The effect of domain geometry and long-range dispersal on the motion of hybrid zones

Ian Patrick Letter Restuccia

University College

Supervised by Alison M. Etheridge and Julien Berestycki

Thesis submitted for the degree of

DOCTOR OF PHILOSOPHY

Department of Statistics

University of Oxford

July 2022

"The most important step anyone can take is the next one. Always the next one."

- Dalinar Kholin.

Declaration

I hereby declare that, except where specific reference is made to the work of others, the contents of this dissertation are original. The results of Chapter 2 have been obtained in a joint effort with Alison Etheridge and Mitchell Gooding. The results of Chapter 3 have been obtained in a joint effort with Kimberly Becker and Alison Etheridge. This thesis is submitted to the University of Oxford for the degree of Doctor of Philosophy and has not been submitted in whole or in part for consideration for any other degree or qualification in this, or any other university.

Ian Patrick Letter Restuccia

July 2022

Acknowledgements

First, I would like to thank my supervisors, Alison Etheridge and Julien Berestycki. Especially Alison, with whom I worked more closely during these four years. I count myself lucky for her detailed feedback, support, guidance and patience.

I thank Christina Goldschmidt and Nic Freeman for their careful reading of this thesis. I greatly appreciate the time and effort they put into it, and I hope it was worth it.

Thanks to the Probability group for being an amazing place to work. Thanks to Adrian, Bastian, Andy, Alex and Bohao for sharing an office with me and for all the (not so) mathematical discussions we had. Thanks to Terence and Kim, for their support and collaborations with me. I would also like to thank Mitch. While you may not have been a student when I was, our time-delayed collaboration is a big part of this thesis.

Thanks to Campari, especially to Robin and Maria, for believing in me all those times I did not. Thanks to Alex for wasting our afternoons talking about random Wikipedia facts and for his comments on this thesis. Thanks to my friends from Math; especially Jeremy, Tommaso, Eliana and Vicky. Thanks to my brothers, Rola, Juanchi and Dierk. Thanks to Felipe, Martin, Kele, Giselle and Marcelo. Thanks to Javier, Camila, Arturo and Azocar. Thanks to Macarena. Thanks to my friends from Beauchef and el DIM, especially to Las*Crêmes. Thanks to the Smash community. Thanks to all the friends I have not mentioned here due to space, I hope they know I have not forgotten them.

I am incredibly grateful for Catalina's support, love and company. Travelling across the world to grow with you has been the most fun adventure I could have ever imagined. I am excited about all the adventures the future will bring us.

Finally, thanks to my family. To John and Mirtha for educating, loving and withstanding me through high and low. Thanks to Gala, for her kisses that woke me up for uncountable years, and her excitement every time she heard my voice on the phone.

For everyone mentioned here, even implicitly, please know this degree is all thanks to you.

This work was supported by the ANID/Doctorado en el extranjero doctoral scholarship, grant number 2018-72190055.

Abstract

We consider two different population models that relate to hybrid zones. In both cases individuals in the population can be of one of three types: aa which may be fitter than AA, and both fitter than the aA heterozygotes. The *hybrid zone* is the region separating a subpopulation consisting entirely of aa individuals from one consisting of AA individuals.

First, we investigate the interplay between the motion of the hybrid zone and the shape of the habitat, both with and without genetic drift (corresponding to stochastic and deterministic models respectively). In the deterministic model, we investigate the effect of a wide opening and provide some explicit sufficient conditions under which the spread of the advantageous type is halted, and complementary conditions under which it sweeps through the whole population. As a standing example, we are interested in the outcome of the advantageous population passing through an isthmus. We also identify rather precise conditions under which genetic drift breaks down the structure of the hybrid zone, complementing previous work that identified conditions on the strength of genetic drift under which the structure of the hybrid zone is preserved.

The second set of results regard the study of how hybrid zones are affected by long-range dispersal of the individuals in the population, assuming both homozygotes are equally fit. We ignore genetic drift and use a deterministic model, the fractional Allen-Cahn equation to study the population. We show not only that under a suitable rescaling the motion of hybrid zones converges to motion under mean curvature flow, but also that there is a family of possible rescalings that gives this result. To overcome technical difficulties arising from the heavy-tailed nature of the stable distribution, we couple ternary branching stable motions to ternary branching Brownian motions subordinated by truncated stable subordinators, and use this coupling to quantify how the large jumps of the stable motion affect our main result.

