfrac{1}{4}v_1(s,y) + f(v_1(s,y)) \leq \tfrac{5}{4}\mathbb{1}_{v_1(s,y)\geq \frac{1+\lambda}{4}}.$$

It follows that

$$\begin{aligned}\mathbb{E}v_1(t\wedge\tilde{\sigma},x) \leq & \mathbb{E}\int_{-\infty}^{\infty} e^{-(t\wedge\tilde{\sigma})/4} G_{t\wedge\tilde{\sigma}}(x-y)v_0(y)dy \\ & + \mathbb{E}\int_0^{t\wedge\tilde{\sigma}}\int_{-\infty}^{\infty} e^{-((t\wedge\tilde{\sigma})-s)/4} G_{(t\wedge\tilde{\sigma})-s}(x-y)\tfrac{5}{4}\mathbb{1}_{v_1(s,y)\geq\frac{1+\lambda}{4}}dyds.\end{aligned}$$

By Fubini's theorem and Lemma 3.7,

$$\mathbb{E}\int_{-\infty}^0 v_1(t\wedge\tilde{\sigma},x)dx = \int_{-\infty}^0 \mathbb{E}v_1(t\wedge\tilde{\sigma},x)dx.$$

Therefore,

$$\begin{aligned}\mathbb{E}\int_{-\infty}^0 v_1(t\wedge\tilde{\sigma},x)dx \leq & \int_{-\infty}^0 \mathbb{E}\int_{-\infty}^{\infty} e^{-(t\wedge\tilde{\sigma})/4} G_{t\wedge\tilde{\sigma}}(x-y)v_0(y)dydx \\ & + \int_{-\infty}^0 \mathbb{E}\int_0^{t\wedge\tilde{\sigma}}\int_{-\infty}^{\infty} e^{-((t\wedge\tilde{\sigma})-s)/4} G_{(t\wedge\tilde{\sigma})-s}(x-y)\tfrac{5}{4}\mathbb{1}_{v_1(s,y)\geq\frac{1+\lambda}{4}}dydsdx \\ = & \mathbb{E}e^{-(t\wedge\tilde{\sigma})/4}\int_{-\infty}^{\infty}\left(\int_{-\infty}^0 G_{t\wedge\tilde{\sigma}}(x-y)dx\right)v_0(y)dy \\ & + \tfrac{5}{4}\mathbb{E}\int_0^{t\wedge\tilde{\sigma}} e^{-1/4((t\wedge\tilde{\sigma})-s)}\int_{-\infty}^{\infty}\left(\int_{-\infty}^0 G_{(t\wedge\tilde{\sigma})-s}(x-y)dx\right)\mathbb{1}_{v_1(s,y)\geq\frac{1+\lambda}{4}}dyds\end{aligned} \tag{3.22}$$

by Fubini's theorem. Now since $0 \leq v_0 \leq 1$,

$$\begin{aligned} e^{-(t\wedge\tilde{\sigma})/4}\int_{-\infty}^{\infty}\left(\int_{-\infty}^0 G_{t\wedge\tilde{\sigma}}(x-y)dx\right)v_0(y)dy = & e^{-(t\wedge\tilde{\sigma})/4}\int_{-\infty}^{\infty}\mathbb{P}_y\left(B_{t\wedge\tilde{\sigma}}\leq 0\right)v_0(y)dy \\ \leq & \int_{-\infty}^0 v_0(y)dy + e^{-(t\wedge\tilde{\sigma})/4}\int_0^{\infty}\mathbb{P}_y\left(B_{t\wedge\tilde{\sigma}}\leq 0\right)dy \\ \leq & \int_{-\infty}^0 v_0(y)dy + e^{-(t\wedge\tilde{\sigma})/4}\int_0^{\infty} e^{-y^2/(2(t\wedge\tilde{\sigma}))}dy \\ = & \epsilon^{-1/2}\int_{-\infty}^0 u_0(y)dy + e^{-(t\wedge\tilde{\sigma})/4}\tfrac{1}{2}\sqrt{2\pi(t\wedge\tilde{\sigma})} \\ \leq & C + \sqrt{\pi}e^{-1/2}\end{aligned}$$

by (3.7). Similarly,

$$\begin{aligned}\int_{-\infty}^{\infty}\left(\int_{-\infty}^{0}G_{(t\wedge\tilde{\sigma})-s}(x-y)dx\right)\mathbb{1}_{v_1(s,y)\geq\frac{1+\lambda}{4}}dy =& \int_{-\infty}^{\infty}\mathbb{P}_y\left(B_{(t\wedge\tilde{\sigma})-s}\leq 0\right)\mathbb{1}_{v_1(s,y)\geq\frac{1+\lambda}{4}}dy\\ \leq& \int_{-\infty}^{0}\mathbb{1}_{v_1(s,y)\geq\frac{1+\lambda}{4}}dy+\int_{0}^{\infty}e^{-y^2/(2((t\wedge\tilde{\sigma})-s))}dy\\ \leq& \int_{-\infty}^{0}\tfrac{16}{(1+\lambda)^2}v_1(s,y)^2dy+\tfrac{1}{2}\sqrt{2\pi((t\wedge\tilde{\sigma})-s)},\end{aligned}$$

since $\mathbb{1}_{v_1(s,y)\geq\frac{1+\lambda}{4}}\leq\frac{16}{(1+\lambda)^2}v_1(s,y)^2$. Substituting into (3.22), it follows that

$$\begin{aligned}&\mathbb{E}\int_{-\infty}^{0}v_1(t\wedge\tilde{\sigma},x)dx\\ &\leq C+\sqrt{\pi}e^{-1/2}+\tfrac{5}{4}\mathbb{E}\int_{0}^{t\wedge\tilde{\sigma}}e^{-((t\wedge\tilde{\sigma})-s)/4}\left(\int_{-\infty}^{0}\tfrac{16}{(1+\lambda)^2}v_1(s,y)^2dy+\tfrac{1}{2}\sqrt{2\pi((t\wedge\tilde{\sigma})-s)}\right)ds.\end{aligned}\tag{3.23}$$

Since $a^2\leq 2(a-b)^2+2b^2$ for $a,b\in\mathbb{R}$, we have

$$\int_{-\infty}^{0}v(s,y)^2dy\leq 2\int_{-\infty}^{\infty}(v(s,y)-v^{\delta,\lambda}(s,y))^2dy+2\int_{-\infty}^{0}v^{\delta,\lambda}(s,y)^2dy.\tag{3.24}$$

We now bound each term on the right hand side of (3.24). For the first term, by the definition of $v^{\delta,\lambda}$,

$$\begin{aligned}\int_{-\infty}^{\infty}(v(s,y)-v^{\delta,\lambda}(s,y))^2dy =& \int_{-\infty}^{\infty}\left(\int_{-\infty}^{\infty}\rho^{\delta}(z)(v(s,y)-v(s,y-z))dz\right)^2dy\\ \leq& \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\rho^{\delta}(z)(v(s,y)-v(s,y-z))^2dzdy\\ \leq& \beta'\end{aligned}\tag{3.25}$$

for $s<\sigma$, where the second inequality follows by Jensen's inequality and the last inequality follows from the definition of $\sigma_3$. Then for the second term on the right hand side of (3.24),

$$\begin{aligned}\int_{-\infty}^{0}v^{\delta,\lambda}(s,y)^2dy\leq& 2\int_{-\infty}^{\infty}(v^{\delta,\lambda}(s,y)-m_{\eta(v_s^{\delta})}(y))^2dy+2\int_{-\infty}^{0}m_{\eta(v_s^{\delta})}(y)^2dy\\ \leq& 2\|S(v_s^{\delta})\|_{H^1}^2+2\left(\int_{-\infty}^{0}m(y)^2dy+|\eta(v_s^{\delta})|\right)\\ \leq& 2\beta^2+3\epsilon^{-1/2-\kappa}\end{aligned}$$

for $s < \sigma$, for $\epsilon$ sufficiently small, by the definition of $\sigma_2$ and $\sigma_4$. Therefore, substituting into (3.24), for $s < \sigma$, for $\epsilon$ sufficiently small,

$$\int_{-\infty}^{0} v(s,y)^2 dy \leq 2\beta' + 2(2\beta^2 + 3\epsilon^{-1/2-\kappa}) \leq 8\epsilon^{-1/2-\kappa}. \tag{3.26}$$

By substituting into (3.23), it follows that

$$\begin{aligned}
\mathbb{E}\int_{-\infty}^{0} v_1(t\wedge\tilde{\sigma},x)dx \leq& \tfrac{5}{4}\mathbb{E}\int_{0}^{t\wedge\tilde{\sigma}} e^{-((t\wedge\tilde{\sigma})-s)/4}\left(\tfrac{128}{(1+\lambda)^2}\epsilon^{-1/2-\kappa} + \tfrac{1}{2}\sqrt{2\pi((t\wedge\tilde{\sigma})-s)}\right)ds \\
&+ C + \sqrt{\pi}e^{-1/2} \\
\leq& \tfrac{5}{4}\mathbb{E}\int_{0}^{\infty} e^{-r/4}\left(\tfrac{128}{(1+\lambda)^2}\epsilon^{-1/2-\kappa} + \tfrac{1}{2}\sqrt{2\pi r}\right)ds + C + \sqrt{\pi}e^{-1/2} \\
\leq& \tfrac{650}{(1+\lambda)^2}\epsilon^{-1/2-\kappa}
\end{aligned}$$

for $\epsilon$ sufficiently small. Therefore for $t > 0$, and for $\epsilon$ sufficiently small,

$$\mathbb{E}\int_{-\infty}^{0} v^{\epsilon}(t\wedge\sigma,x)dx \leq \tfrac{650}{(1+\lambda)^2}\epsilon^{-1/2-\kappa}.$$

Since $f_\lambda(1-u) = -f_{-\lambda}(u)$ for $0 \leq u \leq 1$, $v_2(t,x) = 1 - v(\epsilon^{3/2+2\gamma}t,x)$ solves the SPDE (3.19) with initial condition $v_2(0,\cdot) = 1 - v_0(\cdot)$. Therefore, by the same argument,

$$\mathbb{E}\int_{0}^{\infty}(1 - v^{\epsilon}(t\wedge\sigma,x))dx \leq \tfrac{650}{(1+\lambda)^2}\epsilon^{-1/2-\kappa}.$$

It follows that for $t > 0$, and for $\epsilon$ sufficiently small,

$$\mathbb{E}\int_{-\infty}^{\infty} v^{\epsilon}(t\wedge\sigma,x)(1 - v^{\epsilon}(t\wedge\sigma,x))dx \leq \tfrac{1300}{(1+\lambda)^2}\epsilon^{-1/2-\kappa}.$$

We can now use Lemma 3.9 to complete the proof. We have

$$\begin{aligned}
&\mathbb{E}\int_{-\infty}^{\infty} v^{\epsilon}(t\wedge\sigma,x)(1 - v^{\epsilon}(t\wedge\sigma,x))dx \\
&\qquad\geq \mathbb{E}\mathbb{1}_{\sigma_1\leq\min(\sigma_2,\sigma_3,\sigma_4,t)}\epsilon^{-1/2-2\kappa}\mathbb{1}_{\int_{-\infty}^{\infty} v^{\epsilon}(t\wedge\sigma,x)(1-v^{\epsilon}(t\wedge\sigma,x))dx\geq\epsilon^{-1/2-2\kappa}} \\
&\qquad\geq \epsilon^{-1/2-2\kappa}\mathbb{P}\left(\sigma_1 \leq \min(\sigma_2,\sigma_3,\sigma_4,t)\right)
\end{aligned}$$

by Lemma 3.9. Therefore

$$\mathbb{P}\left(\sigma_1 \leq \min(\sigma_2, \sigma_3, \sigma_4, t)\right) \leq \tfrac{1300}{(1+\lambda)^2}\epsilon^{\kappa}$$

and the result follows, as $\lambda \in (-1, 1)$. $\square$

### 3.2.2 Control of $\sigma_2$

We shall use the following lemma from [20].

**Lemma 3.10.** *[Lemma 5.5 in [20]] There exists a constant $C < \infty$ such that if dist$(v_i, M) \leq \beta_1$ for $v_i \in L^2 + m$ and for $i = 1, 2$ then $|\eta(v_1) - \eta(v_2)| \leq C\|v_1 - v_2\|_{L^2}$.*

We can now prove the following lemma.

**Lemma 3.11.**

$$\mathbb{P}\left(\sigma_2 \leq t, \sigma_2 \leq \min(\sigma_1, \sigma_3, \sigma_4), \|S(v^{\delta,\lambda}_{\sigma_2})\|_{H^1} \neq \beta\right) = 0$$

*Proof.* By Lemma 3.3, for $t < \infty$ fixed, there exist a constant $K$ and a random variable $X_0$ such that $X_0 < \infty$ a.s. and for any $s \leq t + 1$, $v^{\epsilon,\lambda}_s(x) \leq e^{-Kx^2}$ for $x \leq -X_0$ and $1 - v^{\epsilon,\lambda}_s(x) \leq e^{-Kx^2}$ for $x \geq X_0$. Now suppose $t_n \leq t + 1$ for all $n$ and $t_n \to t_0$ as $n \to \infty$. For $x \leq -X_0 - \delta$,

$$|v^{\delta,\lambda}_{t_n}(x) - v^{\delta,\lambda}_{t_0}(x)| \leq e^{-K(x+\delta)^2}$$

and for $x \geq X_0 + \delta$, $|v^{\delta,\lambda}_{t_n}(x) - v^{\delta,\lambda}_{t_0}(x)| \leq e^{-K(x-\delta)^2}$. Also, $\nabla v^{\delta,\lambda}_s = \nabla\rho^\delta * v^{\epsilon,\lambda}_s$ and $\|\nabla\rho^\delta\|_\infty < \infty$ since $\rho$ is smooth and compactly supported. Hence for $x \leq -X_0 - \delta$,

$$|\nabla v^{\delta,\lambda}_{t_n}(x) - \nabla v^{\delta,\lambda}_{t_0}(x)| \leq 4\delta\|\nabla\rho^\delta\|_\infty e^{-K(x+\delta)^2}$$

and for $x \geq X_0 + \delta$,

$$|\nabla v^{\delta,\lambda}_{t_n}(x) - \nabla v^{\delta,\lambda}_{t_0}(x)| \leq 4\delta\|\nabla\rho^\delta\|_\infty e^{-K(x-\delta)^2}.$$

Therefore by dominated convergence, on $\{X_0 < \infty\}$, $\|v^{\delta,\lambda}_{t_n} - v^{\delta,\lambda}_{t_0}\|_{H^1} \to 0$ as $n \to \infty$.

On $\{X_0 < \infty\} \cap \{\sigma_2 \leq \min(t, \sigma_4)\}$, we have that $\text{dist}(v^{\delta,\lambda}_s, M) \leq \beta$ and $|\eta(v^{\delta,\lambda}_s)| \leq \epsilon^{-1/2-\kappa}$ for all $s < \sigma_2$. It follows that there exists a sequence $(t_n)$ such that $t_n \to \sigma_2$ as $n \to \infty$, $t_n \leq t+1$

for all $n$, $\|v_{t_n}^{\delta,\lambda} - m_{\eta(v_{t_n}^{\delta,\lambda})}\|_{L^2} \leq \beta$ and $\eta(v_{t_n}^{\delta,\lambda}) \to \eta$ as $n \to \infty$ for some $\eta \in \mathbb{R}$. Hence

$$\begin{aligned}\|v_{\sigma_2}^{\delta,\lambda} - m_\eta\|_{L^2} \leq& \|v_{t_n}^{\delta,\lambda} - v_{\sigma_2}^{\delta,\lambda}\|_{L^2} + \|m_\eta - m_{\eta(v_{t_n}^{\delta,\lambda})}\|_{L^2} + \beta \\ \leq& \|v_{t_n}^{\delta,\lambda} - v_{\sigma_2}^{\delta,\lambda}\|_{L^2} + \|\nabla m\|_{L^2}|\eta - \eta(v_{t_n}^{\delta,\lambda})| + \beta.\end{aligned}$$

Letting $n \to \infty$, it follows that $\|v_{\sigma_2}^{\delta,\lambda} - m_\eta\|_{L^2} \leq \beta$ and hence $\operatorname{dist}(v_{\sigma_2}^{\delta,\lambda}, M) \leq \beta$.

On $\{X_0 < \infty\} \cap \{\sigma_2 \leq \min(t, \sigma_4)\}$, by the definition of $\sigma_2$, $\|v_{\sigma_2 - 1/n}^{\delta,\lambda} - m_{\eta(v_{\sigma_2-1/n}^{\delta,\lambda})}\|_{H^1} < \beta$ for all $n$. By Lemma 3.10, since, for $\epsilon$ sufficiently small, $\operatorname{dist}(v_s^{\delta,\lambda}, M) \leq \beta_1$ for all $s \leq \sigma_2$, we have

$$\|m_{\eta(v_{\sigma_2-1/n}^{\delta,\lambda})} - m_{\eta(v_{\sigma_2}^{\delta,\lambda})}\|_{L^2} \leq C\|\nabla m\|_{L^2}\|v_{\sigma_2-1/n}^{\delta,\lambda} - v_{\sigma_2}^{\delta,\lambda}\|_{L^2}$$

and similarly

$$\|\nabla m_{\eta(v_{\sigma_2-1/n}^{\delta,\lambda})} - \nabla m_{\eta(v_{\sigma_2}^{\delta,\lambda})}\|_{L^2} \leq C\|\Delta m\|_{L^2}\|v_{\sigma_2-1/n}^{\delta,\lambda} - v_{\sigma_2}^{\delta,\lambda}\|_{L^2}.$$

Hence writing

$$\|v_{\sigma_2}^{\delta,\lambda} - m_{\eta(v_{\sigma_2}^{\delta,\lambda})}\|_{H^1} \leq \|v_{\sigma_2-1/n}^{\delta,\lambda} - v_{\sigma_2}^{\delta,\lambda}\|_{H^1} + \|m_{\eta(v_{\sigma_2-1/n}^{\delta,\lambda})} - m_{\eta(v_{\sigma_2}^{\delta,\lambda})}\|_{H^1} + \beta$$

and letting $n \to \infty$, since $\|v_{\sigma_2-1/n}^{\delta,\lambda} - v_{\sigma_2}^{\delta,\lambda}\|_{H^1} \to 0$ as $n \to \infty$, we have $\|S(v_{\sigma_2}^{\delta})\|_{H^1} \leq \beta$.

We now define two events. Let

$$\begin{aligned}A_1 = \{\exists (t_n)_n \text{ s.t. } & t_n \to \sigma_2 \text{ as } n \to \infty, \\ & t_n \leq t+1 \;\; \forall n, \\ & \|S(v_{t_n}^{\delta})\|_{H^1} \geq \beta \;\; \forall n \text{ and} \\ & \operatorname{dist}(v_{t_n}^{\delta,\lambda}, M) \leq \beta_1 \;\; \forall n\}.\end{aligned}$$

Also let

$$\begin{aligned}A_2 = \{\exists (t_n)_n \text{ s.t. } & t_n \to \sigma_2 \text{ as } n \to \infty, \\ & t_n \leq t+1 \;\; \forall n, \text{ and}\end{aligned}$$

$$\operatorname{dist}(v^{\delta,\lambda}_{t_n}, M) \geq \beta_1 \ \ \forall n\}.$$

Then by the definition of $\sigma_2$, $\{\sigma_2 \leq t\} \subset A_1 \cup A_2$. On $\{\sigma_2 \leq t\} \cap A_1 \cap \{X_0 < \infty\}$,

$$\begin{aligned}\|S(v^{\delta}_{\sigma_2})\|_{H^1} \geq & \|S(v^{\delta,\lambda}_{t_n})\|_{H^1} - \|S(v^{\delta,\lambda}_{\sigma_2}) - S(v^{\delta,\lambda}_{t_n})\|_{H^1} \\ \geq & \beta - \|v^{\delta,\lambda}_{\sigma_2} - v^{\delta,\lambda}_{t_n}\|_{H^1} - \|m_{\eta(v^{\delta,\lambda}_{\sigma_2})} - m_{\eta(v^{\delta,\lambda}_{t_n})}\|_{H^1} \\ \geq & \beta - \|v^{\delta,\lambda}_{\sigma_2} - v^{\delta,\lambda}_{t_n}\|_{H^1} - C\|\nabla m\|_{H^1}\|v^{\delta,\lambda}_{\sigma_2} - v^{\delta,\lambda}_{t_n}\|_{H^1}\end{aligned}$$

by Lemma 3.10. Letting $n \to \infty$, it follows that $\|S(v^{\delta,\lambda}_{\sigma_2})\|_{H^1} \geq \beta$.

On $\{\sigma_2 \leq t\} \cap A_2 \cap \{X_0 < \infty\}$, we have that for all $\eta \in \mathbb{R}$, $n \in \mathbb{N}$, $\|v^{\delta}_{t_n} - m_\eta\|_{L^2} \geq \beta_1$. Therefore

$$\|v^{\delta,\lambda}_{\sigma^2} - m_\eta\|_{L^2} \geq \|v^{\delta,\lambda}_{t_n} - m_\eta\|_{L^2} - \|v^{\delta,\lambda}_{\sigma^2} - v^{\delta,\lambda}_{t_n}\|_{L^2} \geq \beta_1 - \|v^{\delta,\lambda}_{\sigma^2} - v^{\delta,\lambda}_{t_n}\|_{L^2}.$$

Letting $n \to \infty$, it follows that $\operatorname{dist}(v^{\delta,\lambda}_{\sigma^2}, M) \geq \beta_1$. □

We have that

$$dv^{\delta,\lambda}_t(z) = \epsilon^{-3/2-2\gamma} b^\delta_t(z)dt + d\mu^\delta_t(z),$$

where

$$b^\delta_t(z) = \Delta v^{\delta,\lambda}_t(z) + \langle f_\lambda(v^{\epsilon,\lambda}_t(\cdot)), \rho^\delta(z-\cdot)\rangle$$

and

$$\mu^\delta_t(z) = \epsilon^{-1/2}\int_0^t \int_{-\infty}^{\infty} \rho^\delta(z-y)\sqrt{v^{\epsilon,\lambda}_t(y)(1-v^{\epsilon,\lambda}_t(y))}W(dy,ds).$$

The following lemma is the same as Lemma 5.1 in [20]. Recall the definition of $\mathcal{H}$ from (3.10).

**Lemma 3.12.** *For $t < \sigma$,*

$$d\mathcal{H}^p(v^\delta_t) = p\mathcal{H}^{p-1}(v^{\delta,\lambda}_t)\langle dv^{\delta,\lambda}_t, D\mathcal{H}(\cdot, v^{\delta,\lambda}_t)\rangle + \tfrac{1}{2}\epsilon^{-1}\int_{-\infty}^{\infty} V^{p,\delta}(y, v^{\delta,\lambda}_t)v^{\epsilon,\lambda}_t(y)(1-v^{\epsilon,\lambda}_t(y))dydt,$$

*where*

$$V^{p,\delta}(y,v) = p(p-1)\mathcal{H}^{p-2}(v)(V^\delta_1(y,v))^2 + p\mathcal{H}^{p-1}(v)V^\delta_2(y,v)$$

*and*

$$V_1^\delta(y,v) = (D\mathcal{H}(\cdot,v) * \rho^\delta)(y) \ \textit{and}\ V_2^\delta(y,v) = \langle(-\Delta - f'(v(\cdot)))\rho^\delta(\cdot - y), \rho^\delta(\cdot - y)\rangle.$$

*Proof.* The proof is essentially the same as the proof of Lemma 5.1 in [20]. First, by the integration by parts formula, for $t < \sigma$,

$$\begin{aligned} d\langle \Delta v_t^{\delta,\lambda}, v_t^{\delta,\lambda}\rangle =& 2\langle \Delta v_t^{\delta,\lambda}, dv_t^{\delta,\lambda}\rangle + \int_{-\infty}^{\infty} dz d[\Delta\mu^\delta(z), \mu^\delta(z)]_t \\ =& 2\langle \Delta v_t^{\delta,\lambda}, dv_t^{\delta,\lambda}\rangle + \epsilon^{-1}\int_{-\infty}^{\infty} \langle \Delta\rho^\delta(\cdot - y), \rho^\delta(\cdot - y)\rangle v_t^{\epsilon,\lambda}(y)(1 - v_t^{\epsilon,\lambda}(y))dydt. \end{aligned}$$

Also, for $t < \sigma$, by Itô's formula,

$$\begin{aligned} dF_0(v_t^\delta(z)) =& F_0'(v_t^\delta(z))dv_t^\delta(z) + \tfrac{1}{2}F_0''(v_t^\delta(z))d[\mu^\delta(z), \mu^\delta(z)]_t \\ =& F_0'(v_t^\delta(z))dv_t^\delta(z) + \tfrac{1}{2}\epsilon^{-1}F_0''(v_t^\delta(z))\int_{-\infty}^{\infty}(\rho^\delta(z-y))^2 v_t(y)(1 - v_t(y))dy \cdot dt. \end{aligned}$$

Therefore, for $t < \sigma$,

$$\begin{aligned} d\mathcal{H}(v_t^\delta) =& d\left(-\tfrac{1}{2}\langle \Delta v_t^\delta, v_t^\delta\rangle + \int_{-\infty}^{\infty} F_0(v_t^\delta(z))dz\right) \\ =& \langle dv_t^\delta, D\mathcal{H}(\cdot, v_t^\delta)\rangle + \tfrac{1}{2}\epsilon^{-1}\int_{-\infty}^{\infty} V_2^\delta(y, v_t^\delta)v_t(y)(1 - v_t(y))dy \cdot dt. \end{aligned}$$

By Itô's formula again,

$$d\mathcal{H}^p(v_t^{\delta,\lambda}) = p\mathcal{H}^{p-1}(v_t^{\delta,\lambda})d\mathcal{H}(v_t^{\delta,\lambda}) + \tfrac{1}{2}p(p-1)\mathcal{H}^{p-2}(v_t^{\delta,\lambda})d[\mathcal{H}(v^{\delta,\lambda}), \mathcal{H}(v^{\delta,\lambda})]_t.$$

The martingale part of $d\mathcal{H}(v_t^{\delta,\lambda})$ is given by

$$\langle d\mu_t^\delta, D\mathcal{H}(\cdot, v_t^{\delta,\lambda})\rangle = \epsilon^{-1/2}\int_{-\infty}^{\infty} V_1^\delta(y, v_t^{\delta,\lambda})\sqrt{v_t(y)(1 - v_t(y))}W(dydt)$$

for $t < \sigma$. Therefore, for $t < \sigma$,

$$d[\mathcal{H}(v^{\delta,\lambda}), \mathcal{H}(v^{\delta,\lambda})]_t = \epsilon^{-1}\int_{-\infty}^{\infty}(V_1^\delta(y, v_t^{\delta,\lambda}))^2 v_t(y)(1 - v_t(y))dy \cdot dt.$$

The result follows. ☐

**Lemma 3.13.** *There exists a constant* $C < \infty$ *such that a.s., for* $t < \sigma$*,*

$$\begin{aligned}\int_{-\infty}^{\infty} V^{p,\delta}(y, v_t^{\delta,\lambda}) v_t^{\epsilon,\lambda}(y)(1 - v_t^{\epsilon,\lambda}(y))dy \leq & p(p-1)\mathcal{H}^{p-2}(v_t^{\delta,\lambda})\|D\mathcal{H}(\cdot, v_t^{\delta,\lambda})\|_{L^2}^2 \\ & + Cp\mathcal{H}^{p-1}(v_t^{\delta,\lambda})\delta^{-3}\epsilon^{-1/2-2\kappa}.\end{aligned}$$

*Proof.* By Lemma 3.7, almost surely

$$\begin{aligned}\int_{-\infty}^{\infty}(V_1^{\delta}(y, v_t^{\delta,\lambda}))^2 v_t^{\epsilon,\lambda}(y)(1 - v_t^{\epsilon,\lambda}(y))dy \leq & \int_{-\infty}^{\infty}((D\mathcal{H}(\cdot, v_t^{\delta,\lambda}) * \rho^{\delta})(y))^2 dy \\ \leq & \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \rho^{\delta}(z)(D\mathcal{H}(y-z, v_t^{\delta,\lambda}))^2 dz dy \\ = & \|D\mathcal{H}(\cdot, v_t^{\delta,\lambda})\|_{L^2}^2, \end{aligned} \tag{3.27}$$

where the second inequality follows by Jensen's inequality. Also, by Lemma 3.7, almost surely,

$$V_2^{\delta}(y, v_t^{\delta,\lambda}) \leq \langle -\Delta\rho^{\delta}, \rho^{\delta}\rangle + \|f'\|_{[0,1]}\|\rho^{\delta}\|_{L^2}^2 \leq C\delta^{-3},$$

where $C$ is a constant. The result follows by the definition of $\sigma_1$. ☐

Note that $b_t^{\delta}(z) = -D\mathcal{H}(z, v_t^{\delta,\lambda}) + R_t^{\delta}(z) + Q_t^{\delta}(z)$, where $R_t^{\delta}(z) = \langle f_0(v_t^{\epsilon,\lambda}(\cdot)), \rho^{\delta}(z - \cdot)\rangle - f_0(v_t^{\delta,\lambda}(z))$ and $Q_t^{\delta}(z) = \langle (f_\lambda - f_0)(v_t^{\epsilon,\lambda}(\cdot)), \rho^{\delta}(z - \cdot)\rangle$.

**Lemma 3.14.** *For* $t < \sigma$*,*

$$\|R_t^{\delta}\|_{L^2}^2 \leq 4\|f_0'\|_{[0,1]}^2\beta'$$

*Proof.* Since $\int_{-\infty}^{\infty}\rho^{\delta}(z)dz = 1$, we can write this as

$$R_t^{\delta}(z) = \int_{-\infty}^{\infty}\rho^{\delta}(y)(f_0(v_t^{\epsilon,\lambda}(z-y)) - f_0(v_t^{\epsilon,\lambda}(z)))dy + f_0(v_t^{\epsilon,\lambda}(z)) - f_0\left(\int_{-\infty}^{\infty}\rho^{\delta}(y)v_t^{\epsilon,\lambda}(z-y)dy\right).$$

Using the fact that $(a+b)^2 \leq 2(a^2+b^2)$, and then applying Jensen's inequality and the mean value theorem together with Lemma 3.7, it follows that

$$(R_t^{\delta}(z))^2 \leq 2\left(\int_{-\infty}^{\infty}\rho^{\delta}(y)(f_0(v_t^{\epsilon,\lambda}(z-y)) - f_0(v_t^{\epsilon,\lambda}(z)))dy\right)^2$$

$$
\begin{aligned}
&+2\left(f_0(v_t^{\epsilon,\lambda}(z))-f_0\left(\int_{-\infty}^{\infty}\rho^\delta(y)v_t^{\epsilon,\lambda}(z-y)dy\right)\right)^2\\
\leq&2\int_{-\infty}^{\infty}\rho^\delta(y)(f_0(v_t^{\epsilon,\lambda}(z-y))-f_0(v_t^{\epsilon,\lambda}(z)))^2dy\\
&+2\|f_0'\|_{[0,1]}^2\left(v_t^{\epsilon,\lambda}(z)-\int_{-\infty}^{\infty}\rho^\delta(y)v_t^{\epsilon,\lambda}(z-y)dy\right)^2\\
\leq&2\int_{-\infty}^{\infty}\rho^\delta(y)\|f_0'\|_{[0,1]}^2(v_t^{\epsilon,\lambda}(z-y)-v_t^{\epsilon,\lambda}(z))^2dy\\
&+2\|f_0'\|_{[0,1]}^2\int_{-\infty}^{\infty}\rho^\delta(y)(v_t^{\epsilon,\lambda}(z-y)-v_t^{\epsilon,\lambda}(z))^2dy\\
=&4\|f_0'\|_{[0,1]}^2\int_{-\infty}^{\infty}\rho^\delta(y)(v_t^{\epsilon,\lambda}(z-y)-v_t^{\epsilon,\lambda}(z))^2dy,
\end{aligned}
$$

where the third inequality also uses the mean value theorem, Lemma 3.7 and Jensen's inequality. Therefore, for $t<\sigma$, we have

$$
\|R_t^\delta\|_{L^2}^2\leq 4\|f_0'\|_{[0,1]}^2\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\rho^\delta(y)(v_t^{\epsilon,\lambda}(z-y)-v_t^{\epsilon,\lambda}(z))^2dydz\leq 4\|f'\|_{[0,1]}^2\beta'
$$

by the definition of $\sigma_3$. ☐

**Lemma 3.15.** *For $t<\sigma$,*

$$
\|Q_t^\delta\|_{L^2}^2\leq\lambda^2\epsilon^{-1/2-2\kappa}
$$

*Proof.* Using Jensen's inequality, and the fact that the integrand is bounded by 1, we can write

$$
\begin{aligned}
(Q_t^\delta(z))^2=&\left(\int_{-\infty}^{\infty}\rho^\delta(y-z)(f_\lambda(v_t^{\epsilon,\lambda}(y))-f_0(v_t^{\epsilon,\lambda}(y)))dy\right)^2\\
=&\lambda^2\left(\int_{-\infty}^{\infty}\rho^\delta(y-z)(1-v_t^{\epsilon,\lambda}(y))v_t^{\epsilon,\lambda}(y)dy\right)^2\\
\leq&\lambda^2\int_{-\infty}^{\infty}(\rho^\delta(y-z)(1-v_t^{\epsilon,\lambda}(y))v_t^{\epsilon,\lambda}(y))^2dy\\
\leq&\lambda^2\int_{-\infty}^{\infty}\rho^\delta(y-z)(1-v_t^{\epsilon,\lambda}(y))v_t^{\epsilon,\lambda}(y))dy.
\end{aligned}
$$

Therefore, for $t<\sigma$, we have

$$
\|Q_t^\delta\|_{L^2}^2\leq\lambda^2\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\rho^\delta(y-z)(1-v_t^{\epsilon,\lambda}(y))v_t^{\epsilon,\lambda}(y))dydz\leq\lambda^2\epsilon^{-1/2-2\kappa}
$$

by the definition of $\sigma_1$. ☐

We can now prove the main result of this subsection.

**Proposition 3.16.**

$$\lim_{\epsilon \downarrow 0} \mathbb{P}\left(\sigma_2 \leq t, \sigma_2 \leq \min(\sigma_1, \sigma_3, \sigma_4)\right) = 0.$$

*Proof.* Recall that $b_t^\delta(z) = -D\mathcal{H}(z, v_t^{\delta,\lambda}) + R_t^\delta(z) + Q_t^\delta(z)$. Hence for $t < \sigma$,

$$\begin{aligned}
&p\mathcal{H}^{p-1}(v_t^{\delta,\lambda})\langle dv_t^{\delta,\lambda}, D\mathcal{H}(\cdot, v_t^{\delta,\lambda})\rangle \\
&\qquad =p\mathcal{H}^{p-1}(v_t^{\delta,\lambda})\epsilon^{-3/2-2\gamma}\left\langle D\mathcal{H}(\cdot, v_t^{\delta,\lambda}), -D\mathcal{H}(\cdot, v_t^{\delta,\lambda}) + R_t^\delta(\cdot) + Q^\delta(\cdot)\right\rangle dt + d\mu_t^\delta,
\end{aligned}$$

where

$$\mu_t^\delta = p\int_0^t \mathcal{H}^{p-1}(v_s^{\delta,\lambda})\langle D\mathcal{H}(\cdot, v_s^{\delta,\lambda}), d\mu_s^\delta(\cdot)\rangle.$$

Therefore, by Lemmas 3.12 and 3.13, and since $|\langle D\mathcal{H}(\cdot, v_s^{\delta,\lambda}), R_s^\delta(\cdot)\rangle| \leq \frac{1}{2}(\|D\mathcal{H}(\cdot, v_s^{\delta,\lambda})\|_{L^2}^2 + \|R_s^\delta\|_{L^2}^2)$, and the same bound for $|\langle D\mathcal{H}(\cdot, v_s^{\delta,\lambda}), Q_s^\delta(\cdot)\rangle|$, we have that, for $t \leq \sigma$,

$$\begin{aligned}
&\mathcal{H}^p(v_t^{\delta,\lambda}) + p\epsilon^{-3/2-2\gamma}\int_0^t \mathcal{H}^{p-1}(v_s^{\delta,\lambda})\|D\mathcal{H}(\cdot, v_s^{\delta,\lambda})\|_{L^2}^2 ds \\
&\leq\mathcal{H}^p(v_0^{\delta,\lambda}) + \tfrac{p}{2}\epsilon^{-3/2-2\gamma}\int_0^t \mathcal{H}^{p-1}(v_s^{\delta,\lambda})(2\|D\mathcal{H}(\cdot, v_s^{\delta,\lambda})\|_{L^2}^2 + \|R_s^\delta\|_{L^2}^2 + \|Q_s^\delta\|_{L^2}^2)ds + \mu_t^\delta \\
&\quad + \tfrac{1}{2}\epsilon^{-1}\left(p(p-1)\int_0^t \mathcal{H}^{p-2}(v_s^{\delta,\lambda})\|D\mathcal{H}(\cdot, v_s^{\delta,\lambda})\|_{L^2}^2 ds + Cp\delta^{-3}\epsilon^{-1/2-2\kappa}\int_0^t \mathcal{H}^{p-1}(v_s^{\delta,\lambda})ds\right).
\end{aligned}$$

Hence, rearranging and using Lemmas 3.14 and 3.15, for $t \leq \sigma$,

$$\begin{aligned}
&\mathcal{H}^p(v_t^{\delta,\lambda}) + p\epsilon^{-3/2-2\gamma}\int_0^t \mathcal{H}^{p-1}(v_s^{\delta,\lambda})\|D\mathcal{H}(\cdot, v_s^{\delta,\lambda})\|_{L^2}^2 ds \\
&\leq\mathcal{H}^p(v_0^{\delta,\lambda}) + \tfrac{p}{2}\epsilon^{-3/2-2\gamma}\int_0^t \mathcal{H}^{p-1}(v_s^{\delta,\lambda})\left(4\|f'\|_{[0,1]}^2\beta' + \lambda^2\epsilon^{-1/2-\kappa}\right) ds + \mu_t^\delta \\
&\quad + \tfrac{1}{2}\epsilon^{-1}p(p-1)\int_0^t \mathcal{H}^{p-2}(v_s^{\delta,\lambda})\|D\mathcal{H}(\cdot, v_s^{\delta,\lambda})\|_{L^2}^2 ds + Cp\delta^{-3}\epsilon^{-3/2-2\kappa}\int_0^t \mathcal{H}^{p-1}(v_s^{\delta,\lambda})ds.
\end{aligned}$$

We now need to show that $(\mu_{t\wedge\sigma}^\delta)_t$ is a martingale. As in the proof of Lemma 3.12, we can calculate

$$\begin{aligned}
[\mu_{t\wedge\sigma}^\delta]_t =&\epsilon^{-1}p^2\int_0^{t\wedge\sigma}\int_{-\infty}^{\infty} \mathcal{H}^{2(p-1)}(v_s^{\delta,\lambda})V_1^\delta(y, v_s^{\delta,\lambda})^2 v_t(y)(1-v_t(y))dyds \\
\leq&\epsilon^{-1}p^2\int_0^{t\wedge\sigma} \mathcal{H}^{p-1}(v_s^{\delta,\lambda})^2\|D(\cdot, v_s^{\delta,\lambda})\|_{L^2}^2 ds
\end{aligned}$$

$$\leq \epsilon^{-1}p^2\int_0^{t\wedge\sigma}(c_2\beta^2)^{p-1}\cdot c_2(\beta+\|\Delta S(v_s^{\delta,\lambda})\|_{L^2})ds,$$

where the second line follows by (3.27) and the third line follows for $\epsilon$ sufficiently small by the definition of $\sigma_2$ and Theorem 3.6. Now

$$\Delta S(v_s^{\delta,\lambda}) = \Delta\rho^\delta * v_s - \Delta m_{\eta(v_s^{\delta,\lambda})}.$$

We have that $\|\Delta m_{\eta(v_s^{\delta,\lambda})}\|_{L^2} = \|\Delta m\|_{L^2} < \infty$. Also, since $\rho$ is smooth and compactly supported, $\|\Delta\rho^\delta\|_\infty < \infty$. By Lemma 3.3, there exist a constant $K$ and a random variable $X_0$ with $\mathbb{E}X_0 < \infty$ such that for any $s \leq t$,

$$v_s(x) \leq e^{-Kx^2} \text{ for } x \leq -X_0$$

and

$$1 - v_s(x) \leq e^{-Kx^2} \text{ for } x \geq X_0.$$

Therefore for $x \leq -X_0 - \delta$,

$$|\Delta\rho^\delta * v_s(x)| \leq \delta\|\Delta\rho^\delta\|_\infty e^{-K(x+\delta)^2}.$$

Also since $\Delta\rho^\delta * v_s(x) = -\Delta\rho^\delta * (1-v_s)(x)$, we have that for $x \geq X_0 + \delta$,

$$|\Delta\rho^\delta * v_s(x)| \leq \delta\|\Delta\rho^\delta\|_\infty e^{-K(x-\delta)^2}.$$

Therefore since $\mathbb{E}X_0 < \infty$, we have $\sup_{s\leq t}\mathbb{E}\|\Delta S(v_s^{\delta,\lambda})\|_{L^2} < \infty$. Hence $(\mu^\delta_{t\wedge\sigma})_t$ is a martingale and in particular $\mathbb{E}\mu^\delta_{t\wedge\sigma} = 0$.