Contents

1

Introduction

Population genetics is a branch of science that studies the genetic composition within populations. It is assumed that organisms have their traits encoded in the form of *genes*. Physically, genes are regions of extremely long molecules called DNA, copies of which are in each cell of an organism. Variations in the DNA result in variations of a trait in an organism. An *allele* is a possible form of a gene. The composition of the alleles in a population changes over time due to biological mechanisms that choose some of these traits to be inherited in the generation of new individuals from old ones, that is, the process of reproduction. It is a natural question to try to understand how this process can change the distribution of genes and alleles in a population over time. One of the roles of population genetics is to manage the interplay between biology and mathematics to arrive at a reasonable model that incorporates biologically interesting factors, while also retaining mathematical tractability. Models that try to capture too many biological details rapidly become mathematically intractable. On the other hand, too many simplifications may result in a model from which many conclusions can be established, but with no biological relevance. This trade-off is an important challenge for our understanding of evolution.

Since not all individuals possess the same number of copies of genetic material, nor are those copies inherited in the same fashion during reproduction, we need some language to identify what type of population we focus on studying. It could be the case that each organism in the population has a single copy of each gene, this is called a *haploid* organism. Often these individuals reproduce by *asexual reproduction*, passing down a copy of their genetic material to their offspring. However, there exist more complex and interesting mechanisms for generating new individuals. For example, we could think of two organisms that generate a new one by a combination of their genes by some rule, this being deterministic or stochastic. This is called *sexual reproduction*. While so far we have only referred to haploid individuals, a notable number of sexually reproducing life forms have two copies of their DNA in each cell, one inherited from each parent. We call these organisms *diploid* individuals. Note that each pair of genes carried by a diploid individual could be exactly the same, or the genes could differ in some way. These organisms are called *homozygote* in the former, and *heterozygote* in the latter case. Some genetic compositions may be advantageous to the organism carrying them, for example, by conferring resistance to extreme temperatures or by giving a particular size that allows easier access to food. We say any organism with a survival advantage is *fitter*. Even if a particular genotype is fitter it could still disappear from the population, given random fluctuations from the reproductive events or segregation given by the spatial distribution.

In this thesis, we are interested in the interplay between selection and spatial distribution. We will study two models of populations which are distributed in space and that are under the effect of selection against heterozygotes, and give results that describe the evolution of said populations. Of particular novelty is our study of the effect of domain geometry and long-range dispersal in said models.

We proceed to give a short exposition of some well known models in population genetics. This serves to develop the mathematical language we will use, while also placing our contributions in the correct context. The following exposition is based on [Eth11] and [Eth19]. However, some notation and definitions were changed, following [EFP17].

1.1 The Moran model and the Wright-Fisher diffusion

We start with one of the simplest model in population genetics, originally due to Moran [Mor58]. This model assumes a haploid population of constant size N, with no spatial structure, that reproduces asexually.

Definition 1.1 (The neutral Moran model). The neutral Moran model is the model in which N haploid individuals evolve according to the following dynamics. We run a Poisson process of rate $\binom{N}{2}$. At each point of the Poisson process, a reproduction event occurs, in which two individuals are sampled uniformly from the population. The first individual dies, and the second produces a single offspring that inherits its parent's genetic material.

We could equivalently define the model by running a Poisson process for each pair of individuals. That is, we label the population with numbers $1, \dots, N$, and associate to each pair $1 \leq i < j \leq N$ a Poisson process of rate 1. At each point of the Poisson process associated to the pair (i, j), we choose one of the two individuals uniformly, who dies. The remaining individual reproduces asexually, and the offspring is given the label of the individual that perished. By standard Poisson thinning this gives a process with the same dynamics as the one described above.

The model, as we have written it, may seem of little relevance. But this is just because we are lacking an interesting genetic trait associated to each individual. We now suppose each individual has associated to it one of two possible alleles, which we call a and A. Then, during reproduction, each offspring copies the type (a or A) from its parent. Let p_t^N denote the proportion of individuals that are of type a in the population. Suppose a reproduction event occurs at time t. Then,

$$p_t^N := \begin{cases} p_{t-}^N + \frac{1}{N} & \text{with probability } \frac{N}{N-1} p_{t-}^N (1 - p_{t-}^N), \\ p_{t-}^N - \frac{1}{N} & \text{with probability } \frac{N}{N-1} p_{t-}^N (1 - p_{t-}^N), \\ p_{t-}^N & \text{with probability } 1 - \frac{2N}{N-1} p_{t-}^N (1 - p_{t-}^N). \end{cases}$$

Recalling that reproduction events arrive according to a Poisson process of rate $\binom{N}{2}$, for small time a Taylor expansion yields

$$p_{t+h}^N := \begin{cases} p_t^N + \frac{1}{N} & \text{with probability } \frac{N^2}{2} p_t^N (1 - p_t^N) h + \mathcal{O}(h^2), \\ p_t^N - \frac{1}{N} & \text{with probability } \frac{N^2}{2} p_t^N (1 - p_t^N) h + \mathcal{O}(h^2), \\ p_t^N & \text{with probability } 1 - N^2 p_t^N (1 - p_t^N) h + \mathcal{O}(h^2). \end{cases} \tag{1.1}$$