Therefore taking expectations,

$$\begin{aligned}
&\mathbb{E}\mathcal{H}^p(v^\delta_{t\wedge\sigma}) + \tfrac{p}{2}\epsilon^{-3/2-2\gamma}\mathbb{E}\int_0^{t\wedge\sigma}\mathcal{H}^{p-1}(v_s^{\delta,\lambda})\|D\mathcal{H}(\cdot,v_s^{\delta,\lambda})\|^2_{L^2}ds \\
&\leq \mathcal{H}^p(v_0^{\delta,\lambda}) + \tfrac{p}{2}\left(4\|f'\|^2_{[0,1]}\beta' + \lambda^2\epsilon^{-1/2-2\kappa}\right)\epsilon^{-3/2-2\gamma}\mathbb{E}\int_0^{t\wedge\sigma}\mathcal{H}^{p-1}(v_s^{\delta,\lambda})ds \\
&\quad + \tfrac{1}{2}\epsilon^{-1}p(p-1)\mathbb{E}\int_0^{t\wedge\sigma}\mathcal{H}^{p-2}(v_s^{\delta,\lambda})\|D\mathcal{H}(\cdot,v_s^{\delta,\lambda})\|^2_{L^2}ds + Cp\delta^{-3}\epsilon^{-3/2-2\kappa}\mathbb{E}\int_0^{t\wedge\sigma}\mathcal{H}^{p-1}(v_s^{\delta,\lambda})ds.
\end{aligned} \tag{3.28}$$

**Lemma 3.17.** *There exists a constant* $C_2 < \infty$ *such that* $\mathcal{H}(v_0^{\delta,\lambda}) \le C_2\delta$.

*Proof.* By Theorem 3.6 and (3.8), for $\epsilon$ sufficiently small, $\mathcal{H}(v_0) \le c_2\delta$. Hence

$$\begin{aligned}
\mathcal{H}(v_0^{\delta,\lambda}) \le& \tfrac{1}{2}\left|\int_{-\infty}^{\infty}\Big(|\nabla v_0^{\delta,\lambda}|^2(x) - |\nabla v_0|^2(x)\Big)\,dx\right| + \left|\int_{-\infty}^{\infty}\Big(F_0(v_0^{\delta,\lambda}(x)) - F_0(v_0(x))\Big)\,dx\right| + c_2\delta \\
\le& C\int_{-\infty}^{\infty}\Big|\nabla v_0^{\delta,\lambda}(x) - \nabla v_0(x)\Big|\,dx + \|F_0'\|_{[0,1]}\int_{-\infty}^{\infty}\Big|v_0^{\delta,\lambda}(x) - v_0(x)\Big|\,dx + c_2\delta \\
\le& C\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\rho^{\delta}(z)\int_{x}^{x-z}|\Delta v_0(y)|dydzdx \\
&+ \|F_0'\|_{[0,1]}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\rho^{\delta}(z)\int_{x}^{x-z}|\nabla v_0(y)|dydzdx + c_2\delta \\
\le& C\delta\|\Delta v_0\|_{L^1} + \|F_0'\|_{[0,1]}\delta\|\nabla v_0\|_{L^1} + c_2\delta,
\end{aligned}$$

where the second line follows by (3.7), as does the result from the last line. ☐

Since $\mathcal{H}(v) \ge 0$, taking $p = 1$ in (3.28) gives

$$\mathbb{E}\int_0^{t\wedge\sigma}\|D\mathcal{H}(\cdot, v_s^{\delta,\lambda})\|_{L^2}^2 ds \le 2C_2\epsilon^{3/2+2\gamma}\delta + 4\|f'\|_{[0,1]}^2\beta' t + \lambda^2\epsilon^{-1/2-2\kappa}t + 2Cp\delta^{-3}\epsilon^{2\gamma-2\kappa}t.$$

Recall that $\beta' = \epsilon^{-3/10+4\gamma/5-5\kappa}$, $\delta = \epsilon^{1/10+2\gamma/5}$ and $\lambda \le \epsilon^{1+2\gamma}$; hence $\epsilon^{2\gamma}\delta^{-3} = \epsilon^{-3/10+4\gamma/5}$. It follows that

$$\mathbb{E}\int_0^{t\wedge\sigma}\|D\mathcal{H}(\cdot, v_s^{\delta,\lambda})\|_{L^2}^2 ds \le C_3(\epsilon^{-3/10+4\gamma/5-5\kappa} + \epsilon^{2l}). \tag{3.29}$$

Now set

$$A_p := \mathbb{E}\int_0^{t\wedge\sigma}\mathcal{H}^{p-1}(v_s^{\delta,\lambda})\|D\mathcal{H}(\cdot, v_s^{\delta,\lambda})\|_{L^2}^2 ds.$$

Then by (3.28),

$$\begin{aligned}
A_p \le& \tfrac{2}{p}C_2^p\epsilon^{3/2+2\gamma+p(1/10+2\gamma/5)} + \Big(4\|f'\|_{[0,1]}^2\beta' + \lambda^2\epsilon^{-1/2-2\kappa}\Big)\,\mathbb{E}\int_0^{t\wedge\sigma}\mathcal{H}^{p-1}(v_s^{\delta,\lambda})ds \\
&+ (p-1)\epsilon^{1/2+2\gamma}A_{p-1} + 2C\delta^{-3}\epsilon^{2\gamma-2\kappa}\mathbb{E}\int_0^{t\wedge\sigma}\mathcal{H}^{p-1}(v_s^{\delta,\lambda})ds.
\end{aligned} \tag{3.30}$$

By Hölder's inequality and then (3.29),

$$\begin{aligned}
A_{p-1} \le& (A_p)^{(p-2)/(p-1)}\mathbb{E}\left[\int_0^{t\wedge\sigma}\|D\mathcal{H}(\cdot, v_s^{\delta,\lambda})\|_{L^2}^2 ds\right]^{1/(p-1)} \\
\le& C_4(\epsilon^{(-3/10+4\gamma/5-5\kappa)/(p-1)} + \epsilon^{2l/(p-1)}))(A_p)^{(p-2)/(p-1)}.
\end{aligned}$$

Since $\beta \to 0$ as $\epsilon \to 0$, for $\epsilon$ sufficiently small, by Theorem 3.6, for all $t < \sigma$,

$$\mathcal{H}(v_t^{\delta,\lambda}) \leq C_5 \|D\mathcal{H}(\cdot, v_t^{\delta,\lambda})\|_{L^2}^2.$$

Therefore

$$\mathbb{E}\int_0^{t\wedge\sigma} \mathcal{H}^{p-1}(v_s^{\delta,\lambda})ds \leq C_5 \mathbb{E}\int_0^{t\wedge\sigma} \mathcal{H}^{p-2}(v_s^{\delta,\lambda})\|D\mathcal{H}(\cdot, v_s^{\delta,\lambda})\|_{L^2}^2 ds = C_5 A_{p-1}.$$

Substituting into (3.30),

$$A_p \leq \tfrac{2}{p} C_2^p \epsilon^{3/2+2\gamma+p(1/10+2\gamma/5)} + C_6(\epsilon^{p(-3/10+4\gamma/5)/(p-1)} + \epsilon^{2l/(p-1)})(A_p)^{(p-2)/(p-1)}.$$

Initially, assume $-3/10 + 4\gamma/5 \leq 2l$ and let $\tilde{A}_p = \epsilon^{-p(-3/10+4\gamma/5)} A_p$. Then

$$\tilde{A}_p \leq \tfrac{2}{p} C_2^p \epsilon^{3/2+2\gamma+2p(1-\gamma)/5} + C_7(\tilde{A}_p)^{(p-2)/(p-1)}.$$

It follows that either $\tilde{A}_p \leq \frac{4}{p} C_2^p \epsilon^{3/2+2\gamma+2p(1-\gamma)/5}$, or $\tilde{A}_p \leq 2C_7(\tilde{A}_p)^{(p-2)/(p-1)}$. The second possibility implies that $\tilde{A}_p \leq (2C)^{p-1}$. Therefore

$$A_p \leq C \max(\epsilon^{p(-3/10+4\gamma/5)}, \epsilon^{3/2+2\gamma+p(1/10+2\gamma/5)}).$$

Similarly, if $-3/10 + 4\gamma/5 > 2l$, using the same reasoning we find that

$$A_p \leq C \max(\epsilon^{2lp}, \epsilon^{3/2+2\gamma+p(1/10+2\gamma/5)}).$$

which allows us to conclude that

$$A_p \leq C \max(\epsilon^{2lp}, \epsilon^{p(-3/10+4\gamma/5)}, \epsilon^{3/2+2\gamma+p(1/10+2\gamma/5)}).$$

Now by (3.28) again,

$$\begin{aligned}\epsilon^{3/2+2\gamma}\mathbb{E}\mathcal{H}^p(v_{t\wedge\sigma}^{\delta}) \leq& \tfrac{2}{p} C_2^p \epsilon^{3/2+2\gamma+p(1/10+2\gamma/5)} + \left(4\|f'\|_{[0,1]}^2 \beta' + \lambda^2 \epsilon^{-1/2-2\kappa}\right) \mathbb{E}\int_0^{t\wedge\sigma} \mathcal{H}^{p-1}(v_s^{\delta,\lambda})ds \\ &+ (p-1)\epsilon^{1/2+2\gamma} A_{p-1} + 2C\delta^{-3}\epsilon^{2\gamma-\kappa}\mathbb{E}\int_0^{t\wedge\sigma} \mathcal{H}^{p-1}(v_s^{\delta,\lambda})ds\end{aligned}$$

$$
\begin{aligned}
&\leq \tfrac{2}{p} C_2^p \epsilon^{3/2+2\gamma+p(1/10+2\gamma/5)} + C_6(\epsilon^{p(-3/10+4\gamma/5)/(p-1)} + \epsilon^{2lp/(p-1)})(A_p)^{(p-2)/(p-1)} \\
&\leq C \max(\epsilon^{2lp}, \epsilon^{p(-3/10+4\gamma/5)}, \epsilon^{3/2+2\gamma+p(1/10+2\gamma/5)}).
\end{aligned}
$$

Now by Theorem 3.6 and Lemma 3.11,

$$\mathbb{E}\mathcal{H}^p(v^\delta_{t\wedge\sigma}) \geq (c_1\beta^2)^p \mathbb{P}\left(\sigma_2 \leq t, \sigma_2 \leq \min(\sigma_1, \sigma_3, \sigma_4)\right).$$

Therefore, taking

$$\beta^2 = \begin{cases} \epsilon^{2l-\kappa} & \text{if } 2l < \frac{1}{10} + \frac{2\gamma}{5} \\ \epsilon^{1/10+2\gamma/5-\kappa} & \text{if } 2l \geq \frac{1}{10} + \frac{2\gamma}{5} \end{cases}$$

we have $\beta \to 0$ as $\epsilon \to 0$ and

$$\mathbb{P}\left(\sigma_2 \leq t, \sigma_2 \leq \min(\sigma_1, \sigma_3, \sigma_4)\right) \leq C\epsilon^{p\kappa-3/2-2\gamma}.$$

Take $p = (3/2+2\gamma)/\kappa + 1$. Then $\mathbb{P}\left(\sigma_2 \leq t, \sigma_2 \leq \min(\sigma_1, \sigma_3, \sigma_4)\right) \to 0$ as $\epsilon \to 0$. □

### 3.2.3 Control of $\sigma_3$

The main result of this subsection is the following.

**Proposition 3.18.**

$$\lim_{\epsilon\downarrow 0} \mathbb{P}\left(\sigma_3 \leq \min(\sigma_1, \sigma_2, \sigma_4, t)\right) = 0.$$

Before we begin the proof, we need some estimates on the fundamental solution to the Laplace's equation. We shall denote this by $G_t(x)$ in this section, but we note that it is the same function as $p_t(x)$, the standard Gaussian distribution function. Suppose $y_1, y_2 \in \mathbb{R}$ with $|y_1 - y_2| \leq \delta$ and $t > 0$. Then

$$
\begin{aligned}
&\int_{-\infty}^{\infty} (G_t(x-y_1) - G_t(x-y_2))^2 dx \\
&\qquad = \int_{-\infty}^{\infty} \tfrac{1}{2\pi t} \left( e^{-(x-y_1)^2/t} - 2e^{-((x-y_1)^2+(x-y_2)^2)/(2t)} + e^{-(x-y_2)^2/t} \right) dx \\
&\qquad = \tfrac{1}{2\pi t} \left( 2\sqrt{\pi t} - 2\int_{-\infty}^{\infty} e^{-((x-\frac{1}{2}(y_1+y_2))^2 + \frac{1}{4}(y_1-y_2)^2)/t} dx \right) \\
&\qquad = \tfrac{1}{\sqrt{\pi t}} (1 - e^{-\frac{1}{4t}(y_1-y_2)^2})
\end{aligned}
$$

$$\leq \tfrac{1}{\sqrt{\pi t}}(1 - e^{-\delta^2/(4t)}), \tag{3.31}$$

since $|y_1 - y_2| \leq \delta$. Suppose $z \geq 0$, $y \in \mathbb{R}$ and $t > 0$. If $y - \frac{1}{2}z \geq 0$ then

$$\begin{aligned}\int_{-\infty}^{0} |G_t(x-y) - G_t(x+z-y)|dx =& \int_{-\infty}^{0} (G_t(x+z-y) - G_t(x-y))dx \\ =& \mathbb{P}_{y-z}\left(B_t \leq 0\right) - \mathbb{P}_y\left(B_t \leq 0\right) \\ =& \mathbb{P}_y\left(B_t \in [0,z]\right). \end{aligned} \tag{3.32}$$

If $y - \frac{1}{2}z < 0$ then

$$\begin{aligned}&\int_{-\infty}^{0} |G_t(x-y) - G_t(x+z-y)|dx \\ =& \int_{-\infty}^{y-\frac{1}{2}z} (G_t(x+z-y) - G_t(x-y))dx + \int_{y-\frac{1}{2}z}^{0} (G_t(x-y) - G_t(x+z-y))dx \\ =& 2\mathbb{P}_y\left(B_t \leq y + \tfrac{1}{2}z\right) - 2\mathbb{P}_y\left(B_t \leq y - \tfrac{1}{2}z\right) + \mathbb{P}_y\left(B_t \leq 0\right) - \mathbb{P}_y\left(B_t \leq z\right) \\ \leq& 2\mathbb{P}_0\left(|B_t| \leq \tfrac{1}{2}z\right) \end{aligned} \tag{3.33}$$

since $z \geq 0$ and so $\mathbb{P}_y\left(B_t \leq 0\right) \leq \mathbb{P}_y\left(B_t \leq z\right)$. We also need the following estimate. For $x \in \mathbb{R}$, $z \geq 0$ and $t > 0$,

$$\begin{aligned}&\int_{-\infty}^{\infty} |G_t(x-y) - G_t(x+z-y)|dy \\ =& \int_{x+\frac{1}{2}z}^{\infty} (G_t(x+z-y) - G_t(x-y))dy + \int_{-\infty}^{x+\frac{1}{2}z} (G_t(x-y) - G_t(x+z-y))dy \\ =& \mathbb{P}_{x+z}\left(B_t \geq x + \tfrac{1}{2}z\right) - \mathbb{P}_x\left(B_t \geq x + \tfrac{1}{2}z\right) + \mathbb{P}_x\left(B_t \leq x + \tfrac{1}{2}z\right) - \mathbb{P}_{x+z}\left(B_t \leq x + \tfrac{1}{2}z\right) \\ =& 2\mathbb{P}_0\left(|B_t| \leq \tfrac{1}{2}z\right). \end{aligned} \tag{3.34}$$

We also need the following lemma.

**Lemma 3.19.** *For $t < \infty$,*

$$\mathbb{P}\left(\sigma_3 \leq t, \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \rho^{\delta}(z)(v^{\epsilon,\lambda}_{\sigma_3}(y) - v^{\epsilon,\lambda}_{\sigma_3}(y+z))^2 dydz < \beta'\right) = 0.$$

*Proof.* The proof is similar to the proof of Lemma 3.9. By Lemma 3.3, there exist $K < \infty$ and a random variable $X_0$ such that for all $s \le t+1$ we have $v_s(x) \le e^{-Kx^2}$ for all $x \le -X_0$ and $1 - v_s(x) \le e^{-Kx^2}$ for all $x \ge X_0$. By the definition of $\sigma_3$, on $\{\sigma_3 \le t\}$, there exist times $t_1, t_2, \ldots$ such that $t_n \le t+1$ for all $n$, $t_n \to \sigma_3$ as $n \to \infty$ and

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \rho^\delta(z)(v_{t_n}^{\epsilon,\lambda}(y) - v_{t_n}^{\epsilon,\lambda}(y+z))^2 dydz \ge \beta'.$$

On $\{X_0 < \infty\} \cap \{\sigma_3 \le t\}$, since for $|z| \le \delta$, $(v_{t_n}^{\epsilon,\lambda}(y) - v_{t_n}^{\epsilon,\lambda}(y+z))^2 \le \mathbb{1}_{y \in [-X_0-\delta, X_0+\delta]} + e^{-2K(|y|-\delta)^2}$, by dominated convergence,

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \rho^\delta(z)(v_{t_n}^{\epsilon,\lambda}(y) - v_{t_n}^{\epsilon,\lambda}(y+z))^2 dydz \to \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \rho^\delta(z)(v_{\sigma_3}^{\epsilon,\lambda}(y) - v_{\sigma_3}^{\epsilon,\lambda}(y+z))^2 dydz$$

as $n \to \infty$. Hence on $\{X_0 < \infty\} \cap \{\sigma_3 \le t\}$,

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \rho^\delta(z)(v_{\sigma_3}^{\epsilon,\lambda}(y) - v_{\sigma_3}^{\epsilon,\lambda}(y+z))^2 dydz \ge \beta'.$$

The result follows since $X_0 < \infty$ almost surely. □

*Proof of Proposition 3.18.* As in the proof of Proposition 3.8, let $v_1(t,x) = v^{\epsilon,\lambda}(\epsilon^{3/2+2\gamma}t, x)$. Take $z \in [0,\delta]$. Then using the integral form in (3.20) and since $(a+b+c)^2 \le 3(a^2+b^2+c^2)$,

$$\begin{aligned}
&(v_1(t,x) - v_1(t,x+z))^2 \\
&\le 3\left(\int_{-\infty}^{\infty} e^{-t/4} G_t(x-y)(v_0(y) - v_0(y+z))dy\right)^2 \\
&+ 3\left(\int_0^t \int_{-\infty}^{\infty} e^{-(t-s)/4}(G_{t-s}(x-y) - G_{t-s}(x+z-y))\left(\tfrac{1}{4}v_1(s,y) + f_\lambda(v_1(s,y))\right) dyds\right)^2 \\
&+ 3\epsilon^{1/2+2\gamma}(M_t(x) - M_t(x+z))^2,
\end{aligned}$$

where $M_t(y)$ is defined in (3.21). As in the proof of Proposition 3.8, let $\tilde{\sigma} = \epsilon^{-3/2-2\gamma}\sigma$. Then, since $|f(v_1(s,y))| \le \max(1, 1+\lambda)v_1(s,y)$ for all $s \ge 0$, $y \in \mathbb{R}$ a.s.,

$$\begin{aligned}
&\mathbb{E}(v_1(t\wedge\tilde{\sigma}, x) - v_1(t\wedge\tilde{\sigma}, x+z))^2 \\
&\le 3\mathbb{E}\left(\int_{-\infty}^{\infty} e^{-(t\wedge\tilde{\sigma})/4} G_{t\wedge\tilde{\sigma}}(x-y)(v_0(y) - v_0(y+z))dy\right)^2
\end{aligned}$$

$$
\begin{aligned}
&+ 3\left(\tfrac{5}{4}+|\lambda|\right)^2 \mathbb{E}\bigg(\int_0^{t\wedge\tilde{\sigma}} \int_{-\infty}^{\infty} e^{-(t\wedge\tilde{\sigma}-s)/4} \left|G_{t\wedge\tilde{\sigma}-s}(x-y) - G_{t\wedge\tilde{\sigma}-s}(x+z-y)\right| v_1(s,y)dyds\bigg)^2 \\
&+ 3\epsilon^{1/2+2\gamma}\mathbb{E}\int_0^{t\wedge\tilde{\sigma}}\int_{-\infty}^{\infty} e^{-(t\wedge\tilde{\sigma}-s)/2}(G_{t\wedge\tilde{\sigma}-s}(x-y) - G_{t\wedge\tilde{\sigma}-s}(x+z-y))^2 v_1(s,y)(1-v_1(s,y))dyds.
\end{aligned}
$$

By Jensen's inequality (and bounding $|\lambda|$ by 1), it follows that

$$
\begin{aligned}
&\mathbb{E}(v_1(t\wedge\tilde{\sigma},x) - v_1(t\wedge\tilde{\sigma},x+z))^2 \\
&\quad\leq 3\mathbb{E}\int_{-\infty}^{\infty} G_{t\wedge\tilde{\sigma}}(x-y)(v_0(y)-v_0(y+z))^2dy \\
&\qquad + 64\mathbb{E}\bigg[\int_0^{t\wedge\tilde{\sigma}} e^{-(t\wedge\tilde{\sigma}-s)/4}\left(\int_{-\infty}^{\infty} |G_{t\wedge\tilde{\sigma}-s}(x-y) - G_{t\wedge\tilde{\sigma}-s}(x+z-y)|dy\right) \\
&\qquad\qquad\qquad \left(\int_{-\infty}^{\infty} |G_{t\wedge\tilde{\sigma}-s}(x-y) - G_{t\wedge\tilde{\sigma}-s}(x+z-y)| v_1(s,y)^2dy\right)ds\bigg] \\
&\qquad + 3\epsilon^{1/2+2\gamma}\mathbb{E}\int_0^{t\wedge\tilde{\sigma}}\int_{-\infty}^{\infty} e^{-(t\wedge\tilde{\sigma}-s)/2}(G_{t\wedge\tilde{\sigma}-s}(x-y) - G_{t\wedge\tilde{\sigma}-s}(x+z-y))^2 \\
&\qquad\qquad\qquad\qquad v_1(s,y)(1-v_1(s,y))dyds.
\end{aligned}
$$

Integrating over $x$ and applying Fubini's theorem and (3.34),

$$
\begin{aligned}
&\int_{-\infty}^{0}\mathbb{E}(v_1(t\wedge\tilde{\sigma},x) - v_1(t\wedge\tilde{\sigma},x+z))^2dx \\
&\quad\leq 3\mathbb{E}\int_{-\infty}^{0}\int_{-\infty}^{\infty} G_{t\wedge\tilde{\sigma}}(x-y)(v_0(y)-v_0(y+z))^2dydx \\
&\qquad + 64\mathbb{E}\bigg[\int_0^{t\wedge\tilde{\sigma}} e^{-(t\wedge\tilde{\sigma}-s)/4} 2\mathbb{P}_0\left(|B_{t\wedge\tilde{\sigma}-s}| \leq \tfrac{1}{2}z\right) \\
&\qquad\qquad \left(\int_{-\infty}^{\infty}\left(\int_{-\infty}^{0} |G_{t\wedge\tilde{\sigma}-s}(x-y) - G_{t\wedge\tilde{\sigma}-s}(x+z-y)|dx\right) v_1(s,y)^2dy\right)ds\bigg] \\
&\qquad + 3\epsilon^{1/2+2\gamma}\mathbb{E}\bigg[\int_0^{t\wedge\tilde{\sigma}}\int_{-\infty}^{\infty} e^{-(t\wedge\tilde{\sigma}-s)/2}\left(\int_{-\infty}^{0}(G_{t\wedge\tilde{\sigma}-s}(x-y) - G_{t\wedge\tilde{\sigma}-s}(x+z-y))^2dx\right) \\
&\qquad\qquad\qquad\qquad v_1(s,y)(1-v_1(s,y))dyds\bigg] \\
&\quad\leq 3\mathbb{E}\int_{-\infty}^{0}\int_{-\infty}^{\infty} G_{t\wedge\tilde{\sigma}}(x-y)(v_0(y)-v_0(y+z))^2dydx \\
&\qquad + 128\mathbb{E}\bigg[\int_0^{t\wedge\tilde{\sigma}} e^{-(t\wedge\tilde{\sigma}-s)/4} \mathbb{P}_0\left(|B_{t\wedge\tilde{\sigma}-s}| \leq \tfrac{1}{2}z\right) \\
&\qquad\qquad \left(\int_{-\infty}^{\infty}\left(\mathbb{1}_{y\geq z/2}\mathbb{P}_y\left(B_{t\wedge\tilde{\sigma}-s}\in[0,z]\right) + \mathbb{1}_{y<z/2}2\mathbb{P}_0\left(|B_{t\wedge\tilde{\sigma}-s}|\leq\tfrac{1}{2}z\right)\right)v_1(s,y)^2dy\right)ds\bigg] \\
&\qquad + 3\epsilon^{1/2+2\gamma}\mathbb{E}\int_0^{t\wedge\tilde{\sigma}}\int_{-\infty}^{\infty} e^{-(t\wedge\tilde{\sigma}-s)/2}\frac{\left(1-e^{-\delta^2/(4(t\wedge\tilde{\sigma}-s))}\right)}{\sqrt{\pi(t\wedge\tilde{\sigma}-s)}} v_1(s,y)(1-v_1(s,y))dyds, \qquad (3.35)
\end{aligned}
$$

where the second inequality follows by (3.32), (3.33) and (3.31), since $|z| \le \delta$. We now bound each term on the right hand side of (3.35). For the first term, since $\int_{-\infty}^{\infty} G_{t\wedge\tilde{\sigma}}(x-y)dx = 1$,

$$\begin{aligned}\mathbb{E}\int_{-\infty}^{0}\int_{-\infty}^{\infty} G_{t\wedge\tilde{\sigma}}(x-y)(v_0(y)-v_0(y+z))^2 dydx <& \int_{-\infty}^{\infty}\left(\int_{y}^{y+z}\nabla v_0(z')dz'\right)^2 dy\\ \le& \int_{-\infty}^{\infty} z\int_{y}^{y+z}\nabla v_0(z')^2 dz'dy\\ =& z^2\int_{-\infty}^{\infty}\nabla v_0(y)^2 dy\\ \le& \delta^2\|\nabla v_0\|_{L^2}^2, \end{aligned} \tag{3.36}$$

where the second inequality follows by Jensen's inequality. For the second term on the right hand side of (3.35), note first that for $s < \tilde{\sigma}$,

$$\int_{-\infty}^{z/2} v_1(s,y)^2 dy \le \tfrac{1}{2}\delta + \int_{-\infty}^{0} v_1(s,y)^2 dy \le 10\epsilon^{-1/2-2\kappa}$$

for $\epsilon$ sufficiently small, by the same argument as for (3.26). Hence for $s < \tilde{\sigma}$,

$$\begin{aligned}\int_{-\infty}^{z/2} 2\mathbb{P}_0\left(|B_{t\wedge\tilde{\sigma}-s}| \le \tfrac{1}{2}z\right) v_1(s,y)^2 dy \le& 20\epsilon^{-1/2-2\kappa}\mathbb{P}_0\left(|B_{t\wedge\tilde{\sigma}-s}| \le \tfrac{1}{2}z\right)\\ \le& 20\epsilon^{-1/2-2\kappa}\min\left(\frac{z}{\sqrt{2\pi(t\wedge\tilde{\sigma}-s)}}, 1\right).\end{aligned}$$

Also since $0 \le v_1(s,y) \le 1$ for all $s \ge 0, y \in \mathbb{R}$ a.s.,

$$\begin{aligned}\int_{z/2}^{\infty}\mathbb{P}_y\left(B_{t\wedge\tilde{\sigma}-s} \in [0,z]\right) v_1(s,y)^2 dy \le& \frac{z}{\sqrt{2\pi(t\wedge\tilde{\sigma}-s)}}\int_{z}^{\infty} e^{-(y-z)^2/(2(t\wedge\tilde{\sigma}-s))}dy + \tfrac{1}{2}z\\ =& z.\end{aligned}$$

Therefore

$$\begin{aligned}&\mathbb{E}\Bigg[\int_{0}^{t\wedge\tilde{\sigma}} e^{-(t\wedge\tilde{\sigma}-s)/4}\mathbb{P}_0\left(|B_{t\wedge\tilde{\sigma}-s}| \le \tfrac{1}{2}z\right)\\ &\qquad\left(\int_{-\infty}^{\infty}\left(\mathbb{1}_{y\ge z/2}\mathbb{P}_y\left(B_{t\wedge\tilde{\sigma}-s}\in[0,z]\right) + \mathbb{1}_{y<z/2}2\mathbb{P}_0\left(|B_{t\wedge\tilde{\sigma}-s}| \le \tfrac{1}{2}z\right)\right) v_1(s,y)^2 dy\right) ds\Bigg]\\ &\le \mathbb{E}\left[\int_{0}^{t\wedge\tilde{\sigma}} e^{-(t\wedge\tilde{\sigma}-s)/4}\min\left(\frac{z}{\sqrt{2\pi(t\wedge\tilde{\sigma}-s)}}, 1\right)\left(z + 20\epsilon^{-1/2-2\kappa}\min\left(\frac{z}{\sqrt{2\pi(t\wedge\tilde{\sigma}-s)}}, 1\right)\right) ds\right]\end{aligned}$$

$$
\begin{aligned}
&< \int_0^\infty e^{-s/4} \min\left(\frac{z}{\sqrt{2\pi s}}, 1\right) \left(z + 20\epsilon^{-1/2-2\kappa} \min\left(\frac{z}{\sqrt{2\pi s}}, 1\right)\right) ds \\
&\leq \left(\delta^2 \int_0^\infty e^{-s/4} \frac{1}{\sqrt{2\pi s}} ds + 20\epsilon^{-1/2-2\kappa} \int_0^\infty e^{-s/4} \min\left(\frac{\delta}{\sqrt{2\pi s}}, 1\right)^2 ds\right) \\
&\leq \left(\delta^2 \int_0^\infty e^{-s/4} \frac{1}{\sqrt{2\pi s}} ds + 20\epsilon^{-1/2-2\kappa} \left(\delta^2 + \delta^2 \int_{\delta^2}^\infty e^{-s/4} \frac{1}{2\pi s^{1-\frac{1}{2}\kappa} (\delta^2)^{\frac{1}{2}\kappa}} ds\right)\right) \\
&\leq C\epsilon^{-1/2-2\kappa} \delta^{2-\kappa}
\end{aligned}
\tag{3.37}
$$

for $\epsilon$ sufficiently small, for some constant $C$, since $\int_0^\infty e^{-s/4} s^{-a} ds < \infty$ if $a < 1$. Finally, for the third term on the right hand side of (3.35), by the definition of $\sigma_1$ and then since $1 - e^{-a} \leq a$ for $a \geq 0$,

$$
\begin{aligned}
&\epsilon^{1/2+2\gamma} \mathbb{E} \int_0^{t \wedge \tilde{\sigma}} \int_{-\infty}^\infty e^{-(t\wedge\tilde{\sigma}-s)/2} \frac{1}{\sqrt{\pi(t \wedge \tilde{\sigma} - s)}} \left(1 - e^{-\delta^2/(4(t\wedge\tilde{\sigma}-s))}\right) v_1(s,y)(1 - v_1(s,y)) dy ds \\
&\leq \epsilon^{1/2+2\gamma} \mathbb{E} \int_0^{t\wedge\tilde{\sigma}} e^{-(t\wedge\tilde{\sigma}-s)/2} \frac{1}{\sqrt{\pi(t\wedge\tilde{\sigma}-s)}} (1 - e^{-\delta^2/(4(t\wedge\tilde{\sigma}-s))}) \epsilon^{-1/2-2\kappa} ds \\
&\leq \epsilon^{2\gamma-2\kappa} \int_0^\infty e^{-s/2} \frac{1}{\sqrt{\pi s}} \min(1, \delta^2/(4s)) ds \\
&\leq \epsilon^{2\gamma-2\kappa} \left(\int_0^{\delta^2} \frac{1}{\sqrt{\pi s}} ds + \tfrac{1}{4}\delta^2 \int_{\delta^2}^\infty e^{-s/2} \frac{1}{\sqrt{\pi} s^{1-\kappa} (\delta^2)^{\kappa+\frac{1}{2}}} ds\right) \\
&\leq C\epsilon^{2\gamma-2\kappa} \delta^{1-2\kappa}
\end{aligned}
\tag{3.38}
$$

for some constant $C$, since $\int_0^\infty e^{-s/2} s^{-a} ds < \infty$ if $a < 1$. Substituting the bounds in (3.36), (3.37) and (3.38) into (3.35),

$$
\int_{-\infty}^0 \mathbb{E}(v_1(t\wedge\tilde{\sigma}, x) - v_1(t\wedge\tilde{\sigma}, x+z))^2 dx \leq 3\delta^2 \|\nabla v_0\|_{L^2}^2 + C(\epsilon^{-1/2-2\kappa}\delta^{2-\kappa} + \epsilon^{2\gamma-2\kappa}\delta^{1-2\kappa}).
$$

By the same argument applied to $v_2(t,x) = 1 - v(\epsilon^{3/2+2\gamma} t, x)$,

$$
\int_0^\infty \mathbb{E}(v_1(t\wedge\tilde{\sigma}, x) - v_1(t\wedge\tilde{\sigma}, x+z))^2 dx \leq 3\delta^2 \|\nabla v_0\|_{L^2}^2 + C(\epsilon^{-1/2-2\kappa}\delta^{2-\kappa} + \epsilon^{2\gamma-2\kappa}\delta^{1-2\kappa}).
$$

Therefore for $z \in [0, \delta]$, for $\epsilon$ sufficiently small, by (3.7),

$$
\int_{-\infty}^\infty \mathbb{E}(v_1(t\wedge\tilde{\sigma}, x) - v_1(t\wedge\tilde{\sigma}, x+z))^2 dx \leq 3C(\epsilon^{-1/2-2\kappa}\delta^{2-\kappa} + \epsilon^{2\gamma-2\kappa}\delta^{1-2\kappa}),
$$

and hence by symmetry the same result holds for $z \in [-\delta, 0]$. Therefore, by Fubini's theorem,

$$\begin{aligned}\mathbb{E}\int_{-\infty}^{\infty} \rho^{\delta}(z) \int_{-\infty}^{\infty} (v_1(t\wedge\tilde{\sigma}, x) - v_1(t\wedge\tilde{\sigma}, x+z))^2 dxdz \leq & 3C(\epsilon^{-1/2-2\kappa}\delta^{2-\kappa} + \epsilon^{2\gamma-2\kappa}\delta^{1-2\kappa}) \\ \leq & 6C\epsilon^{4\gamma/5-3/10-4\kappa},\end{aligned}$$

since $\delta = \epsilon^{1/10+2\gamma/5}$. We can now use Lemma 3.19 to complete the proof. We have

$$\begin{aligned}&\mathbb{E}\int_{-\infty}^{\infty} \rho^{\delta}(z) \int_{-\infty}^{\infty} (v^{\epsilon,\lambda}(t\wedge\sigma, x) - v^{\epsilon,\lambda}(t\wedge\sigma, x+z))^2 dxdz \\ &\qquad \geq \mathbb{E}\mathbb{1}_{\sigma_3 \leq \min(\sigma_1,\sigma_2,\sigma_4,t)}\beta' \mathbb{1}_{\int_{-\infty}^{\infty}\rho^{\delta}(z)\int_{-\infty}^{\infty}(v^{\epsilon,\lambda}(t\wedge\sigma,x)-v^{\epsilon,\lambda}(t\wedge\sigma,x+z))^2dxdz\geq\beta'} \\ &\qquad \geq \beta'\,\mathbb{P}\left(\sigma_3 \leq \min(\sigma_1,\sigma_2,\sigma_4,t)\right),\end{aligned}$$

by Lemma 3.19. Therefore since $\beta' = \epsilon^{4\gamma/5-3/10-5\kappa}$,

$$\mathbb{P}\left(\sigma_3 \leq \min(\sigma_1,\sigma_2,\sigma_4,t)\right) \leq 6C\epsilon^{\kappa},$$

and the result follows. □

## 3.3 Proof of Proposition 3.5

As in the proof of Proposition 3.8, let $v_1(t,x) = v^{\epsilon,\lambda}(\epsilon^{3/2+2\gamma}t, x)$ and let $v_0(x) = u_0(\epsilon^{1/2}x)$. Then $v_1$ solves the SPDE (3.19), which we can write in integral form as

$$\begin{aligned}v_1(t,x) = &\int_{-\infty}^{\infty} G_t(x-y)v_0(y)dy + \int_0^t \int_{-\infty}^{\infty} G_{t-s}(x-y)f_{\lambda}(v_1(s,y))dyds \\ &+ \epsilon^{1/4+\gamma}\int_0^t\int_{-\infty}^{\infty} G_{t-s}(x-y)\sqrt{v_1(s,y)(1-v_1(s,y)}W(dyds). \qquad (3.39)\end{aligned}$$

The proof of Proposition 3.5 will rely on the following lemma. We abuse notation slightly and treat time as if it were a spatial dimension, which means we have expressions like $|(t_1,x_1) - (t_2,x_2)|$, meaning $\sqrt{(t_1-t_2)^2 - (x_1-x_2)^2}$.

**Lemma 3.20.** *Take $T > 0$. There exist constants $A$ and $k > 1$ such that for $M \in \mathbb{R}$, for*

*$x_1, x_2 \in [M, M+1]$ and $t_1, t_2 \in [0,T]$,*

$$\mathbb{E}(v_1(t_1,x_1) - v_1(t_2,x_2))^{2k} \le Ae^{-M/A}|(t_1,x_1)-(t_2,x_2)|^3.$$

The proof of Lemma 3.20 uses the following result. For $t<0$, we let $G_t \equiv 0$.

**Lemma 3.21** ((2.6) in [32])**.** *For $T<\infty$, $\kappa \in (0,1/4)$, there exists a constant $c$ such that for all $s,t \in [0,T]$,*

$$\int_0^\infty \int_{-\infty}^\infty |G_{t-r}(x-z) - G_{s-r}(y-z)|^2 drdz \le c|(t,x)-(s,y)|^{2\kappa}.$$

*Proof of Lemma 3.20.* By Lemma 3.3, there exist a constant $K$ and a random variable $X_0$ such that for all $t \le T$ we have $v_1(t,x) \le e^{-Kx^2}$ for $x \le -X_0$ and $1 - v_1(t,x) \le e^{-Kx^2}$ for $x \ge X_0$.

We assume that $M>0$ (the proof for $M<0$ is the same). Take $t_1 \le t_2 \in [0,T]$ and $x_1, x_2 \in [M, M+1]$. Then by (3.39),

$$\begin{aligned}
&v_1(t_1,x_1) - v_1(t_2,x_2)\\
&\quad = \int_{-\infty}^\infty (G_{t_1}(x_1-y) - G_{t_2}(x_2-y))v_0(y)dy\\
&\qquad + \int_0^\infty \int_{-\infty}^\infty (G_{t_1-s}(x_1-y) - G_{t_2-s}(x_2-y))f_\lambda(v_1(s,y))dyds\\
&\qquad + \epsilon^{1/4+\gamma} \int_0^\infty \int_{-\infty}^\infty (G_{t_1-s}(x_1-y) - G_{t_2-s}(x_2-y))\sqrt{v_1(s,y)(1-v_1(s,y)}W(dyds).
\end{aligned}$$

Therefore for $k \in \mathbb{N}$, since $(a_1+a_2+a_3)^{2k} \le 3^{2k-1}(a_1^{2k}+a_2^{2k}+a_3^{2k})$,

$$\begin{aligned}
&(v_1(t_1,x_1) - v_1(t_2,x_2))^{2k}\\
&=3^{2k-1}\left(\int_{-\infty}^\infty (G_{t_1}(x_1-y) - G_{t_2}(x_2-y))v_0(y)dy\right)^{2k}\\
&\quad + 3^{2k-1}\left(\int_0^\infty \int_{-\infty}^\infty (G_{t_1-s}(x_1-y) - G_{t_2-s}(x_2-y))f_\lambda(v_1(s,y))dyds\right)^{2k}\\
&\quad + 3^{2k-1}\epsilon^{1/4+\gamma}\left(\int_0^\infty \int_{-\infty}^\infty (G_{t_1-s}(x_1-y) - G_{t_2-s}(x_2-y))\sqrt{v_1(s,y)(1-v_1(s,y)}W(dyds)\right)^{2k}.
\end{aligned} \tag{3.40}$$

We shall bound the expectation of each term on the right hand side of (3.40). For the first term on the right hand side of (3.40), we write

$$\begin{aligned}\left(\int_{-\infty}^{\infty}(G_{t_1}(x_1-y)-G_{t_2}(x_2-y))v_0(y)dy\right)^{2k}&\\ \leq&2^{2k-1}\left(\int_{-\infty}^{\infty}(G_{t_1}(x_1-y)-G_{t_1}(x_2-y))v_0(y)dy\right)^{2k}\\ &+2^{2k-1}\left(\int_{-\infty}^{\infty}(G_{t_1}(x_2-y)-G_{t_2}(x_2-y))v_0(y)dy\right)^{2k}.\end{aligned} \tag{3.41}$$

We can bound the first term on the right hand side of (3.41) by writing

$$\begin{aligned}\left|\int_{-\infty}^{\infty}(G_{t_1}(x_1-y)-G_{t_1}(x_2-y))v_0(y)dy\right|&\\ \leq&\int_{-\infty}^{\infty}G_{t_1}(x_1-y)|v_0(y)-v_0(y+x_2-x_1)|dy\\ \leq&\|\nabla v_0\|_\infty|x_2-x_1|\int_{-\infty}^{R+1}G_{t_1}(x_1-y)dy\\ \leq&\|\nabla v_0\|_\infty|x_2-x_1|\left(e^{-(M-R-1)^2/(2t_1)}\mathbb{1}_{M\geq R-1}+\mathbb{1}_{M<R-1}\right)\\ \leq&\|\nabla v_0\|_\infty|x_2-x_1|\left(e^{-(M-R-1)^2/(2T)}\mathbb{1}_{M\geq R-1}+\mathbb{1}_{M<R-1}\right),\end{aligned}$$

where the second line follows by the mean value theorem and since $v_0(z_1)=v_0(z_2)$ if $z_1,z_2\geq R$ by (3.9), the third line follows since $x_1\geq M$ and the last line since $t_1\leq T$. For the second term on the right hand side of (3.41), by the semigroup property of $G$,

$$\begin{aligned}\left|\int_{-\infty}^{\infty}(G_{t_1}(x_2-y)-G_{t_2}(x_2-y))v_0(y)dy\right|&\\ =&\left|\int_{-\infty}^{\infty}G_{t_1}(x_2-y)\int_{-\infty}^{\infty}G_{t_2-t_1}(y-z)(v_0(y)-v_0(z))dzdy\right|\\ \leq&\|\nabla v_0\|_\infty\int_{-\infty}^{\infty}G_{t_1}(x_2-y)\mathbb{E}_y\left[|B_{t_2-t_1}-y|\right]dy\\ =&\|\nabla v_0\|_\infty\sqrt{t_2-t_1}\sqrt{\frac{2}{\pi}},\end{aligned}$$

where the first inequality follows by the mean value theorem and the last line follows by Brownian scaling and since $\mathbb{E}|B_1|=\sqrt{\frac{2}{\pi}}$. We can also bound the first term on the right hand side of (3.41) by writing

$$\begin{aligned}
&\left|\int_{-\infty}^{\infty} G_{t_1}(x_2-y)\int_{-\infty}^{\infty} G_{t_2-t_1}(y-z)(v_0(y)-v_0(z))dzdy\right| \\
&\qquad\qquad\qquad\qquad \le \int_{-\infty}^{\infty} G_{t_1}(x_2-y)\mathbb{P}_y\left(\min(y,B_{t_2-t_1})\le R\right)dy \\
&\qquad\qquad\qquad\qquad \le \int_{-\infty}^{R} G_{t_1}(x_2-y)dy+\int_{-\infty}^{\infty} G_{t_1}(x_2-y)\mathbb{P}_y\left(B_{t_2-t_1}\le R\right)dy \\
&\qquad\qquad\qquad\qquad \le e^{-(M-R)^2/(2t_1)}\mathbb{1}_{M\ge R}+\mathbb{1}_{M<R}+\mathbb{P}_{x_2}\left(B_{t_2}\le R\right) \\
&\qquad\qquad\qquad\qquad \le 2\left(e^{-(M-R)^2/(2T)}\mathbb{1}_{M\ge R}+\mathbb{1}_{M<R}\right),
\end{aligned}$$

where the first inequality follows since $|v_0(y)-v_0(z)|\le \mathbb{1}_{\min(y,z)\le R}$ by (3.9) and the third inequality follows since $x_2\ge M$ and by the semigroup property of $G$, and the last inequality holds since $t_1,t_2\le T$. Substituting into (3.41), we have that

$$\begin{aligned}
&\left(\int_{-\infty}^{\infty}(G_{t_1}(x_1-y)-G_{t_2}(x_2-y))v_0(y)dy\right)^{2k} \\
&\quad \le 2^{2k-1}\|\nabla v_0\|_\infty^{2k}|x_2-x_1|^{2k}\left(e^{-(M-R-1)^2/(2T)}\mathbb{1}_{M\ge R-1}+\mathbb{1}_{M<R-1}\right)^{2k} \\
&\qquad + 2^{2k}\left(e^{-(M-R)^2/(2T)}\mathbb{1}_{M\ge R}+\mathbb{1}_{M<R}\right)\|\nabla v_0\|_\infty^{2k-1}|t_2-t_1|^{(2k-1)/2}\left(\frac{2}{\pi}\right)^{(2k-1)/2} \\
&\quad \le A e^{-M/A}|(t_1,x_1)-(t_2,x_2)|^3 \qquad (3.42)
\end{aligned}$$

for some constant $A=A(k)<\infty$, if $k\ge 4$.

We now bound the expectation of the second term on the right hand side of (3.40). Since $(a_1+a_2+a_3)^{2k}\le 3^{2k-1}(a_1^{2k}+a_2^{2k}+a_3^{2k})$ and

$$|f(_\lambda v_1(s,y))|\le \min(1,(1+|\lambda|)v_1(s,y),(1+|\lambda|)(1-v_1(s,y)))\le 2\min(1,v_1(s,y),1-v_1(s,y)),$$

then, almost surely, for all $s\ge 0$ and $y\in\mathbb{R}$,

$$\begin{aligned}
&\mathbb{E}\left(\int_0^\infty\int_{-\infty}^{\infty}(G_{t_1-s}(x_1-y)-G_{t_2-s}(x_2-y))f_\lambda(v_1(s,y))dyds\right)^{2k} \\
&\qquad \le 6^{2k-1}\mathbb{E}\left(\int_0^\infty\int_{X_0}^{\infty}|G_{t_1-s}(x_1-y)-G_{t_2-s}(x_2-y)|e^{-Ky^2}dyds\right)^{2k} \\
&\qquad\quad + 6^{2k-1}\mathbb{E}\left(\int_0^\infty\int_{-X_0}^{X_0}|G_{t_1-s}(x_1-y)-G_{t_2-s}(x_2-y)|dyds\right)^{2k}
\end{aligned}$$

$$+\,6^{2k-1}\mathbb{E}\bigg(\int_0^\infty\int_{-\infty}^{-X_0}|G_{t_1-s}(x_1-y)-G_{t_2-s}(x_2-y)|e^{-Ky^2}dyds\bigg)^{2k}. \tag{3.43}$$

We begin by bounding the second term on the right hand side. By Jensen's inequality,

$$\begin{aligned}\bigg(\int_0^\infty\int_{-X_0}^{X_0}&|G_{t_1-s}(x_1-y)-G_{t_2-s}(x_2-y)|dyds\bigg)^2\\ &\le 2TX_0\int_0^\infty\int_{-\infty}^\infty(G_{t_1-s}(x_1-y)-G_{t_2-s}(x_2-y))^2dyds\\ &\le 2TX_0\,c|(t_1,x_1)-(t_2,x_2)|^{1/4},\end{aligned}$$

for some constant $c<\infty$, by Lemma 3.21. We also have that

$$\begin{aligned}\int_0^\infty\int_{-X_0}^{X_0}&|G_{t_1-s}(x_1-y)-G_{t_2-s}(x_2-y)|dyds\\ &\le\int_0^\infty\int_{-\infty}^{X_0}(G_{t_1-s}(x_1-y)+G_{t_2-s}(x_2-y))dyds\\ &\le\int_0^\infty\left(e^{-(M-X_0)^2/(2(t_1-s))}\mathbb{1}_{s<t_1}+e^{-(M-X_0)^2/(2(t_2-s))}\mathbb{1}_{s<t_2}\right)\mathbb{1}_{M\ge X_0}\\ &\qquad\qquad+\mathbb{1}_{M<X_0}(\mathbb{1}_{s<t_1}+\mathbb{1}_{s<t_2})ds\\ &\le 2T\left(e^{-(M-X_0)^2/(2T)}\mathbb{1}_{M\ge X_0}+\mathbb{1}_{M<X_0}\right),\end{aligned}$$

where the second inequality holds since $x_1,x_2\ge M$ and the last inequality holds since $t_1,t_2\le T$. Hence

$$\begin{aligned}\mathbb{E}\bigg(\int_0^\infty\int_{-X_0}^{X_0}&|G_{t_1-s}(x_1-y)-G_{t_2-s}(x_2-y)|dyds\bigg)^{2k}\\ \le\mathbb{E}\Big[&2T\left(e^{-(M-X_0)/(2T)}\mathbb{1}_{M\ge X_0+1}+\mathbb{1}_{M<X_0+1}\right)(2T)^{(2k-1)/2}\\ &\qquad X_0^{(2k-1)/2}\,c^{(2k-1)/2}|(t_1,x_1)-(t_2,x_2)|^{(2k-1)/8}\Big]\\ \le A&e^{-M/A}|(t_1,x_1)-(t_2,x_2)|^3\end{aligned}\tag{3.44}$$

for some constant $A=A(k)<\infty$, if $k\ge 13$, by Lemma 3.3. For the first term on the right hand side of (3.43), by Jensen's inequality,

$$\bigg(\int_0^\infty\int_{X_0}^\infty|G_{t_1-s}(x_1-y)-G_{t_2-s}(x_2-y)|e^{-Ky^2}dyds\bigg)^2$$

$$\begin{aligned}
\leq & T\int_0^\infty \left(\int_{X_0}^\infty |G_{t_1-s}(x_1-y)-G_{t_2-s}(x_2-y)|e^{-Ky^2}dy\right)^2 ds\\
\leq & T\left(\int_{-\infty}^\infty e^{-2Ky^2}dy\right)\int_0^\infty\int_{-\infty}^\infty (G_{t_1-s}(x_1-y)-G_{t_2-s}(x_2-y))^2 dyds\\
\leq & T\sqrt{\tfrac{2\pi}{4K}}c|(t_1,x_1)-(t_2,x_2)|^{1/4},
\end{aligned}$$

where the second inequality follows by Cauchy-Schwarz and the last inequality follows by Lemma 3.21. We also have that

$$\begin{aligned}
&\int_0^\infty\int_{X_0}^\infty |G_{t_1-s}(x_1-y)-G_{t_2-s}(x_2-y)|e^{-Ky^2}dyds\\
&\qquad\leq \int_0^\infty\int_{X_0}^\infty (G_{t_1-s}(x_1-y)+G_{t_2-s}(x_2-y))e^{-Ky^2}dyds\\
&\qquad\leq 2T.
\end{aligned}$$

If $M \geq 2X_0$ then

$$\begin{aligned}
\int_{X_0}^\infty G_{t_1-s}(x_1-y)e^{-Ky^2}dy \leq & e^{-KM^2/4}\int_{M/2}^\infty G_{t_1-s}(x_1-y)dy+\int_{X_0}^{M/2} G_{t_1-s}(x_1-y)dy\\
\leq & e^{-KM^2/4}+\mathbb{P}_{x_1}\left(B_{t_1-s}\leq M/2\right)\\
\leq & e^{-KM^2/4}+e^{-M^2/(8T)}, \qquad (3.45)
\end{aligned}$$

where the last inequality holds since $x_1 \geq M$ and $t_1 \leq T$. Therefore

$$\begin{aligned}
&\int_0^\infty\int_{X_0}^\infty |G_{t_1-s}(x_1-y)-G_{t_2-s}(x_2-y)|e^{-Ky^2}dyds\\
&\qquad\leq 2T(e^{-KM^2/4}+e^{-M^2/(8T)})\mathbb{1}_{M\geq 2X_0}+2T\mathbb{1}_{M<2X_0}.
\end{aligned}$$

It follows that

$$\begin{aligned}
&\mathbb{E}\left(\int_0^\infty\int_{X_0}^\infty |G_{t_1-s}(x_1-y)-G_{t_2-s}(x_2-y)|e^{-Ky^2}dyds\right)^{2k}\\
&\leq \mathbb{E}\Big[\Big(2T(e^{-KM^2/4}+e^{-M^2/(8T)})+2T\mathbb{1}_{M<2X_0}\Big)\\
&\qquad T^{(2k-1)/2}\left(\tfrac{2\pi}{4K}\right)^{(2k-1)/4}c^{(2k-1)/2}|(t_1,x_1)-(t_2,x_2)|^{(2k-1)/8}\Big]\\
&\leq Ae^{-M/A}|(t_1,x_1)-(t_2,x_2)|^3 \qquad (3.46)
\end{aligned}$$

for some constant $A = A(k) < \infty$, if $k \geq 13$, by Lemma 3.3. By the same argument (slightly simplified), we have that

$$\mathbb{E}\bigg(\int_0^\infty \int_{-\infty}^{-X_0} |G_{t_1-s}(x_1-y) - G_{t_2-s}(x_2-y)|e^{-Ky^2}dyds\bigg)^{2k} \leq Ae^{-M/A}|(t_1,x_1)-(t_2,x_2)|^3. \tag{3.47}$$

By substituting (3.44), (3.46) and (3.47) into (3.43), we have that for $k \geq 13$, there exists a constant $A(k)$ such that

$$\begin{aligned}\mathbb{E}\bigg(\int_0^\infty \int_{-\infty}^{\infty} (G_{t_1-s}(x_1-y) - G_{t_2-s}(x_2-y))f_\lambda(v_1(s,y))dyds\bigg)^{2k}&\\ \leq 6^{2k}Ae^{-M/A}|(t_1,x_1)-(t_2,x_2)|^3.&\end{aligned} \tag{3.48}$$

Finally, we bound the expectation of the last term on the right hand side of (3.40). By the Burkholder-Davis-Gundy inequality,

$$\begin{aligned}
&\mathbb{E}\bigg(\int_0^\infty \int_{-\infty}^{\infty} (G_{t_1-s}(x_1-y) - G_{t_2-s}(x_2-y))\sqrt{v_1(s,y)(1-v_1(s,y))}W(dyds)\bigg)^{2k}\\
&\leq \mathbb{E}\bigg(\int_0^\infty \int_{-\infty}^{\infty} (G_{t_1-s}(x_1-y) - G_{t_2-s}(x_2-y))^2 v_1(s,y)(1-v_1(s,y))dyds\bigg)^{k}\\
&\leq c^{k-1}|(t_1,x_1)-(t_2,x_2)|^{(k-1)/4}\mathbb{E}\int_0^\infty \int_{-\infty}^{\infty} (G_{t_1-s}(x_1-y) - G_{t_2-s}(x_2-y))^2\\
&\qquad\qquad v_1(s,y)(1-v_1(s,y))dyds\\
&\leq c^{k-1}|(t_1,x_1)-(t_2,x_2)|^{(k-1)/4}\bigg(\mathbb{E}\int_0^\infty \int_{X_0}^{\infty} (G_{t_1-s}(x_1-y) - G_{t_2-s}(x_2-y))^2 e^{-Ky^2}dyds\\
&\qquad + \mathbb{E}\int_0^\infty \int_{-X_0}^{X_0} (G_{t_1-s}(x_1-y) - G_{t_2-s}(x_2-y))^2 dyds\\
&\qquad + \mathbb{E}\int_0^\infty \int_{-\infty}^{-X_0} (G_{t_1-s}(x_1-y) - G_{t_2-s}(x_2-y))^2 e^{-Ky^2}dyds\bigg),
\end{aligned} \tag{3.49}$$

where the second inequality holds by Lemma 3.21 and the last inequality follows by the definition of $X_0$. Since $\|G_t\|_\infty = (2\pi t)^{-1/2}$,

$$\int_0^\infty \int_{-X_0}^{X_0} (G_{t_1-s}(x_1-y) - G_{t_2-s}(x_2-y))^2 dyds$$

$$
\begin{aligned}
&\leq 2\int_0^\infty \int_{-X_0}^{X_0} \left( \frac{1}{\sqrt{2\pi(t_1-s)}} G_{t_1-s}(x_1-y) + \frac{1}{\sqrt{2\pi(t_2-s)}} G_{t_2-s}(x_2-y) \right) dyds \\
&\leq 2\int_0^{t_1} \frac{1}{\sqrt{2\pi(t_1-s)}} \left( e^{-(M-X_0)^2/(2T)} + \mathbb{1}_{\leq X_0} \right) ds \\
&\quad + 2\int_0^{t_2} \frac{1}{\sqrt{2\pi(t_2-s)}} \left( e^{-(M-X_0)^2/(2T)} + \mathbb{1}_{M\leq X_0} \right) ds \\
&\leq \sqrt{\frac{2T}{\pi}} \left( e^{-(M-X_0)/(2T)} + \mathbb{1}_{M\leq X_0+1} \right),
\end{aligned}
$$

where the second inequality follows since $x_1, x_2 \geq M$ and $t_1, t_2 \leq T$. Similarly,

$$
\begin{aligned}
&\int_0^\infty \int_{X_0}^\infty (G_{t_1-s}(x_1-y) - G_{t_2-s}(x_2-y))^2 e^{-Ky^2} dyds \\
&\leq 2\int_0^\infty \int_{X_0}^\infty \left( \frac{1}{\sqrt{2\pi(t_1-s)}} G_{t_1-s}(x_1-y) + \frac{1}{\sqrt{2\pi(t_2-s)}} G_{t_2-s}(x_2-y) \right) e^{-Ky^2} dyds \\
&\leq 2\Big( \int_0^{t_1} \frac{1}{\sqrt{2\pi(t_1-s)}} (e^{-KM^2/4} + e^{-M^2/(8T)} + \mathbb{1}_{M<2X_0}) ds \\
&\qquad\qquad + \int_0^{t_2} \frac{1}{\sqrt{2\pi(t_2-s)}} (e^{-KM^2/4} + e^{-M^2/(8T)} + \mathbb{1}_{M<2X_0}) ds \Big) \\
&\leq \sqrt{\frac{2T}{\pi}} (e^{-KM^2/4} + e^{-M^2/(8T)} + \mathbb{1}_{M<2X_0}),
\end{aligned}
$$

where the second inequality follows by (3.45). By the same argument,

$$
\begin{aligned}
&\int_0^\infty \int_{-\infty}^{-X_0} (G_{t_1-s}(x_1-y) - G_{t_2-s}(x_2-y))^2 e^{-Ky^2} dyds \\
&\qquad\qquad\qquad \leq \sqrt{\frac{2T}{\pi}} (e^{-KM^2/4} + e^{-M^2/(8T)} + \mathbb{1}_{M<2X_0}).
\end{aligned}
$$

Therefore substituting into (3.49) and by Lemma 3.3, for $k \geq 13$, there exists a constant $A(k)$ such that

$$
\begin{aligned}
&\mathbb{E}\Big( \int_0^\infty \int_{-\infty}^\infty (G_{t_1-s}(x_1-y) - G_{t_2-s}(x_2-y)) \sqrt{v_1(s,y)(1-v_1(s,y))} W(dyds) \Big)^{2k} \\
&\leq A e^{-M/A} |(t_1,x_1) - (t_2,x_2)|^3. \qquad (3.50)
\end{aligned}
$$

Combining the bounds in (3.42), (3.48) and (3.50) with (3.40), the result follows. □

**Corollary 3.22.** *Take $T < \infty$. There exist constants $C < \infty$ and $a > 0$ such that the following* *holds. For $M \in \mathbb{R}$, for $x_1, x_2 \in [M, M+1]$ and $t_1, t_2 \in [0, T]$,*

$$|v_1(t_1, x_1) - v_1(t_2, x_2)| \le Y_M |(t_1, x_1) - (t_2, x_2)|^a,$$

*where $Y_M$ is a random variable with $\mathbb{E} Y_M^{2k} \le C A e^{-M/A}$.*

*Proof.* By the form of Kolmogorov's continuity criterion given in Corollary 1.2 in [42], this is a direct consequence of Lemma 3.20. □

We can now prove Proposition 3.5.

*Proof of Proposition 3.5.* For $\alpha \in (0,1)$, $\|\cdot\|_{H^\alpha}$ is equivalent to $|\cdot|_{H^\alpha}$ given by

$$|g|_{H^\alpha}^2 = \|g\|_{L^2}^2 + \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \frac{|g(y) - g(z)|^2}{|y - z|^{2\alpha}} dy dz.$$

(see (9.97) in [20]). It follows from Corollary 3.22 and the definition of $v^{\delta,\lambda}$ that for $\alpha \in (0,1)$ sufficiently small, $\mathbb{E} \sup_{t \le T} |v_t^{\epsilon,\lambda} - v_t^{\delta,\lambda}|_{H^\alpha}^2 < \infty$.

Now note that

$$v_t = (v_t^{\epsilon,\lambda} - v_t^{\delta,\lambda}) + S(v_t^{\delta,\lambda}) + m_{\eta(v_t^{\delta,\lambda})}.$$

For $t \le \sigma$, $S(v_t^{\delta,\lambda}) \in H^1 \subset H^\alpha$. For any $\eta \in \mathbb{R}$, $m - m_\eta \in H^1 \subset H^\alpha$. The result follows. □

## 3.4 Analysis of $\zeta(v)$

In this section we summarise some deterministic results about the map $\zeta$ from [20] which we will need in the next section. We also prove an additional result about $\zeta$, Proposition 3.24, which is not needed in [20].

The main results we need in the next section are summarised in the following theorem.

**Theorem 3.23** (Theorems 7.2 and 7.3 and Lemma 9.5 in [20]). *There exists $\beta_2 > 0$ such that the map $\zeta(v)$ is twice Fréchet differentiable on $\{v : dist(v, M) \le \beta_2\}$. There exists a constant $C < \infty$ such that for $v \in L^2 + m$ with $dist(v, M) \le \beta_2$ and $\zeta(v) = 0$,*

$$\|D\zeta(v) - D\zeta(m)\|_{L^2} \le C \sqrt{\text{dist}(v, M)}. \tag{3.51}$$

*Moreover,*

$$D\eta(y,v) = -\frac{\nabla m_{\eta(v)}(y)}{\langle v, \Delta m_{\eta(v)}\rangle}, \tag{3.52}$$

*and hence, for* $\eta \in \mathbb{R}$*,*

$$D\zeta(y,m_\eta) = -\frac{\nabla m_\eta(y)}{\|\nabla m\|_{L^2}^2}. \tag{3.53}$$

*For* $v \in L^2 + m$ *with dist*$(v,M) \leq \beta_2$ *and* $\zeta(v) = 0$*,*

$$\int_{-\infty}^{\infty} |D^2\zeta(y,y,v) - D^2\zeta(y,y,m)|dy \leq C\sqrt{\mathrm{dist}(v,M)}. \tag{3.54}$$

*Finally,*

$$\int_{-\infty}^{\infty} |D^2\zeta(y,y,m)|^2 dy < \infty. \tag{3.55}$$

*Proof.* The statements (3.51) and (3.53) are in Theorem 7.2 in [20]. The statement (3.52) is in Lemma 9.5 in [20]. The next statement, (3.54), follows by (9.47,i) in the proof of Theorem 7.3 in [20]. The last statement, (3.55), follows by (7.9) in Theorem 7.3 in [20]. ☐

The result we wish to prove is very similar to these results.

**Proposition 3.24.** *There exists a constant* $C < \infty$ *and a* $\beta_2 > 0$ *such that for* $v \in L^2 + m$ *with* $\mathrm{dist}(v,M) \leq \beta_2$ *and* $\zeta(v) = 0$*,*

$$\|D\zeta(v) - D\zeta(m)\|_{L^1} \leq C\sqrt{\mathrm{dist}(v,M)}. \tag{3.56}$$

In order to prove this result, we build on several results proven in [20]. First we have an explicit expression for the Fréchet derivative of $\zeta(v)$.

**Proposition 3.25** (Proposition 9.2 in [20])**.** *The Frećhet derivative of* $\zeta(v)$ *for* $\{v \in L^2 + m : \mathrm{dist}(v,M) \leq \beta_2\}$ *and* $y \in \mathbb{R}$ *is given by*

$$D\zeta(y,v) = e^{\int_0^\infty \Lambda(s,v)ds} D\eta(y,v) + \int_0^\infty \Xi(s,y,v)e^{\int_s^\infty \Lambda(r,v)dr}ds, \tag{3.57}$$

*where*

$$\Lambda(t,v) := \frac{\varphi_1(t,v)\tilde{\varphi}_2(t,v) - \tilde{\varphi}_1(t,v)\varphi_2(t,v)}{(\tilde{\varphi}_1(t,v))^2} \tag{3.58}$$

$$\Xi(t,y,v) := \frac{\psi(t,y,v)\tilde{\varphi}_1(t,v) - \tilde{\psi}(t,y,v)\varphi_1(t,v)}{(\tilde{\varphi}_1(t,v))^2}, \tag{3.59}$$

*and where, for* $k = 1, 2$

$$\begin{aligned} \varphi_k(t,v) &:= \langle \Delta v_t + f_0(v_t), \nabla^k m_{\eta_t} \rangle \qquad (3.60)\\ \tilde{\varphi}_k(t,v) &:= \langle v_t, \nabla^{k+1} m_{\eta_t} \rangle \\ \psi(t,y,v) &:= \langle q_t(y,\cdot,v), \{\Delta + f_0'(v_t)\} \nabla m_{\eta_t} \rangle \qquad (3.61)\\ \tilde{\psi}(t,y,v) &:= \langle q_t(y,\cdot,v), \nabla^2 m_{\eta_t} \rangle. \qquad (3.62) \end{aligned}$$

*In these expressions,* $\eta_t = \eta(v_t)$, $q_t(x,y,v)$ *is the fundamental solution of* $\frac{d}{dt} - \Delta - f_0'(v_t(z,v))$ *and finally,* $v_t(x) = v_t(x,v)$ *is the solution to* $\frac{\partial u}{\partial t} = \Delta u + f_0(u)$ *with initial condition* $v$*, as in the definition of* $\zeta$ *in Theorem 3.6.*

Second, we collect a series of bounds proven in [20].

**Proposition 3.26.** *There exists* $\beta_2 > 0$ *such that the following estimates hold for all* $\{v : \mathrm{dist}(v, M) \leq \beta_2\}$ *with constants* $C, c, \bar{c} < \infty$*:*

$$\begin{aligned} \|S(v_t)\|_{L^2} &\leq C\sqrt{\mathrm{dist}(v,M)}\exp(-\bar{c}t), \qquad (3.63)\\ |\varphi_k(t,v)| &\leq C\sqrt{\mathrm{dist}(v,M)}\exp(-\bar{c}t) \qquad (3.64)\\ |\tilde{\varphi}_k(t,v)| &\leq C \qquad (3.65)\\ |\tilde{\varphi}_1(t,v)| &\geq c \qquad (3.66)\\ |\Lambda(t,v)| &\leq C\sqrt{\mathrm{dist}(v,M)}\exp(-ct) \qquad (3.67)\\ \left|\frac{d\eta_t}{dt}\right| &\leq C\sqrt{\mathrm{dist}(v,M)}\exp(-ct) \qquad (3.68) \end{aligned}$$

*recalling that* $S(v) := v - m_{\eta(v)}$ *is one of the Fermi coordinates of* $v$*. Furthermore, we may take* $\bar{c} = 1/6$*.*

*Proof.* Most of these bounds are exactly given in Funaki, but $\bar{c}$ is not explicit, so we shall prove this here. Before that, we summarise where the other inequalities come from. The first equation, (3.63), is item (iii) in Corollary 9.1 of [20]. The next three bounds are from Proposition 9.1 in [20]. This proposition has an additional assumption on $v$, and additional results, however, the additional assumption is only necessary for proving these additional results, as testified by Remark 9.2 of [20]. The fifth result is a corollary of the previous three and (3.58). The final bound is Equation 9.15 in [20].

Using the definition of $\varphi$ in (3.60), and recalling that $\Delta m + f_0(m) = 0$ it is not too hard to see that

$$\begin{aligned}|\varphi_k(t)| =& \left|\langle \Delta v_t - \Delta m_{\eta_t}, \nabla^k m_{\eta_t}\rangle + \langle f_0(v_t) + \Delta m_{\eta_t}, \nabla^k m_{\eta_t}\rangle\right| \\ \leq& \left|\langle v_t - m_{\eta_t}, \nabla^{k+2} m_{\eta_t}\rangle\right| + \left|\langle f_0(v_t) - f_0(m_{\eta_t}), \nabla^k m_{\eta_t}\rangle\right| \\ \leq& \|S(v_t)\|_{L^2} \|\nabla^{k+2} m_{\eta_t}\|_{L^2} + \|S(v_t)\|_{L^2} \|f_0\|_{L^\infty} \|\nabla^k m_{\eta_t}\|_{L^2}, \\ \leq& C\|S(v_t)\|_{L^2},\end{aligned}$$

and so we just need to show the explicit bound for (3.63). We shall do this by following this constant back through the lemmas and corollaries of [20] until we can calculate this value. We will break this into several parts.

### Part 1, Corollary 9.1

First, let us state Corollary 9.1 of [20].

**Corollary 3.27** (Corollary 9.1 of [20]). *There exist $C, \bar{c}_{cor9.1}, \bar{c}$ and $\beta_2 > 0$ such that for all $\{v : \operatorname{dist}(v, M) \leq \beta_2\}$,*

*(i)* $\|S(v_t)\|_{L^2} \leq C \operatorname{dist}(v, M) e^t$, *for* $t \geq 0$,

*(ii)* $\|S(v_t)\|_{H^1} \leq C\sqrt{\operatorname{dist}(v, M)} e^{-\bar{c}_{cor9.1} t}$, *for* $t \geq 1$,

*(iii)* $\|S(v_t)\|_{L^2} \leq C\sqrt{\operatorname{dist}(v, M)} e^{-\bar{c} t}$, *for* $t \geq 0$.

It is clear from its definition, that $\|v\|_{H^1} \geq \|v\|_{L^2}$. Hence, we may take $\bar{c} = \bar{c}_{cor9.1}$. Now we seek an explicit expression for $\bar{c}_{cor9.1}$. Corollary 9.1 (ii) relies upon Corollary 9.1 (i), Lemma 9.1 (ii) and Lemma 9.1 (iii).

### Part 2, Lemma 9.1

Again, let us state the relevent parts of Lemma 9.1 of [20].

**Lemma 3.28** (Lemma 9.1 of [20]). *There exist* $C, \bar{c}_{lem9.1} > 0$ *and* $\beta > 0$ *such that*

*(ii)* $\mathcal{H}(v_t) \leq \frac{C}{t}\operatorname{dist}(v, M)$ *for all* $t \in [0,1]$ *if* $\operatorname{dist}(v, M) \leq 1$*,*

*(iii)* $\mathcal{H}(v_t) \leq e^{-\bar{c}_{lem9.1}(t-s)}\mathcal{H}(v_s)$ *for* $t \geq s \geq 0$ *if* $\mathcal{H}(v_s) \leq \beta$*.*

By Lemma 9.1 (ii), $\mathcal{H}(v_1)$ can be made sufficiently small by taking $\operatorname{dist}(v, M)$ sufficiently small. Applying this to Lemma 9.1 (iii),

$$c_1\|S(v_t)\|_{H^2}^2 \leq \mathcal{H}(v_t) \leq e^{-\bar{c}_{lem9.1}(t-1)}\mathcal{H}(v_1)\operatorname{dist}(v, M),$$

where we have added the lower bound from (3.11). This gives Corollary 9.1 (ii), with $\bar{c}_{cor9.1} = \frac{1}{2}\bar{c}_{lem9.1}$

To prove Lemma 9.1 (iii), we note that by Theorem 3.6,

$$\frac{d}{dt}\mathcal{H}(v_t) = -\|D\mathcal{H}(\cdot, v)\|_{L^2}^2 \leq -\frac{c_1}{c_2}\mathcal{H}(v_t), \tag{3.69}$$

if $\|S(v_t)\|_{H^1} \leq \beta_1$, and therefore, by Lemma 9.1 (ii), if $\mathcal{H}(v_s) \leq \beta$ for some $\beta$. We will calculate these constants slightly more directly, $c_1$ in particular, than by appealing to Theorem 3.6, but the idea is the same.

### Part 3, Theorem 3.1, upper bound

It is a simple rearrangement of the definition of $\mathcal{H}(v)$ to see that

$$\begin{aligned}
\mathcal{H}(v) \leq &\int_{\mathbb{R}} \frac{1}{2}|\nabla v(y)|^2 + F_0(v(y))dy \\
= &\int_{\mathbb{R}} -\frac{1}{2}m_\eta(y)\Delta m_\eta(y) + \frac{1}{2}|\nabla S(y)|^2 - S(y)\Delta m_\eta(y) + F_0(v(y))dy \\
= &\frac{1}{2}\|\nabla s\|_{L^2}^2 + \int_{\mathbb{R}} F_0(S(y) + m_\eta(y)) - F_0(m_\eta(y)) - F_0'(m_\eta(y))S(y)dy \\
= &\frac{1}{2}\|\nabla s\|_{L^2}^2 + \int_{\mathbb{R}} \frac{1}{2}F_0''(\xi)S(y)^2dy
\end{aligned}$$

for some $\xi$ between $m_\eta(y)$ and $v(y)$, where we have used Taylor's theorem. Both these terms are bounded by 0 and 1, so $\xi \in [0,1]$. We can calculate explicitly $F_0''(x) = -f_0'(x) = 6x^2 - 6x + 1$, so $\max_{x\in[0,1]} |F_0''(x)| = 1$. Hence,

$$\begin{aligned}\mathcal{H}(v) \leq& \frac{1}{2}\|\nabla s\|_{L^2}^2 + \frac{1}{2}\|s\|_{L^2}^2 \\ =& \frac{1}{2}\|s\|_{H^1}^2,\end{aligned} \tag{3.70}$$

and hence, $c_2 = \frac{1}{2}$.

**Part 3, Theorem 3.1, lower bound.**

The bound in Theorem 3.6 is for $\|D\mathcal{H}(\cdot, v)\|_{L^2}^2$ in terms of $\|s\|_{H^2}^2$, however, we only need a bound in terms of $\|s\|_{H^1}^2$ in order to satisfy the inequality in (3.69). We do this in two steps.

Let $\mathcal{A} = -\Delta - f_0'(m(y))$, and note that 0 is an eigenvalue of this operator. Denote the second eigenvalue by $\mu$. Hence,

$$\|\mathcal{A}s\|_{L^2}^2 \geq \mu^2 \|s\|_{L^2}^2. \tag{3.71}$$

Let $0 < a \leq \frac{1}{2}$. Recalling that $\|s\|_{H^1}^2 = \|s\|_{L^2}^2 + \|\nabla s\|_{L^2}^2$ and integrating the final term by parts, we can also directly calculate,

$$\begin{aligned}\|\mathcal{A}s\|_{L^2}^2 =& \|\Delta s\|_{L^2}^2 + \|f_0'(m)s\|_{L^2}^2 + 2\langle f_0'(m)s, \Delta s\rangle + a\|s\|_{H^1}^2 - a\|s\|_{L^2}^2 + a\langle s, \Delta s\rangle \\ \geq& a\|s\|_{H^1}^2 + \|\Delta s\|_{L^2}^2 - \langle |2f_0'(m) + a|, |s\Delta s|\rangle - a\|s\|_{L^2}^2 \\ \geq& a\|s\|_{H^1}^2 + \|\Delta s\|_{L^2}^2 - (2-a)\|s\|_{L^2}\|\Delta s\|_{L^2} - a\|s\|_{L^2}^2 \\ \geq& a\|s\|_{H^1}^2 + \|\Delta s\|_{L^2}^2 - \tfrac{(2-a)^2}{4}\|s\|_{L^2}^2 - \|\Delta s\|_{L^2}^2 - a\|s\|_{L^2}^2 \\ =& a\|s\|_{H^1}^2 - ((1-a/2)^2 + a)\|s\|_{L^2}^2,\end{aligned}$$

where the third line holds by the upper bound on $a$ and recalling that $f_0'(x) \in [-1, 1/2]$ if $x \in [0,1]$. Bounding a proportion $\alpha \in [0,1]$ of $\|\mathcal{A}s\|_{L^2}^2$ with this bound, and the rest with the eigenvalue bound (3.71) gives,

$$\|\mathcal{A}s\|_{L^2}^2 \geq \alpha a\|s\|_{H^1}^2 - \alpha(1 + a^2/4)\|s\|_{L^2}^2 + (1-\alpha)\mu^2\|s\|_{L^2}^2.$$

Choosing $\alpha$ such that the $\|s\|_{L^2}^2$ terms cancel means choosing $\alpha = \frac{\mu^2}{1+a^2/4+\mu^2}$, and hence,

$$\|\mathcal{A}s\|_{L^2}^2 \geq \frac{a\mu^2}{1 + a^2/4 + \mu^2}\|s\|_{H^1}^2. \tag{3.72}$$

We also need an upper bound on this term. Similarly to before, we find that

$$\begin{aligned}\|\mathcal{A}s\|_{L^2}^2 \leq& \|\Delta s\|_{L^2}^2 + \|(2f_0'(m))s\Delta s\|_{L^2} + \|f_0'(m))s\|_{L^2}^2 \\ \leq& \|\Delta s\|_{L^2}^2 + \|s\|_{L^2}^2 + \|\Delta s\|_{L^2}^2 + \|s\|_{L^2}^2 \\ =& 2\|s\|_{H^2}^2. \end{aligned} \tag{3.73}$$

Now we consider $\|D\mathcal{H}(\cdot, v)\|_{L^2}^2 = \|\mathcal{A} + V\|_{L^2}^2$, where

$$V = F_0'(v) - F'(m_\eta(y)) - F_0''(m_\eta(y))S(y).$$

Note that, by applying Taylor's theorem again, for some $\xi \in [0,1]$,

$$\|V\|_{L^2} \leq \frac{1}{2}\|F_0'''(\xi)s^2\|_{L^2} \leq 3\|s\|_{L^2}, \tag{3.74}$$

as $F_0'''(x) = -f_0''(x) = 12x - 6$, so $\max_{x\in[0,1]} \|F_0'''(x)| = 6$, and noting also that $S(y) \in [0,1]$, and hence $|S(y)|^4 \leq |S(y)|^2$. Hence,

$$\begin{aligned}\big|\|D\mathcal{H}(\cdot, v)\|_{L^2}^2 - \|\mathcal{A}s\|_{L^2}^2\big| \leq& 2\|\mathcal{A}s\|_{L^2}^2\|V\|_{L^2} + \|V\|_{L^2}^2 \\ \leq& 3\|s\|_{L^2}(4\|s\|_{H^2}^2 + 3\|s\|_{L^2}), \end{aligned} \tag{3.75}$$

where we have used the bounds (3.74) and (3.73). Theorem 3.6 says that this should hold for all $v \in H^1 + m : \|s\|_{H^1} \leq \beta_1$. For such $v$, we must also have that $\|s\|_{L^2} \leq \beta_1$, so we choose $\beta_1$ small enough so that the bound in (3.75) is less than, say $\|s\|_{H^1}^2/200$. Combining this with (3.72) allows us to conclude that

$$\|D\mathcal{H}(\cdot, v)\|_{L^2}^2 \geq \frac{a\mu^2}{1 + a^2/4 + \mu^2}\|s\|_{H^1}^2 - \|s\|_{H^1}^2/200, \tag{3.76}$$

and hence, all that remains is to calculate this bound explicitly.

To do so, we calculate the eigenvalues of $\mathcal{A}$. The eigenvalues $k$ will solve the ODE $\mathcal{A}u = ku$. As we have an explicit expression for $f_0'(m(y))$, we can solve this ODE explicitly:

$$u(x) = C_1 e^{-x\sqrt{1+k}} \frac{3 + 2(1+e^x)^2 k - 3\sqrt{1+k} + 3e^{2x}(1 + \sqrt{1+k}))}{(1+e^x)^2}$$

$$+ C_2 e^{x\sqrt{1+k}} \frac{3 + 2(1+e^x)^2 k + 3\sqrt{1+k} - 3e^{2x}(-1+\sqrt{1+k}))}{(1+e^x)^2},$$

and hence, we find three eigenvalues: 0, with eigenfunction $\frac{e^x}{(1+e^x)^2}$; 3/4, with eigenfunction $\frac{-3e^{x/2}(e^x-1)}{(1+e^x)^2}$; and 1, with eigenfunction $1 - \frac{6e^x}{(1+e^x)^2}$. Substituting $\mu = 3/4$ into (3.72), then optimising for $a \in (0, 1/2]$ gives that the bound is optimised at $a = \frac{1}{2}$, at a value of 9/52=0.17307... Applying this to (3.76), we may safely say that $\|D\mathcal{H}(\cdot, v)\|_{L^2}^2 > \frac{1}{6}\|s\|_{H^1}^2$. Hence, as in (3.69),

$$\frac{d}{dt}\mathcal{H}(v_t) = -\|D\mathcal{H}(\cdot, v)\|_{L^2}^2 \leq -\frac{1}{6}\|s\|_{H^1}^2 \leq -\frac{1}{3}\mathcal{H}(v_t),$$

by (3.70) in the last inequality. Hence, $\bar{c}_{lem9.1} = 1/3$, giving $\bar{c}_{cor9.1} = 1/6 = \bar{c}$. □

The final ingredient is a bound on $q_t(x, y, v)$, which is the fundamental solution to the linerised version of (3.2), which we introduced in Proposition 3.25. To be more precise, we use the characterisation of this term from [20].