More formally, we take $f : \mathbb{R} \to \mathbb{R}$ to be a twice differentiable function. Define the *generator* of p_t^N acting on f by

$$\mathcal{L}^N f(p) := \frac{\partial}{\partial t}\mathbb{E}\left[f(p_t^N)|p_0^N = p\right]\Big|_{t=0}.$$

Then (1.1) translates to:

$$\begin{aligned}\mathcal{L}^N f(p) &= \frac{N^2}{2}p(1-p)(f(p+\tfrac{1}{N}) - f(p)) + \frac{N^2}{2}p(1-p)(f(p-\tfrac{1}{N}) - f(p)),\\ &= \frac{1}{2}p(1-p)f''(p) + \mathcal{O}\left(\frac{1}{N}\right).\end{aligned}$$

Supposing p_0^N converges to $p \in \mathbb{R}$, standard scaling limit arguments for diffusions (see Chapter 4 of [EK86]) give that, when N goes to infinity, $(p_t^N)_{t\geq 0}$ converges in law in the Skorohod topology on $\mathcal{D}([0,\infty),[0,1])$, the space of càdlàg paths in $[0,1]$. The limit p_t solves the stochastic differential equation (SDE),

$$dp_t = \sqrt{p_t(1-p_t)}dW_t, \tag{1.2}$$

with initial condition $p_0 = p$, where $(W_t)_{t\geq 0}$ denotes a standard Brownian motion.

Definition 1.2 (Wright-Fisher diffusion). We call the (weak) solution of equation (1.2) the Wright-Fisher diffusion.

1.2 Kingman's coalescent and duality

Although the Moran model and the Wright-Fisher diffusion may seem simple models, we can already ask interesting and meaningful questions about them. For example, we can start thinking about describing *genealogies*. That is, suppose that we take, at some time $t > 0$, a sample of individuals in a population evolving in time according to the Wright-Fisher diffusion. How can we describe the common ancestors of the sample of individuals as we go back in time? From the Wright-Fisher diffusion itself, the question may be hard to answer, since we have lost track of the notion of individuals. However, the individuals are the building blocks of the Moran model. So we start by defining the genealogy in that model and take N to infinity. As we will see, this will give us information on the genealogies in the Wright-Fisher diffusion.

For this, recall that we can see the Moran model as a process in which, for each pair of individuals (i,j) in the population, we run a rate 1 Poisson process $(E_{i,j}(t))$. At a point t of $E_{i,j}$, one individual (either i or j) dies and is replaced by an offspring of the surviving individual. Since the Poisson process is reversible, the dynamics of the Moran model are easily seen to be reversible. From the backwards in time point of view, for each pair of individuals we have a Poisson process $\widehat{E}_{i,j}$. At a point t of $\widehat{E}_{i,j}$, we have that i and j *coalesce* into a single common ancestor. What we have roughly described is known as Kingman's coalescent, which we now formally define.

Definition 1.3 (Kingman's coalescent [Kin82]). A k-coalescent is a continuous-time Markov chain on $\mathcal{E}_k$, the space of equivalence relations on $1,\dots,k$, whose transition rates $(q_{\xi,\eta} : \xi,\eta \in \mathcal{E}_k)$ are given by

$$q_{\xi,\eta} = \begin{cases} 1 & \text{if } \eta \text{ is obtained by coalescing two equivalence classes of } \xi, \\ 0 & \text{otherwise.} \end{cases}$$

Kingman's coalescent is then the process taking values in $\mathcal{E}_{\mathbb{N}}$, the equivalence relations on $\mathbb{N}$, such that, for each k, its restriction to $1,\dots,k$ is the k-coalescent. For convenience, we assume the initial condition is the trivial partition into singletons.

From the description of the backwards in time Moran model in terms of $\widehat{E}_{i,j}$, it follows that the k-coalescent describes the genealogies of a sample of k individuals in the Moran model, in the sense that i and j being in the same block of the k-coalescent at t is interpreted as the individuals with label i and j having a common ancestor at time t before the present.

We can now establish a well-known relationship between the Wright-Fisher diffusion and Kingman's coalescent.

Proposition 1.4. *Let $(p_t)_{t\geq 0}$ be a Wright-Fisher diffusion and $(n_t)_{t\geq 0}$ be the number of blocks in a k-coalescent. Then, there is a coupling such that, for all $p \in [0,1]$ and $k \in \mathbb{N}$,*

$$\mathbb{E}_p\left[p_t^k\right] = \mathbb{E}^k\left[p^{n_t}\right], \tag{1.3}$$

where $\mathbb{E}_p$ denotes the law of the Wright-Fisher diffusion with $p_0 = p$ and $\mathbb{E}^k$ denotes the law of the k-coalescent.