**Lemma 3.29** (Lemma 9.3 of [20]). *Recall that* $v_t(x, v) = v_t(x)$ *is the solution to* (3.2) *with initial condition* $v$. *For* $0 < t$,

$$q_t(x, y, v) = E_{t,x,y}\left[\exp\left(\int_0^t f_0'\left(v_r\left(B_r, v\right)\right) dr\right)\right] p_t(x, y), \tag{3.77}$$

*where the expectation is taken with respect to the Brownian motion,* $B_t$, *starting at* $\frac{x}{\sqrt{2}}$ *and pinned to* $\frac{y}{\sqrt{2}}$ *at time* $t$, *and where* $p_t(x, y)$ *is the heat kernel.*

*Proof.* This is just an application of the Feynman-Kac formula. □

Now we can find the required bound on $q_t(x, y, v)$. Roughly, we have an explicit expression for $f_0$, and a simple calculation shows that the integrand in (3.77) is positive only when the pinned Brownian motion is near the origin. Hence, to bound $q_t(x, y, t)$, we just need to bound the amount of time the pinned Brownian motion spends in this region where the integral is positive.

**Lemma 3.30.** *There exists* $\beta_2 > 0$ *small enough such that if* $\operatorname{dist}(v, M) \leq \beta_2$, *then, for* $t > 0$, *there exists constants* $C > 0$ *and* $c < 1/6$, *which do not depend on* $t$, *such that*

$$q_t(x, y, v) < Ce^{ct} p_t(x, y).$$

*Proof.* Initially, consider the set $X_a^r := \{x \in \mathbb{R} : f_0'(v_r(x)) \geq a\}$ for some $a \in [-1, 1/2]$. Recalling that $f_0'(u) = -6u^2 + 6u - 1$, a simple calculation and the trivial bounds $0 \leq v, m \leq 1$ also allows us to write,

$$f_0'(v_r) \leq f_0'(m_{\eta(v_r)}) + 30|v_r - m_\eta(v_r)| = f_0'(m_{\eta(v_r)}) + 30|S(v_r)|.$$

Let $Y_\epsilon^r := \{x \in \mathbb{R} : |S(v_r)(x)| > \epsilon\}$. By Chebyshev's inequality, $Leb(Y_\epsilon^r) \leq \frac{\|S(v_r)\|_{L^2}^2}{\epsilon^2} \leq \epsilon^{-2}\beta_2^2$. Hence,

$$\begin{aligned} X_a^r \subset & Y_\epsilon^r \cup \{x \in \mathbb{R} : f_0'(m_{\eta(v_r)}) \geq a - 30\epsilon\} \\ = & Y_\epsilon^r \cup \left[\eta(v_r) \pm \log\left(\frac{2 + 30\epsilon - a + \sqrt{3(1 - 2a + 60\epsilon)}}{1 + a - 30\epsilon}\right)\right], \end{aligned}$$

where we abuse interval notation slightly, and we have calculated the second line using the explicit formula for $m$. By Theorem 3.6, and the assumption $\operatorname{dist}(v, M) \leq \beta_2$, $\eta(v_r)$ converges as $r \to \infty$, there exists some finite $\bar{\eta} = \bar{\eta}(v)$ such that $\bar{\eta} \geq |\eta(v_r)|$ for all $r$. If necessary, we can take $\beta_2$ small enough so that $\bar{\eta} < \epsilon$. Hence, if we define

$$I_\epsilon^a = \left[0 \pm \left(\epsilon + \log\left(\frac{2 + 30\epsilon - a + \sqrt{3(1 - 2a + 60\epsilon)}}{1 + a - 30\epsilon}\right)\right)\right],$$

we still have that $X_a^r \subset I_\epsilon^a \cup Y_\epsilon^r$, and thus in order to bound the amount of time the pinned Brownian motion spends in the sets $X_a^r$, we can consider the amount of time it spends in the sets $I_\epsilon^a \cup Y_\epsilon^r$. As we can choose $\beta_2$ to be arbitrarily small, the vast majority of this time will be spent in the interval $I_\epsilon^a$, rather than $Y_\epsilon^r$. The pinned Brownian motion which spends the most time in $I_\epsilon^a$ is the one that begins and terminates in the center of this interval, so we may assume our Brownian motion begins and terminates at 0, in order to find an upper bound.

In this case, the worst place for $Y_\epsilon^r$ to be would be, firstly, in the complement of $I_\epsilon^a$, as otherwise it does not increase the union of $Y_\epsilon^r$ and $I_\epsilon^a$, and secondly, as close to the interval as possible, as the closer a set is to the terminal points of the Brownian motion, the more likely the Brownian motion is to be there. Hence, if $\Gamma_A(t, x, y)$ is the amount of time a Brownian motion from $x$ pinned to $y$ at time $t$ spends in the set $A$, and we define

$$Z_\epsilon^a := \left[0 \pm \left(\epsilon + \beta_2^2\epsilon^{-2} + \log\left(\frac{2 + 30\epsilon - a + \sqrt{3(1 - 2a + 60\epsilon)}}{1 + a - 30\epsilon}\right)\right)\right],$$

we have that $\Gamma_{X_a^r}\left(t, \frac{x}{\sqrt{2}}, \frac{y}{\sqrt{2}}\right) \leq \Gamma_{Z_\epsilon^a}(t,0,0)$.

Unfortunately, approximating $f_0'(v_r(x))$ with a single set is not quite enough for this method to prove the bound we are looking for (see Appendix A, in particular, Figure A.1, for an illustration of the single set approximation, and why this approximation is insufficient). We can improve this approximation by adding more sets. If $\{a_i\}_{i=0}^{n+1}$ is a finite strictly decreasing sequence in $[-1,1/2]$ such that $a_0 = \frac{1}{2}$ and $a_{n+1} = -1$, by the same reasoning as above,

$$\sum_{i=0}^{n} a_i \Gamma_{X_{a_i}^r \setminus X_{a_{i+1}}^r}\left(t, \frac{x}{\sqrt{2}}, \frac{y}{\sqrt{2}}\right) \leq \sum_{i=0}^{n} a_i \Gamma_{Z_\epsilon^{a_i} \setminus Z_\epsilon^{a_{i+1}}}(t,0,0).$$

We also define $\{R_i\}_{i=0}^{n+1}$ by $Z_\epsilon^{a_i} = [-R_i, R_i]$. We will choose these $a_i$ explicitly later. The important thing to observe at this point is that we may write

$$\begin{aligned}
q_t(x,y,v) =& E_{t,x,y}\left[\exp\left(\int_0^t f_0'\left(v_r\left(B_r, v\right)\right) dr\right)\right] p_t(x,y) \\
\leq& E_{t,x,y}\left[\exp\left(\int_0^t \sum_{i=0}^{n} a_i \mathbb{1}_{\left\{B_r \in X_{a_i}^r \setminus X_{a_{i+1}}^r\right\}} dr\right)\right] p_t(x,y) \\
\leq& E_{t,0,0}\left[\exp\left(\int_0^t \sum_{i=0}^{n} a_i \mathbb{1}_{\left\{B_r \in Z_\epsilon^{a_i} \setminus Z_\epsilon^{a_{i+1}}\right\}} dr\right)\right] p_t(x,y) \\
=& E_{t,0,0}\left[e^{t/2} \exp\left(-\int_0^t \sum_{i=1}^{n} (\tfrac{1}{2} - a_i) \mathbb{1}_{\left\{B_r \in Z_\epsilon^{a_i} \setminus Z_\epsilon^{a_{i+1}}\right\}} dr\right)\right] p_t(x,y).
\end{aligned}$$

Next, we wish to remove the condition that the Brownian motion is pinned at time $t$. We can do this relatively easily by differentiating the joint cumulative distribution function of our expression and the final location of an unpinned Brownian motion, $W_t$:

$$\begin{aligned}
q_t(x,y,v) \leq& E_{t,0,0}\left[e^{t/2} \exp\left(-\int_0^t \sum_{i=1}^{n} (\tfrac{1}{2} - a_i) \mathbb{1}_{\left\{B_r \in Z_\epsilon^{a_i} \setminus Z_\epsilon^{a_{i+1}}\right\}} dr\right)\right] p_t(x,y) \\
=& \frac{d}{dz} E\left[\mathbb{1}_{W_t \geq z} e^{t/2} \exp\left(-\int_0^t \sum_{i=1}^{n} (\tfrac{1}{2} - a_i) \mathbb{1}_{\left\{W_r \in Z_\epsilon^{a_i} \setminus Z_\epsilon^{a_{i+1}}\right\}} dr\right)\right]\Bigg|_{z=0} p_t(x,y) \\
=:& \, g(t) e^{t/2} p_t(x,y), \qquad (3.78)
\end{aligned}$$

where the expectation is now being taken with respect to $W_t$, and we have defined a new function $g$. We are done if we can prove that $g(t) \leq Ce^{\tilde{c}t}$ for some $\tilde{c} < 1/6 - 1/2$. We can do this by applying again a version of the Feynman-Kac formula.

If we take the the Laplace transform of $g(t)$, we may exchange the expectation and limit, as the integrand is positive:

$$\begin{aligned}\mathcal{L}_t\left[g(t);\lambda\right] &= \int_0^\infty e^{-\lambda t}\frac{d}{dz}\,E\left[\mathbb{1}_{W_t\geq z}\exp\left(-\int_0^t\sum_{i=1}^n(\tfrac{1}{2}-a_i)\mathbb{1}_{\left\{W_r\in Z_\epsilon^{a_i}\setminus Z_\epsilon^{a_i+1}\right\}}dr\right)\right]\Bigg|_{z=0} dt\\ &=\frac{d}{dz}\,E\left[\mathbb{1}_{W_t\geq z}\int_0^\infty e^{-\lambda t}\exp\left(-\int_0^t\sum_{i=1}^n(\tfrac{1}{2}-a_i)\mathbb{1}_{\left\{W_r\in Z_\epsilon^{a_i}\setminus Z_\epsilon^{a_i+1}\right\}}dr\right)dt\right]\Bigg|_{z=0.}\end{aligned} \tag{3.79}$$

We now apply Kac's formula, a one-dimensional version of the Feynman-Kac formula (see [25], Theorem 4.9, p271) to (3.79). In slightly less generality than presented in [25], Kac's formula says that if $k:\mathbb{R}\to[0,\infty)$ is piecewise continuous, and $f:\mathbb{R}\to\mathbb{R}$, then an expression of the form

$$u(x) := E_x\left[f(W_t)\int_0^\infty e^{-\lambda t}e^{-\int_0^t k(W_r)dr}dt\right],$$

where the expectation is taken with respect to a Brownian motion started at $x$, satisfies the differential equation

$$\lambda u + k(x)u = \frac{1}{2}u'' + f(x). \tag{3.80}$$

Recalling that we have denoted the radii of the sets $Z_\epsilon^{a_{i+1}}$ by $R_i$, the unique bounded solution of (3.80) with $k=\sum_{i=1}^n(\frac{1}{2}-a_i)\mathbb{1}_{\left\{x\in Z_\epsilon^{a_i}\setminus Z_\epsilon^{a_i+1}\right\}}$ and $f=\mathbb{1}_{x>z}$ with $-R_1<z<R_1$ is

$$u(x)=\begin{cases} A_n e^{-x\sqrt{2(\lambda+\frac{1}{2}-a_n)}} & x\geq R_n\\ \vdots & \vdots\\ A_2\cosh\left(x\sqrt{2(\lambda+\frac{1}{2}-a_1)}\right)+B_2\sinh\left(x\sqrt{2(\lambda+\frac{1}{2}-a_1)}\right) & R_1\leq x<R_2\\ A_1\cosh\left(x\sqrt{2\lambda}\right)+B_1\sinh\left(x\sqrt{2\lambda}\right) & z\leq x<R_1\\ C_1\cosh\left(x\sqrt{2\lambda}\right)+D_1\sinh\left(x\sqrt{2\lambda}\right)+\frac{1}{\lambda} & -R_1\leq x<z\\ C_2\cosh\left(x\sqrt{2\lambda+\frac{1}{2}-a_1}\right)+D_2\sinh\left(x\sqrt{2\lambda+\frac{1}{2}-a_1}\right)+\frac{1}{\lambda+\frac{1}{2}-a_1} & -R_2\leq x<-R_1\\ \vdots & \vdots\\ C_n e^{x\sqrt{2(\lambda+\frac{1}{2}-a_n)}}+\frac{1}{\lambda+\frac{1}{2}-a_n} & x<-R_n,\end{cases}$$

where $\{A_i\}_{i=1}^n$, $\{C_i\}_{i=1}^n$, $\{B_i\}_{i=1}^{n-1}$ and $\{D_i\}_{i=1}^{n-1}$ are constants chosen so that $u(x)$ and $u'(x)$ are continuous. This expression is obviously quite complicated, but we only want an expression for $\frac{d}{dz}u(0)\big|_{z=0}$, so we don't need to completely solve for $u$. In fact, assuming $z \leq 0$, $u(0)$ is simply $A_1$. We still don't need a fully explicit expression for $A_1$, but we need to know how it depends on $z$, in order to differentiate it.

To do this, we notice that it is not too difficult to write $A_{n-1}$ and $B_{n-1}$ in terms of $A_n$ by considering $u(r_n)$ and $u'(r_n)$ and exploiting the Pythagorean hyperbolic identity. For example, if we define $\phi(x, y, z) = x\cosh(z) + y\sinh(z)$, we find

$$\begin{aligned} A_{n-1} =& A_n e^{-R_n\sqrt{2(\lambda+\frac{1}{2}-a_n)}}\phi\left(1, \sqrt{\frac{\lambda+\frac{1}{2}-a_n}{\lambda+\frac{1}{2}-a_{n-1}}}, R_n\sqrt{2(\lambda+\tfrac{1}{2}-a_{n-1})}\right), \\ B_{n-1} =& -B_n e^{-R_n\sqrt{2(\lambda+\frac{1}{2}-a_n)}}\phi\left(\sqrt{\frac{\lambda+\frac{1}{2}-a_n}{\lambda+\frac{1}{2}-a_{n-1}}}, 1, R_n\sqrt{2(\lambda+\tfrac{1}{2}-a_{n-1})}\right). \end{aligned} \tag{3.81}$$

The coefficient of $A_n$ in this expression is just a function of $\lambda$ and constants, in particular, not a function of $z$. Using the same ideas, we can find and expression for $A_{i-1}$ and $B_{i-1}$ in terms of $A_i$ and $B_i$, and hence $A_n$, until we can express $A_1$ and $B_1$ in terms of $A_n$. We won't write all these expressions here, but we do take note of a number of their important properties.

1. We know that the expressions will be of the form $A_1 = aA_n$ and $B_1 = bA_n$, where $a$ and $b$ are functions of $\lambda$ and the constants $\{R_1\}_{i=1}^n$ and $\{a_1\}_{i=1}^n$, but not $z$.

2. We can do the same iteration for the $C_i$ and $D_i$ constants, however, there will be an additional additive constant term in their expression, resulting in expressions $C_1 = cC_n + \bar{c}$ and $D_1 = dC_n + \bar{d}$, again, where $c$, $\bar{c}$, $d$ and $\bar{d}$ are functions of $\lambda$ and the constants $\{R_1\}_{i=1}^n$ and $\{a_1\}_{i=1}^n$, and not $z$.

3. In fact, the differences between the positive and negative halves of $u$ only affect the additional constant terms, and we find that $a = c$ and $b = -d$, where the negative is because sinh changes sign with a negative argument.

We now consider the continuity conditions of $u$ and $u'$ at $z$. Using the Pythagorean hyperbolic identity again, we find that

$$A_1 = C_1 + \frac{1}{\lambda}\cosh\left(z\sqrt{2\lambda}\right) \text{ and } B_1 = D_1 - \frac{1}{\lambda}\sinh\left(z\sqrt{2\lambda}\right).$$

Hence, we find

$$A_1 = \frac{\frac{ac}{\lambda}\sinh\left(z\sqrt{2\lambda}\right) - ac\bar{d} + ad\bar{c} + \frac{ad}{\lambda}\cosh\left(z\sqrt{2\lambda}\right)}{ad - bc}.$$

We can differentiate this expression with respect to $z$, and then set $z$ to 0, yielding

$$\left.\frac{d}{dz}A_1\right|_{z-0} = \sqrt{\frac{2}{\lambda}}\frac{ac}{ad-bc} = -\frac{1}{\sqrt{2\lambda}}\frac{a}{b} = \mathcal{L}_t\left[g(t);\lambda\right], \tag{3.82}$$

where we have used observation 3 in the second equality.

This expression is still quite a complicated function of $\lambda$ (recalling observation 1), so we have no hope of inverting our Laplace transform explicitly. Fortunately, we don't need an explicit formula for $g(t)$; we are only interested in its asymptotics. In particular, the final value theorem for Laplace transforms says that

$$\lim_{t\to\infty} f(t) = \lim_{\lambda\to 0}\lambda\mathcal{L}_t\left[f(t);\lambda\right]$$

if all the poles of $\lambda\mathcal{L}_t\left[f(t);\lambda\right]$ are in the left half-plane. If this limit were finite, this would show that $f(t)$ was bounded asymptotically, which is not quite strong enough. We want to show that that there is some $\tilde{c} < 1/6 - 1/2$ such that $g(t)e^{-\tilde{c}t}$ is asymptotically bounded. So we can use the 'frequency shifting' property of Laplace transforms, to write

$$\lim_{t\to\infty} g(t)e^{-\tilde{c}t} = \lim_{\lambda\to 0}\lambda G(\lambda+\tilde{c}),$$

if all the poles of $\lambda G(\lambda+\tilde{c})$ are in the left hand plane. In other words, we are done if there exists a $\tilde{c} < 1/6 - 1/2$ such that all the poles of $G(\lambda)$ have real part less than $\tilde{c}$.

We now numerically solve this system of equations to investigate $G(\lambda)$. Figure 3.1 shows a plot of $G(\lambda)$ when $\epsilon = 10^{-4}$, $\beta = 10^{-6}$ and $a_i = 1/2 - 0.03i$ for $i = 1$ to 49; see Appendix A for the R code which plots this graph. The rightmost pole is clearly identified, and the data giving this plot indicates that the pole is no greater than -0.359. This is certainly less than $1/6 - 1/2 = -1/3$, so we may conclude that

$$\lim_{t\to\infty} e^{-0.359t}g(t) = \lim_{\lambda\to 0}\lambda G\left(\lambda - 0.359\right) = 0.$$

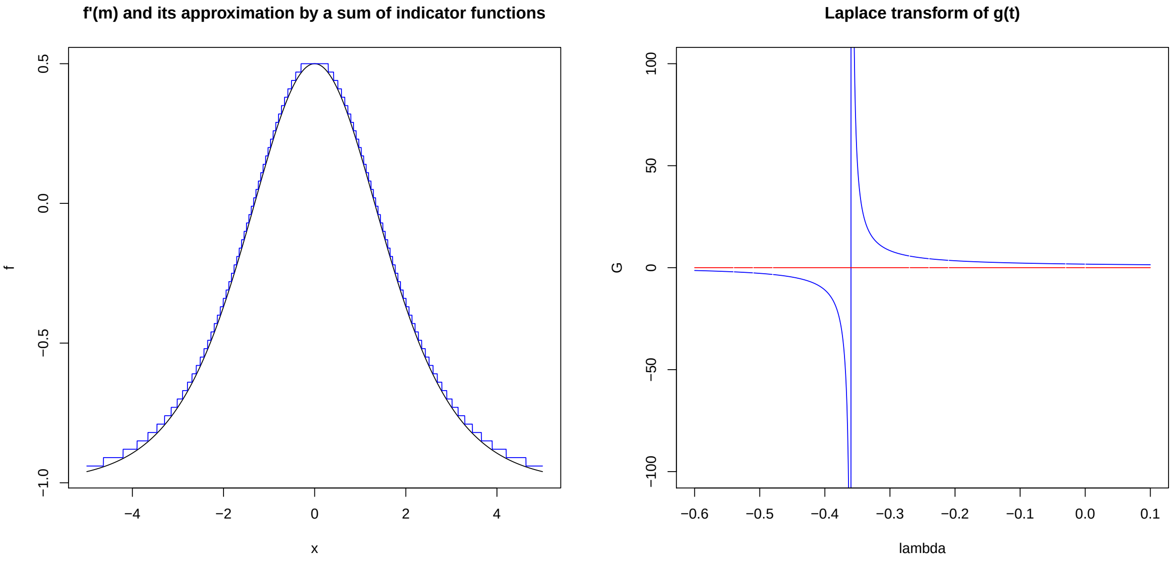

Figure 3.1: *Left*: Comparison of the approximation $k$ (in blue) to $f'(m)$ (in black) when $\epsilon = 10^{-4}$, $\beta = 10^{-6}$ and $a_j = 1/2 - 0.03j$ for $j = 1$ to 49. *Right*: A plot of $G(\lambda)$, for the same parameters. The real part is in blue, and the imaginary part is in red (i.e., it is a purely real function).

Thus, there exists a sufficiently large constant $C$ such that $Ce^{0.359t} \geq g(t)$. Recalling the definition of $g(t)$, and substituting this into (3.78) give us that

$$\begin{aligned} q_t(x,y,v) \leq& e^{t/2} g(t) p_t(x,y) \\ \leq& C e^{0.141t} p_t(x,y) \\ <& C e^{t/6} p_t(x,y). \end{aligned}$$

$\square$

We now have all we need to prove the main result of this section.

*Proof of Proposition 3.24.* In the notation of Proposition 3.24, as $m_\eta$ satisfies the PDE (3.2), by (3.60), for all $t \geq 0$, $\eta \in \mathbb{R}$ and $k = 1$ or 2, $\varphi_k(t, m_\eta) = 0$. Thus, by (3.58), $\Lambda(t, m_\eta) = 0$. Similarly, as $\{\Delta + f_0'(m_\eta)\}\nabla m_\eta = \nabla(\Delta(m_\eta) + f_0(m_\eta))$, by (3.61), for all $t \geq 0$ and $\eta, y \in \mathbb{R}$, $\psi(t, y, m_\eta) = 0$. Thus, by (3.59), $\Xi(t, y, m_\eta) = 0$. Applying (3.57), we see that

$$\begin{aligned} \|D\zeta(v) - D\zeta(m)\|_{L^1} \leq& \left\| e^{\int_0^\infty \Lambda(s,v)ds} D\eta(v) - e^{\int_0^\infty \Lambda(s,m)ds} D\eta(m) \right\|_{L^1} \\ &+ \left\| \int_0^\infty \Xi(s,v) e^{\int_s^\infty \Lambda(r,v)dr} ds - \int_0^\infty \Xi(s,m) e^{\int_s^\infty \Lambda(r,m)dr} ds \right\|_{L^1} \end{aligned}$$

$$\leq \left| e^{\int_0^\infty \Lambda(s,v)ds} - 1 \right| \|D\eta(v)\|_{L^1} + \|D\eta(v) - D\eta(m)\|_{L^1}$$
$$+ \left\| \int_0^\infty \Xi(s,v) e^{\int_s^\infty \Lambda(r,v)dr} ds \right\|_{L^1} \tag{3.83}$$

We note that $\|D\eta(v)\|_{L^1}$ is bounded, by applying (3.52) and noting that the denominator is strictly positive, as $v \geq 0$ and $\Delta m_{\eta(v)} > 0$, and the numerator is $\int_{-\infty}^{\infty} \left|\nabla m_{\eta(v)}(x)\right| dx = m(\infty) - m(-\infty) = 1$ by (3.3). We can verify that the integrand is nonnegative using the explicit formula for $m$.

Next, we note that $|1 - e^x| \leq |x|e^{|x|}$, so applying (3.67) gives,

$$\left| e^{\int_0^\infty \Lambda(s,v)ds} - 1 \right| \leq C\sqrt{\operatorname{dist}(v,M)} e^{C\sqrt{\operatorname{dist}(v,M)}} \leq C_1 \sqrt{\operatorname{dist}(v,M)}, \tag{3.84}$$

taking $\beta_2 \geq \operatorname{dist}(v,M)$ smaller than 1, if necessary, for the second inequality. We now turn to the second term in (3.83). First note that, by (3.52),

$$D\eta(v) - D\eta(m) = C_2\langle m, \Delta m\rangle \nabla m_{\eta(v)}(y) - C_2\langle v, \Delta m_{\eta(v)}\rangle \nabla m(y)$$

where $C_2 = (\langle v, \Delta m_{\eta(v)}\rangle \langle m, \Delta m\rangle)^{-1}$. Again, the denominator of $C_2$ is strictly positive, as $v \geq 0$ and $\Delta m_{\eta(v)} > 0$. Thus, we may write

$$\begin{aligned}
\|D\eta(v) - D\eta(m)\|_{L^1} \leq & C_2\langle m, \Delta m\rangle \left\|\nabla m_{\eta(v)} - \nabla m\right\|_{L^1} \\
& + C_2 \|\nabla m(y)\|_{L^1} \left|\langle m, \Delta m\rangle - \langle v, \Delta m_{\eta(v)}\rangle\right| \\
\leq & C_2\langle m, \Delta m\rangle \\
& \left| \int_{-\infty}^{\eta(v)/2} \nabla m_{\eta(v)}(x) - \nabla m(x) dx + \int_{\eta(v)/2}^{\infty} \nabla m(x) - \nabla m_{\eta(v)}(x) dx \right| \\
& + C_2 \|\nabla m(y)\|_{L^1} \left(|\langle \Delta m, m - m_{\eta(v)}\rangle| + |\langle m - m_{\eta(v)}, \Delta m_{\eta(v)}\rangle| \right. \\
& \left. + |\langle m_{\eta(v)} - v, \Delta m_{\eta(v)}\rangle|\right) \\
\leq & 2C_2\langle m, \Delta m\rangle \left|m(-\eta(v)/2) - m(\eta(v)/2)\right| \\
& + C_2 \|\nabla m(y)\|_{L^1} \Big(|\eta(v)| \, \|\nabla m\|_{L^\infty} \left(\|\Delta m\|_{L^1} + \left\|\Delta m_{\eta(v)}\right\|_{L^1}\right) \\
& + \left\|m_{\eta(v)} - v\right\|_{L^2} \left\|\Delta m_{\eta(v)}\right\|_{L^2}\Big) \\
\leq & C_3|\eta(v)| + C_3 \|S(v)\|_{L^2},
\end{aligned} \tag{3.85}$$

where we have exploited the symmetry of $\nabla m$ and the fact that $\nabla m$ has a single maximum at 0 for the first term of the second inequality, and integration by parts has been used in the second term of the same inequality to shift the Laplacian from one factor in the first integrand to the other. The mean value theorem and the Cauchy-Schwarz inequality has been used in the third inequality. We can use the explicit expression for $m$ to verify that $\|\nabla m(y)\|_{L^1}$, $\|\nabla m\|_{L^\infty}$, $\|\Delta m\|_{L^2}$ and $\langle m, \Delta m\rangle$ are all finite, as well as applying the mean value theorem again to the first term of the third inequality, which gives the final inequality, as well as recalling the definition of $S(v)$.

Note that by the assumption $\zeta(v) = 0$,

$$|\eta(v)| = |\eta(v) - \zeta(v)| = \left|\int_0^\infty \frac{d\eta_t}{dt}dt\right| \le C\sqrt{\operatorname{dist}(v, M)} \tag{3.86}$$

by (3.68). Substituting this back into (3.85) and applying (3.63) gives

$$\|D\eta(v) - D\eta(m)\|_{L^1} \le C_4\sqrt{\operatorname{dist}(v, M)}. \tag{3.87}$$

We consider the final term in (3.83).

$$\begin{aligned}\left\|\int_0^\infty \Xi(s,y,v)e^{\int_s^\infty \Lambda(r,v)dr}ds\right\|_{L^1} &\le \int_{-\infty}^\infty\int_0^\infty |\Xi(s,y,v)|\left(\left|e^{\int_s^\infty \Lambda(r,v)dr} - 1\right| + 1\right)dsdy \\ &\le \left(C_1\sqrt{\operatorname{dist}(v, M)} + 1\right)\int_{-\infty}^\infty\int_0^\infty |\Xi(s,y,v)|\,dsdy \end{aligned}\tag{3.88}$$

by similar reasoning which gave (3.84). Using the definition of $\Xi(s, y, v)$ in (3.59), and by (3.64), (3.65) and (3.66), we see that

$$\begin{aligned}|\Xi(s,y,v)| \le& \frac{|\psi(t,y,v)\tilde{\varphi}_1(t,v)| + |\tilde{\psi}(t,y,v)\varphi_1(t,v)|}{(\tilde{\varphi}_1(t,v))^2} \\ \le& \frac{C}{c^2}\left(|\psi(t,y,v)| + \sqrt{\operatorname{dist}(v, M)}e^{-\bar{c}t}|\tilde{\psi}(t,y,v)|\right)\end{aligned}$$

Substituting this back into (3.88), and again assuming $\beta_2 < 1$, gives

$$\begin{aligned}\left\|\int_0^\infty \Xi(s,y,v)e^{\int_s^\infty \Lambda(r,v)dr}ds\right\|_{L^1} \le& C_5\int_{-\infty}^\infty\int_0^\infty |\psi(s,y,v)|\,dsdy \\ &+ C_5\sqrt{\operatorname{dist}(v, M)}\int_{-\infty}^\infty\int_0^\infty |e^{-\bar{c}t}\tilde{\psi}(s,y,v)|dsdy. \end{aligned}\tag{3.89}$$

We shall consider these integrals separately. By definition of $\psi(s,y,v)$ in (3.61) and Lemma 3.30,

$$\begin{aligned}\int_{-\infty}^{\infty}\int_{0}^{\infty}|\psi(s,y,v)|\,dsdy =&\int_{-\infty}^{\infty}\int_{0}^{\infty}\int_{-\infty}^{\infty} q_t(y-x)\left|\nabla^3 m_{\eta_t}(x)+f_0'(v_t(x))\nabla m_{\eta_t}(x)\right| dxdsdy\\ \leq& C\int_{-\infty}^{\infty}\int_{0}^{\infty}\int_{-\infty}^{\infty} e^{ct}p_t(y-x)\left|\nabla^3 m_{\eta_t}(x)+f_0'(v_t(x))\nabla m_{\eta_t}(x)\right| dxdsdy\end{aligned}$$

for some $c<1/6$. As the integrand is positive, we may switch the order of integration. Integrating with respect to $y$ first allows us to integrate out the heat kernel. Recall also, that by definition of $m$, $\nabla^3 m_{\eta_t} = -\nabla(f_0(m_{\eta_t})) = -f_0'(m_{\eta_t})\nabla m_{\eta_t}$. Thus, we have

$$\begin{aligned}\int_{-\infty}^{\infty}\int_{0}^{\infty}|\psi(s,y,v)|\,dsdy \leq& \int_{0}^{\infty}\int_{-\infty}^{\infty} e^{ct}\left|f_0'(m_{\eta_t}(x))-f_0'(v_t(x))\right||\nabla m_{\eta_t}(x)|dxds\\ \leq& \int_{0}^{\infty}\int_{-\infty}^{\infty} e^{ct}\|f_0''\|_{[0,1]}\,|S(v_t(x))|\,|\nabla m_{\eta_t}(x)|dxds\\ \leq& \|f_0''\|_{[0,1]}\int_{0}^{\infty} e^{ct}\|S(v_t)\|_{L^2}\|\nabla m_{\eta_t}\|_{L^2}ds\\ \leq& \|f_0''\|_{[0,1]}\int_{0}^{\infty} C\sqrt{\operatorname{dist}(v,M)}e^{(c-\bar{c})t}\|\nabla m\|_{L^2}ds\\ =& C_6\sqrt{\operatorname{dist}(v,M)},\end{aligned} \tag{3.90}$$

where the mean value theorem gives the first inequality, Cauchy-Schwarz gives the second inequality, the fact that the $L^2$ norm is not affected by translations and (3.63) gives the third inequality, and the final line performs the simple integral, recalling that $c-\bar{c}<0$, and collects the constants into one.

We now consider the second term in (3.89). By the definition of $\tilde{\psi}(s,y,v)$ in (3.62) and Lemma 3.30,

$$\begin{aligned}\int_{-\infty}^{\infty}\int_{0}^{\infty} e^{-\bar{c}t}\left|\tilde{\psi}(s,y,v)\right| dsdy =& \int_{-\infty}^{\infty}\int_{0}^{\infty}\int_{-\infty}^{\infty} e^{(c-\bar{c})t}q_t(y-x)\left|\Delta m_{\eta_t}(x)\right| dxdsdy\\ \leq& C\int_{-\infty}^{\infty}\int_{0}^{\infty}\int_{-\infty}^{\infty} e^{(c-\bar{c})t}p_t(y-x)\left|\Delta m_{\eta_t}(x)\right| dxdsdy.\end{aligned}$$

Again, as the integrand is positive, we may switch the order of integration, and integrating with respect to $y$ first allows us to integrate out the heat kernel. Again, the $L^1$ norm is not affected by translations, so we find

$$\int_{-\infty}^{\infty}\int_{0}^{\infty} e^{-\bar{c}t}\left|\tilde{\psi}(s,y,v)\right| dsdy \leq \int_{0}^{\infty}\int_{-\infty}^{\infty} e^{(c-\bar{c})t}\left|\Delta m_{\eta_t}(x)\right| dxds$$

$$=\int_0^\infty e^{(c-\bar{c})t}\left\|\Delta m_{\eta_t}(x)\right\|_{L^1} ds$$

$$=(\bar{c}-c)^{-1}\left\|\Delta m_{\eta_t}(x)\right\|_{L^1}, \tag{3.91}$$

which is a constant, again, recalling that $c-\bar{c}<0$. Substituting (3.90) and (3.91) into (3.89) gives

$$\left\|\int_0^\infty \Xi(s,y,v)e^{\int_s^\infty \Lambda(r,v)dr}ds\right\|_{L^1} \le C_7\sqrt{\operatorname{dist}(v,M)},$$

and substituting this, (3.87) and (3.84) into (3.83) gives the desired result. □

## 3.5 Proofs of Proposition 3.4 and Theorem 3.2

We note two more results from [20]. The first allows us to translate the Fréchet derivative of $\zeta$ in the natural way. The second result, however, is central to this section.

**Lemma 3.31** (Lemma 7.1 in [20]). *For $v\in L^2+m$ with dist$(v,M)\le\beta_2$ and any $y_1,y_2\in\mathbb{R}$,*

$$D\zeta(y_1,m_{\eta(v)})=D\zeta(y_1-\eta(v),m), \tag{3.92}$$

$$D^2\zeta(y_1,y_2,m_{\eta(v)})=D^2\zeta(y_1-\eta(v),y_2-\eta(v),m) \tag{3.93}$$

*and*

$$D^2\zeta(y_1,y_2,v(\cdot))=-D^2\zeta(-y_1,-y_2,-v(-\cdot)). \tag{3.94}$$

**Theorem 3.32** (Theorem 7.4 in [20]). *There exists $\beta_2>0$ such that for $v\in L^2+m$ with dist$(v,M)\le\beta_2$, $D\zeta(\cdot,v)\in\cap_{\delta>0}H^{2-\delta}$. If in addition $\|v\|_\infty\le 1$ and $v\in H^\delta+m$ for some $\delta>0$ then*

$$\langle D\zeta(\cdot,v),\Delta v+f_0(v)\rangle=0.$$

Let us now demonstrate the significance of this theorem. Let $\tau=\sigma\wedge\inf\{t\ge 0: v_t^{\epsilon,\lambda}\notin H^\alpha+m\}$. Let $\xi_t^\epsilon=\epsilon^{1/2}\zeta(v_{t\wedge\tau}^{\epsilon,\lambda})$. Then by Itô's formula,

$$\xi_t^\epsilon=\xi_0^\epsilon+\epsilon^{-1-2\gamma}\int_0^{t\wedge\tau}\langle D\zeta(v_s^{\epsilon,\lambda}),\Delta v_s^\epsilon+f_\lambda(v_s^{\epsilon,\lambda})\rangle ds$$

$$
\begin{aligned}
&+\int_0^{t\wedge\tau}\int_{-\infty}^{\infty} D\zeta(y, v_s^{\epsilon,\lambda}), \sqrt{v_s^{\epsilon,\lambda}(y)(1-v_s^{\epsilon,\lambda}(y))}W(dyds) \\
&+\tfrac{1}{2}\epsilon^{-1/2}\int_0^{t\wedge\tau}\int_{-\infty}^{\infty} D^2\zeta(y,y,v_s^{\epsilon,\lambda})v_t^{\epsilon,\lambda}(y)(1-v_t^{\epsilon,\lambda}(y))dyds. \qquad (3.95)
\end{aligned}
$$

We shall consider each term in (3.95) separately. First we note that for $t < \tau$, by the definition of $\sigma_2$,

$$
\operatorname{dist}(v_t^{\epsilon,\lambda}, M) \le \beta + \|v_t^{\epsilon,\lambda} - v_t^{\delta,\lambda}\|_{L^2}.
$$

By the definition of $v^{\delta,\lambda}$ and then by Jensen's inequality, for $t < \tau$,

$$
\begin{aligned}
\|v_t^{\epsilon,\lambda} - v_t^{\delta,\lambda}\|_2^2 =& \int_{-\infty}^{\infty}\left(\int_{-\infty}^{\infty}\rho^\delta(z)(v_t^{\epsilon,\lambda}(y) - v_t^{\epsilon,\lambda}(y-z))dz\right)^2 dy \\
\le& \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\rho^\delta(z)(v_t^{\epsilon,\lambda}(y) - v_t^{\epsilon,\lambda}(y-z))^2 dzdy \\
\le& \beta'
\end{aligned}
$$

by the definition of $\sigma_3$. Therefore, for $t < \tau$,

$$
\operatorname{dist}(v_t^{\epsilon,\lambda}, M) \le \beta + (\beta')^{1/2} \le 2\epsilon^{2\gamma/5-3/20-3\kappa}. \qquad (3.96)
$$

For the first integral on the right hand side of (3.95), note that for $\epsilon$ sufficiently small, for all $t < \tau$, $\operatorname{dist}(v_t^{\epsilon,\lambda}, M) \le \beta_2$ and $v_t^{\epsilon,\lambda} \in H^\alpha + m$. Hence by Theorem 3.32,

$$
\int_0^{t\wedge\tau}\langle D\zeta(v_s^{\epsilon,\lambda}), \Delta v_s^{\epsilon,\lambda} + f_0(v_s^{\epsilon,\lambda})\rangle ds = 0. \qquad (3.97)
$$

This crucial observation was made by Katzenberger in a similar situation in [26]. As $f_\lambda(v) = f_0(v) + \lambda v(1-v)$, first integral on the right hand side of (3.95) becomes,

$$
\int_0^{t\wedge\tau}\langle D\zeta(v_s^{\epsilon,\lambda}), \Delta v_s^{\epsilon,\lambda} + f_\lambda(v_s^{\epsilon,\lambda})\rangle ds = \lambda\int_0^{t\wedge\tau}\langle D\zeta(v_s^{\epsilon,\lambda}), v_s^{\epsilon,\lambda}(1-v_s^{\epsilon,\lambda})\rangle ds
$$

The next three lemmas investigate each of the three integrals on the right hand side of (3.95).

**Lemma 3.33.** *For $\epsilon$ sufficiently small, for all $t < \tau$,*

$$
\frac{1}{2}\epsilon^{-1/2}\left|\int_{-\infty}^{\infty} D^2\zeta(y,y,v_t^{\epsilon,\lambda})v_t^{\epsilon,\lambda}(y)(1-v_t^{\epsilon,\lambda}(y))dy\right| \le \epsilon^{\gamma/5-13/40-2\kappa}.
$$

*Proof.* For $t < \tau$,

$$\begin{aligned}
&\bigg|\int_{-\infty}^{\infty} D^2\zeta(y,y,v_t^{\epsilon,\lambda})v_t^{\epsilon,\lambda}(y)(1-v_t^{\epsilon,\lambda}(y))dy \\
&\qquad\qquad\qquad - \int_{-\infty}^{\infty} D^2\zeta\left(y,y,m_{\eta(v_t^{\epsilon,\lambda})}\right) m_{\eta(v_t^{\epsilon,\lambda})}(y)(1-m_{\eta(v_t^{\epsilon,\lambda})}(y))dy\bigg| \\
&\leq \int_{-\infty}^{\infty} |D^2\zeta(y,y,v_t^{\epsilon,\lambda}) - D^2\zeta(y,y,m_{\eta(v_t^{\epsilon,\lambda})})|dy \\
&\quad + \int_{-\infty}^{\infty} |D^2\zeta(y,y,m_{\eta(v_t^{\epsilon,\lambda})})||v_t^{\epsilon,\lambda}(y)(1-v_t^{\epsilon,\lambda}(y)) - m_{\eta(v_t^{\epsilon,\lambda})}(y)(1-m_{\eta(v_t^{\epsilon,\lambda})}(y))|dy \\
&\leq \int_{-\infty}^{\infty} |D^2\zeta(y,y,v_t^{\epsilon}) - D^2\zeta(y,y,m_{\eta(v_t^{\epsilon})})|dy + \int_{-\infty}^{\infty} |D^2\zeta(y,y,m_{\eta(v_t^{\epsilon,\lambda})})||v_t^{\epsilon}(y) - m_{\eta(v_t^{\epsilon,\lambda})}(y)|dy
\end{aligned}$$

since $|\frac{d}{dx}(x(1-x))| \leq 1$ for $x \in [0,1]$. By (3.93) and (3.55)

$$\left(\int_{-\infty}^{\infty} |D^2\zeta(y,y,m_{\eta(v_t^{\epsilon,\lambda})})|^2 dy\right)^{1/2} = \left(\int_{-\infty}^{\infty} |D^2\zeta(y,y,m)|^2 dy\right)^{1/2} := C' < \infty.$$

Therefore for $t < \tau$, by Cauchy's inequality and (3.54),

$$\begin{aligned}
&\bigg|\int_{-\infty}^{\infty} D^2\zeta(y,y,v_t^{\epsilon,\lambda})v_t^{\epsilon,\lambda}(y)(1-v_t^{\epsilon,\lambda}(y))dy \\
&\qquad\qquad\qquad - \int_{-\infty}^{\infty} D^2\zeta\left(y,y,m_{\eta(v_t^{\epsilon,\lambda})}\right) m_{\eta(v_t^{\epsilon,\lambda})}(y)(1-m_{\eta(v_t^{\epsilon,\lambda})}(y))dy\bigg| \\
&\leq C\sqrt{\operatorname{dist}(v_t^{\epsilon,\lambda},M)} + \left(\int_{-\infty}^{\infty} |D^2\zeta(y,y,m_{\eta(v_t^{\epsilon,\lambda})})|^2 dy\right)^{1/2} \left(\int_{-\infty}^{\infty} |v_t^{\epsilon,\lambda}(y) - m_{\eta(v_t^{\epsilon,\lambda})}(y)|^2 dy\right)^{1/2} \\
&\leq (C+C')\sqrt{\operatorname{dist}(v_t^{\epsilon,\lambda},M)} \\
&\leq 2(C+C')\epsilon^{\gamma/5-3/40-3\kappa/2},
\end{aligned}$$

where the last line follows by (3.96).