Let us sketch a proof. One can argue in terms of the description of p_t as a limit of the Moran model and n_t being the corresponding genealogy. Suppose we have N individuals in a Moran model and each one is of type a at time 0 with probability p_0 and of type A at time 0 with probability $1 - p_0$. Take a sample of size k in the population at time t. That sample is going to be entirely of type a if and only if all ancestors at time 0 are of type a. So $\mathbb{E}^k\left[p^{n_t}\right]$ is the probability of the k sample being entirely of type a. On the other hand, p_t^N is exactly the proportion of individuals of type a in the population at time t. So, the probability of the sample of size k being entirely of type a is also given by $\mathbb{E}_p\left[(p_t^N)^k\right]$, which if we take N to infinity is $\mathbb{E}_p\left[p_t^k\right]$. From this, (1.3) follows.

Formulae of the form given by (1.3) are known as duality relationships.

Definition 1.5 (Duality). Let X, Y be Markov processes on Polish state spaces E and F, respectively. 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 $x \in X$, $y \in Y$ and $t \geq 0$,

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

where $\mathbb{E}_x$ is the law for X under which $X_0 = x$ and $\mathbb{E}^y$ is the law for Y under which $Y_0 = y$. We say X is dual to Y (and vice-versa). The function H is called the duality function.

We refer to [SK14] for a review of duality relationships. In this language, Proposition 1.4 says that the Wright-Fisher diffusion and Kingman's coalescent are dual with duality function $H(x, y) = x^y$. When the duality function is of this particular form it is called a *moment* duality. Although it is not true that a moment duality implies a relationship between genealogies [Tay09], moment duality is still a way of studying our forward in time models with what can be understood as backward in time models.

While the models and structures presented so far may seem simple, they form the building blocks of a large literature on different mathematical models in population genetics. In subsequent sections, we immerse ourselves in the basics of spatial structures and selection, but certainly, there is much more that these models can incorporate. The ideas behind how the Wright-Fisher diffusion is derived from the Moran model can be extended to incorporate other factors, such as seedbanks [Bla+20], random environment [CV19], recombination

[EPB16] and mutation [DK99]. Kingman's coalescent can also be modified to incorporate interesting phenomena. One can consider dormancy given by hibernation [Cor+20], populations that lose the ability to exchange genetic material (event known as *speciation*) [FLS20], or multiple mergers that serve to incorporate more general reproduction mechanisms in the forward in time model [Pit99], [Sag99], [DK99].

1.3 Adding spatial structure

So far all the models we have talked about lack any sort of spatial structure. When we assume that individuals are equally likely to have offspring or die, we are implicitly assuming the population is well mixed enough to neglect the effects of the spatial distribution of the individuals. This is not always a good assumption in nature, as the habitat of some populations can cover entire countries or continents. So it is imperative that our models can be modified to incorporate effects resulting from the spatial distribution of the population.

A simple way of doing this is to subdivide the spaces into *demes*, which can be thought of as islands of population. Demes sit at the vertices of a graph, which for simplicity we assume to be $\mathbb{Z}^{\mathrm{d}}$, with migration occurring along the edges of the graph. If we assume the dynamics in each deme to be given by the Moran model we obtain an interesting process defined on $\mathbb{Z}^{\mathrm{d}}$. This is known as the structured Moran model.

Definition 1.6 (Structured Moran model). Let $N \in \mathbb{N}$ and let $(m_{i,j})_{i,j\in\mathbb{Z}^{\mathrm{d}}}$ be a collection of non-negative numbers (which we will interpret as the migration rate from deme i to deme j).

Suppose we start with a population in $\mathbb{Z}^{\mathrm{d}}$ with N individuals in each deme. The structured Moran model is described by the following dynamics. For each $i \in \mathbb{Z}^{\mathrm{d}}$ we have a Poisson process, E_i, of rate $\binom{N}{2}$ and for each $i, j \in \mathbb{Z}^{\mathrm{d}}$ that are nearest-neighbours in $\mathbb{Z}^{\mathrm{d}}$ (denoted $i \sim j$) we have a Poisson process, E_{ij}, of rate $Nm_{i,j}$.

At a point t of E_i we sample two individuals uniformly at random from the deme i. The first one dies and is replaced by a copy of the second individual. At a point t of E_{ij} we sample an individual uniformly from deme i and another one uniformly from deme j. The first individual dies and is replaced by a copy of the second individual (at deme i).