As $m(y)(1-m(y)) = m(-y)(1-m(-y))$ for all $y \in \mathbb{R}$, by (3.94)

$$\int_{-\infty}^{\infty} D^2\zeta(y,y,m)m(y)(1-m(y))dy = 0.$$

Then by (3.93)

$$\begin{aligned}
\int_{-\infty}^{\infty} D^2\zeta(y,y,m_{\eta(v_t^{\epsilon,\lambda})})m_{\eta(v_t^{\epsilon,\lambda})}(y)(1-m_{\eta(v_t^{\epsilon,\lambda})}(y))dy =& \int_{-\infty}^{\infty} D^2\zeta(y,y,m)m(y)(1-m(y))dy \\
=& 0.
\end{aligned}$$

The result follows. $\square$

Let $M_t = \int_0^{t\wedge\tau}\int_{-\infty}^{\infty} D\zeta(y, v_s^{\epsilon,\lambda}), \sqrt{v_s^{\epsilon,\lambda}(y)(1-v_s^{\epsilon,\lambda}(y))}W(dyds)$. Then $M$ is a martingale with quadratic variation given by

$$[M]_t = \int_0^{t\wedge\tau} \langle D\zeta(v_s^{\epsilon,\lambda})^2, v_s^{\epsilon,\lambda}(1-v_s^{\epsilon,\lambda})\rangle ds.$$

**Lemma 3.34.** *For $\epsilon$ sufficiently small, for $t<\tau$,*

$$\left|\langle D\zeta(v_t^{\epsilon,\lambda})^2, v_t^{\epsilon,\lambda}(1-v_t^{\epsilon,\lambda})\rangle - \frac{1}{\|\nabla m\|_{L^2}^4}\langle |\nabla m|^2, m(1-m)\rangle\right| \le \epsilon^{\gamma/5-3/40-3\kappa/2}.$$

*Proof.* For $t<\tau$,

$$\begin{aligned}
&\left|\langle D\zeta(v_t^{\epsilon,\lambda})^2, v_t^{\epsilon,\lambda}(1-v_t^{\epsilon,\lambda})\rangle - \langle D\zeta(m_{\eta(v_t^{\epsilon,\lambda})})^2, m_{\eta(v_t^{\epsilon,\lambda})}(1-m_{\eta(v_t^{\epsilon,\lambda})})\rangle\right| \\
&\qquad\le \left|\langle D\zeta(v_t^{\epsilon,\lambda})^2 - D\zeta(m_{\eta(v_t^{\epsilon,\lambda})})^2, v_t^{\epsilon,\lambda}(1-v_t^{\epsilon,\lambda})\rangle\right| \\
&\qquad\quad + \left|\langle D\zeta(m_{\eta(v_t^{\epsilon,\lambda})})^2, v_t^{\epsilon,\lambda}(1-v_t^{\epsilon,\lambda}) - m_{\eta(v_t^{\epsilon,\lambda})}(1-m_{\eta(v_t^{\epsilon,\lambda})})\rangle\right| \\
&\qquad\le \left|\langle D\zeta(v_t^{\epsilon,\lambda}) - D\zeta(m_{\eta(v_t^{\epsilon,\lambda})}), (D\zeta(v_t^{\epsilon,\lambda}) + D\zeta(m_{\eta(v_t^{\epsilon,\lambda})}))v_t^{\epsilon,\lambda}(1-v_t^{\epsilon,\lambda})\rangle\right| \\
&\qquad\quad + \langle D\zeta(m_{\eta(v_t^{\epsilon,\lambda})})^2, |v_t^{\epsilon,\lambda} - m_{\eta(v_t^{\epsilon,\lambda})}|\rangle \\
&\qquad\le \|D\zeta(v_t^{\epsilon,\lambda}) - D\zeta(m_{\eta(v_t^{\epsilon,\lambda})})\|_{L^2}\|(D\zeta(v_t^{\epsilon,\lambda}) + D\zeta(m_{\eta(v_t^{\epsilon,\lambda})}))\|_{L^2} \\
&\qquad\quad + \|D\zeta(m_{\eta(v_t^{\epsilon,\lambda})})^2\|_{L^2}\|v_t^{\epsilon,\lambda} - m_{\eta(v_t^{\epsilon,\lambda})}\|_{L^2},
\end{aligned}$$

where the second inequality holds since $|\frac{d}{dx}(x(1-x))| \le 1$ for $x\in[0,1]$ and the third inequality follows by Cauchy's inequality. By (3.51) and (3.96), for $\epsilon$ sufficiently small,

$$\|D\zeta(v_t^{\epsilon,\lambda}) - D\zeta(m_{\eta(v_t^{\epsilon,\lambda})})\|_{L^2} \le C\sqrt{\operatorname{dist}(v_t^{\epsilon,\lambda}, M)} \le 2C\epsilon^{\gamma/5-3/40-3\kappa/2}.$$

Also, by (3.53), $\|D\zeta(m_{\eta(v_t^{\epsilon,\lambda})})\|_{L^2} = \|D\zeta(m)\|_{L^2} < \infty$ and $\|D\zeta(m_{\eta(v_t^{\epsilon,\lambda})})^2\|_{L^2} = \|D\zeta(m)^2\|_{L^2} < \infty$. Hence

$$\begin{aligned}
&\left|\langle D\zeta(v_t^{\epsilon,\lambda})^2, v_t^{\epsilon,\lambda}(1-v_t^{\epsilon,\lambda})\rangle - \langle D\zeta(m_{\eta(v_t^{\epsilon,\lambda})})^2, m_{\eta(v_t^{\epsilon,\lambda})}(1-m_{\eta(v_t^{\epsilon,\lambda})})\rangle\right| \\
&\qquad\le 2C\epsilon^{\gamma/5-3/40-3\kappa/2}(2C\epsilon^{\gamma/5-3/40-3\kappa/2} + \|D\zeta(m)\|_{L^2}) + \|D\zeta(m)^2\|_{L^2}2\epsilon^{2\gamma/5-3/20-3\kappa}
\end{aligned}$$

by (3.96). By (3.53),

$$\langle D\zeta(m_{\eta(v_t^{\epsilon,\lambda})})^2, m_{\eta(v_t^{\epsilon,\lambda})}(1-m_{\eta(v_t^{\epsilon,\lambda})})\rangle = \frac{1}{\|\nabla m\|_{L^2}^4}\langle |\nabla m|^2, m(1-m)\rangle.$$

The result follows since $\gamma > 3/8$. □

**Lemma 3.35.** *For $\epsilon$ sufficiently small, for $t < \tau$,*

$$\left|\langle D\zeta(v_t^{\epsilon,\lambda}), v_t^{\epsilon,\lambda}(1-v_t^{\epsilon,\lambda})\rangle - 1\right| \leq \epsilon^{\gamma/5-3/40-3\kappa/2}.$$

*Proof.* For $t < \tau$,

$$\begin{aligned}
&\left|\langle D\zeta(v_t^{\epsilon,\lambda}), v_t^{\epsilon,\lambda}(1-v_t^{\epsilon,\lambda})\rangle - \langle D\zeta(m_{\eta(v_t^{\epsilon,\lambda})}), m_{\eta(v_t^{\epsilon,\lambda})}(1-m_{\eta(v_t^{\epsilon,\lambda})})\rangle\right| \\
&\qquad\leq \left|\langle D\zeta(v_t^{\epsilon,\lambda}) - D\zeta(m_{\eta(v_t^{\epsilon,\lambda})}), v_t^{\epsilon,\lambda}(1-v_t^{\epsilon,\lambda})\rangle\right| \\
&\qquad\quad + \left|\langle D\zeta(m_{\eta(v_t^{\epsilon,\lambda})}), v_t^{\epsilon,\lambda}(1-v_t^{\epsilon,\lambda}) - m_{\eta(v_t^{\epsilon,\lambda})}(1-m_{\eta(v_t^{\epsilon,\lambda})})\rangle\right| \\
&\qquad\leq \left\langle \left|D\zeta(v_t^{\epsilon,\lambda}) - D\zeta(m_{\eta(v_t^{\epsilon,\lambda})})\right|, 1\right\rangle + \langle D\zeta(m_{\eta(v_t^{\epsilon,\lambda})}), |v_t^{\epsilon,\lambda} - m_{\eta(v_t^{\epsilon,\lambda})}|\rangle \\
&\qquad\leq \|D\zeta(v_t^{\epsilon,\lambda}) - D\zeta(m_{\eta(v_t^{\epsilon,\lambda})})\|_{L^1} + \|D\zeta(m_{\eta(v_t^{\epsilon,\lambda})})\|_{L^2}\|v_t^{\epsilon,\lambda} - m_{\eta(v_t^{\epsilon,\lambda})}\|_{L^2},
\end{aligned}$$

where the second inequality holds since $|\frac{d}{dx}(x(1-x))| \leq 1$ for $x \in [0,1]$ and the third inequality follows by Cauchy's inequality. By (3.56) and (3.92),

$$\|D\zeta(v_t^{\epsilon,\lambda}) - D\zeta(m_{\eta(v_t^{\epsilon,\lambda})})\|_{L^1} \leq C\sqrt{\operatorname{dist}(v_t^{\epsilon,\lambda}, M)} \leq 2C\epsilon^{\gamma/5-3/40-3\kappa/2}.$$

Also, by (3.53), $\|D\zeta(m_{\eta(v_t^{\epsilon,\lambda})})\|_{L^2} = \|D\zeta(m)\|_{L^2} < \infty$. Hence

$$\begin{aligned}
&\left|\langle D\zeta(v_t^{\epsilon,\lambda}), v_t^{\epsilon,\lambda}(1-v_t^{\epsilon,\lambda})\rangle - \langle D\zeta(m_{\eta(v_t^{\epsilon,\lambda})}), m_{\eta(v_t^{\epsilon,\lambda})}(1-m_{\eta(v_t^{\epsilon,\lambda})})\rangle\right| \\
&\qquad\qquad\leq 2C\epsilon^{\gamma/5-3/40-3\kappa/2}(1+\|D\zeta(m)\|_{L^2})
\end{aligned}$$

by (3.96). By (3.53),

$$\langle D\zeta(m_{\eta(v_t^{\epsilon,\lambda})}), m_{\eta(v_t^{\epsilon,\lambda})}(1-m_{\eta(v_t^{\epsilon,\lambda})})\rangle = \frac{1}{\|\nabla m\|_{L^2}^2}\langle \nabla m, m(1-m)\rangle. \tag{3.98}$$

Let $g(x) = x(1-x)$, and note that by (3.3) and the chain rule,

$$\Delta m = -f_0(m) = m(1-m)(1-2m) = g(m)g'(m) = \nabla g(m).$$

Integrating gives $\nabla m = m(1-m)$, and substituting this into (3.98) allows us to calculate

$$\langle D\zeta(m_{\eta(v_t^{\epsilon,\lambda})}), m_{\eta(v_t^{\epsilon,\lambda})}(1-m_{\eta(v_t^{\epsilon,\lambda})})\rangle = \frac{\langle \nabla m, \nabla m\rangle}{\|\nabla m\|_{L^2}^2} = 1.$$

The result follows since $\gamma > 3/8$. □

*Proof of Proposition 3.4.* Note that $\xi_0^\epsilon = \epsilon^{1/2}\zeta(v_0)$. By (9.46) in [20], there exists a constant $C$ such that if $\operatorname{dist}(v, M) \leq \beta_1$ then

$$|\eta(v) - \zeta(v)| \leq C\sqrt{\operatorname{dist}(v, M)}. \tag{3.99}$$

By (3.8), for $\epsilon$ sufficiently small we have that $\|v_0 - m\|_{L^2} \leq \beta_1$. Hence by Lemma 3.10,

$$|\zeta(v_0)| \leq |\zeta(v_0) - \eta(v_0)| + |\eta(v_0)| \leq 2C\beta_1^{1/2}$$

Let $a = \frac{1}{\|\nabla m\|_{L^2}^4}\langle |\nabla m|^2, m(1-m)\rangle$. By (3.95),

$$\begin{aligned}
\mathbb{E}(\xi_t^\epsilon)^2 \leq& 4(\xi_0^\epsilon)^2 + 4\mathbb{E}M_t^2 + 4\mathbb{E}\bigg(\epsilon^{-1-2\gamma}\lambda\int_0^{t\wedge\tau}\langle D\zeta(v_s^{\epsilon,\lambda}), v_s^{\epsilon,\lambda}(1-v_s^{\epsilon,\lambda})\rangle ds\bigg)^2 \\
&+ 4\mathbb{E}\bigg(\tfrac{1}{2}\epsilon^{-1/2}\int_0^{t\wedge\tau}\int_{-\infty}^{\infty} D^2\zeta(y,y,v_s^\epsilon)v_t(y)(1-v_t(y))dyds\bigg)^2 \\
\leq& 8C^2\beta_1\epsilon + 4\mathbb{E}[M]_t + 4\lambda^2t^2\epsilon^{-2-4\gamma}\left(1+\epsilon^{\gamma/5-3/40-3\kappa/2}\right)^2 + 4t^2\epsilon^{2\gamma/5-13/20-4\kappa} \\
\leq& 8C^2\beta_1\epsilon + 4t\left(\epsilon^{\gamma/5-3/40-3\kappa/2}+a\right) + 4\nu^2t^2\epsilon^{2\varrho-2-4\gamma}\left(1+\epsilon^{\gamma/5-3/40-3\kappa/2}\right)^2 \\
&+ 4t^2\epsilon^{2\gamma/5-13/20-4\kappa} \\
\leq& 5tA
\end{aligned} \tag{3.100}$$

for

$$A = \begin{cases} a & \text{if } \varrho > 1+2\gamma \\ a+\nu & \text{if } \varrho = 1+2\gamma, \end{cases}$$

where the second inequality follows by Lemma 3.33 and Lemma 3.35 and the third inequality by Lemma 3.34 and our assumption on $\lambda$, and the last line holds for $\epsilon$ sufficiently small, since $\gamma > 13/8$.

By Lemma 3.3, for $t < \infty$ fixed, there exist a constant $K$ and a random variable $X_0$ such that $X_0 < \infty$ a.s. and for any $s \leq t+1$, $v_s(x) \leq e^{-Kx^2}$ for $x \leq -X_0$ and $1 - v_s(x) \leq e^{-Kx^2}$ for $x \geq X_0$. As at the start of the proof of Lemma 3.11, on $\{X_0 < \infty\}$, if $t_n \leq t+1$ for all $n$ and $t_n \to t_0$ as $n \to \infty$ then $\|v^\delta_{t_n} - v^\delta_{t_0}\|_{L^2} \to 0$ and $\|v_{t_n} - v_{t_0}\|_{L^2} \to 0$ as $n \to \infty$.

On $\{X_0 < \infty\} \cap \{\tau = \sigma_4 \leq t\} \cap \{\sigma_4 < \sigma_2\}$, there exists $(t_n)_n$ such that $t_n < \min(t+1, \sigma_2)$ for all $n$, $t_n \to \sigma_4$ as $n \to \infty$ and $|\eta(v^\delta_{t_n})| \geq \epsilon^{-1/2-\kappa}$ for all $n$. Then $\text{dist}(v^\delta_{t_n}, M) \leq \beta$ for all $n$ and $\text{dist}(v^\delta_{\sigma_4}, M) \leq \beta$ by the definition of $\sigma_2$. Hence for $\epsilon$ sufficiently small, by Lemma 3.10,

$$|\eta(v^\delta_{t_n}) - \eta(v^\delta_{\sigma_4})| \leq C\|v^\delta_{t_n} - v^\delta_{\sigma_4}\|_{L^2} \to 0$$

as $n \to \infty$. Therefore $|\eta(v^\delta_{\sigma_4})| \geq \epsilon^{-1/2-\kappa}$. By (3.25), $\|v_s - v^\delta_s\|^2_{L^2} \leq \beta'$ for $s < \sigma_3$ and therefore $\|v_{\sigma_4} - v^\delta_{\sigma_4}\|^2_{L^2} \leq \beta'$. For $\epsilon$ sufficiently small, it follows by Lemma 3.10 that $|\eta(v_{\sigma_4})| \geq \frac{1}{2}\epsilon^{-1/2-\kappa}$. Then by (3.99), for $\epsilon$ sufficiently small, it follows that $|\zeta(v_{\sigma_4})| \geq \frac{1}{4}\epsilon^{-1/2-\kappa}$. Therefore, since $X_0 < \infty$ a.s.,

$$\mathbb{P}\left(\tau = \sigma_4 \leq t, \sigma_4 < \sigma_2, |\xi^\epsilon_{\sigma_4}| < \tfrac{1}{4}\epsilon^{-\kappa}\right) = 0.$$

We now have

$$\begin{aligned}\mathbb{E}(\xi^\epsilon_t)^2 \geq& \mathbb{E}(\xi^\epsilon_t)^2 \mathbb{1}_{\tau=\sigma_4\leq t, \sigma_4<\sigma_2, |\xi^\epsilon_{\sigma_4}|\geq \frac{1}{4}\epsilon^{-\kappa}} \\ \geq& \tfrac{1}{16}\epsilon^{-2\kappa}\mathbb{P}\left(\tau = \sigma_4 \leq t, \sigma_4 < \sigma_2\right).\end{aligned}$$

Therefore by (3.100), $\mathbb{P}\left(\tau = \sigma_4 \leq t, \sigma_4 < \sigma_2\right) \to 0$ as $\epsilon \to 0$. Now

$$\mathbb{P}\left(\sigma_4 \leq \min(\sigma_1, \sigma_2, \sigma_3, t), \sigma_4 < \sigma_2\right) \leq \mathbb{P}\left(\tau = \sigma_4 \leq t, \sigma_4 < \sigma_2\right) + \mathbb{P}\left(\exists s \leq t \wedge \sigma : v^\epsilon_s \notin H^\alpha + m\right)$$

which goes to 0 as $\epsilon \to 0$ by Proposition 3.5.

We can now complete the proof. We have

$$\mathbb{P}\left(\sigma \leq t\right) \leq \mathbb{P}\left(\sigma_1 \leq \min(\sigma_2, \sigma_3, \sigma_4, t)\right) + \mathbb{P}\left(\sigma_2 \leq \min(\sigma_1, \sigma_3, \sigma_4, t)\right)$$

$$+ \mathbb{P}\left(\sigma_3 \le \min(\sigma_1, \sigma_2, \sigma_4, t)\right) + \mathbb{P}\left(\sigma_4 \le \min(\sigma_1, \sigma_2, \sigma_3, t), \sigma_4 < \sigma_2\right).$$

The result then follows by Propositions 3.8, 3.16 and 3.18. ☐

*Proof of Theorem 3.2.* The result follows from Propositions 3.4 and 3.5, and Lemmas 3.33 and 3.34 in the same way as the proof of Theorem 8.1 in [20]. We give a brief sketch.

First, notice that

$$\begin{aligned}
\sup_{\substack{v\in L^2+m\\ \mathrm{dist}(v,M)\le\beta_2}} \left\|v - m_{\zeta(v)}\right\|_{L^2} \le & \sup_{\substack{v\in L^2+m\\ \mathrm{dist}(v,M)\le\beta_2}} \|S(v)\|_{L^2} + \left\|m_{\eta(v)} - m_{\zeta(v)}\right\|_{L^2} \\
\le & \sup_{\substack{v\in L^2+m\\ \mathrm{dist}(v,M)\le\beta_2}} \beta_2 + |\eta(v) - \zeta(v)| \, \|\nabla m\|_{L^2} \\
\le & \sup_{\substack{v\in L^2+m\\ \mathrm{dist}(v,M)\le\beta_2}} \beta_2 + C\sqrt{\mathrm{dist}(v,M)} \, \|\nabla m\|_{L^2} \\
< & \infty
\end{aligned}$$

applying (3.63) and using the mean value theorem in the first inequality, and applying (3.86) in the third inequality. Hence,

$$\begin{aligned}
\left\|\bar{u}^\epsilon(t,\cdot) - \chi_{\xi_t^\epsilon}\right\|_{L^2} \le & \left\|\bar{u}^\epsilon(t,\cdot) - m(\epsilon^{-1/2}(\cdot - \xi_t^\epsilon))\right\|_{L^2} + \left\|m(\epsilon^{-1/2}(\cdot - \xi_t^\epsilon)) - \chi_{\xi_t^\epsilon}\right\|_{L^2} \\
\le & \epsilon^{1/4}\left\|v^{\epsilon,\lambda} - m_{\zeta(v^{\epsilon,\lambda})}\right\|_{L^2} + \epsilon^{1/4}\left\|m - \chi_0\right\|_{L^2} \\
\le & C\epsilon^{1/4},
\end{aligned}$$

for $0 < t < \sigma$. Thus, by Proposition 3.4, the first claim of the theorem holds.

Recall that we set $a = \frac{1}{\|\nabla m\|_{L^2}^4}\langle|\nabla m|^2, m(1-m)\rangle$. Applying Proposition 3.5 to (3.95) allows us to replace the minimum of $t$ and $\tau$ in this expression with simply $t$. Then, taking $\epsilon \to 0$ in and applying Lemmas 3.33, 3.34 and 3.35 yields

$$d\xi_t = \nu \mathbb{1}_{\{\varrho = 1+2\gamma\}} dt + a dW_s$$

where $\xi_t := \lim_{\epsilon\to 0} \xi_t^\epsilon$. The second claim of the theorem follows from this observation and standard arguments. ☐

# Chapter 4

# Asymptotics of the planar branching and annihilating random walk

In this chapter, we turn our attention to the branching annihilating random walk (BARW). The aim of this chapter is to prove Theorem 4.1, which, informally, gives a way to construct a ternary branching Brownian motion (TBBM) from a BARW for an arbitrarily long time, with an arbitrarily high probability, if the branching rate of the BARW is sufficiently small. This alone is a useful step in the proof of Conjecture 1.10, as the number of particles in a branching Brownian motion does not decrease, let alone go extinct. However, we shall see that the coupling actually gets better the smaller the branching rate is, which suggests further that survival should be possible for arbitrarily small branching rates. We discuss how one might go from this result to a proof of the second part of Conjecture 1.10 in the final chapter of this thesis, Chapter 5.

For simplicity, we shall only consider the case where the BARW starts from a single initial particle, but these arguments can be straightforwardly extended to a BARW starting with any finite number of particles.

Before we can state the theorem precisely, we need some notation. Let $W$ be a TBBM and $\xi$ be a BARW. We shall label the particles in the TBBM and the BARW with the set of Ulam-Harris labels, $\mathcal{U} = \{\emptyset\} \cup \bigcup_{k=1}^{\infty} \{1, 2, 3\}^k$, and we denote the set of particles existing in the TBBM $W$ (resp. BARW $\xi$) at time $t$ by $\mathcal{U}_t(W)$ (resp. $\mathcal{U}_t(\xi)$). Furthermore, if $l, l' \in \mathcal{U}$, we shall use the notation that $l \preceq l'$ if $l$ is an ancestor of $l'$, i.e., $l' = (l, l'')$ for some $l'' \in \mathcal{U}$.

The following event will quantify how we measure the similarity between a BARW and a

TBBM. We denote

$$\mathcal{D}_x^\epsilon(T,\sigma,\kappa) = \Bigg\{ \exists \text{ a bijection } f : \mathcal{U}_T(\xi) \to \mathcal{U}_T(W) \text{ such that } \forall l \in \mathcal{U}, f(l) \preceq l \\ \text{and} \sup_{t\in[0,T]} |\xi_t^l - W_t^l| \le \epsilon \Bigg\}, \tag{4.1}$$

where $\xi$ is a BARW with branching rate $\sigma_n$ started from a single initial particle at $x$, and $W$ is a TBBM with branching rate $\kappa$ also started from a single initial particle at $x$. This definition may look a bit strange, as the BARW is being compared directly to the TBBM. However, these events are defined in terms of the processes before we do our rescaling. Our scaling parameter will be $n$, and we shall be applying a standard diffusive scaling. We do not scale the annihilation rate, which we fix to be 1, but we do scale the branching rate; let $\sigma_n = \frac{\sigma(\log n)^3}{n}$. We will scale the parameter $\epsilon$ from (4.1) in such a way that it *grows* with $n$, but slower than $\sqrt{n}$, so that it becomes small after diffusive rescaling.

We can now state the theorem we wish to prove in this chapter.

**Theorem 4.1.** *Fix $T > 0$ and $x \in \mathbb{Z}^2$. For $n \in \mathbb{N}$ define $\sigma_n = \frac{(\log n)^3}{n}$ and $\epsilon = 4\sqrt{n}\frac{\log\log n}{\log n}$. Then there exists a constant $K$ and a sequence $\kappa_n$ such that $n\kappa_n \to K$ with the following property.*

*There exists a sequence of couplings of BARWs, $\xi^{(n)}$, with branching rate $\sigma_n$ started from a single initial particle at $x$, and TBBMs, $W^{(n)}$, with branching rate $\kappa_n$ also started from a single initial particle at $x$, such that, as $n \to \infty$ the event $\mathcal{D}_x^\epsilon(nT, \sigma_n, \kappa_n)$ occurs with probability $1 - \mathcal{O}\left((\log n)^{-1/4}\right)$.*

This theorem says that, after scaling time by $n$ and space by $\sqrt{n}$, the BARW and the TBBM are close, up until time $T$ (in rescaled time). As $n\kappa_n \to K$, the branching rate of the TBBM converges to a constant in rescaled time. The branching rate of the BARW grows like $(\log n)^3$ in rescaled time, however, these extra branches, with high probability, occur between immediate sibling particles, so can be safely suppressed.

The expected number of particles in a TBBM with branching rate $K$, after time $T$ is $e^{2KT}$, so we can choose $T$ to control the expected number of particles at the time the coupling breaks down. The distribution of the number of particles in a TBBM is quite well understood, see for example [5], so $T$ could be chosen to give finer control over the number of particles seen at the time the coupling breaks down.

The proof of this theorem follows a method in [14]. In this paper, the authors couple a binary branching Brownian motion to the branching and coalescing dual process of a spatial $\Lambda$-Fleming-Viot process. Our work makes a number of deviations from this proof, two significant ones being that our branching Brownian motion is ternary, rather than binary, and the BARW involves annihilation, rather than coalescence. There are also some more subtle differences which require more work to adapt. In particular, the BARW lives in discrete space, where as the dual to the spatial $\Lambda$-Fleming Viot process lives in continuous space.

To prove this theorem, we introduce an object which we will call a *branching vine*. This object essentially evolves in the same way as a BARW, but with some of the annihilation events suppressed, and some branching events ignored. The events which cause the branching vine to deviate from the BARW will be rare, so we will be able to couple these two processes with high probability. We then couple the branching vine to a ternary branching Brownian motion.

Before we can prove Theorem 4.1, we need to adapt a result in [10] from a result about the first collision time of particles performing random walks, to one concerning the first annihilation time, when annihilation is not instantaneous. This will be the goal of Section 4.1. We then turn to the proof of Theorem 4.1 in Section 4.2.

## 4.1 Asymptotics of first annihilation time for delayed planar random walks

In this section, we investigate the asymptotics of 'non-annihilation' events of a coalescing or annihilating random walk.

In [10], Cox, Merle and Perkins derive an asymptotic expression for the non-collision probabilities of symmetric random walks. Their results easily apply to coalescing and annihilating random walks which coalesce or annihilate immediately upon contact, at least up until the first coalescence or annihilation event. The main goal of this section is to extend their results to coalescing and annihilating random walks which do not immediately coalesce or annihilate upon contact. In particular, we wish to prove Theorem 4.2, which is an extension of Proposition 1.3 of [10]. We also prove Lemma 4.16, which gives a coupling between non-annihilating and standard random walks. We use both Theorem 4.2 and Lemma 4.16 in Section 4.2.

We will call walks which do not coalesce or annihilate immediately *delayed* random walks, and the original random walks, when necessary, will be called *instantaneous*.

Lemmas 4.4, 4.8, 4.9, 4.10 and 4.12 are identical to Claims 9.3, 9.5, 9.6, 9,7 and 9.9 of [10], respectively, except that they concern delayed random walks, rather than instantaneous walks. We use these claims from [10] in our proofs, but rather than state them again, we refer the reader to their counterpart concerning delayed random walks.

We also mention at this stage that we will occasionally give otherwise arbitrary constants names, such as $C_{1.3}$. This helps identify where certain constants come from in our calculations, but the precise names we give may be unexpected. This is because, if we do name our constants, we will give them the same subscript that the corresponding constant was given in [10]. This section should be self-contained, in that it should not be necessary to refer to [10] to understand the proofs, but should the reader be interested in consulting this paper, this choice of labels may help connect this work.

### 4.1.1 Definitions and statement of theorem

To begin, we state some definitions from Section 9 of [10]. In almost all cases, the notation we will use is the same as that used in [10]. Let $\{x_i\}_{i=1}^N$ be the initial positions of a finite collection of particles living in the planar lattice $\mathbb{Z}^2$. These particles perform independent rate 1 continuous time simple random walks; we shall denote the probability measure describing this behaviour by $P_{x_1 \ldots x_N}$. Let $\{B_i\}_{i=1}^N$ be the random walks described by these particles, and $B = (B^1, \ldots, B^N)$. Similarly, let

$$Y = (Y^1, \ldots, Y^{\frac{1}{2}N(N-1)}) = (B^1 - B^2, \ldots, B^1 - B^n, B^2 - B^3, \ldots, B^2 - B^n, \ldots, B^{N-1} - B^N)$$

be a vector containing the distance between each pair of particles, with initial positions denoted $y_1, \ldots, y_{\frac{1}{2}N(N-1)})$. For $k \in [N]$, we shall also use the non-collision events,

$$D_{[t',t]}^{(k)} = \{\forall 1 \leq i < j \leq k, \forall s \in [t', t], B_s^i \neq B_s^j\},$$

and the simplifications of this notation $D_t^{(k)} = D_{[0,t]}^{(k)}$, $A_{[t',t]} = D_{[t',t]}^{(2)}$ and $A_t = D_t^{(2)}$.

We also define $\binom{N}{2}$ i.i.d. $\exp(\mu)$ random variables, denoted $e_{1,2}, \ldots, e_{N-1,N}$, where $\mu$ is the rate at which particles at the same location coalesce or annihilate. Define

$$E^{(k)}_{[t',t]} = \left\{ \forall 1 \leq i < j \leq k, \int_t^{t'} \mathbb{1}(B^i_s = B^j_s) ds < e_{i,j} \right\}$$

and similarly, $E^{(k)}_t = E^{(k)}_{[0,t]}$. These are the analogues of $D^{(k)}_{[t',t]}$ and $D^{(k)}_t$ for the delayed random walks. Similarly, set $F_{[t',t]} = E^{(2)}_{[t',t]}$ and $F_t = E^{(2)}_t$, which will be the analogues of $A_{[t',t]}$ and $A_t$. We may refer to events such as $E^{(k)}_t$ and $F_t$ as delayed non-collision events.

We wish to extend to our setting Proposition 1.3 of [10], which gives the asymptotics of the first collision time of a set of independent random walks. We use the notation $f(n) \sim g(n)$ if $f$ equals $g$ asymptotically, i.e., $\lim\limits_{n\to\infty} \frac{f(n)}{g(n)} = 1$. If $f$ and $g$ are functions of multiple variables, a subscript may be used to indicate the variable which is being taken to infinity. Specifically, we wish to prove the following.

**Theorem 4.2.** *Fix $N \in \mathbb{N}$, $N \geq 2$ and let $x_i \in \mathbb{Z}^2$ for $i = 1, \ldots, N$.There exists a constant $C_N(x_1, \ldots, x_N) > 0$ depending only on $x_1, \ldots, x_N$ such that*

$$P\left(E^{(N)}_t\right) \underset{t\to\infty}{\sim} \frac{C_N(x_1, \ldots, x_N)}{(\log t)^{\frac{N(N-1)}{2}}}.$$

*Furthermore, there exists a constant $C_{1.3}$ depending only on the transition kernel and $N$ such that*

$$C_N(x_1, \ldots, x_N) \leq C_{1.3} \left( \max_{1\leq i<j\leq N} (\log|x_i - x_j|)^{N(N-1)/2} + 1 \right).$$

The theorem in [10] gives the asymptotics for $D^{N(N-1)/2}_t$, rather than $E^{N(N-1)/2}_t$, and consequently includes the additional assumption that the initial locations of each particle are distinct, but is otherwise the same, albeit with a different constant.

We use this theorem in Section 4.2, but specifically for $N = 3$. As shall become evident in the proof, once generalised beyond $N = 2$, the remaining cases essentially follow in the same way.

We break the proof of Theorem 4.2 up into three subsections. In Subsection 4.1.2, we specialise to the case where we have only two particles, and prove asymptotic results for the probabilities of certain delayed collision and delayed non-collision events. As a delayed collision necessitates a collision, many upper bounds on our delayed collision events follow precisely from the equivalent lemma in [10] for the instantaneous walks. However, proving lower bounds needs more work. This subsection will have the most changes from the corresponding lemmas in [10].

In Subsection 4.1.3, we return to the general $n \geq 2$ case, building on the results for two particles in order to prove Theorem 4.2. As the results for the delayed walks have the same asymptotics as the instantaneous walks, many of these lemmas can be adapted trivially, or almost trivially. Hence, we omit the details of some of these proofs.

Finally, we present the proofs of Theorem 4.2 and Lemma 4.16, in their own subsection, Subsection 4.1.4.

Before finishing this subsection, we state a standard kernel estimate. A proof of (4.2) can be found in [10], Lemma 9.2, but it is essentially a restatement of a classical result, and could be found in many textbooks, such as [28]. We denote the two-dimensional Gaussian kernel by $\bar{p}_t(x) = \frac{1}{2\pi t}\exp(-\frac{|x|^2}{2t})$, and we recall from Definition 1.7 that we denote the transition kernel for the BARW by $m$. Recall further from Subsection 1.2.2 that we are immediately specialising to the case where $m$ is the simple symmetric random walk transition kernel, i.e., $m_{ij} = 1/4$ if and only if $|i-j| = 1$. Finally, we abuse notation slightly and use $m_t$ to denote the transition density for the continuous time simple symmetric random walk, i.e., for $t \geq 0$, $x, y \in \mathbb{Z}^2$, $m_t(|x-y|)$ is the probability that a simple symmetric random walk starting at $x$ is at $y$ at time $t$.

**Lemma 4.3** (Lemma 9.2 of [10])**.** *There exists a constant $C_{9.2} > 0$ such that for all $t > 0$, and any $x \in \mathbb{Z}^2$,*

$$|m_t(x) - \bar{p}_t(x)| \leq \frac{C_{9.2}}{t^{\frac{3}{2}}}. \tag{4.2}$$

### 4.1.2 Two random walks

In this subsection we focus on estimates involving only two particles, a natural first step.

The corresponding section in [10], concerning only two random walks, starts with the last-exit from the origin formula for random walks. This formula partitions the probability space on the last time the walk left the origin. In words, this says that at time $t > 0$, a random walk starting at $y$ either has not yet reached the origin, happens to be at the origin at that exact time, or it reached the origin, but left at some time $s < t$, and then never returned before time $t$.

Symbolically, for a random walk $Y$ with transition kernel $p$, started at $y \in \mathbb{Z}$, and recalling that $A_t$ is the non-collision event before time $t$ for two random walks starting $y$ apart, the last-exit formula is

$$1 = P_y\left(\forall s \leq t : Y_s \neq 0\right) + p_t(y) + \int_0^t p_s(y) P_0\left(\{\forall r < t-s, Y_r \neq 0\}\right) ds.$$

The first term on the right side of the last-exit formula is $P_{x_1,x_2}\left(A_t\right)$ if $x_1 - x_2 = y$, so it is clear that this approach is useful for bounding such non-collision events. It is not so helpful for delayed non-collision events, as delayed random walks could meet many times before coalescing or annihilating. Our adaptations of the lemmas proved in Section 9.2 of [10] will build on the results for the instantaneous random walks, in combination with a 'first return' formula, Lemma 4.7.

However, before getting into these details, we prove a simpler lemma. As this lemma bounds the delayed collision probability (as opposed to the non-collision probability which we will usually focus on) for two walks, the equivalent bound in [10] gives us the result easily. Recall that $A_t$ is the non-collision before time $t$ event for a pair of random walks, and that $F_t$ is the corresponding event for delayed walks.

**Lemma 4.4** (Modification of Claim 9.3 of [10])**.** *There exists a constant $C_{9.3}$ such that for any $x_1, x_2 \in \mathbb{Z}^2$, for any $t > 1$,*

$$P_{x_1,x_2}\left(\left(F_t\right)^{\complement}\right) \leq \frac{C_{9.3}}{\log t}\left(1 + \left(\log(t) - 2\log|x_2 - x_1|\right)^+\right).$$

*Proof.* Notice that $A_t \subset F_t$, as the delayed random walk cannot annihilate or coalesce until its particles occupy the same location. Hence,

$$\begin{aligned} P_{x_1,x_2}\left(\left(F_t\right)^{\complement}\right) \leq& 1 - P_{x_1,x_2}\left(A_t\right) \\ =& P_{x_1,x_2}\left(\left(A_t\right)^{\complement}\right) \\ \leq& \frac{C_{9.3}}{\log t}\left(1 + \left(\log(t) - 2\log|x_2 - x_1|\right)^+\right) \end{aligned}$$

where the last line is from Claim 9.3 of [10]. □

We now aim to prove Lemma 4.6, which gives an asymptotic estimate of the delayed non-collision probability. Our Lemma 4.6 is quite different from the corresponding lemma in [10], so we state it explicitly here. But first, we introduce some new notation.

We may write $P_{0,0}\left(F_t\right)$, as this is the the probability that two delayed random walks, both starting at the origin, never coalesce. The corresponding expression for instantaneous random walks, $P_{0,0}\left(A_t\right)$, is redundant, as if the walks both start at the origin, they will coalesce immediately, and this probability is always 0. To get around this, Cox, Merle and Perkins introduce some new notation,

$$\hat{P}(A_t) = \sum_{i \in \mathbb{Z}^2} m_{0i} P_i\left(A_t\right),$$

which is the probability that the random walks starting at the same place make a transition immediately, and once they have separated, they don't meet before time $t$.

**Lemma 4.5** (Claim 9.4 of [10])**.** *There exists a constant $C_{9.4}$ such that for every $t \geq e$,*

$$\left|\hat{P}(A_t) - \frac{2\pi}{\log t}\right| = \frac{C_{9.4}}{\log^{3/2} t}.$$

Our analogue of this lemma is much weaker. However, it is strong enough, as we benefit from Claim 9.4 of [10] itself when proving Lemma 4.8.

**Lemma 4.6.** *As $t \to \infty$,*

$$P_{0,0}\left(F_t\right) = \mathrm{o}\left(\frac{\log\log t}{\log t}\right).$$

Before proving this lemma, we introduce the aforementioned claim, the 'first-return' formula.

**Lemma 4.7** (First-return formula)**.** *Let $\tau(t) = t - t\log^{-2}(t)$. For $t > e^e$, it holds that*

$$P_{0,0}\left(F_t\right) \leq \frac{1}{1+\mu}\left(\hat{P}\left(A_{t-\log(t)}\right) + P_{0,0}\left(F_{t-\tau(t)-\log(t)}\right)\right) + \mathcal{O}\left(\log^{-2}(t)\right).$$

*Proof.* As this proof only concerns a single pair of particles starting only from the origin, we will drop the subscript from the probability measure for this proof. Conditioning on the first time the particles separate, again on the first time they return, and using the Markov property,

$$P(F_t) = P(\text{never seperate nor annihilate}) + \int_0^t P(\text{separate at } s\text{, never return})ds$$

$$
\begin{aligned}
&+\int_0^t\int_s^t P(\text{seperate at } s\text{, return at } r\text{, continue to not annihilate})drds\\
=&e^{-t(\mu+1)}+\frac{1}{(1+\mu)}\int_0^t(1+\mu)e^{-s(1+\mu)}\hat{P}(A_{t-s})ds\\
&+\int_0^t\int_s^t e^{-s(1+\mu)}\hat{P}(\text{return for the first time at } r-s)P(F_{t-r})drds, \qquad (4.3)
\end{aligned}
$$

where we can obviously cancel the constant coefficients in the second term in the second equation (and have done so for the third term), but we have left them there to make it clear that one factor is from the density function for the exponential variable describing the time of the first event, annihilation or migration, and one factor is the probability that that first event is migration. Now consider only the last integral, which we shall split into three parts,

$$
\begin{aligned}
I_1+I_2+I_3=&\int_0^{\tau(t)}\int_s^{\tau(t)+s} e^{-s(1+\mu)}\hat{P}(\text{return for the first time at } r-s)P(F_{t-r})drds\\
&+\int_0^{\tau(t)}\int_{\tau(t)+s}^t e^{-s(1+\mu)}\hat{P}(\text{return for the first time at } r-s)P(F_{t-r})drds\\
&+\int_{\tau(t)}^t\int_s^t e^{-s(1+\mu)}\hat{P}(\text{return for the first time at } r-s)P(F_{t-r})drds
\end{aligned}
$$

We can split the integral here because $t>e^e>e$. The integral $I_1$ is the most substantial integral, in terms of area being integrated over, and absolute magnitude, which we shall now show. As $P(F_t)$ decreases monotonically in $t$, we can see that

$$
\begin{aligned}
I_1&\leq\int_0^{\tau(t)}e^{-s(1+\mu)}\int_s^{\tau(t)+s}\hat{P}(\text{return for the first time at } r-s)drP\left(F_{t-\tau(t)-s}\right)ds\\
&=\int_0^{\tau(t)}e^{-s(1+\mu)}\hat{P}(\text{return at some time in } [0,\tau(t)])P\left(F_{t-\tau(t)-s}\right)ds\\
&=\int_0^{\tau(t)}e^{-s(1+\mu)}(1-\hat{P}\left(A_{\tau(t)}\right))P\left(F_{t-\tau(t)-s}\right)ds. \qquad (4.4)
\end{aligned}
$$

In $I_2$, we exploit the fact that the probability that the first return to the origin is at time $t$ must be less than the probability of being at the origin at all at time $t$.

$$
\begin{aligned}
I_2&\leq\int_0^{\tau(t)}e^{-s(1+\mu)}\int_{\tau(t)+s}^t m_{r-s}(0)drds\\
&\leq\int_{\tau(t)}^t m_r(0)dr
\end{aligned}
$$

$$
\begin{aligned}
&\leq \int_{\tau(t)}^{t} \bar{p}_r(0) + \frac{C_{9.2}}{r^{3/2}} dr \\
=&\frac{\log(t) - \log\tau(t)}{2\pi} + \mathcal{O}\left(\frac{1}{\sqrt{t}}\right) \\
=&-\frac{\log(1-\log^{-2}(t))}{2\pi} + \mathcal{O}\left(\frac{1}{\sqrt{t}}\right) \\
=&\mathcal{O}\left(\log^{-2}(t)\right),
\end{aligned} \tag{4.5}
$$

where we use estimate (4.2) in the fourth line. The final term, $I_3$, corresponds to large times, so the exponential factor in the integrand renders this part of the integral tiny.

$$
I_3 \leq \int_{\tau(t)}^{t} e^{-s(1+\mu)} ds = o\left(e^{-t(1+\mu)}\right). \tag{4.6}
$$

Substituting (4.4), (4.5), and (4.6) into (4.3) gives,

$$
\begin{aligned}
P(F_t) \leq & \int_0^t e^{-s(1+\mu)} \hat{P}(A_{t-s}) ds \\
&+ \int_0^{\tau(t)} e^{-s(1+\mu)} (1 - \hat{P}\left(A_{\tau(t)}\right)) P\left(F_{t-\tau(t)-s}\right) ds + o\left(\log^{-2}(t)\right)
\end{aligned} \tag{4.7}
$$

We conclude by splitting our integrals at time $\log t$. Before this time, we crudely bound the probabilities that appear in our integrals by 1, and otherwise, we again exploit the fact that $P(F_t)$ and $\hat{P}(A_{t-s})$ decrease monotonically in $t$. Hence, (4.7) becomes

$$
\begin{aligned}
P(F_t) \leq & \int_0^{\log(t)} e^{-s(1+\mu)} \hat{P}(A_{t-s}) ds + \int_0^{\log(t)} e^{-s(1+\mu)} (1 - \hat{P}\left(A_{\tau(t)}\right)) P\left(F_{t-\tau(t)-s}\right) ds \\
&+ \int_{\log(t)}^{t} e^{-s(1+\mu)} \hat{P}(A_{t-s}) ds + \int_{\log(t)}^{\tau(t)} e^{-s(1+\mu)} P\left(F_{t-\tau(t)-s}\right) ds + \mathcal{O}\left(\log^{-2}(t)\right) \\
\leq & \hat{P}\left(A_{t-\log(t)}\right) \int_0^{\log(t)} e^{-s(1+\mu)} ds + P\left(F_{t-\tau(t)-\log(t)}\right) \int_0^{\log(t)} e^{-s(1+\mu)} ds \\
&+ \int_{\log(t)}^{t} e^{-s(1+\mu)} ds + \int_{\log(t)}^{\tau(t)} e^{-s(1+\mu)} ds + \mathcal{O}\left(\log^{-2}(t)\right) \\
= & \frac{1}{1+\mu} (1 - e^{-\log(t)(1+\mu)}) \left(\hat{P}\left(A_{t-\log(t)}\right) + P\left(F_{t-\tau(t)-\log(t)}\right)\right) \\
&+ \frac{1}{1+\mu} \left(2e^{-\log(t)(1+\mu)} - e^{-t(1+\mu)} - e^{-\tau(t)(1+\mu)}\right) + \mathcal{O}\left(\log^{-2}(t)\right) \\
\leq & \frac{1}{1+\mu} \left(\hat{P}\left(A_{t-\log(t)}\right) + P\left(F_{t-\tau(t)-\log(t)}\right)\right) + \mathcal{O}\left(\log^{-2}(t)\right)
\end{aligned}
$$

This is our result. Note that as $t > e^e$, $t - \tau(t) - \log(t) > 0$. This bound is not sharp, but it is sufficient. ☐

*Proof of Lemma 4.6.* Again, we drop the subscript from the probability measure for this proof. We apply Lemma 4.7 iteratively. In particular, we apply it $N = \lceil c \log\log t \rceil$ times, where $c = \log^{-1}(1+\mu)$. Let $g(t) = g_1(t) = t - \tau(t) - \log(t) = t\log^{-2} t - \log t$ and $g_n(t) = (\overbrace{g \circ \ldots \circ g}^{n})(t)$. When $t$ is sufficiently large, $g_1(t) \sim \frac{t}{\log^2 t}$, and hence $g_n(t) \sim \frac{t}{\log^{2n} t}$. Thus, for sufficiently large $t$, $g_n(t) < g_{n-1}(t) < \ldots < g(t) < t$. Applying Lemma 4.7 N times,

$$
\begin{aligned}
P(F_t) \leq & \frac{1}{1+\mu}\hat{P}\left(A_{t-\log(t)}\right) + \frac{1}{(1+\mu)^2}\hat{P}\left(A_{g_1(t)-\log(g_1(t))}\right) + \ldots \\
& + \frac{1}{(1+\mu)^N}\hat{P}\left(A_{g_{N-1}(t)-\log(g_{N-1}(t))}\right) + \frac{1}{(1+\mu)^N}P\left(F_{g_N(t)}\right) + \mathcal{O}\left(\frac{\log\log t}{\log^2 t}\right) \\
\leq & N\hat{P}\left(A_{g_{N-1}(t)-\log(t)}\right) + \frac{1}{(1+\mu)^{c\log\log t}}P\left(F_{g_N(t)}\right) + \mathcal{O}\left(\frac{\log\log t}{\log^2 t}\right) \\
\leq & (c\log\log t + 1)\hat{P}\left(A_{g_{N-1}(t)-\log(t)}\right) + \frac{1}{\log t} + \mathcal{O}\left(\frac{\log\log t}{\log^2 t}\right) \\
= & (c\log\log t + 1)\left(\frac{\pi}{\log(g_{N-1}(t) - \log(t))} + \mathcal{O}\left(\log^{-\frac{3}{2}}(g_{N-1}(t) - \log(t))\right)\right) + \mathcal{O}\left(\log^{-1} t\right) \\
= & o\left(\frac{\log\log t}{\log(t) - 2N\log\log t}\right) + \mathcal{O}\left(\frac{\log\log t}{(\log(t) - 2N\log\log t)^{\frac{3}{2}}}\right) + o\left(\log^{-1} t\right) \\
= & o\left(\frac{\log\log t}{\log t}\right)
\end{aligned}
$$

where in the second inequality we have used that $A_t$ is a decreasing event in $t$, and $g_n$ is a decreasing sequence, we merely bound $F_{g_N(t)}$ by 1 in the third inequality, and we have applied Claim 9.4 of [10] in the fourth equation. The rest of the calculation is just working out the order of the remaining expressions. ☐

We use the same ideas as in the first-return formula to extend the results of Claim 9.5 of [10] to the following lemma, which gives an asymptotic expansion of the probability of non-collision during the interval $[t, 2t]$.

**Lemma 4.8** (Modification of Claim 9.5 of [10]). *Fix $x_1, x_2 \in \mathbb{Z}$. It holds that,*

$$\left| P_{x_1,x_2}\left(F_{[t,2t]}\right) - 1 + \frac{\log 2}{\log t} \right| = \mathcal{O}\left(\log^{-\frac{3}{2}}(t)\right).$$

*In particular, the order of this expression does not depend on $x_1$ or $x_2$.*

*Proof.* Recall the notation $y = x_1 - x_2$, and $P_y(\cdot) = P_{x_1,x_2}(\cdot)$. As in the proof of Lemma 4.7, conditioning on where the particles are at time $t$, and again on the first time they return after $t$, and using the Markov property,

$$\begin{aligned}P_y\left(F_{[t,2t]}\right) =& P_y\,(\text{at the origin at time } t\text{, but then never annihilate})\\ &+ P_y\,(\text{never reach the origin in [t,2t]})\\ &+ \int_0^t P_y\,(\text{reach origin for the first time after } t \text{ at } s\text{, continue to not annihilate})ds.\end{aligned}$$

Thus, by applying (4.2), the Markov property at time $s$ and rearranging, we have that

$$\left|P_y\left(F_{[t,2t]}\right) - P_y\left(A_{[t,2t]}\right)\right| \tag{4.8}$$

$$\leq \int_0^t P_y\,(\text{reach origin for the first time after } t \text{ at } s)P_{0,0}\,(F_{t-s})ds + o\left(t^{-1}\right). \tag{4.9}$$

We split this integral into two parts as we did in the proof of Lemma 4.7.

$$\begin{aligned}I_1 + I_2 =& \int_0^{\tau(t)} P_y\,(\text{reach origin the first time after } t \text{ at } s)P_{0,0}\,(F_{t-s})ds\\ &+ \int_{\tau(t)}^t P_y\,(\text{reach origin for the first time after } t \text{ at } s)P_{0,0}\,(F_{t-s})ds.\end{aligned}$$

We can split the integral here because $t > e$. Again, the integral $I_1$ is the most substantial integral. As $P(F_t)$ decreases monotonically in $t$, we can see that

$$\begin{aligned}I_1 \leq& \int_0^{\tau(t)} P_y\,(\text{reach origin for the first time after } t \text{ at } s)dsP_{0,0}\left(F_{t-\tau(t)}\right)\\ \leq& P_y\,(\text{reach origin at some time in } [t,2t])P_{0,0}\left(F_{t-\tau(t)}\right)\\ =&(1 - P_y\left(A_{[t,2t]}\right))P_{0,0}\left(F_{t-\tau(t)}\right)\\ =&\mathcal{O}\left(\log^{-\frac{3}{2}} t\right) o\left(\frac{\log\log(t\log^{-2} t)}{\log(t\log^{-2} t)}\right)\\ =&\mathcal{O}\left(\log^{-\frac{3}{2}} t\right)\end{aligned} \tag{4.10}$$

by Claim 9.5 of [10] and Lemma 4.6.

In $I_2$, we again exploit the fact that the probability that the first return to the origin is at time $t$ must be less than the probability of being at the origin at all at time $t$, which we can estimate with (4.2). Thus,

$$\begin{aligned} I_2 \leq& \int_{\tau(t)}^{t} m_s(y) ds \\ \leq& \int_{\tau(t)}^{t} \frac{1}{2\pi s} + \frac{C_{9.2}}{s^{3/2}} ds \\ =& \frac{\log(t) - \log\tau(t)}{\pi} + \mathcal{O}\left(\frac{1}{\sqrt{t}}\right) \\ =& o\left(\log^{-2}(t)\right). \end{aligned} \tag{4.11}$$

Finally, we observe that, for sufficiently large $t$, by Claim 9.5 of [10] we may write,

$$\left| P_{x_1,x_2}\left(A_{[t,2t]}\right) - 1 + \frac{\log 2}{\log t} \right| = \mathcal{O}\left(\log^{-\frac{3}{2}}(t)\right). \tag{4.12}$$

Hence, substituting (4.10) and (4.11) into (4.8) and comparing the resulting inequality to 4.12 gives our result. ☐

The final result of this subsection gives a bound for the increments of $P_{0,x}(F_t)$ in space.

**Lemma 4.9** (Modification of Claim 9.6 of [10])**.** *There is a constant $C_{9.6}$ so that if $x$, $y$ are such that $|x| \geq \sqrt{t}\log^{-1}(t)$, $|x-y| \leq 2\sqrt{t}\log^{-2}(t)$ and $t$ is sufficiently large, then*

$$|P_{0,x}(F_t) - P_{0,y}(F_t)| \leq \frac{C_{9.6}}{\log^{\frac{3}{2}}(t)}.$$

*Proof.* Again, as Lemma 4.7, we condition on where the particles are at time $t$, and again on the first time they return after $t$, and use the Markov property to show that

$$\begin{aligned} P_{0,x}(F_t) =& P_{0,x}(\text{never meet}) \\ &+ \int_0^t P_{0,x}(\text{reach origin for the first time at } s\text{, continue to not annihilate}) ds. \\ =& P_{0,x}(A_t) + \int_0^t P_{0,x}(\text{reach origin for the first time at } s) P_{0,0}(F_{t-s}) ds. \end{aligned}$$

Thus,

$$|P_{0,x}\left(F_t\right) - P_{0,x}\left(A_t\right)| \leq \int_0^t P_{0,x}\left(\text{return for the first time at } s\right) P_{0,0}\left(F_{t-s}\right) ds.$$

Splitting the integral again, as before, gives

$$\begin{aligned} I = I_1 + I_2 = &\int_0^{\tau(t)} P_{0,x}\left(\text{return for the first time at } s\right) P_{0,0}\left(F_{t-s}\right) ds \\ &+ \int_{\tau(t)}^t P_{0,x}\left(\text{return for the first time at } s\right) P_{0,0}\left(F_{t-s}\right) ds. \end{aligned}$$

As $P(F_t)$ decreases monotonically in $t$, we can see that

$$\begin{aligned} I_1 \leq &\int_0^{\tau(t)} P_{0,x}\left(\text{return for the first time at } s\right) ds P_{0,0}\left(F_{t-\tau(t)}\right) \\ =&P_{0,x}\left(\text{return at some time before } \tau(t)\right) P_{0,0}\left(F_{t-\tau(t)}\right) \\ =&P_{0,x}\left(\left(A_{\tau(t)}\right)^{\complement}\right) P_{0,0}\left(F_{t-\tau(t)}\right) \\ =&\mathcal{O}\left(\frac{\log\log t}{\log t}\right) o\left(\frac{\log\log(t\log^{-2} t)}{\log(t\log^{-2} t)}\right) \\ =&\mathcal{O}\left(\log^{-\frac{3}{2}}(t)\right) \end{aligned}$$

by Claim 9.3 of [10] and Lemma 4.6, where we have used the assumption that $|x| \geq \sqrt{t}\log^{-1}(t)$ in the application of Claim 9.3 of [10]. Note that by the assumptions on $x$ and $y$ in the lemma, $|y| \geq \sqrt{t}\log^{-1}(t)(1 - 2\log^{-1}(t))$, so for sufficiently large $t$, Claim 9.3 of [10] can also be applied when this same calculation is repeated for $P_{0,y}\left(F_t\right)$.

In $I_2$, we again exploit the fact that the probability of the first return to the origin at time $t$ must be less than the probability of being at the origin at all at time $t$, which we can estimate with (4.2). Thus,

$$\begin{aligned}
I_2 \leq& \int_{\tau(t)}^{t} m_s(x) ds \\
\leq& \int_{\tau(t)}^{t} \bar{p}_s(x) ds + \int_{\tau(t)}^{t} \frac{C_{9.2}}{s^{\frac{3}{2}}} ds \\
\leq& \int_{\tau(t)}^{t} \frac{1}{2\pi s} ds + C\left(\tau(t)^{-\frac{1}{2}} - t^{-\frac{1}{2}}\right) \\
=& \frac{\log t - \log \tau(t)}{2\pi} + \mathcal{O}\left(t^{-\frac{1}{2}}\right) \\
=& o\left(\log^{-2} t\right).
\end{aligned}$$

This estimate did not depend on $x$, so it will hold when repeated for $P_{0,y}(F_t)$. Thus, we have that

$$|P_{0,x}(F_t) - P_{0,x}(A_t)| = \mathcal{O}\left(\log^{-\frac{3}{2}}(t)\right),$$

and the same holds for $P_{0,y}(F_t)$. Finally, we observe that, by Claim 9.6 of [10]

$$|P_{0,x}(A_t) - P_{0,y}(A_t)| = \mathcal{O}\left(\log^{-\frac{3}{2}}(t)\right),$$

and the result follows. □

### 4.1.3 Greater than two random walks

In this subsection we generalise to $N$ particles. When the context is clear, we will drop the subscript for $P_{x_1,\ldots,x_n}()$, and the exponent $N$ from the notation $D_t^{(N)}$. As we have done most of the hard work of the adaptation in the previous section, most of the results follow from the lemmas in the previous section in the same way as they do in [10]. As in the previous subsection, we get our first lemma as a cheap consequence of the equivalent lemma in [10].

**Lemma 4.10** (Modification of Claim 9.7 of [10])**.** *There exists a positive constant $\alpha$ depending only on $N$ such that for any $t \geq 2$, for any $x_1, \ldots, x_n$ ,*

$$P_{x_1,\ldots,x_n}(E_t) \geq t^{-\alpha}.$$

*Proof.* We get this lemma immediately from Claim 9.7 of [10], as in order to annihilate or coalesce, the particles must meet. Thus,

$$P_{x_1,\ldots,x_n}\left(E_t\right) \geq P_{x_1,\ldots,x_n}\left(D_t\right) \geq t^{-\alpha}.$$

□

We introduce notation for the running minimum of the differences between our $n$ random walks, $Z_t = \inf_{k\in\{1,\ldots,\frac{1}{2}n(n-1)\}} |Y_t^k|$. Lemma 9.8 of [10] essentially only concerns $Z_t$, which is a function of the underlying continuous time random walk $Y$, which is the same for both the delayed and instantaneous random walks. So Lemma 9.8 of [10] needs almost no adaptation, but we shall state it here for completeness. This lemma gives an asymptotic bound in probability on how quickly the running minimum grows, conditional on non-collision.

**Lemma 4.11.** *Suppose* $h(t)$ *is positive and satisfies* $\lim_{t\to\infty} h(t)^{\frac{3}{2}} \log t = 0$ *and* $\lim_{t\to\infty} th(t)^2 = +\infty$. *For* $t > 0$, *we define* $T_h(t) = \inf\{s \geq 0 : Z_s \geq \sqrt{t}h(t)\}$. *There exists a constant* $C_{9.8}$ *depending only on* $N$ *and* $h$ *such that for any distinct starting points and any* $t > 0$, $P\left(T_h(t) > t\sqrt{h(t)}\middle| A_t\right) \leq C_{9.8}t^{-2}$.

*Proof.* The main part of the proof of this lemma in [10] proves that

$$P\left(T_h(t) > t\sqrt{h(t)}\right) \leq t^{-\beta} \tag{4.13}$$

for any $\beta > 0$. This is done by breaking up the interval $[0, t\sqrt{h(t)}]$ into on the order of $h(t)^{\frac{3}{2}}$ pieces, and bounding the probability that the running minimum over one interval is less than $\sqrt{t}h(t)$, given that it was in the previous interval.

We need not prove this again here, as $Z_t$ is a function of the underlying continuous time random walk $Y$, which is the same for both the delayed and instantaneous random walks. To conclude this proof, all that we must do is use Bayes' theorem on $P\left(T_h(t) > t\sqrt{h(t)}|E_t\right)$ and apply (4.13) and Lemma 4.10. □

Lemma 9.9 of [10] also needs little modification. This lemma bounds the probability that the random walks are particularly close at time $t$, given non-collision. All we need to do to adapt this proof is to condition on delayed non-collision, which essentially follows, at least in the initial part of the proof of Lemma 9.9 of [10], from the fact that, like $D_t$, $E_t$ is a decreasing event in $t$, i.e., that $E_t \subset E_s$ whenever $t > s$. The second part of the proof holds exactly the same for $E_t$, but it is slightly less straightforward to see this, so we shall include an outline of this proof.

**Lemma 4.12** (Modification of Lemma 9.9 of [10])**.** *There exists a constant $C_{9.9}$ depending only on $N$ such that for all initial points and $t > e$,*

$$P\left(Z_t \le \frac{\sqrt{t}}{\log t}|E_t\right) \le \frac{C_{9.9}}{(\log t)^2}.$$

*Proof.* Let $h(t) = \frac{1}{\log t}$. As mentioned, like the events $D_t$, the $E_t$ are decreasing events, so

$$\{E_t, T_h(t) \le t(\log t)^{-\frac{1}{2}}\} \subset \{E_{T_h(t)}, T_h(t) \le t(\log t)^{-\frac{1}{2}}\}, \tag{4.14}$$

holds for $E_t$ as the equivalent statement held for $D_t$ in Lemma 9.9 of [10]. Applying Lemma 4.11 to (4.14), we have that for $t \ge e$,

$$\begin{aligned} P&\left(Z_t \le \frac{\sqrt{t}}{\log t}\middle| E_t\right) \\ &\le C_{9.8}t^{-2} + \frac{P\left(Z_t \le \frac{\sqrt{t}}{\log t}, T_h(t) \le t(\log t)^{-\frac{1}{2}}, E_t\right)}{P(E_t)} \\ &\le C_{9.8}t^{-2} + \frac{P\left(Z_t \le \frac{\sqrt{t}}{\log t}, T_h(t) \le t(\log t)^{-\frac{1}{2}}, E_{T_h(t)}\right)}{P(E_t)} \\ &= C_{9.8}t^{-2} + P\left(Z_t \le \frac{\sqrt{t}}{\log t}|T_h(t) \le t(\log t)^{-\frac{1}{2}}, E_{T_h(t)}\right)\frac{P\left(D_{T_h(t)}, T_h(t) \le t(\log t)^{\frac{1}{2}}\right)}{P(E_t)}. \end{aligned} \tag{4.15}$$

Using the strong Markov property at time $T_h(t)$, we have

$$\begin{aligned} P&\left(Z_t \le \frac{\sqrt{t}}{\log t}\middle| T_h(t) \le t(\log t)^{-\frac{1}{2}}, E_{T_h(t)}\right) \\ &\le \sup_{\substack{y_1,\dots,y_{\frac{1}{2}n(n-1)} \in (B(0,\sqrt{t}(\log t)^{-1}))^{\complement} \\ s\in[t-t(\log t)^{-\frac{1}{2}},t]}} P_{y_1,\dots,y_{\frac{1}{2}n(n-1)}}\left(Z \le \sqrt{t}(\log t)^{-1}\right) \\ &\le \sup_{y_1\in\mathbb{Z}, s\in[t-t(\log t)^{-\frac{1}{2}},t]} \frac{n(n-1)}{2} P_{y_1}\left(|Y_s^1| \le \sqrt{t}(\log t)^{-1}\right) \end{aligned}$$

$$\leq C(\log t)^{-2} \tag{4.16}$$

where (4.14) was used in the last line. Applying the strong Markov property at time $T_h(t)$ again then gives

$$\begin{aligned}\frac{P(E_t)}{P\left(E_t, T_h(t) \leq t\sqrt{h(t)}\right)} \geq & P\left(E_t | E_t, T_h(t) \leq t\sqrt{h(t)}\right) \\ \geq & \inf_{x_1,\dots,x_n \in B(0,\sqrt{t}h(t))^{\complement}} P_{x_1,\dots,x_n}(E_t) \\ \geq & 1 - \sup_{x_1 \in B(0,\sqrt{t}h(t))^{\complement}} \frac{n(n-1)}{2} P_{x_1}\left(F_t^{\complement}\right),\end{aligned}$$

where we have used a union bound in the last line. We may apply Lemma 4.4 here to find that

$$\begin{aligned}\frac{P(E_t)}{P\left(E_t, T_h(t) \leq t\sqrt{h(t)}\right)} \geq & 1 - \frac{C}{\log t}\left(1 + \log(t) - 2\log\left(\frac{\sqrt{t}}{\log t} \vee 1\right)\right) \\ \geq & 1 - \frac{C}{\log t}(1 + 2\log\log t) \\ \geq & \frac{1}{2} \end{aligned} \tag{4.17}$$

for sufficiently large $t$. Combining (4.15), (4.16) and (4.17), we get our desired result for sufficiently large $t$, and hence for any $t > e$ by choosing $C_{9.9}$ sufficiently large. □

Similarly, Lemma 9.10 of [10], which bounds the probability of collision within a certain time interval given non-collision up to the beginning of that interval, relies upon Lemmas 9.3 and 9.9 of [10], as well as elementary probability relations, so can be proven exactly the same way for our purposes. We will omit the proof of this lemma.

**Lemma 4.13** (Modification of Lemma 9.10 of [10])**.** *Fix $\delta \in (0,1)$ and let $g : (1,\infty) \to (0,\infty)$ be such that $g(t)^{-\delta} \to \infty$ and $t - g(t) \to \infty$ when $t \to \infty$. There exists a constant $C_{9.10} \geq 0$ depending on $g$, such that for any $t$ with $g(t) \geq 2e$ and for any starting points $x_1, \dots, x_n$,*

$$P\left(E_t^{\complement} | E_{g(t)}\right) \leq \frac{C_{9.10}(\log t - \log(g(t)) + 2\log\log g(t)}{\log t}.$$

The next lemma, a modification of Claim 9.11 of [10], is quite a significant one. Like

Lemma 4.16, one of the main lemmas of this section, this lemma gives a coupling between non-colliding and independent random walks, and Lemma 4.16 is essentially just a refinement of this Lemma 4.14. As such, we will include most of the detail of the proof of Lemma 4.14, despite the fact that it is essentially unchanged from the proof of Claim 9.11 in [10]. Furthermore, as we outline the details of the proof of Lemma 4.14, to avoid repetition, the proof of Lemma 4.16 simply focuses on how the proof of Lemma 4.14 differs from Lemma 4.16.

**Lemma 4.14.** *For sufficiently large $t$, there exists a probability measure $\tilde{Q}^t$ on the set of $2n$-tuplets of walk paths such that the following holds. Let $(B, \tilde{B})$ be a $2n$-tuplet of walks under $\tilde{Q}^t$, both started at $x_1, \dots, x_n$. Let $Y$, respectively $\tilde{Y}$ be the corresponding $n(n-1)/2$-tuplet of differences. Then, for sufficiently large $\lambda \geq 8$,*

1. *the distribution of $B$ under $\tilde{Q}^t$ is $P_{x_1,\dots,x_n}\left(\cdot | D_{t(\log t)^\lambda}\right)$, i.e., the first $n$ co-ordinates are walks which do not collide up to $t(\log t)^\lambda$,*
2. *the distribution of $\tilde{B}$ under $\tilde{Q}^t$ is $P_{x_1,\dots,x_n}(\cdot)$, i.e., the last $n$ co-ordinates are independent walks.*
3. $\tilde{Q}^t\left(|Y_t - \tilde{Y}_t|_\infty > \frac{2\sqrt{t}}{(\log t)^2}\right) \leq \frac{C_{9.11}}{(\log t)^2}$.

*Proof.* Initially, make no assumptions on $\lambda$. The idea is that we will split up our time interval into two parts, $[0, \tau]$ and $[\tau, t]$, where $\tau$ will equal $(\log t)^{-\lambda}$, but for now consider $\tau$ unknown. We construct our coupling of $B$ and $\tilde{B}$, both started at $x_1, \dots, x_n$, so that $\tilde{B}$ evolves according to $P_{x_1,\dots,x_n,}(\cdot)$, and $B$ evolves according to $P_{x_1,\dots,x_n,}(\cdot)$ independently over the interval $[0, \tau]$, and the increments of $B$ and $\tilde{B}$ coincide for the rest of the interval, for $[\tau, t]$. So by construction, the first two assertions of the claim hold, and we just need to show the final assertion, and we do this by choosing the right time $\tau$. We want $\tau$ as big as we can, but small enough so that, while the walks are acting independently, they don't get too far away from each other, which we achieve by preventing them from going too far from their common initial points by time $\tau$ (as of course the difference at time $\tau$ will be the same at time $t$, by the coincidence of the increments of the walks after this time).

Again, without yet choosing $\lambda$, we may apply Lemma 4.13 with $g(t) = t(\log t)^{-\lambda}$ , to obtain that for a constant $C_{166}$ depending only on $n$ and sufficiently large $t$,

$$P\left(E_t^{\complement}|E_{t(\log t)^{-\lambda}}\right) \le \frac{C_{166}\log\log t}{\log t}.$$

Applying Markov's inequality, and noting that the variance of the simple random walk has variance which grows with the square root of time, allows us to see that

$$\begin{aligned} P_{y_1,\dots,y_{n(n-1)/2}}\left(\sup_{j\in[n(n-1)/2]}\left|Y^j_{t(\log t)^{-\lambda}}-y_j\right| > \sqrt{t}(\log t)^{-\lambda}\right) & \\ \le & \frac{n(n-1)}{2}(\log t)^{-\lambda/2}E\left[\frac{\left(Y^1_{t(\log t)^{-\lambda}}-y_1\right)^2}{t(\log t)^{-\lambda}}\right] \\ \le & C_{167}(\log t)^{-\lambda/2}. \end{aligned} \tag{4.18}$$

We just need to prove a similar bound on the walk conditioned not to collide.

Choose $\beta \in (0,1)$ such that $2C_{9.10}(1-\beta) \le 1$ and applying Lemma 4.13 with $g(t) = t^\beta$ gives that for $t \ge (2e)^{1/\beta}$,

$$P\left(E_t^{\complement}|E_{t^\beta}\right) \le \frac{1}{2}. \tag{4.19}$$

We want to repeatedly apply of (4.19), until we have $P(E_t) \ge P(E_{2e})\frac{1}{2^l}$ for some $l$. A simple calculations shows that we need $l = \lfloor (\log\beta)^{-1}\log(\log(1+\log 2)-\log t)\rfloor$. Rearranging this gives, for constant $k \ge \frac{\log 2}{\log\beta}$,

$$P(E_t) \ge C_{165}(\log t)^{-k}. \tag{4.20}$$

Constants $C_{165}$ and $k$ do not depend on $t$, so (4.20) holds for any $t > 0$. Choose $\lambda \ge 2k+4$. Combining (4.20) and (4.21), and by our choice of $\lambda$, we have

$$P_{y_1,\dots,y_{n(n-1)/2}}\left(\sup_{j\in[n(n-1)/2]}\left|Y^j_{t(\log t)^{-\lambda}}-y_j\right| > \sqrt{t}(\log t)^{-\lambda}\,\middle|\, E_{(\log t)^{-\lambda}}\right) \le C_{168}(\log t)^{-2}. \tag{4.21}$$

Finally, combining (4.18) and (4.21) gives the third assertion, and consequently, the result. □

We come to the final lemma that we need. As above, the proof of this lemma is very similar to the proof of Lemma 9.12 of [10], and is quite long and technical, so we just sketch out the main differences. This proof is an analogue of Lemma 4.8, but for greater than 2 particles.

**Lemma 4.15.** *There exists a constant $C_{9.12}$, depending only on $n$, such that for any distinct $x_1 \dots x_n$, for any sufficiently large $t \geq \max_{j\in[n(n-1)/2]} |y_j|^4$,*

$$\forall i \in \left[\frac{n(n-1)}{2}\right], \qquad \left| P\left(F_{2t}^{(i)} | E_t\right) - 1 + \frac{\log 2}{\log t}\right| \leq \frac{C_{9.12}}{(\log t)^{3/2}}, \tag{4.22}$$

$$\forall (i_1, i_2) \in \left[\frac{n(n-1)}{2}\right]^2, i_1 \neq i_2 \qquad P\left(\bigcap_{j=1}^{2} \left(F_{2t}^{(i_j)}\right)^{\complement} | E_t\right) \leq \frac{C_{9.12}}{(\log t)^{3/2}}, \tag{4.23}$$

$$\left| P(E_{2t}|E_t) - 1 + \frac{n(n-1)\log 2}{2\log t}\right| \leq \frac{C_{9.12}}{(\log t)^{3/2}}. \tag{4.24}$$

*Sketch of proof of Lemma 4.15.* We need to show these three bounds. The third, (4.24), is a simple consequence of the first two, nothing new is needed here.

We outline the adaptation of the second, (4.23), here. Follow the proof of (4.24) in [10] up until Equation (176). Until this point, the proof only argues that w.l.o.g., we can look at a particular pair of particles, and then apply Lemma 9.9 of [10]. We may apply our version, Lemma 4.12, and also apply the Markov property at time $t$. As the proof notes, we then need to establish Equation (176), i.e.,

$$\sup_{\xi \in \{y \in (\mathbb{Z}^2)^{n(n-1)/2} : \min_{i\in[n(n-1)/2]} |y_i| \geq \frac{\sqrt{t}}{\log t}\}} P_\xi\left(\left(A_t^{(1)}\right)^{\complement} \cap \left(A_t^{(i_2)}\right)^{\complement}\right) \leq \frac{C_{176}}{(\log t)^{3/2}}.$$

The paper then goes on to establish this, but as we have observed, an upper bound on collision is also an upper bound on coalescence/annihilation, so in establishing Equation (176), [10] gives us our version for free.

All that remains is to show (4.22), which is also not too hard to adapt given the work we have done already. The key to showing (4.22), is to show

$$P\left(F^{(1)\,\complement}_{[t,2t]}\,\middle|\,E_{g(t)}\right)-\frac{C_{172}}{\log^{\frac{3}{2}}t}\le P\left(F^{(1)\,\complement}_{[t,2t]}\,\middle|\,E_{t}\right)\le P\left(F^{(1)\,\complement}_{[t,2t]}\,\middle|\,E_{g(t)}\right)\left(1-\frac{C_{166}\log\log t}{\log t}\right)^{-1}, \quad (4.25)$$

for $g(t)=t(\log t)^{-\lambda}$. The upper bound in (4.25) can be found the same way as in [10], by an application of Lemma 4.13. The lower bound, again, is essentially free by the same reasoning as for (4.23), but it is slightly harder to see, so we'll be a touch more explicit. Let

$$\Upsilon_t=\{y\in(\mathbb{Z}^2)^{n(n-1)/2}:\min_{i\in[n(n-1)/2]}|y_i|\ge g(t)\}$$

By the same elementary reasoning as in [10], we have that

$$P\left(F^{(1)\,\complement}_{[t,2t]}\,\middle|\,E_{t}\right)\ge P\left(F^{(1)\,\complement}_{[t,2t]}\,\middle|\,E_{g(t)}\right)-P\left(F^{(1)\,\complement}_{[t,2t]}\cap E^{\complement}_{[g(t),t]}\,\middle|\,E_{g(t)}\right).$$

Fix $j\in[n(n-1)/2]$. Again, following [10], we may use Lemma 4.12 at time $g(t)$ and the Markov property at time $g(t)$ to show that,

$$P\left(F^{(1)\,\complement}_{[t,2t]}\cap E^{\complement}_{[g(t),t]}\,\middle|\,E_{g(t)}\right)\le\frac{n(n-1)}{2}\left(\frac{C_{9.9}}{\log^2 2}+\sup_{\mathcal{Y}\in\Upsilon_t}P_{\mathcal{Y}}\left(F^{(j)\,\complement}_{[0,t-g(t)]}\cap F^{(1)\,\complement}_{[t-g(t),2t-g(t)]}\right)\right).$$

At this point our proof diverges from the proof in [10]. In the paper, the proof goes on to show that this bound is of low enough order to give the bound in (4.25). To justify this, it appeals again to the last-exit formula. This doesn't matter to us, because now that the expression is not conditional on $E_{g(t)}$, we can use our observation that in order to annihilate, the particles must collide, and can see that

$$\begin{aligned}P\left(\left(F^{(1)}_{[t,2t]}\right)^{\complement}\cap E^{\complement}_{[g(t),t]}\,\middle|\,E_{g(t)}\right)&\\ \le\frac{n(n-1)}{2}\left(\frac{C_{9.9}}{\log^2 2}+\sup_{\mathcal{Y}\in\Upsilon_t}P_{\mathcal{Y}}\left(\left(A^{(j)}_{[0,t-g(t)]}\right)^{\complement}\cap\left(A^{(1)}_{[t-g(t),2t-g(t)]}\right)^{\complement}\right)\right).&\end{aligned}$$

Thus, at this point, the bound for the delayed random walk is inherited from the corresponding bound for the instantaneous random walk.

Once (4.25) has been established, the rest of the proof of (4.22) follows by straightforward applications of lemmas which we have shown analogues of above. This gives the lemma. □

### 4.1.4 Proof of Theorem 4.2

Now that we have proven the necessary lemmas, we outline the proofs of the main results of this section, Theorem 4.2 and Lemma 4.16.

*Proof of Theorem 4.2.* Define $f(t) = (\log t)^{\frac{n(n-1)}{2}} P(E_t)$, and $k(t) = \max\{i \in \mathbb{N} : 2^i \leq t\}$. As for $D_t$, $E_t$ are decreasing events, so we can see that

$$\frac{f(t)}{f(2^{k(t)})} = \left(\frac{\log t}{\log 2^{k(t)}}\right)^{\frac{n(n-1)}{2}} P(E_t | E_{2^{k(t)}})$$

which tends to 1 as $t \to \infty$ by (4.24). Therefore, we just need to show that the limit of the sequence $(f(2^m))_{m \in \mathbb{N}}$ is positive and finite in order to show the first assertion of Theorem 4.2.

Let $T$ be a time sufficiently large to apply Lemma 4.15, and let $m_0 = \inf\{m \in \mathbb{N} : 2^m \geq T\}$. In particular, $T \geq \max_{j \in [n(n-1)/2]} |y_j|^4$. If $m, m' \in \mathbb{N}$ with $m \geq m_0$,

$$f(2^{m+m'}) = f(2^m) \prod_{i=0}^{m'-1} \frac{(m+i+1)^{n(n-1)/2}}{(m+i)^{n(n-1)/2}} P(E_{2^{m+i+1}} | E_{2^{m+i}}).$$

As $m \geq m_0$, we may apply (4.24) to the second term in the product to deduce that uniformly in $x_1, \dots, x_n$,

$$\begin{aligned} f(2^{m+m'}) = f(2^m) \prod_{i=0}^{m'-1} &\left(1 + \frac{n(n-1)}{2} \frac{1}{(m+i)} + \mathcal{O}\left((m+i)^{-2}\right)\right) \\ &\times \left(1 - \frac{n(n-1)}{2} \frac{1}{(m+i)} + \mathcal{O}\left((m+i)^{-3/2}\right)\right), \end{aligned}$$

where the first term in the product comes from a second order expansion of $(1+x)^a$ in $x$ near 0, for $x = \frac{1}{m+i}$ and $a = n(n-1)/2$. Hence, the sequence $\log(f(2^m))$ is Cauchy, and as $f(2^{m_0}) > 0$, we see by applying the above with $m = m_0$, $m' \to \infty$ that the limit of $(f(2^m))$ is positive. Moreover, $f(2^{m_0}) \leq (\log T)^{n(n-1)/2} \leq c_n \left(\max_{j \in [n(n-1)/2]} (4 \log |y_j|)^{n(n-1)/2} + 1\right)$, so the second part of Theorem 4.2 holds. $\square$

We finish this subsection with a sketched proof of our adapted version of Claim 9.13 of [10].

**Lemma 4.16.** *For any sufficiently large $t$, there exists a probability measure $\tilde{Q}^t$ on the set of $2n$-tuplets of walk paths such that the following holds. Let $(B, \tilde{B})$ be a $2n$-tuplet of walks under*

*$\tilde{Q}^t$, and both started at $x_1, \ldots, x_n$. Let $Y$, respectively $\tilde{Y}$ be the corresponding $n(n-1)/2$-tuplet of differences. Then, for $\lambda \geq 8$,*

1. *the distribution of $B$ under $\tilde{Q}^t$ is $P_{x_1,\ldots,x_n,}(\cdot|E_t)$, i.e., the first $n$ co-ordinates are walks which do not annihilate up to $t$,*

2. *the distribution of $\tilde{B}$ under $\tilde{Q}^t$ is $P_{x_1,\ldots,x_n}(\cdot)$, i.e., the last $n$ co-ordinates are independent walks,*

3. $\tilde{Q}^t\left(|Y_t - \tilde{Y}_t|_\infty > \frac{2\sqrt{t}}{(\log t)^{\lambda/4}}\right) \leq \frac{C_{9.13} \log\log t}{\log t}$.