We assume again that, under the structured Moran model, each individual has one of two possible alleles a or A. Since in the Moran model the proportion of a alleles converges in the large population limit, it is natural to ask if, in this structured version that quantity also has a limit. Let us fix $i \in \mathbb{Z}^{\mathrm{d}}$, and denote by $p_t^{N,i}$ the proportion of individuals of type a at deme i at time t. Similarly to (1.1) we can write

$$p_{t+h}^{N,i} := \begin{cases} p_t^{N,i} + \frac{1}{N} & \text{w.p. } \frac{N^2}{2} p_t^{N,i}(1-p_t^{N,i})h + N\sum_{j:j\sim i} m_{j,i} p_t^{N,j}(1-p_t^{N,i})h + \mathcal{O}(h^2), \\ p_t^{N,i} - \frac{1}{N} & \text{w.p. } \frac{N^2}{2} p_t^{N,i}(1-p_t^{N,i})h + N\sum_{j:j\sim i} m_{j,i} p_t^{N,i}(1-p_t^{N,j})h + \mathcal{O}(h^2), \\ p_t^{N,i} & \text{otherwise.} \end{cases} \tag{1.4}$$

As before we take $f : \mathbb{R} \to \mathbb{R}$ to be a twice differentiable function. Again, we define the generator of $p_t^{N,i}$ acting on f by:

$$\mathcal{L}^{N,i} f(p^i) := \frac{\partial}{\partial t} \mathbb{E}\left[f(p_t^{N,i}) | p_0^{N,i} = p^i \right] \Big|_{t=0}.$$

We can use (1.4) and a Taylor expansion to obtain, at least formally, that

$$\begin{aligned} \mathcal{L}^{N,i} f(p^i) &= \frac{N^2}{2} p^i(1-p^i)(f(p^i + \tfrac{1}{N}) - f(p^i)) + \frac{N^2}{2} p^i(1-p^i)(f(p^i - \tfrac{1}{N}) - f(p^i)) \\ &\quad + N \sum_{j:i\sim j} m_{j,i} \left(p^j(1-p^i)(f(p+\tfrac{1}{N}) - f(p)) + p^i(1-p^j)(f(p-\tfrac{1}{N}) - f(p)) \right) \\ &= \frac{1}{2} p^i(1-p^i) f''(p^i) + \sum_{j:j\sim i} m_{j,i} \left(p^j - p^i \right) f'(p) + \mathcal{O}\left(\frac{1}{N} \right). \end{aligned}$$

While this computation is certainly not rigorous, it serves to illustrate that it is not surprising that the structured Moran model converges to what is known as Kimura's stepping stone model.

Definition 1.7 (Kimura's stepping stone model [Kim53]). We suppose we have a population in demes located at the vertices of $\mathbb{Z}^{\mathrm{d}}$. Each individual can have two possible types a and A. We say that the proportion of the allele a in each deme, $(p_i(t))_{i\in\mathbb{Z}^{\mathrm{d}}}$ evolves under Kimura's stepping stone model if:

$$dp_i(t) = \sum_{j:j\sim i} m_{j,i}(p_j(t) - p_i(t))dt + \sqrt{p_i(t)(1-p_i(t))} dW_i(t), \tag{1.5}$$

where $\{W_i\}_{i\in\mathbb{Z}^{\mathrm{d}}}$ are independent Brownian motions.

In other words, under Kimura's stepping stone model the proportions of a alleles evolve as a system of interacting Wright-Fisher diffusions.

From this, it should also be natural to ask if we can obtain an expression for the genealogy of a sample, analogous to Kingman's coalescent for the Wright-Fisher diffusion. Since at each deme we are seeing Wright-Fisher dynamics we expect that, when ancestral lineages are in the same deme, they behave like Kingman's coalescent. The difference here is that we also need to take into account migration. Putting these two things together gives the structured coalescent.

Definition 1.8 (Structured coalescent). The structured coalescent $(\overline{n}(t)) = (n_i(t))_{i\in\mathbb{Z}^{\mathrm{d}}}$ is a continuous time Markov Chain taking values in $\mathbb{N}^{\mathbb{Z}^{\mathrm{d}}}$. The transition rates are given by:

$$\forall i \in \mathbb{Z}^{\mathrm{d}},\ n_i \to n_i - 1 \text{ at rate } \binom{n_i}{2},$$
$$\forall i, j \in \mathbb{Z}^{\mathrm{d}},\ n_i \to n_i - 1 \text{ and } n_j \to n_j + 1 \text{ at rate } n_i m_{ji}.$$

To solidify the relationship between the structured coalescent and Kimura's stepping stone model, and also mirror what happens with Kingman's coalescent and the Moran model, we note they are in a moment duality.