*Sketch of Proof.* The proof of the third assertion in this claim is essentially the same as the proof of Lemma 4.14, so all we will do here is show how the coupling is constructed.

Fix a sufficiently large $t$. As before, we construct our coupling of $B$ and $\tilde{B}$, both started at $x_1, \ldots, x_n$, so that $\tilde{B}$ evolves according to $P_{x_1,\ldots,x_n,}(\cdot)$ on the interval $[0,t]$, and $B$ evolves according to $P_{x_1,\ldots,x_n,}(\cdot)$, but on the shorter interval $[0, t(\log t)^{-\lambda}]$. We now explain how to evolve $B$ for the rest of this interval. Define $\bar{B}$ on $[0,t]$ as

$$\bar{B}_s = \begin{cases} B_s & 0 \leq s \leq t(\log t)^{-\lambda} \\ \tilde{B}_s - \tilde{B}_{t(\log t)^{-\lambda}} + \bar{B}_{t(\log t)^{-\lambda}} & s > t(\log t)^{-\lambda} \end{cases}$$

that is, $\bar{B}$ is $B$ until time $t(\log t)^{-\lambda}$, after which it evolves according to the increments of $\tilde{B}$. Furthermore, define the delayed non-collision event for $\bar{B}$,

$$\bar{E}_t = \left\{ \forall 1 \leq i < j \leq N, \int_0^t \mathbb{1}\left(\bar{B}^i_s = \bar{B}^j_s\right) ds < e_{i,j} \right\}.$$

On $\bar{E}_t$, set $B = \bar{B}$, and otherwise continue evolving $B$ under $P_{x_1,\ldots,x_n,}(\cdot|E_t)$, independently of $\tilde{B}$, on the interval $[t(\log t)^{-\lambda}, t]$. So by construction, the first two assertions of the claim hold.

□

## 4.2 Proof of Theorem 4.1

In this section, we prove the main theorem of this chapter, Theorem 4.1. In the following subsection, we state some necessary estimates. In Subsection 4.2.3, we introduce the branching vine, and prove the necessary couplings, however, before we can do that, in Subsection 4.2.2, we introduce a necessary component for the definition of the branching vine, which we call a vine.

### 4.2.1 Initial estimates

In this section, we construct some estimates for the interactions between precisely three particles, all of which start at the origin. In the notation of the previous section, we are taking $N = 3$ and our probability measure is $P_{0,0,0}(\cdot)$. As this is the only probability measure we use until the last lemma in this section, we shall drop the subscript until we reach that lemma.

Let $c$ be a constant greater than 5. The choice of the value of this constant will become clear as we proceed with our calculations. We define

$$\gamma_n = \tfrac{1}{(\log n)^c}, \tag{4.26}$$

and recall from the Section 4.1, that $Y_s^i$ for $i = 1$ to 3 are the three differences between three simple random walks, and $E_s^3$ is the event that none of these three walks annihilate before time $s$. Define

$$\begin{aligned}
\tau^{an} &= \inf\{s > 0 : (E_s^3)^{\complement} \text{ occurs}\}\\
\tau^{div} &= \inf\{s > 0 : |Y_s^1| \wedge |Y_s^2| \wedge |Y_s^3| \geq \sqrt{n}\gamma_n\}\\
\tau^{over} &= n\gamma_n\\
\tau^{type} &= \tau^{an} \wedge \tau^{div} \wedge \tau^{over}.
\end{aligned}$$

In words, $\tau^{an}$ is the time of the first annihilation event for three simple random walks, $\tau^{div}$ is the first time that the minimum distance between each pair of the random walks exceeds $\sqrt{n}\gamma_n$, $\tau^{over}$ is just a fixed amount of time, and $\tau^{type}$ is the first time that either of the previous three events occur. If $\tau^{type} = \tau^{over}$, we say the particles have 'overshot', as they have failed to annihilate or diverge in the given time. We prove two lemmas regarding these quantities.

First, conditional on non-annihilation up to time $n\gamma_n$, the probability of not diverging is low.

**Lemma 4.17.** *It holds that*

$$P\left(\tau^{div} \geq n\gamma_n \middle| E_{n\gamma_n}^3\right) = \mathcal{O}\left(\frac{1}{(\log n)^{\frac{1}{2}}}\right).$$

*Proof.* Let $\tilde{Y}_1$, $\tilde{Y}_2$ and $\tilde{Y}_3$ be independent random walks starting at the origin conditioned on not annihilating. Then, using some sufficient but rather unsophisticated estimates, we have

$$\begin{aligned} P\left(\tau^{div} \geq n\gamma_n \middle| E^3_{n\gamma_n}\right) =& P\left(\forall s < n\gamma_n, |\tilde{Y}^1_s| \vee |\tilde{Y}^2_s| \vee |\tilde{Y}^3_s| \leq \sqrt{n}\gamma_n\right) \\ \leq& P\left(|\tilde{Y}^1_{n\gamma_n}| \vee |\tilde{Y}^2_{n\gamma_n}| \vee |\tilde{Y}^3_{n\gamma_n}| \leq \sqrt{n}\gamma_n\right) \\ \leq& 3P\left(|\tilde{Y}^1_{n\gamma_n}| \leq \sqrt{n}\gamma_n\right). \end{aligned} \tag{4.27}$$

Next, we exploit the coupling we derived in the previous section, Lemma 4.16. For $\lambda \geq 8$,

$$\begin{aligned} P\left(|\tilde{Y}^1_{n\gamma_n}| \leq \sqrt{n}\gamma_n\right) \leq& P\left(|Y^1_{n\gamma_n} - \tilde{Y}_{n\gamma_n}| > \frac{2\sqrt{n\gamma_n}}{(\log(n\gamma_n))^{\frac{\lambda}{4}}}\right) \\ &+ P\left(|\tilde{Y}^1_{n\gamma_n}| \leq \sqrt{n}\gamma_n \middle| |Y^1_{n\gamma_n} - \tilde{Y}^1_{n\gamma_n}| \leq \frac{2\sqrt{n\gamma_n}}{(\log(n\gamma_n))^{\frac{\lambda}{4}}}\right) \\ \leq& \frac{C_{9.13}\log\log(n\gamma_n)}{\log(n\gamma_n)} \\ &+ P\left(|Y^1_{n\gamma_n}| \leq \sqrt{n}\gamma_n + \frac{2\sqrt{n\gamma_n}}{(\log(n\gamma_n))^{\frac{\lambda}{4}}} \middle| |Y^1_{n\gamma_n} - \tilde{Y}^1_{n\gamma_n}| \leq \frac{2\sqrt{n\gamma_n}}{(\log(n\gamma_n))^{\frac{\lambda}{4}}}\right). \end{aligned} \tag{4.28}$$

We can easily remove the conditioning from this last term.

$$\begin{aligned} P&\left(|Y^1_{n\gamma_n}| \leq \sqrt{n}\gamma_n + \frac{2\sqrt{n\gamma_n}}{(\log(n\gamma_n))^{\frac{\lambda}{4}}} \middle| |Y^1_{n\gamma_n} - \tilde{Y}^1_{n\gamma_n}| \leq \frac{2\sqrt{n\gamma_n}}{(\log(n\gamma_n))^{\frac{\lambda}{4}}}\right) \\ &\leq \frac{P\left(|Y^1_{n\gamma_n}| \leq \sqrt{n}\gamma_n + \frac{2\sqrt{n\gamma_n}}{(\log(n\gamma_n))^{\frac{\lambda}{4}}}\right)}{1 - P\left(|Y^1_{n\gamma_n} - \tilde{Y}^1_{n\gamma_n}| \geq \frac{2\sqrt{n\gamma_n}}{(\log(n\gamma_n))^{\frac{\lambda}{4}}}\right)} \\ &\leq 2P\left(|Y^1_{n\gamma_n}| \leq \sqrt{n}\gamma_n + \frac{2\sqrt{n\gamma_n}}{(\log(n\gamma_n))^{\frac{\lambda}{4}}}\right) \end{aligned} \tag{4.29}$$

where the last line follows by Lemma 4.16, for sufficiently large $n$. Applying Lemma 4.3 allows us to estimate

$$P\left(|Y^1_{n\gamma_n}| \leq \sqrt{n}\gamma_n + \frac{2\sqrt{n\gamma_n}}{(\log(n\gamma_n))^{\frac{\lambda}{4}}}\right) = 2\int_0^{\sqrt{n}\gamma_n + \frac{2\sqrt{n\gamma_n}}{(\log(n\gamma_n))^{\frac{\lambda}{4}}}} m_{n\gamma_n}(x)dx$$

$$
\begin{aligned}
&\leq 2\int_0^{\sqrt{n\gamma_n}+\frac{2\sqrt{n\gamma_n}}{(\log(n\gamma_n))^{\frac{\lambda}{4}}}} \bar{p}_{n\gamma_n}(x)dx + 2\int_0^{\sqrt{n\gamma_n}+\frac{2\sqrt{n\gamma_n}}{(\log(n\gamma_n))^{\frac{\lambda}{4}}}} |\bar{p}_{n\gamma_n}(x) - m_{n\gamma_n}(x)|dx \\
&\leq \operatorname{erf}\left(\frac{\sqrt{n\gamma_n}+\frac{2\sqrt{n\gamma_n}}{(\log(n\gamma_n))^{\frac{\lambda}{4}}}}{\sqrt{2n\gamma_n}}\right) + \frac{2C_{9.2}}{(n\gamma_n)^{3/2}}\left(\sqrt{n\gamma_n}+\frac{2\sqrt{n\gamma_n}}{(\log(n\gamma_n))^{\frac{\lambda}{4}}}\right) \\
&\leq \mathcal{O}\left((\log n)^{\frac{-c}{2}\vee\frac{-\lambda}{4}}\right) + \mathcal{O}\left(n^{-1}\right) \qquad (4.30)
\end{aligned}
$$

as the term in the error function is of order $(\log n)^{-\frac{c}{2}\vee-\frac{\lambda}{4}}$, so considering the asymptotic expansion of the error function at zero gives the last inequality. As $c > 5$ and $\lambda > 8$, combining (4.27), (4.28), (4.29) and (4.30) gives the lemma. □

We get immediate but important corollaries.

**Corollary 4.18.** *It holds that*

$$
P\left(\tau^{type} = \tau^{over}\right) = \mathcal{O}\left(\frac{1}{(\log n)^{\frac{7}{2}}}\right).
$$

*Proof.* In order to overshoot, the particles must not annihilate nor diverge before time $\tau^{over}$.

$$
P\left(\tau^{type} = \tau^{over}\right) = P\left(E^3_{n\gamma_n}\right)P\left(\forall s < n\gamma_n, |Y^1_s| \vee |Y^2_s| \vee |Y^3_s| \leq \sqrt{n\gamma_n} \,\middle|\, E^3_{n\gamma_n}\right).
$$

By Theorem 4.2, $P(E_{n\gamma_n}) \sim \frac{1}{(\log n)^3}$, so combined with Lemma 4.17, we have our result. □

**Corollary 4.19.** *It holds that*

$$
P\left(\tau^{type} = \tau^{div}\right) \sim \frac{K}{(\log n)^3},
$$

*for some constant $K$.*

*Proof.* In order to diverge, the particles must not annihilate nor overshoot. Conditional on not annihilating, diverging is just the complement of overshooting, so we have

$$
P\left(\tau^{type} = \tau^{div}\right) = P\left(E^3_{n\gamma_n}\right)P\left(\left(\forall s < n\gamma_n, |Y^1_s| \vee |Y^2_s| \vee |Y^3_s| \leq \sqrt{n\gamma_n}\right)^{\complement} \,\middle|\, E^3_{n\gamma_n}\right).
$$

Again, by Theorem 4.2, $P\left(E^3_{n\gamma_n}\right) \sim \frac{K}{(\log n)^3}$ for some constant $K$,and by Lemma 4.17,

$$P\Big( \left(\forall s < n\gamma_n, |Y^1_s| \vee |Y^2_s| \vee |Y^3_s| \leq \sqrt{n}\gamma_n\right)^{\complement} \Big| E^3_{n\gamma_n}\Big) \sim 1.$$

Hence the result follows. ☐

**Corollary 4.20.** *It holds that*

$$1 - P\left(\tau^{type} = \tau^{an}\right) \sim \frac{K}{(\log n)^3}$$

*for some constant $K$.*

*Proof.* If the particles don't overshoot, and don't diverge, then they must annihilate. Thus, the corollary is a simple consequence of Corollaries 4.18 and 4.19. ☐

**Remark 4.21.** Recalling that we have used the notation $\sim$ to indicate asymptotic equality, if $\kappa_n = (\log n)^3 P\left(\tau^{type} \neq \tau^{an}\right)$, we know that the limit $K = \lim_{n\to\infty} \kappa_n$ exists. These $\kappa_n$ will ultimately become a coefficient in the branching rate of the branching Brownian motions we wish to couple to our BARW, so knowing how they grow with $n$ is very helpful.

The next two lemmas approximate our random walks by an appropriate Brownian motion. We quote a standard result about coupling a Brownian motion with a simple random walk from [29]. The result we quote below is specific to our needs; the actual result in the textbook is more general.

**Theorem 4.22** (Theorem 3.4.2 of [29])**.** *There exist $C, a > 0$ and a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ on which are defined a standard two-dimensional Brownian motion $W$, a simple symmetric two-dimensional continuous time random walk $B$ such that for all $n \in \mathbb{N}$ and all $1 \leq r \leq T^{1/4}$,*

$$P\left(\sup_{0\leq s\leq T} |W_s - B_s| \geq rT^{1/4}\log(T)^{1/2}\right) \leq Ce^{-ar}.$$

We also need to quantify the idea that if an annihilation event occurs between our three particles, the non-annihilating particle is unlikely to have moved very far away from its initial position. Recall again that $E^3_t$ is the event that three random walks do not annihilate before time $t$.

**Lemma 4.23.** *With probability 1, two of the three particles, $B^1$, $B^2$ and $B^3$, annihilate (not necessarily before $\tau^{over}$ though). Let $i$ be the index of the particle which does not annihilate. Then,*

$$P_{0,0,0}\left(|B^i_{n\gamma_n}| \geq \sqrt{n\sqrt{\gamma_n}}\,\middle|\,(E^3_{n\gamma_n})^{\complement}\right) = \mathrm{o}\left(n^{-1}\right),$$

*which gives an asymptotic for the probability that $B^i$ is more than $\sqrt{n\sqrt{\gamma_n}}$ away from its initial position at time $n\gamma_n$, conditional on the other two particles annihilating each other before this time.*

*Proof.* By Bayes' theorem,

$$P\left(|B^i_{n\gamma_n}| \geq \sqrt{n\sqrt{\gamma_n}}\,\middle|\,(E^3_{n\gamma_n})^{\complement}\right) \leq \frac{P\left(|B^1_{n\gamma_n}| \geq \sqrt{n}\gamma_n\right)}{P\left((E^3_{n\gamma_n})^{\complement}\right)}. \tag{4.31}$$

Once we have no longer conditioned on annihilation, $B_i$ is just a random walk, so we may assume, without loss of generality, that $i = 1$. Taking $r = n^{1/8}$ and $T = n\gamma_n$ in Theorem 4.22, we can couple the random walk to a Brownian motion $(W_t)_{0\leq t\leq T}$ such that for some $C, a > 0$,

$$P\left(\sup_{0\leq s\leq n\gamma_n} |W_s - B_s| \geq (n\gamma_n)^{3/8}\sqrt{\log n\gamma_n}\right) \leq Ce^{-an^{1/8}}.$$

Thus,

$$\begin{aligned}
P\left(|B^1_{n\gamma_n}| \geq \sqrt{n\sqrt{\gamma_n}}\right) \leq & P\left(|W^1_{n\gamma_n}| \geq \sqrt{n\sqrt{\gamma_n}} - (n\gamma_n)^{3/8}\sqrt{\log n\gamma_n}\right) + \mathcal{O}\left(e^{-an^{1/8}}\right) \\
\leq & P\left(Z \geq \gamma_n^{-1/4}\right) + \mathcal{O}\left(e^{-an^{1/8}}\right) \\
= & P\left(Z \geq (\log n)^{c/4}\right) + \mathcal{O}\left(e^{-an^{1/8}}\right) \\
= & o\left(n^{-1}\right)
\end{aligned}$$

as $c \geq 5$, where $Z$ is a standard normal variable. Applying this and Corollary 4.20 to (4.31) gives the result. □

We need one more lemma. Once separated, it is unlikely the particles will come back together.

**Lemma 4.24.** *Fix $K > 0$. For $x_1, x_2 \in \mathbb{Z}^2$ such that $|x_1 - x_2| > \sqrt{n}\gamma_n$,*

$$P_{x_1,x_2}\left((E^2_{nK})^{\complement}\right) = \mathcal{O}\left(\frac{\log\log n}{\log n}\right).$$

*Proof.* Note that $P_{x_1,x_2}\left((E^2_{nK})^{\complement}\right) \leq P_{x_1,x_2}\left(A^{\complement}_{nK}\right)$, as particles must come into contact in order to annihilate. Let $Y_t$ be a random walk starting at $x_1 - x_2$ corresponding to the difference between the random walks performed by the particles described in the lemma. Thus,

$$P_{x_1,x_2}\left(A_{nK}\right) = P(\forall s \leq Kn, |Y_s| > 0).$$

Again, taking $r = \log n$ and $T = Kn$ in Theorem 4.22, we can couple this continuous time random walk $Y$ to a Brownian motion $W_t$ such that

$$P\left(\sup_{0\leq s\leq Kn} |W_s - Y_s| \geq (Kn)^{1/4} \log n \log(Kn)^{1/2}\right) \leq Cn^{-a}.$$

Thus,

$$P(\forall s \leq Kn, |Y_s| > 0) \geq P\left(\forall s \leq Kn, |W_s| > (Kn)^{1/4} \log n \log(Kn)^{1/2}\right) + \mathcal{O}\left(Cn^{-a}\right).$$

Define the stopping times $T_- = \inf\{t > 0 : |W_t| \leq (Kn)^{1/4} \log n \log(Kn)^{1/2}\}$ and $T^+ = \inf\{t > 0 : |W_t| \geq y + \sqrt{n}\log n\}$. Then

$$\begin{aligned}
&P\left(\forall s \leq Kn, |W_s| > (Kn)^{1/4} \log n \log(Kn)^{1/2}\right)\\
=&P(T_- > Kn)\\
\geq&P\left(\{T_- > T^+\} \cap \{T^+ \geq Kn\}\right)\\
\geq&1 - P\left(T_- < T^+\right) - P\left(T^+ \leq Kn\right)\\
\geq&1 - \frac{\log(y + \sqrt{n}\log n) - \log(y)}{\log(y + \sqrt{n}\log n) - \log((Kn)^{1/4} \log n \log(Kn)^{1/2})} - P\left(\sup_{t<Kn} |W_t - W_0| \geq \log n\right)\\
\geq&1 - \frac{\log(\sqrt{n}(\log n)^{-c} + \sqrt{n}\log n) - \log(\sqrt{n}(\log n)^{-c})}{\log(\sqrt{n}(\log n)^{-c} + \sqrt{n}\log n) - \log((Kn)^{1/4} \log n \log(Kn)^{1/2})} - \sqrt{\frac{2}{\pi Kn}} e^{\frac{-(n\log n)^2}{8Kn}}\\
\geq&1 - \mathcal{O}\left(\frac{\log\log n}{\log n}\right) - \mathcal{O}\left(e^{\frac{-(\log n)^2}{8K}}\right)
\end{aligned}$$

where the third last line uses the scale function for a two-dimensional Bessel process, and as the scale function is monotonic in $y$ (this is easily seen by expressing the fraction as 1- another fraction), we can apply the bound on $y = |x_1 - x_2|$ given in the statement of the lemma.

Furthermore, the second last line bounds the modulus of two-dimensional Brownian motion by twice the maximum of the moduli of two one-dimensional Brownian motions. Putting all this together gives the desired result. □

### 4.2.2 Constructing the Vines

We wish to define the process we shall call a vine that was mentioned at the beginning of this section. While the branching vine, which we introduce next subsection, is essentially a modification of the whole BARW, a single vine will be an approximation to the trajectory of at most three particles. After this definition, we couple the vine to a single Brownian motion.

**Definition 4.25** (Vine). *Fix $x \in \mathbb{Z}^2$. We define a lifetime $h > 0$, and a process $(v_t)_{0\leq t\leq h}$ on $(\mathbb{Z}^2)^3$, which we shall refer to as a* vine. *This process will be a deterministic function of the BARW $\xi$. For each $t \geq 0$, we write $v_t = \left(v_t^1, v_t^2, v_t^3\right)$. We also define $k^* \in \mathbb{N}$ and a sequence $(\tau_k^{br})_{k\leq k^*}$ of stopping times.*

*Set $\tau_0^{br} = 0$ and let $\xi^{\emptyset}$ be a particle starting at $x$ evolving according to only the migration events in the BARW $\xi$ up until time $n(\log n)^{-c}$ (that is, any branching or annihilation events occurring before time $n(\log n)^{-c}$ are suppressed, and we follow the trajectory of an arbitrarily chosen one of the particles that exists at the end of this time interval). Let $\tau_1^{br}$ be the time of the first branching event after $n(\log n)^{-c}$. For $t \leq \tau_1^{br}$, let $v_t^1 = v_t^2 = v_t^3 = \xi_t^{\emptyset}$.*

*We shall continue our definition inductively. For $k \geq 1$, suppose we have defined $(\tau_l^{br})_{l\leq k}$. For $t \in [\tau_k^{br}, \tau_k^{br} + n(\log n)^{-c}]$, denote the locations of the particles created at this branching event by $\xi_t^1$, $\xi_t^2$, and $\xi_t^3$. During this interval, these particles evolve according to the BARW $\xi$, again, suppressing any branching events that may occur. Let $v_t^i$ follow the trajectory of the $\xi_t^i$.*

*Denote the differences between each pair of $\xi_t^1$, $\xi_t^2$ and $\xi_t^3$ by $Y_t^1$, $Y_t^2$ and $Y_t^3$. Define*

$$\begin{aligned}
\tau_k^{an} &= \inf\{s > \tau_k^{br} : (E_s)^{\complement} \text{ occurs}\} \\
\tau_k^{div} &= \inf\{s > \tau_k^{br} : |Y_s^1| \wedge |Y_s^2| \wedge |Y_s^3| \geq \sqrt{n}\gamma_n\} \\
\tau_k^{over} &= \tau_k^{br} + n\gamma_n \\
\tau_k^{type} &= \tau^{an} \wedge \tau^{div} \wedge \tau^{over}
\end{aligned}$$

*If $\tau_k^{type} \neq \tau_k^{an}$ (that is, the sibling particles overshoot or diverge before they annihilate), then* *set $k^*$) $= k$ and $h = \tau_{k^*}^{type}$. Otherwise, two of the sibling particles annihilate at time $\tau_k^{an}$. Only one of $\xi^1$, $\xi^2$ and $\xi^3$ exists, call that particle $\xi$. Let $\tau_{k+1}^{br}$ be the time of the first branching event occurring to $\xi$ after time $\tau_k^{br} + n(\log n)^{-c}$. For $t \in (\tau_k^{an} + n(\log n)^{-c}, \tau_{k+1}^{br}]$, let $v_t^1 = v_t^2 = v_t^3 = \xi_t$. We then continue iteratively for each $k \leq k^*$.*

The notation in this definition has been chosen for consistency with Definition 4.12 of [14]. Definition 4.12 defines the notion of a 'caterpillar', an analogue of our vines with only two components, rather than three.

Should a branching event occur in $\xi$ which is ignored in the vine $(v_t)_{0 \leq t \leq h}$, then the trajectory of the vine and the trajectory of the particles in the BARW could become substantially different. This is not a problem though; rather than attempt to control how bad things can get in this case, we will simply show these deviations are rare. We now show a pair of lemmas necessary for coupling the vine to a Brownian motion.

Herein, we shall rescaled the BARW $\xi$. Specifically, the branching rate will be chosen to be $\sigma_n = \frac{(\log n)^3}{n}$.

**Lemma 4.26.** *We can couple the lifetime, $h$, of a vine to $H$, an exponentially distributed random variable with parameter $\kappa_n/n$ in such a way that for some $\delta > 0$, with probability at least $1 - \mathcal{O}\left(e^{-\delta(\log n)^{1/8}}\right)$,*

$$|h - H| \leq 2n(\log n)^{9/8-c}.$$

*Proof.* Recall from Remark 4.21 that $\kappa_n = (\log n)^3 P\left(\tau^{type} \neq \tau^{an}\right)$. Thus, by the strong Markov property, $(\tau_{k^*}^{type})_{k \geq 1}$ is an i.i.d. sequence and

$$k^* \sim \text{Geom}(\kappa_n (\log n)^{-3}).$$

Hence, for large enough $n$,

$$P\left(k^* \geq (\log n)^{\frac{9}{8}}\right) = (1 - \kappa_n (\log n)^{-3})^{(\log n)^{\frac{9}{8}}} = \mathcal{O}\left(e^{-\delta(\log n)^{1/8}}\right). \tag{4.32}$$

Let $H_k = \tau_k^{br} - (\tau_{k-1}^{br} + n\gamma_n) \sim \text{Exp}(\sigma_n)$. The random variables $H_k$ and $\tau_k^{type}$ are independent, and within each sequence, the variables are identically distributed. Set

$$H = \sum_{k=1}^{k^*} H_k = \tau_{k^*}^{br} - k^* n\gamma_n = h - k^* n\gamma_n - (\tau_{k^*}^{type} - \tau_{k^*}^{br}).$$

As the sum of a geometrically distributed number of exponential variables, $H$ itself has an exponential distribution, with parameter equal to the product of the two constituent distributions' parameters, i.e., $\sigma_n \kappa_n (\log n)^{-3} = \kappa_n / n$. Then, as $\tau_{k^*}^{type} - \tau_{k^*}^{br} \le n\gamma_n$,

$$P\Big(|H - h| \ge n\gamma_n (\log n)^{\frac{9}{8}} + n\gamma_n\Big) = \mathcal{O}\left(e^{-\delta(\log n)^{1/8}}\right).$$

The result follows. □

Like the caterpillar, the vine is unlikely to end with an overshooting event.

**Lemma 4.27.** *It holds that*

$$P\left(\tau_{k^*}^{type} = \tau_{k^*}^{over}\right) = \mathcal{O}\left((\log n)^{\frac{-19}{8}}\right).$$

*Proof.* Observing that,

$$\{\tau_{k^*}^{type} = \tau_{k^*}^{over}\} \subset \{k^* \ge (\log n)^{\frac{9}{8}}\} \bigcup_{k=1}^{(\log n)^{\frac{9}{8}}} \{\tau_{k}^{type} = \tau_{k}^{over}\},$$

and applying Corollary 4.18 and (4.32), we can see that,

$$P\left(\tau_{k^*}^{type} = \tau_{k^*}^{over}\right) = \mathcal{O}\left(e^{-\delta(\log n)^{1/8}}\right) + \mathcal{O}\left((\log n)^{\frac{9}{8} + \frac{-7}{2}}\right) = \mathcal{O}\left((\log n)^{\frac{-19}{8}}\right).$$

□

We can now couple the vine to a Brownian motion. Recall from (4.26) that $c$ is the constant in the definition of $\gamma$, and that all we have assumed about it so far is that it is greater than 5.

**Lemma 4.28.** *Let $(W_t)_{t\ge 0}$ be a two-dimensional Brownian motion. There exists an $r > 0$ such that we can couple $(v_t)_{t\le H}$ with $(W_t)_{t\ge 0}$, in such a way that $(W_t)_{t\ge 0}$ is independent of $h$ and*

$k^*$, *and with probability at least* $1-\mathcal{O}(n^{-r})$, *for* $t \leq h$,

$$|v_t^1 - W_t| \leq 2\sqrt{n}(\log n)^{\frac{9}{8}-\frac{c}{4}}. \tag{4.33}$$

*Proof.* We will firstly consider a modified version of $(v_t^1)_{t\leq h}$, which we shall call $(\hat{v}_t)_{t\leq H}$. This modification is simply the concatenation of the parts of $v_t^1$ where branching events are not suppressed, i.e., if $\hat{k}(t) = \sup\{k \geq 0 : \tau_k^{br} - kn\gamma_n < t\}$,

$$\hat{v}_t = v^1_{t+n\gamma_n\hat{k}(t)} + \sum_{k=1}^{\hat{k}(t)} \left( v^1_{\tau_k^{br}} - v^1_{\tau_k^{br}+n\gamma_n} \right)$$

where the sum translates the start of the next section of unsuppressed walk to the end of the last section of unsuppressed walk. This is well defined as long as $\hat{k}(t) < k^*$, (otherwise $v^1_{\tau_k^{br}+n\gamma_n}$ is not necessarily defined) which occurs as long as $t \leq H$.

By the Markov property, and the translation invariance of the process, this process has the same dynamics as the BARW driving $(v_t^1)_{t\leq h}$, but without the possibility of branching. As there is only one particle on these intervals, there is also no annihilation; $\hat{v}_t$ is just a simple symmetric random walk up until time $H$. Because of this, if $H \leq n^{\frac{9}{8}}$, we may extend $\hat{v}_t$ beyond $H$ to $n^{\frac{9}{8}}$, by allowing it to continue performing a simple symmetric random walk, independent of the walk that it has already performed, and BARW $\xi$.

Taking $r = \frac{9}{8}\log n$ and $T = n^{9/8}$ in Theorem 4.22, we can couple a Brownian motion, $(\hat{W}_t)_{t\geq 0}$, starting at $v^1_{n\gamma_n}$, to $(\hat{v}_t)_{t\leq n^{\frac{9}{8}}}$ such that,

$$P\left( \sup_{0\leq s\leq n^{\frac{9}{8}}} |\hat{W}_s - \hat{v}_s| \geq n^{\frac{9}{32}} \left(\tfrac{9}{8}\log n\right)^{\frac{3}{2}} \right) \leq Cn^{-a},$$

for some $a, C > 0$. Thus, conditioning on $H$, we can see that

$$\begin{aligned}
&P\left(\sup_{0\le s\le H} |\hat{W}_s - \hat{v}_s| \ge n^{\frac{9}{32}}\left(\tfrac{9}{8}\log n\right)^{\frac{3}{2}}\right) \\
&\qquad\qquad \le P\left(\sup_{0\le s\le H} |\hat{W}_s - \hat{v}_s| \ge n^{\frac{9}{32}}\left(\tfrac{9}{8}\log n\right)^{\frac{3}{2}} \middle| H \le n^{\frac{9}{8}}\right) + P\left(H \ge n^{\frac{9}{8}}\right) \\
&\qquad\qquad \le P\left(\sup_{0\le s\le n^{\frac{9}{8}}} |\hat{W}_s - \hat{v}_s| \ge n^{\frac{9}{32}}\left(\tfrac{9}{8}\log n\right)^{\frac{3}{2}}\right) + \exp(\kappa_n \sigma n^{\frac{1}{8}}) \\
&\qquad\qquad = o\left(n^{-a}\right).
\end{aligned}$$

After this calculation, we shall never again consider $\hat{v}_t$ beyond $H$. We artificially extended $\hat{v}_t$ to make it clear that $\hat{W}_t$ is independent of $H$, and thus also $h$.

We now construct the Brownian motion described in the lemma. Let $(\tilde{W}_t)_{t\ge 0}$ be another Brownian motion, independent of $(v_t)_{t\le H}$ and $(\hat{W}_t)_{t\ge 0}$. Let $k(t) = \sup\{k \ge 0 : \tau_k^{br} < t\}$ and define

$$W_t = \begin{cases} \tilde{W}_{t-\tau_{k(t)}^{br}+k(t)n\gamma_n} + v_0^1 - v_{n\gamma_n}^1 + \hat{W}_{\tau_{k(t)}^{br}-k(t)n\gamma_n} & \text{if } 0 \le t - \tau_k^{br} < n\gamma_n \\ \hat{W}_{t-k(t)n\gamma_n} + v_0^1 - v_{n\gamma_n}^1 + \tilde{W}_{k(t)n\gamma_n} & \text{otherwise} \end{cases}$$

where again, the extra constant terms are added only to ensure that each segment of Brownian motion starts where the previous one ended.

Firstly, it is clear that $(W_t)_{t\ge 0}$ is a Brownian motion. It is a concatenation of independent pieces of Brownian motion, at stopping times which are independent of both Brownian motions, as the branching times are independent of the position of the random walk. We may also observe that $(W_t)_{t\ge 0}$ is independent of $h$ and $k^*$. It is clearly true for $(\tilde{W}_t)_{t\ge 0}$, and we took special care to show that it is true for $(\hat{W}_t)_{t\ge 0}$ and $h$. To show the lemma, all that remains is to show the required error bound in the coupling.

By construction, we can control the error gained between $W_t$ and $v^1$ on the intervals $\bigcup_{k=1}^{k^*}[\tau_{k-1}^{br} + n\gamma_n, \tau_k^{br})$. We don't do anything particularly sophisticated to control the error on the remaining intervals, because they are such a small proportion of $h$. By Lemma 4.23, we simply observe that with high probability, $v_t^1$ doesn't travel beyond $\sqrt{n\sqrt{\gamma_n}}$ in a time interval of duration $n\gamma_n$ from where it started the interval. The same is true for $W_t$, which can be seen using the same techniques used in that lemma. Thus, during each interval of the form

$[\tau_k^{br}, \tau_k^{br} + n\gamma_n)$, with probability on the order of $o\left(n^{-1}\right)$, the extra error introduced is less than $2\sqrt{n\sqrt{\gamma_n}}$.

By (4.32) we are unlikely to have more than $(\log n)^{\frac{9}{8}}$ of such intervals, so with probability still of order $o\left(n^{-1}\right)$, the total error introduced in these intervals is less than $\sqrt{n}(\log n)^{\frac{9}{8}-\frac{c}{4}}$. Adding our two sources of error together, and observing that $h \leq H + k^* n\gamma_n$, we can conclude that with probability $1 - o\left(n^{-a\vee -1}\right)$,

$$\sup_{0\leq s\leq h} |\hat{W}_s - \hat{v}_s| \leq 2\sqrt{n}(\log n)^{\frac{9}{8}-\frac{c}{4}}.$$

Taking $r = a \wedge 1$ completes the lemma. □

Note that as $c \geq 5$, the exponent of the error in (4.33) is negative, and hence, the error goes to 0 under a diffusive rescaling.

### 4.2.3 Branching Vines

We now concatenate our vines into a ternary branching object we call a branching vine.

**Definition 4.29** (Branching vine). *Fix $x \in \mathbb{Z}^2$. We define a lifetime $\mathcal{H} > 0$, and a process $\mathcal{V}_t = (v_t^i, i \in \mathcal{U})_{0\leq t\leq \mathcal{H}}$ on $(\mathbb{Z}^2)^{\mathcal{U}}$, which we shall refer to as a* branching vine.

*Start a vine from $x$, driven by BARW $\xi$, denoting this vine $v_t^{\emptyset}$, with lifetime $h^{\emptyset}$ and setting $t^{\emptyset} = 0$. We continue inductively, assuming time $t^j$, the vine $v_t^j$ and its lifetime $h^j$ have been defined for some $j \in \mathcal{U}$. This inductive process will define a process with the genealogical structure of a Galton-Watson process, where each offspring has offspring 0 or 3. Call this process the extended branching vine.*

*If $v_t^j$ terminates in a branching event, for each $i \in \{1, 2, 3\}$, let $v^{(j,i)}$ be a vine started at the location of $(v^j)^i$ at the time $v_t^j$ terminates, which is $t^j + h^j =: t^{(j,i)}$. These vines are driven by independent copies of the BARW $\xi$ started from time $t^{(j,i)}$, and have lifetimes $h^{(j,i)}$.*

*If $v_t^j$ terminates in an overshoot event, the vine is not continued.*

*Set $\mathcal{H}$ to be the time of he earliest overshoot event in the extended branching vine. The branching vine is the truncation of the extended branching vine at time $\mathcal{H}$ (and hence is a purely ternary branching process).*

Again, from the definition, it is important to note that as distinct vines are driven by independent copies of the BARW $\xi$, as soon as two particles from distinct vines meet at the same site, the process driving the vine and the process driving the BARW may deviate from one another. Again, rather than try to control the branching vine under these events, we simply show that these events are unlikely to occur.

**Lemma 4.30.** *Fix $T > 0$. For any $a, r > 0$,*

$$P\big(\mathcal{U}(nT) \not\subseteq \mathcal{U}_{\lfloor a \log\log n \rfloor}\big) = \mathrm{o}\left((\log n)^{-r}\right).$$

*Proof.* Fix $v \in \{1,2,3\}^{\lfloor a \log\log n \rfloor}$. Then by a union bound over $\{1,2,3\}^{\lfloor a \log\log n \rfloor}$,

$$P\left(\exists w \in \{1,2,3\}^{\lfloor a \log\log n \rfloor} \text{ s.t. } t_w \le nT\right) \le 3^{\lfloor a \log\log n \rfloor} P(t_v \le nT).$$

Note that $t_v \le \sum_{i=1}^{\lfloor a \log\log n \rfloor} H_i + R$ where the $(H_i)_{i \ge 1}$ are i.i.d. with $H_1 \sim \mathrm{Exp}(\kappa_n/n)$ and $R$ is the total error incurred by the exponential approximation to the lifetime of the vine. By Lemma 4.26, we know that

$$P\left(R \ge 2n(a \log\log n)(\log n)^{\frac{9}{8}-c}\right) = \mathcal{O}\left((\log\log n)e^{-\delta(\log n)^{1/8}}\right).$$

Hence (if $n$ is sufficiently large that $2n(a\log\log n)(\log n)^{9/8-c} \le nT/2$),

$$P(t_v \le nT) \le P(\mathrm{Po}(\kappa_n T/2) \ge a \log\log n) + \mathcal{O}\left(\log\log n) e^{-\delta(\log n)^{1/8}}\right).$$

By a Chernoff bound argument, if $\Xi \sim \mathrm{Po}(\chi)$, then for $k > \chi$,

$$P(\Xi > k) \le \frac{e^{-\chi}(e\chi)^k}{k^k}.$$

Thus,

$$P(\mathrm{Po}(\kappa_n T/2) \ge a\log\log n) \le \frac{e^{-\kappa_n T/2}(e^{\kappa_n T/2})^{a\log\log n}}{a \log\log n^{a \log\log n}} = o\left(\log\log n^{-a\log\log n}\right).$$

We conclude that, for any $r > 0$,

$$P\big(\mathcal{U}(nT) \not\subseteq \mathcal{U}_{\lfloor a \log\log n \rfloor}\big) = o\left((\log n)^{-r}\right).$$

$\square$

We can now couple the branching vine and the BARW up to a fixed time.

**Lemma 4.31.** *Fix $T > 0$ and $x \in \mathbb{Z}^2$. Let $\xi^{(n)}$ be a BARW starting with a unique particle at $x$. Then $(\mathcal{V}_t)_{0 \le t \le nT}$ and $(\xi)_{0 \le t \le nT}$,* viewed as collections of paths*, are equal with probability at least $1 - \mathcal{O}\left((\log n)^{-1/4}\right)$.*

*Proof.* Set $a = (4 \log 3)^{-1}$. For $j \in \mathcal{U}$, set the number of branching events in $v_{t-t_j}$ before $h_j$ to be $k^*(j)$. By a union bound over $\mathcal{U}_{\lfloor a \log\log n \rfloor}$ and (4.32),

$$P\left(\exists j \in \mathcal{U}_{\lfloor a \log\log n \rfloor} : k^*(j) \ge (\log n)^{9/8}\right) \le 3^{1 + a \log\log n} \mathcal{O}\left(e^{-\delta (\log n)^{1/8}}\right) = \mathcal{O}\left(e^{-\delta (\log n)^{1/8}/2}\right) \tag{4.34}$$

As there are $\frac{1}{2}(3^{n+1} - 1) < 3^{n+1}$ elements in $\mathcal{U}_n$.

Let $(\tau_k^{br}(j))_{k \ge 1}$ denote the sequence of branching times in the vine $v_{t-t_j}$, and similarly define $(\tau_k^{type}(j))_{k \ge 1}$ and $(\tau_k^{over}(j))_{k \ge 1}$. Note that the BARW and the branching vine only differ as collections of paths if either a branching event occurs during a time interval in which the vine ignores branching, or if two different vines meet and an annihilation event occurs. More formally, if $(\mathcal{V}_t)_{t \le nT}$ and $(\xi_t)_{t \le nT}$ differ as collections of paths then one or more of the following events occurs.

1. $\mathcal{U}(nT) \not\subseteq \mathcal{U}_{\lfloor a \log\log n \rfloor}$ or $k^*(j) \ge (\log n)^{9/8}$ for some $j \in \mathcal{U}_{\lfloor a \log\log n \rfloor}$.
2. For some $j \in \mathcal{U}_{\lfloor a \log\log n \rfloor}$ and $k \le (\log n)^{9/8}$, the event $G_1(j,k)$ occurs: one of the vines $(v^j)^1_{t-t_j}$, $(v^j)^2_{t-t_j}$ or $(v^j)^3_{t-t_j}$ is hit by a branching event in the time interval $[\tau_k^{br}(j), \tau_k^{br}(j) + n(\log n)^{-c}]$.
3. For some $j \in \mathcal{U}_{\lfloor a \log\log n \rfloor}$, the event $G_2(j)$ occurs: $\tau^{type}(j) = \tau^{over}(j)$, i.e., the vine $(v^j)_{t-t_j}$ ends in an overshooting event.
4. For some $w \ne v \in \mathcal{U}_{\lfloor a \log\log n \rfloor}$, the event $G_3(v,w)$ occurs: there are $i_1, i_2 \in 1,2,3$ with $(v^j)^{i_1}_{t-t_w} = (v^j)^{i_2}_{t-t_v}$ for some $t \le nT$.

Recall from the definition that branching events hit a single particle with rate $\sigma_n = \frac{(\log n)^3}{n}$. Hence for $k \in \mathbb{N}$ and $j \in \mathcal{U}$, $P(G_1(j,k)) = \mathcal{O}\left((\log n)^{3-c}\right)$.

By Lemma 4.27, $P(G_2(j,k)) = \mathcal{O}\left((\log n)^{-19/8}\right)$.

For $w \neq v \in \mathcal{U}$, let $i = \min\{j \geq 1 : w_j \neq v_j\}$. Then let

$$w \wedge v = \begin{cases} (w_1, \dots, w_{i-1}) & \text{if } i \geq 2, \\ \emptyset & \text{if } i = 1. \end{cases}$$

At time $h_{w\wedge v}$, either $\tau^{type}_{k^*(w\wedge v)}(w\wedge v) = \tau^{over}_{k^*(w\wedge v)}(w\wedge v)$ or $\tau^{type}_{k^*(w\wedge v)}(w\wedge v) = \tau^{div}_{k^*(w\wedge v)}(w\wedge v)$. In the latter case, for $i_1 \neq i_2 \in \{1,2,3\}$, $|x(w\wedge v, i_1) - x(w \wedge v, i_2)| \geq \sqrt{n}(\log n)^{-c}$. Thus, these particles and their descendants which are ancestors of either $(v^j)^{i_1}_{t-t_w}$ or $(v^j)^{i_2}_{t-t_v}$ perform a continuous time random walk started at time $h(w \wedge v)$ with initial displacement at least $(\log n)^{-c}$. Hence by Lemmas 4.24 and again by Lemma 4.27, $P(G_3(w,v)) = \mathcal{O}\left(\frac{\log\log n}{\log n}\right)$.

By a union bound, and using Lemma 4.30 and 4.34, it follows that for $r > 0$,

$$\begin{aligned} &P((\mathcal{V}_t)_{t\leq nT} \neq (\xi_t)_{t \leq nT}) \\ &\qquad \leq o\left((\log n)^{-r}\right) + 3(\log n)^{a\log 3 + 9/8} P(G_1(j,k)) + 9(\log n)^{2a\log 3} P(G_3(w,v)) \\ &\qquad = \mathcal{O}\left((\log n)^{a \log 3 + \frac{9}{8} + 3 - c}\right) + \mathcal{O}\left((\log\log n)(\log n)^{2a\log 3 - 1}\right) \\ &\qquad = \mathcal{O}\left((\log n)^{-1/4}\right), \end{aligned}$$

by our choice of $a = (4\log 3)^{-1}$ and since $c \geq 5$. ☐

The last step before we can prove Theorem 4.1 is to couple a TBBM to the branching vine. As we have already coupled the BARW with the branching vine in previous lemma, Theorem 4.1 follows immediately. As before, denote the set of particles existing in the branching vine $\mathcal{V}$ at time $t$ by $\mathcal{U}_t(\mathcal{V})$.

**Lemma 4.32.** *Fix $T > 0$. There exists an $n \in \mathbb{N}$, such that we can couple the branching vine, $\mathcal{V}$, driven by $\xi_n$ and a TBBM, $W$, with branching rate $\frac{\kappa_n}{n}$ started with a single initial particle at $x$, such that, with probability $1 - \mathcal{O}\left((\log n)^{-1/4}\right)$, the events*

$$\{\forall l \in (\mathcal{U}_{nT}(\mathcal{V}),\ l \in \mathcal{U}_{nT}(W) \text{ and } \sup_{t \in [0,nT]} |\xi^l_t - W^l_t| \leq \epsilon\},$$

*and*

$$\{\forall l \in \mathcal{U}_{nT}(W),\ l \in \mathcal{U}_{nT}(\mathcal{V}) \text{ and } \sup_{t\in[0,nT]} |\xi^l_t - W^l_t| \le \epsilon\}$$

*hold with* $\epsilon = 4\sqrt{n}(\log\log n)(\log n)^{-1}$.

*Proof.* Set $c = 17/2$. We couple $(\mathcal{V}_t)_{0\le t\le nT}$ to a branching Brownian motion with branching rate $\kappa_n/n$. Let $((W^j_t)_{t\ge0}, H_j)_{j\in\mathcal{U}}$ be an i.i.d. sequence, where $(W^j_t)_{t\ge0}$ is a Brownian motion starting at 0 and $H_j \sim \mathrm{Exp}(\kappa_n/n)$ independent of $(W^j_t)_{t\ge0}$. For each $j \in \mathcal{U}$, we couple $(v_{t-t_j})_{t\in[t_j,h_j]}$ to $((W^j_t)_{t\ge0}, H_j)$ as in Lemmas 4.26 and 4.28.

For $j \in \mathcal{U}$, let $A_1(j)$ be the event that both $|(h_j - t_j) - H_j| \le 2n(\log n)^{9/8-c/4} = 2n(\log n)^{-1}$ and for $i = 1, 2, 3$ and $t \in [t_j, h_j]$,

$$|(v^i_{t-t_j} - x_j) - W^j_{t-t_j}| \le 2\sqrt{n}(\log n)^{9/8-c/4} = 2\sqrt{n}(\log n)^{-1}.$$

By Lemmas 4.26 and 4.28, for any $r > 0$, for each $j \in \mathcal{U}$, $P(A_1(j)) \ge 1 - \mathcal{O}\left((\log n)^{-r}\right)$. Hence, taking a union bound over $j \in \mathcal{U}_{\lfloor \log\log n\rfloor}$,

$$P\left(\cap_{j\in\mathcal{U}_{\lfloor \log\log n\rfloor}} A_1(j)\right) \ge 1 - \mathcal{O}\left((\log n)^{\log 3 - r}\right).$$

There is a small discrepancy between the lifetime of the vine and the lifetime of the Brownian motion segment, but we do not know if this occurs at the beginning of the lifetime of the Brownian motion or the end. It costs us little to account for the error at both ends. So, for $j \in \mathcal{U}$, define the event

$$A_2(j) = \left\{ \sup_{t\in[0,2n(\log n)^{-1/8}]} |W^j_t| + \sup_{t\in[H_j - 2n(\log n)^{-31/8}, H_j]} |W^j_t - W^j_{H_j}| \le \sqrt{n}(\log n)^{-1} \right\}.$$

Then by another union bound over $\mathcal{U}_{\lfloor \log\log n\rfloor}$, since for a Brownian motion $(W_t)_{t\ge0}$ started at 0, $P\left(\sup_{t\in[0,2n(\log n)^{-31/8}]} |W_t| \ge \frac{1}{2}\sqrt{n}(\log n)^{-1}\right) = o\left((\log n)^{-r}\right)$, we have that

$$P\left(\cap_{j\in\mathcal{U}_{\lfloor \log\log n\rfloor}} A_2(j)\right) \ge 1 - \mathcal{O}\left((\log n)^{\log 3 - r}\right).$$

By Lemma 4.26, $P\left(\mathcal{U}(nT) \not\subseteq \mathcal{U}_{\lfloor \log\log n\rfloor}\right) = o\left((\log n)^{-r}\right)$.

Define a branching Brownian motion starting at $x$ by $((W^j_t)_{t\ge0}, H^j)_{j\in\mathcal{U}})$ by letting the increments of the initial particle be given by $(W_t^{\emptyset})_{t\geq 0}$ until time $H^{\emptyset}$, when it is replaced by three particles which have lifetimes $H^1$, $H^2$ and $H^3$ and increments given by $(W_t^1)_{t\geq 0}$, $(W_t^2)_{t\geq 0}$ and $(W_t^3)_{t\geq 0}$, and so on. If $\mathcal{U}(nT) \not\subseteq \mathcal{U}_{\lfloor \log\log n\rfloor}$ and $A_1(j) \cap A_2(j)$ occurs for each $j \in \mathcal{U}_{\lfloor \log\log n\rfloor}$, each path in the branching vine stays within distance

$$\epsilon = 4\sqrt{n}(\log\log n)(\log n)^{-1}$$

of some path through the TBBM and vice versa.

Setting $r = \log 3 + 1/4$ gives us a coupling between the branching vine and branching Brownian motion with branching rate $\kappa_n/n$ such that with probability at least $1-\mathcal{O}\left((\log n)^{-1/4}\right)$, up to time $nT$ each particle in the BARW stays within distance $\epsilon$ of some path through the TBBM and vice versa. Fix $n$ sufficiently large so that these bounds hold. $\square$

The proof of Theorem 4.1 is now just a consequence of Lemma 4.32 and Lemma 4.31.

Theorem 4.1 concerns a BARW starting with a single particle. This theorem can be relatively straightforwardly extended to a BARW starting with an odd finite number of particles, however, doing so formally would be rather uninsightful, and notationally tedious. We finish this chapter with an informal description on how this could be done.

First, consider if the initial particles start further than $\sqrt{n\sqrt{\gamma_n}}$ apart from each other. In this case, we could follow exactly the same lines as the proof of Theorem 4.1, as Lemma 4.23 will still hold for any pair of particles, and it would be as if we had started with one particle, and waited until we had the new initial number of particles.

If this was not the case, we could adjust the proof of Theorem 4.1 so that we replace $\gamma_n$ with $C\gamma_n$, where $C$ is large enough so that all particles are within $C\sqrt{n\sqrt{\gamma_n}}$ of each other. As there are only a finite number of particles, $C$ is a constant, and so this modification does not change the asymptotics of the proof.

By arguments similar to Corollary 4.18, with sufficiently high probability, if we allow the BARW to evolve for a duration of $Cn\gamma_n$, these particles will either annihilate until a single particle remains, or if multiple particles remain, they have diverged further than $C\sqrt{n\sqrt{\gamma_n}}$ apart. Hence, we can proceed as if we were in the previous case.

# Chapter 5

# Conclusions and future research

## 5.1 Regarding the spatial $\Lambda$-Fleming-Viot process

We were successfully able to adapt the work of [13] to the asymmetric case in Chapter 2, which allowed us to investigate the long term behaviour of a diploid population where one homozygote had a selective advantage over the other, but both were selected over the heterozygote. When the difference in selection between the homozygotes is small, as expected, the long term behaviour does not differ much from [13], who show that the rescaled boundary between the two homozygotes evolves according to mean curvature flow.

However, when the difference is large enough, we see a new behaviour emerging. Rather than evolve according to mean curvature flow, the boundary simply moves at a constant rate toward the less fit homozygote. Furthermore, the asymmetry can be chosen so that both of these behaviours are present in the limit. This means it is possible for a stationary (albeit unstable) solution to occur, where the fitter homozygote is surrounded by the other homozygote. Mean curvature flow would cause this smaller population to shrink to a point, while the selective advantage would be trying to expand it. This model alone seems too simple to investigate further the long term coexistence between two competing populations, but could be used as a toy example, for other models to build on in the future.

Although it wouldn't change our results, it would be nice to find a way to use the branching Brownian motion dual to the asymmetric Allen-Cahn equation to directly prove Theorem 2.11, rather than relying on the explicit solution to the Allen-Cahn equation. It would not be as biologically relevant, but more complicated selection mechanisms could be used, as opposed to biased majority ternary voting. These mechanisms would correspond to versions of the Allen-Cahn equation with more complicated non-linear terms (i.e., higher order polynomials), so this duality could be used to investigate a more general Allen-Cahn equation. These more general equations won't *a priori* have a simple explicit solution, so a more general way to prove Theorem 2.11 would be needed. However, the rest of the argument should lift pretty smoothly to this more general case once this is proven.

These results on the limiting behaviour of the interacting homozygotes are interesting, but they are deterministic limits, so investigating the fluctuations about these limits is a natural next step in this research. We made progress on this question in Chapter 3. We restricted ourselves to one spatial dimension, but we were able to show that the, appropriately rescaled, hybrid zone between the two zones of differing homozygotes fluctuates like a Brownian motion, with a drift toward the less fit homozygote, proportional to the the strength of the asymmetric selection between the homozygotes.

One way to extend this research would be to increase the magnitude of the asymmetry we analyse. We exploited the work done by [20], and were able to prove our result without changing the functional (3.10) which we used to measure the distance of our solutions from the manifold $M$. However, rather than compare solutions to the asymmetric stochastic Allen-Cahn equation to the stationary solution, it should be possible to compare them directly to an appropriate travelling wave solution.

Doing this would most likely involve repeating the calculations in [20], but for a moving reference frame. A moving reference frame should allow us to use Katzenberger's idea, in (3.97), to eliminate the entire second term in (3.95), the Itô derivative of $\xi_t^\epsilon$, as happens in the symmetric case. We control this term by ensuring our asymmetry, $\lambda$, decays sufficiently quickly as to take this whole term to zero. Without this, we could have proven our result for greater (but still asymptotically decaying) values of asymmetry.

Extending this argument to higher dimensions is very much an area of active research. Clearly, the stochastic Allen-Cahn equation with the type of noise we have considered cannot be extended to more than one dimension, as it cannot be properly defined. Despite this, Barton outlines a heuristic attempt in [3]. Fortunately, the branching and coalescing dual outlined in the proof to Lemma 3.7 can easily be extended to higher dimensions, so this could be used to make the ideas in [3] more rigorous.

## 5.2 Regarding the branching annihilating random walk

There are also many avenues of research remaining for the BARW. Our main result for the BARW is Theorem 4.1, which couples a two-dimensional BARW and a ternary branching Brownian motion for an arbitrary amount of time. This is a promising step in proving part two of Conjecture 1.10; that in two dimensions and starting with an even number of particles, the BARW persists for all time with positive probability for any positive branching rate.

One obvious issue is that our results assume an odd initial number of particles, and the conjecture assumes an even number. However, we believe that showing that the number of particles in a BARW tends to infinity when starting from an odd number of particles, should be equivalent to showing that a BARW starting with an even number of particles persists for all time. We can see this through the duality.

*Heuristic proof.* Starting with $x_i(0) = \frac{1}{2}$ for all $i \in \mathbb{Z}^2$, if $N(t)$ starts from an odd number of particles and tends to infinity, for a particular branching rate, then we would expect

$$\lim_{t\to\infty} \mathbb{E}\Big[\underline{x}(0)^{\underline{N}(t)}\Big] = 0,$$

recalling that $\underline{x}^{\underline{N}} := \prod_{i\in\mathbb{Z}^d} x_i^{N_i}$. Hence, by Lemma 1.8, we would also expect

$$\lim_{t\to\infty} \mathbb{E}\Big[\underline{x}(t)^{\underline{N}(0)}\Big] = 0.$$

Now we define a new initial configuration, identical to $\underline{N}(0)$ but with a single additional particle, denoted $\underline{\bar{N}}(0)$. As $x_i(t) \in [-1, 1]$, for all $i \in \mathbb{Z}^2$, we must have that $\underline{x}(t)^{\underline{\bar{N}}(0)} \leq \underline{x}(t)^{\underline{N}(0)}$, and hence

$$\lim_{t\to\infty} \mathbb{E}\Big[\underline{x}(t)^{\underline{\bar{N}}(0)}\Big] = 0.$$

Thus, applying the duality again gives us that the number of particles in the BARW started from $\underline{\bar{N}}(0)$, which has an even initial number of particles, tends to infinity, and hence survives for all time. The reason this proof is heuristic is because one would need to show that the duality still holds in the limit as $t \to \infty$. $\square$

There are two more key lemmas that would need to be proven in order to prove this conjecture in the affirmative. The first is to show that, when started from a distribution that looks like a ternary branching Brownian motion, the BARW survives for all time. The second is a monotonicity result, that if a BARW has a strictly positive probability of survival for some branching rate, then it has a strictly positive probability of survival for any greater branching rate. It is not clear how the results of this thesis could be used to make progress on this second point. However, if the first additional lemma could be proven, it would be mean that the conjecture holds for sufficiently small branching rates. Combined with the result from [6], that the BARW survives for sufficiently large branching rates, it would be very unusual if the conjecture did not hold for all branching rates.

To prove this former lemma, it is very tempting to use a percolation argument. First, we evolve the BARW until the coupling with ternary branching Brownian motion breaks down. We have a good understanding of the distribution of particles in a branching Brownian motion, and by taking the branching rate to be ever smaller, we can choose the number of particles in our process at this time, so we can have very good control over where these particles are.

The next step is to look at the very outermost frontier of the branching Brownian motion. We define this frontier in such a way that the density of particles is sufficiently low, so that we can divide this frontier in to boxes, in (roughly) an annulus around the edge of the branching Brownian motion. We can then perform a percolation in space and time on boxes of density sufficiently low for the coupling between the BARW and the branching Brownian motion to hold with high probability. If the coupling holds, the BARW cannot go extinct in such box, and hence if the system percolates, the BARW survives for all time.

We can choose the length of the space-time boxes so that each box of low density has a good chance of 'infecting' enough of the empty boxes further from the frontier before the coupling breaks down. So it seems like we have enough freedom to choose the parameters in a way that makes this argument work. However, the one thing were not able to rigorously control was the rate of infection of the frontier from the high density boxes. We can heuristically control it though.

*Heuristic proof.* Heuristically, infection of the frontier from these high density boxes should not overwhelm the infection from the lower density boxes. Although there are many more particles in the high density boxes, these particles are much more likely to annihilate with one another than make it to the frontier. We would expect there to be some 'equilibrium density' for these high density boxes, where the number of branches matches the number of annihilations. In this region of equilibrium density, only particles near the edge would have any chance of escaping, whereas in the region of low density, any particle can escape.

For the rate of branching events to match the rate of annihilation events in this region, we would expect the number of particles per site, and hence the equilibrium density, to be proportional to the branching rate (as we would expect $\sigma n \approx \frac{n(n-1)}{2}$). Hence, if the radius of the high density region is $R$, the infection rate from the high density region is on the order of $\sigma R$. On the other hand, if the low density regions have density $d < \sigma$, but particles from the whole annulus (say of radius $r$) of boxes at the frontier can escape from the frontier, then the infection rate is of order $drm$, where $m$ is the number of sites in the width of the annulus. If we assume the radii of these two regions are of the same order, we should be able control the infection from the high density region by choosing $m \gg \sigma/d$. □

Making this argument, or any similar heuristic, rigorous has proven difficult. It would be helpful if we could make more precise statements about this equilibrium density. A result in [2] seemed like it could be helpful, however we could not adapt it sufficiently. It shows that the distribution of particles in a binary branching and annihilating random walk is the same as a Poisson thinning of a binary branching and coalescing random walk. The distribution of particles for a coalescing and binary branching random walk is not obvious, but at least this process has monotonicity that our BARW does not.

Further evidence in support of the heuristic is that we can observe a relatively stable equilibrium density in simulations of a BARW. Figure 5.1 shows the number of particles over time in a two-dimensional BARW within boxes at various locations. While the density grows rapidly near the frontier, it stays relatively away from the frontier.

Even without this additional lemma, Theorem 4.1 by itself provides strong evidence in support of this conjecture, especially the key observation that the coupling gets *better* as the branching rate decreases.

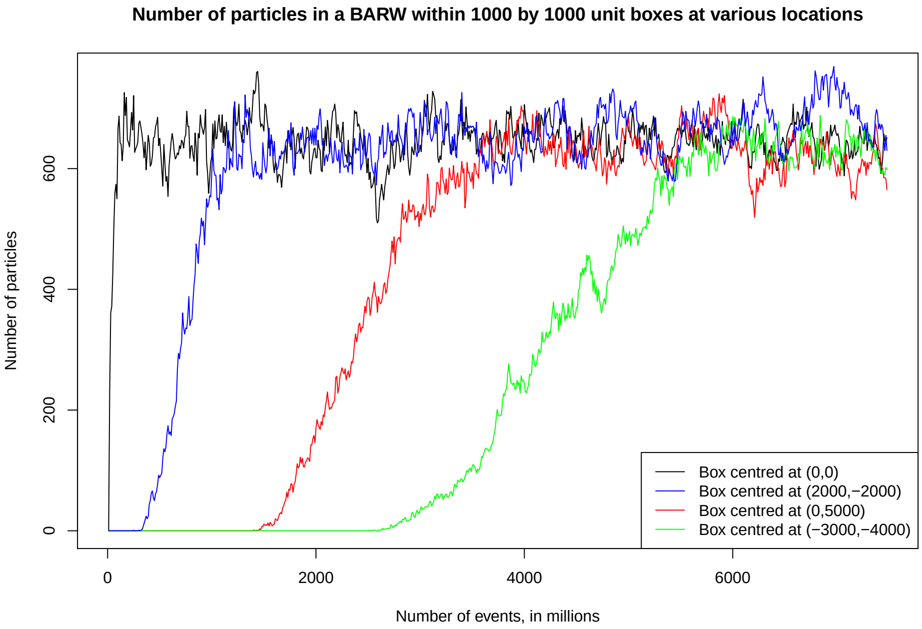

Figure 5.1: The number of particles in a simulation of a BARW within 1000 by 1000 lattice unit boxes, centred at various locations in space. The horizontal axis is the number of events that have occurred, rather than time. An event is any migration, branch or annihilation.

There are the two other cases in Conjecture 1.10: the single and at least three dimensional cases. The ideas regarding the suppression of branching events may be useful when investigating the one-dimensional case, however, as the conjecture in this case is that the BARW does go extinct for sufficiently small branching rates, the strategy of coupling the process to a branching Brownian motion, which never goes extinct, won't be as helpful. This strategy could be helpful in the higher-dimensional case, as again we are trying to prove positive probability of non-extinction. To use this strategy in higher dimensions, though, the estimates from Section 4.1 would have to be revised. However, perhaps a better strategy for this case is to focus on the monotonicity argument, as the transience of the a random walk in greater than three dimensions allows us to very simply prove the conjecture in the case where the branching rate is 0.

Finally, [31] recently generalised and refined Proposition 1.3 of [10] (the analogue of Theorem 4.2) to the correlation functions for an infinite system of instantaneously coalescing planar simple random walks, started with all sites occupied. It would be interesting to see if the methods from Section 4.1 could be used to extend this result to delayed coalescing random walks.

# Appendix A

# R code from Lemma 3.30

We include the R code used to generate the images in the proof of Lemma 3.30.

```r
#This code implements the algorithm described in the proof of Lemma 3.30
#First, we implement a few functions.
#The first two are mathematical functions, one for our function phi, and one which calculates the R_i from a_i
phi<-function(a,b,theta){a*cosh(theta)+b*sinh(theta)}
Rfunction<-function(a,eps,beta){
eps+beta^2/eps^2+log((2+30*eps-a+sqrt(3*(1-2*a+60*eps)))/(1+a-30*eps))}

#This is the main function which implements the algorithm. It outputs a vector of points of G for lambda in [-0.6, 0.1]
Glambda<-function(a,eps,beta){
  #a is a vector describing our set of a_i, and eps and beta are the values for our constants epsilon and beta
  N=length(a) #Must be greater than 1 for this function.
  #Given our a_i, we need to calculate the radii of the approximating Z sets
  R=Rfunction(a,eps,beta)

  #We store lambda as a vector.
  #To simplify expressions, we create a matrix of lambda shifted by 1/2-a_i
  lambda=as.complex((-600:100)/1000)
  n=length(Re(lambda))
  L=lambda
  for(i in 1:N){L=c(L,lambda+1/2-a[i])}
  L=t(matrix(L,n,N+1))

  #We also initialise matrices which will store Ai/An
  A=matrix(rep(0,N*n),N,n)
  B=matrix(rep(0,N*n),N,n)

  #We do the first boundary case of our iteration. These equations
```

```
    correspond to equations (3.81).
  #Note that we have absorbed the exponential factor into the constants A_N and B_N
  A[N,]=phi(1,sqrt(L[N+1,]/L[N,]),R[N]*sqrt(2*L[N,]))
  B[N,]=-phi(sqrt(L[N+1,]/L[N,]),1,R[N]*sqrt(2*L[N,]))

  #Now we perform the iteration
  for(i in (N-1):1){
    #We have three expressions we calculate separately to make the iterations easier to read
    SqrtLl=sqrt(L[i+1,]/L[i,])
    Rili=R[i]*sqrt(2*L[i,])
    RiLi1=R[i]*sqrt(2*L[i+1,])
    #We now calculate the main iterants
    A[i,]=phi(phi(A[i+1,],B[i+1,],RiLi1),-SqrtLl*phi(B[i+1,],A[i+1,],RiLi1),Rili)
    B[i,]=phi(SqrtLl*phi(B[i+1,],A[i+1,],RiLi1),-phi(A[i+1,],B[i+1,],RiLi1),Rili)
  }

  #We can now use equation (3.82) to find G(lambda), which we output from the function.
  -A[1,]/(sqrt(2*lambda)*B[1,])}

#We now run our function for the following input parameters
eps=0.0001
beta=0.000001
a=(49:1)/50*(3/2)-1
lambda=(-600:100)/1000

#Before plotting G, we first visualise the approximation we are using.
#We plot the function we are trying to approximate
x=(-2000:2000)/400
m=1/(1+exp(-x))
f=-1+6*m-6*m^2
pdf('SetApproximation.pdf')
plot(x,f,type='l')
#Next we plot our approximation.
#We calculate the radii of the approximating Z sets
R=Rfunction(a,eps,beta)
N=length(a)
#We construct k, the sum of indicator functions by adding one set at a time
k=1/2*(x<=R[1])*(x>=-R[1])
for(i in 1:(N-1)){
  k=k+a[i]*(x>R[i])*(x<=R[i+1])+a[i]*(x< -R[i])*(x>=-R[i+1])
}
k=k+a[N]*(x>R[N]) +a[N]*(x< -R[N])
```

```r
lines(x,k,col=’blue’)
title("f’(m) and its approximation by a sum of indicator functions")
dev.off()

#We now plot G, on a new set of axes.
pdf(’LaplaceTransform.pdf’)
plot(lambda,Re(Glambda(a,eps,beta)),type=’l’,col=’blue’,ylim=c(-100,100),ylab=’G’,xlab=’lambda’)
title("Laplace transform of g(t)")
lines(lambda,Im(Glambda(a,eps,beta)),col=’red’,)
dev.off()

#Finally, to show that this algorithm has been implemented correctly, we compare it to the output from a known case.
#We change the set of a_i. We would ideally like to use a single a, but need to use two, for the function G to work.
#This approximation is mathematically identical to using a single a=0.
a=c(0,0)

#We again visualise the approximation we are using.
pdf(’SingleSetApproximation.pdf’)
plot(x,f,type=’l’)
r=Rfunction(a,eps,beta)
N=length(a)
#We construct k, the sum of indicator functions by adding one set at a time
k=1/2*(x<=r[1])*(x>=-r[1])
for(i in 1:(N-1)){
  k=k+a[i]*(x>r[i])*(x<=r[i+1])+a[i]*(x< -r[i])*(x>=-r[i+1])
}
k=k+a[N]*(x>r[N]) +a[N]*(x< -r[N])
lines(x,k,col=’blue’)
title("f’(m) and its approximation by a single indicator function")
dev.off()

#We now plot the simple G, on a new set of axes.
pdf(’SimpleLaplaceTransform.pdf’)
plot(lambda,Re(Glambda(a,eps,beta)),col=’blue’,ylim=c(-100,100),ylab=’G’,xlab=’lambda’,pch=4)
title("Laplace transform with simple approximation")
lines(lambda,Im(Glambda(a,eps,beta)),col=’red’,type=’p’,pch=4)

#Finally, we need to add the known formula
a=0
R=r[1]
lambda=as.complex(lambda)
```

```
110 Borodin=(sqrt(1/2/lambda)*(1-exp(-sqrt(2*lambda)*R)*(sqrt(1+2*lambda-2*a)-sqrt(2*lambda))/phi(sqrt(1+2*lambda-2*a),sqrt(2*lambda),sqrt(2*lambda)*R)))
111 lines(lambda,Re(Borodin),col='black')
112 lines(lambda,Im(Borodin),col='cyan')
113 dev.off()
```

As indicated in the code, we can compare the output of this program to a known formula for the moment generating function for the occupation time of a Brownian motion pinned at the origin in the complement of a set centred at 0. Borodin and Salminen calculate an expression for

$$E_x\left[\exp\left(-\gamma\int_0^\tau \mathbb{1}(W_r\in[r,\infty))dr\right);|W_\tau|\in dz\right], \tag{A.1}$$

where $\tau$ is an exponential random variable with parameter $\lambda$ and $W_r$ is a Brownian motion starting at 0, see equation 3.1.4.5 on page 336 of [35]. The expression, for $z, x \le r$ is

$$\frac{\sqrt{\lambda}}{\sqrt{2}}\left(e^{-|z-x|\sqrt{2\lambda}}+e^{-|z-x|\sqrt{2\lambda}}-\frac{2e^{-r\sqrt{2\lambda}}(\sqrt{\lambda+\gamma}-\sqrt{\lambda})}{\sqrt{\lambda}\sinh(r\sqrt{2\lambda})+\sqrt{\lambda+\gamma}\cosh(r\sqrt{2\lambda})}\right). \tag{A.2}$$

Conditioning on the value of $\tau$, (A.1) becomes

$$\begin{aligned}E_x&\left[\int_0^\infty \lambda e^{-\lambda t}\exp\left(-\gamma\int_0^t \mathbb{1}(W_r\in[r,\infty))dr\right);|W_\tau|\in dz\right]\\&=2E_x\left[\int_0^\infty \lambda e^{-\lambda t}\exp\left(-\gamma\int_0^t \mathbb{1}(W_r\in[r,\infty))dr\right);W_\tau\in dz\right]\\&=2\lambda G(\lambda).\end{aligned}$$

Hence, after substituting in $z = x = 0$ into (A.2), we find that

$$G(\lambda)=\frac{1}{\sqrt{2\lambda}}\left(1-\frac{e^{-r\sqrt{2\lambda}}(\sqrt{\lambda+\gamma}-\sqrt{\lambda})}{\sqrt{\lambda}\sinh(r\sqrt{2\lambda})+\sqrt{\lambda+\gamma}\cosh(r\sqrt{2\lambda})}\right), \tag{A.3}$$

which is the expression on line 110 of this r code, with $\gamma = 1/2 - a$.

As can be seen in Figure A.1, these functions are essentially indistinguishable, which gives a strong indication that the algorithm has been implemented correctly. Furthermore, it is obvious that this simple approximation, with a single set, is insufficient for the needs of Lemma 3.30, as the pole for this approximation occurs at sightly less than $-0.2$, when less than $-1/3$ is necessary.

It was thought that this method might be able to prove that $q_t(x; v)$ was bounded by a constant, rather than merely bounding its exponential growth. However, increasingly fine approximations don't seem to give a lower bound to the pole much greater than -0.37. This of course doesn't prove that $q_t(x; v)$ cannot be bounded by a constant, but does suggest it might not be. If the PDE corresponding to the Feynman-Kac formula applied to $q_t(x; v)$ directly could be solved, it would be very interesting to see if that really does grow exponentially, or is bounded.

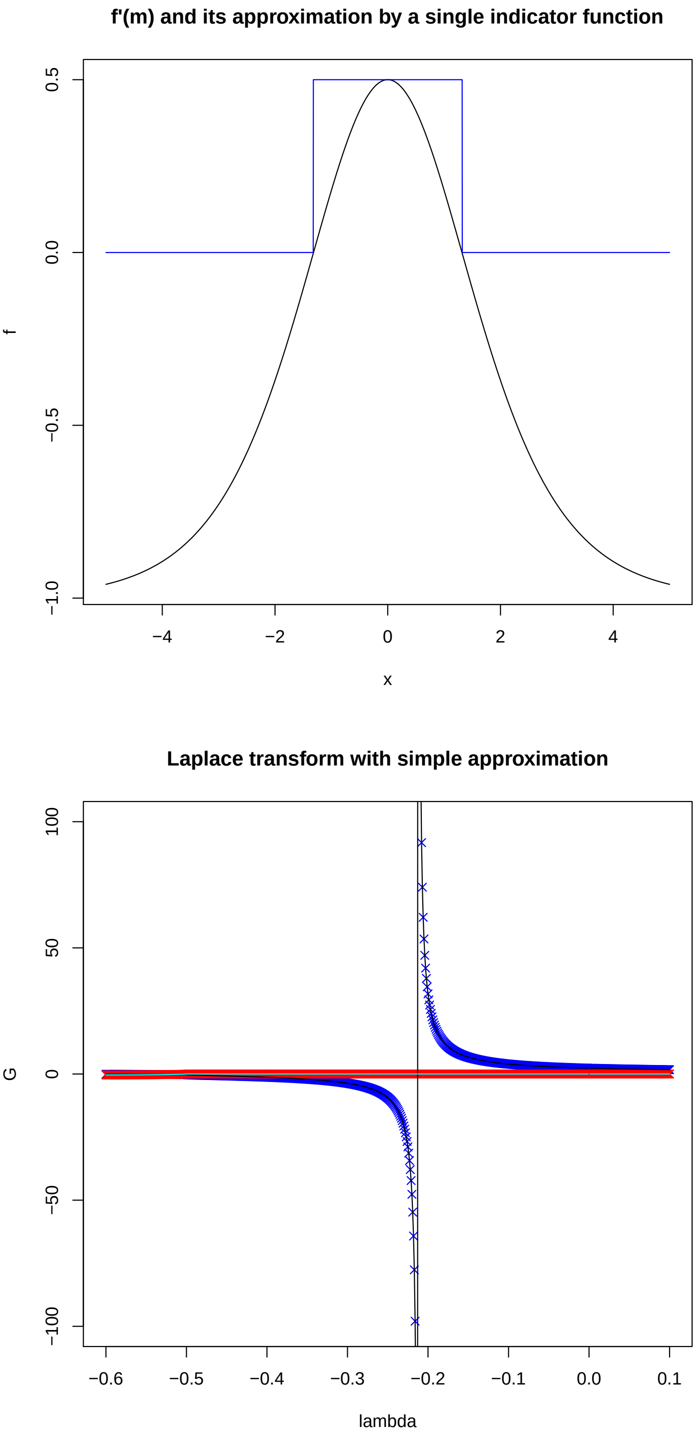

Figure A.1: *Top*: Comparison of the approximation $k = \frac{1}{2}\mathbb{1}(f'(m(x)) \geq 0)$ (in blue) to $f'(m)$ (in black). *Bottom*: Comparison of $G(\lambda)$ calculated using the r function Glambda (plotted in crosses, real part in blue and imaginary part in red), and using (A.3) (in lines, real part in black, imaginary part in cyan), for the same approximation as in the left figure.

# Bibliography

[1] S. M. Allen and J. W. Cahn, *A microscopic theory for antiphase boundary motion and its application to antiphase domain coarsening*, Acta Metallurgica, 27 (1979), pp. 1085–1095.

[2] S. Athreya and J. Swart, *Systems of branching, annihilating, and coalescing particles*, Electron. J. Probab., 17 (2012).

[3] N. H. Barton, *The dynamics of hybrid zones*, Heredity, 43 (1979), p. 341.

[4] N. H. Barton, A. M. Etheridge, and A. Véber, *A new model for evolution in a spatial continuum*, Electron. J. Probab., 15 (2010), pp. 162–216.

[5] J. Berestycki, *Topics on branching Brownian motion*, August 2014.

[6] J. Blath, A. Etheridge, and M. Meredith, *Coexistence in locally regulated competing populations and survival of branching annihilating random walk*, Ann. Appl. Probab., 17 (2007), pp. 1474–1507.

[7] B. M. Bolker and S. W. Pacala, *Spatial moment equations for plant competition: understanding spatial strategies and the advantages of short dispersal*, Amer. Naturalist, 153 (1999), pp. 575–602.

[8] L. Bronsard and R. V. Kohn, *Motion by mean curvature as the singular limit of Ginzburg-Landau dynamics*, J. Differential Equations, 90 (1991), pp. 211–237.

[9] X. Chen, *Generation and propagation of interfaces for reaction-diffusion equations*, J. Differential equations, 96 (1992), pp. 116–141.

[10] J. T. Cox, M. Merle, and E. A. Perkins, *Coexistence in a two-dimensional Lotka-Volterra model*, Electron. J. Probab., 15 (2010), pp. 1190–1266.

[11] E. Dogan-Ciftci and E. J. Allen, *Derivation of several SDE systems in one-and two-locus population genetics*, Stoch. Anal. Appl., 32 (2014), pp. 761–775.

[12] A. Etheridge, *Drift, draft and structure: some mathematical models of evolution*, Banach Center Publications, 1 (2008), pp. 121–144.

[13] A. Etheridge, N. Freeman, and S. Penington, *Branching Brownian motion, mean curvature flow and the motion of hybrid zones*, Electron. J. Probab., 22 (2017), pp. 1–40.

[14] A. Etheridge, N. Freeman, S. Penington, and D. Straulino, *Branching Brownian motion and selection in the spatial Lambda-Fleming-Viot process*, Ann. Appl. Probab., 27 (2017), pp. 2605–2645.

[15] A. Etheridge, A. Véber, and F. Yu, *Rescaling limits of the spatial Lambda-Fleming-Viot process with selection*, arXiv preprint arXiv:1406.5884, (2014).

[16] A. M. Etheridge, *An introduction to superprocesses*, vol. 20, American Mathematical Society Providence, RI, 2000.

[17] S. N. Ethier and T. G. Kurtz, *Markov processes: characterization and convergence*, vol. 282, Wiley, 1986.

[18] L. C. Evans, H. M. Soner, and P. E. Souganidis, *Phase transitions and generalized motion by mean curvature*, Comm. Pure Appl. Math., 45 (1992), pp. 1097–1123.

[19] J. Felsenstein, *A pain in the torus: some difficulties with models of isolation by distance*, Amer. Naturalist, 109 (1975), pp. 359–368.

[20] T. Funaki, *The scaling limit for a stochastic PDE and the separation of phases*, Probab. Theory and Related Fields, 102 (1995), pp. 221–288.

[21] M. Gage and R. Hamilton, *The heat equation shrinking convex plane curves*, J. Differential Geom., 23 (1986), pp. 417–491.

[22] M. A. Grayson, *The heat equation shrinks embedded plane curves to round points*, J. Differential Geom., 26 (1987), pp. 285–314.

[23] G. Huisken, *Flow by mean curvature of convex surfaces into spheres*, J. Differential Geom., 20 (1984), pp. 237–266.

[24] S. Jansen and N. Kurt, *On the notion(s) of duality for Markov processes*, Probability surveys, 11 (2014), pp. 59–120.

[25] I. Karatzas and S. Shreve, *Brownian motion and stochastic calculus*, vol. 113, Springer Science & Business Media, 2012.

[26] G. S. Katzenberger, *Solutions of a stochastic differential equation forced onto a manifold by a large drift*, Ann. Probab., (1991), pp. 1587–1628.

[27] P. Kotelenez, *Comparison methods for a class of function valued stochastic partial differential equations*, Probab. Theory and Related Fields, 93 (1992), pp. 1–19.

[28] G. F. Lawler, *Intersections of random walks*, Springer Science & Business Media, 2013.

[29] G. F. Lawler and V. Limic, *Random walk: a modern introduction*, vol. 123, Cambridge University Press, 2010.

[30] R. H. Liang, *Two Continuum-Sites Stepping Stone Models in Population Genetics with Delayed Coalescence*, PhD thesis, University of California, Berkeley, 2009.

[31] J. Lukins, R. Tribe, and O. Zaboronski, *Multi-point correlations for two dimensional coalescing random walks*, arXiv preprint arXiv:1707.06250, (2017).

[32] C. Mueller and R. B. Sowers, *Random travelling waves for the KPP equation with noise*, J. Funct. Anal., 128 (1995), pp. 439–498.

[33] D. J. Murrell and R. Law, *Heteromyopia and the spatial coexistence of similar competitors*, Ecology Lett., 6 (2003), pp. 48–59.

[34] T. Nagylaki and Y. Lou, *The dynamics of migration–selection models*, in Tutorials in Mathematical Biosciences IV, Springer, 2008, pp. 117–170.

[35] P. Salminen and A. Borodin, *Handbook of Brownian motion: Facts and formulae*, Birkhauser, Boston, 2002.

[36] T. Shiga, *An interacting system in population genetics, II*, J. Math. of Kyoto University, 20 (1980), pp. 723–733.

[37] ——, *Stepping stone models in population genetics and population dynamics*, in Stochastic processes in physics and engineering, Springer, 1988, pp. 345–355.

[38] ——, *Two contrasting properties of solutions for one-dimensional stochastic partial differential equations*, Canad. J. Math., 46 (1994), pp. 415–437.

[39] T. Shiga and A. Shimizu, *Infinite dimensional stochastic differential equations and their applications*, J. Math. of Kyoto University, 20 (1980), pp. 395–416.

[40] R. Tribe, *Large time behavior of interface solutions to the heat equation with Fisher-Wright white noise*, Probab. Theory Related Fields, 102 (1995), pp. 289–311.

[41] A. Véber and A. Wakolbinger, *The spatial Lambda-Fleming–Viot process: An event-based construction and a lookdown representation*, in Annales de l'Institut Henri Poincaré, Probabilités et Statistiques, vol. 51, Institut Henri Poincaré, 2015, pp. 570–598.

[42] J. B. Walsh, *An introduction to stochastic partial differential equations*, Springer, 1986.