Proposition 1.9. *Suppose $(\overline{p}(t))_{t\geq 0} = (p_i(t))_{t\geq 0}$ evolve under the Kimura's stepping stone model. Let $(\overline{n}(t))_{t\geq 0}$ be a structured coalescent. Then, for all $\overline{p}(0) \in [0,1]^{\mathbb{Z}^{\mathrm{d}}}$, $\overline{n}(0) \in \mathbb{N}^{\mathbb{Z}^{\mathrm{d}}}$ with finitely many non-zero components, we have*

$$\mathbb{E}_{\overline{p}(0)}\left[\prod_{i\in\mathbb{Z}^{\mathrm{d}}} (p_i(t))^{n_i(0)}\right] = \mathbb{E}^{\overline{n}(0)}\left[\prod_{i\in\mathbb{Z}^{\mathrm{d}}} (p_i(0))^{n_i(t)}\right], \tag{1.6}$$

where $\mathbb{E}_{\overline{p}(0)}$ denotes the law of Kimura's stepping stone model with initial condition $\overline{p}$ and $\mathbb{E}^{\overline{n}(0)}$ denotes the law of the structured coalescent starting from $\overline{n}(0)$.

While we do not present a proof of Proposition 1.9, an interested reader can see Lemma 3.2.3 in [Eth19]. The essence of the argument is no different from that for Proposition 1.6. Both expectations in (1.6) are the probability that a sample of individuals taken from the population at time t, by sampling $n_i(0)$ individuals from deme i, are all of type a. The left hand side of (1.6) is the probability as seen following the forward in time dynamics, while the right hand side of (1.6) is as seen while following the genealogy backward in time.

While the Wright-Fisher diffusion and Kingman's coalescent are the building blocks of population models without spatial structure, a similar flavour occurs with the structured Moran model and the structured coalescent. The possible deviations and differences from these basic models are enormous. So far we have assumed constant population density in each deme, but we could forget about that condition. In that case we could scale the rates of reproduction and death in such a way that, in the limit, the population densities at the different demes satisfy a system of interacting differential equations, different to (1.5) [SS80], [Smi21]. We could also change the spatial scaling, so that demes are *closer and closer* in the limit. If the scaling is performed in such a way that stochasticity disappears in the limit, we expect the density of individuals to be a solution of a partial differential equation [Coh+99], [DFL86], [Gon11]. Furthermore, one could retain some stochastic behaviour in the limit, and in that case the density of particles is expected to solve a stochastic partial differential equation or follow the law of a superprocess [CP05], [CDP99], [Per02], [Stu03]. There is an enormous literature on the subject, of which we have just cited some notable examples.

1.4 The Spatial Λ-Fleming Viot process

All the spatial models for populations we have mentioned so far share one common feature; they are based on tracking the position of each individual in the population, whether it be in demes or the spatial continuum. If we want to trace a particular allele in one of these populations, computations are inherently hard, since the interaction between different types of individuals forward in time gives information about ancestral lineages backwards in time. In other words, we may need to keep track of the whole population to infer any notion of relatedness between individuals, and this does not sound like a straightforward task.

This is a problem even in simple models of populations evolving in spatial continuum. Wright [Wri43] and Malécot [Mal48] suggested a simple population model of a spatially structured population which allows one to keep track of types of individuals. In their model the generations are discrete. Each individual dies at each generation and is, independently, replaced by a Poisson number of individuals with mean one. These offspring are distributed, independently, according to a symmetric Gaussian distribution around the position of the parent. Wright and Malécot further assume that the individuals start in space as a Poisson random measure which they assumed to be a stationary distribution for the population dynamics. Under those assumptions, Malécot wrote down a recursion for the *probability of identity*; that is, the probability that, if we sample two individuals in the population, they are of the same type (if each individual, independently, mutates at a constant rate to a type never seen in the population). It was Felsenstein [Fel75] who first noticed that the dynamics were incompatible with the assumption of having a Poisson random measure as stationary distribution. A population evolving as Wright and Malécot suggested does not have a stationary distribution in spatial dimensions 1 and 2, since it dies out. This is not solved by changing the space of the population to a compact set: if the population model evolves in the torus, it dies out in finite time. If one further conditions the population to be of constant size N, to avoid extinction, the individuals form clumps of high density. These problems are known as the *pain in the torus*.

All this serves to illustrate that we need to take some care if we want to keep track of individuals in spatial continuum, while also being able to infer relatedness between them. A way of solving this is to suppose that the population density is constant and use a measure to describe the proportion of a particular type in the population. This is the idea behind the Spatial Λ-Fleming Viot process (from now on SLFV for short).

The SLFV was introduced in [Eth08] and rigorously constructed in [BEV10]. While the original constructions allow individuals to have one of an arbitrary number of types, here we focus on the case where each individual can have one of only two types, a and A. At each time t the random function $\{w_t(x) : x \in \mathbb{R}^{\mathrm{d}}\}$ will model the proportion of individuals of type a at position x at time t. The identification

$$\int_{\mathbb{R}^{\mathrm{d}}} (w_t(x)f(x,a) + (1 - w_t(x))f(x,A))dx = \int_{\mathbb{R}^{\mathrm{d}} \times \{a,A\}} f(x,k)M(dx,dk)$$

provides a one-to-one correspondence between the state space of the SLFV and the space $\mathcal{M}_\lambda$

of measures on $\mathbb{R}^{\mathrm{d}} \times \{a, A\}$ with *spatial marginal* Lebesgue measure, which we endow with the topology of vague convergence. Hence, we shall abuse notation and also denote the state space of $(w_t)_{t \geq 0}$ by $\mathcal{M}_\lambda$.

The dynamics of the process are driven by a Poisson point process Π. This Poisson point process encodes *events* of reproduction in the population in different parts of the space. These events are described by a position and a radius, which should be interpreted as defining the area affected by the event, and an impact, that dictates the proportion of individuals in the population that are replaced at each event.

Definition 1.10 (Spatial Λ-Fleming Viot (SLFV)). The Spatial Λ-Fleming Viot process $\{w_t(x); x \in \mathbb{R}^{\mathrm{d}}, t \geq 0\}$ is the $\mathcal{M}_\lambda$-valued process that specifies the proportion of a-allele individuals on position x at time t. Its dynamics are driven by a Poisson point process on $\mathbb{R}_+ \times \mathbb{R}^{\mathrm{d}} \times \mathbb{R}_+ \times (0, 1]$, with intensity $dt \otimes dx \otimes \mu(dr)\nu_r(du)$, where $\mu(dr)$ is a measure on $\mathbb{R}_+$ and $\nu_r(du)$ is a probability measure on $(0, 1]$ for every $r \in (0, \infty)$.

At a point $(t, x, r, u) \in \Pi$, a reproduction event takes place in the closed ball $B(x, r)$. During the event:

1. Select a parental location z from $B(x, r)$ uniformly at random, that is, $z \sim U(B(x, r))$.

2. Select a type k according to the distribution of types z immediately before the event, that is, $k = a$ with probability $w_{t^-}(z)$, and $k = A$ otherwise.

3. For all $y \in B(x, r)$ replace a proportion u of individuals by k. That is,

$$w_t(y) = (1 - u)w_{t^-}(y) + u1_{k=a}.$$

Sites outside $B(x, r)$ are not affected by the reproduction event. That is, for all $y \notin B(x, r)$,

$$w_t(y) = w_{t^-}(y).$$

Remark 1.11. Note that $w_t(x)$ is defined in terms of sampling individuals uniformly. Thus, it will only be defined up to a Lebesgue-null set $\mathcal{N}$. Since we want to avoid these technicalities, it is convenient to extend the definition of $w_t(x)$ to all of $\mathbb{R}^{\mathrm{d}}$. To do this, we set $w_t(x) = 0$ for all $x \in \mathcal{N}$.

Remark 1.12. Sampling parental locations, and then parental types, is convenient for identifying the dual process of branching and coalescing ancestral lineages that we will introduce in Definition 1.13. However, from the perspective of the SLFVS it would be equivalent to sample types independently and uniformly at random from the region affected by the event.

For $B \subseteq \mathbb{R}^{\mathrm{d}}$ we think of $w_t(B)/|B| = \int_B w_t(x)dx/|B|$ as the probability of sampling an individual of type a, if we sample an individual uniformly at random from the set B at time t. In order for this process to be defined we assume

$$\int_0^\infty \int_0^1 ur^{\mathrm{d}} \nu_r(du)\mu(dr) < \infty, \tag{1.7}$$

although condition (1.7) can be relaxed, see [EK19].

We now show that SLFV indeed allows us to write a process of genealogies backwards in time, similar to what we were able to do with the structured Moran model and the structured coalescent.

Our starting point is, again, taking a sample from our population and following the location of their ancestors backwards in time. We fix $x_1, \ldots, x_n \in \mathbb{R}^{\mathrm{d}}$ and $t \geq 0$, and try to compute the joint distribution of $w_t(x_1), w_t(x_2), \ldots, w_t(x_n)$. For that, we trace, backwards in time, the location of the ancestors of the individual sampled at x_i at time t. In other words, we trace backwards until we see an event whose area of effect, $B(x, r)$, intersects at least one of the x_i. Each lineage in $B(x, r)$ has a probability u of being affected by the event. All affected individuals then coalesce and move to the position of the parent, picked uniformly at random from $B(x, r)$. This continues until time 0, giving a system of coalescing random walkers. This system of random walkers has jumps driven by a Poisson point process, which by reversibility has the same intensity as Π. This motivates the following definition.

Definition 1.13 (SLFV dual). Let Π be a Poisson point process on $\mathbb{R}_+ \times \mathbb{R}^{\mathrm{d}} \times \mathbb{R}_+ \times (0, 1]$ with intensity $dt \otimes dx \otimes \mu(dr)\nu_r(du)$, where $\mu(dr)$ is a measure on $\mathbb{R}_+$, $\nu_r(du)$ is a probability measure on $(0, 1]$ for every $r \in (0, \infty)$, and μ, ν_r satisfy (1.7). The particle system dual to the SLFV process, $(\mathcal{P}_t)_{t\geq 0} := (\xi_t^1, \xi_t^2, \ldots, \xi_t^{N(t)})_{t\geq 0}$ is the $\bigcup_{m\geq 1}(\mathbb{R}^{\mathrm{d}})^m$-valued process with dynamics given as follows. At each event $(t, x, r, u) \in \Pi$:

1. For each $\xi_t^i \in B(x,r)$, independently mark it with probability u.

2. If at least one ξ_t^i was marked, we replace all marked particles with a single particle at a position picked uniformly in $B(x,r)$. If no particle was marked, nothing happens.

Note that if we start $(\mathcal{P}_t)_{t\geq 0}$ from at least one particle, it will always have at least one particle.

We proceed to state a duality relationship between the SLFV and its dual. The main difficulty is the fact that $w_t(x)$ is defined only Lebesgue a.e. and so it is not obvious that taking $w_t(x_i)$ in a duality relationship makes sense (i.e. it is not obvious that Remark 1.11 is compatible with having a duality relationship). To address this, instead of taking a deterministic sample $x_1,\ldots,x_n$, we sample according to a distribution, that has density ψ with respect to the Lebesgue measure. Then we can establish the following duality.

Proposition 1.14. *The SLFV is dual to $(\mathcal{P}_t)_{t\geq 0}$ in the following sense. For any $m \in \mathbb{N}$, $\psi \in \mathcal{C}((\mathbb{R}^{\mathrm{d}})^m) \cap L^1((\mathbb{R}^{\mathrm{d}})^m)$, and for any measurable function $w_0(\cdot)$, we have that*

$$\mathbb{E}_{w_0(\cdot)}\left[\int_{(\mathbb{R}^{\mathrm{d}})^m} \psi(x_1,\ldots,x_m)\left\{\prod_{j=1}^{k} w_t(x_j)\right\} dx_1\ldots dx_k\right]$$
$$= \int_{(\mathbb{R}^{\mathrm{d}})^m} \psi(x_1,\ldots,x_m)\mathbb{E}^{(x_1,\ldots,x_m)}\left[\prod_{j=1}^{N(t)} w_0(\xi_j)\right] dx_1\ldots dx_k.$$

In the last equality, the expectation on the left hand side is with respect to the SLFV with initial condition $w(0,\cdot)$, and on the right hand side it is with respect to the dual process with initial condition $(x_1,\ldots,x_m)$.

Remark 1.15. As we have already mentioned, a duality relationship does not necessarily imply a relationship between genealogies. However, [EK19] obtained the SLFV as the limit of an individual-based model while retaining the information about genealogical trees relating individuals. This justifies the interpretation of $(\mathcal{P}_t)_{t\geq 0}$ as the (backwards in time) dynamics of individuals that evolve according to the SLFV.

1.5 Natural selection and hybrid zones

Up to this point, all models presented have considered individuals of all genetic types in the population to be equally fit. Of course, this is not true in nature, as natural selection favours the reproduction of some genetic types over others. This bias in reproduction is not immediate from the genetic type of the individuals. It could be that individuals are fitter because they carry a particular allele, or it could be that the interaction between multiple alleles affects their fitness. And in both cases there could be interaction with the environment or other individuals, influencing the fitness. We do not expect that a single model with selection can explain all selective mechanisms that can be found in nature.

When talking about natural selection we thus need to focus on a particular population and a particular form of selection acting on this population. In our case we will focus on a form of selection that can maintain so-called hybrid zones: narrow geographic regions where two genetically distinct populations are found close together and hybridise, resulting in offspring of mixed ancestry. Hybrid zones are prevalent in nature, with many more than 170 examples being known [BH89], [Bug07], [HS73], [Toe+18]. There are multiple hypotheses for how hybrid zones are maintained. One states that hybrid zones are due to a sharp change in the environment, affecting the fitness of individuals depending on their genetic composition and position in space. Another hypothesis, and the one on which we focus, is that hybrid zones are due to hybrids in the population being less fit than purebred individuals. This could not only explain the formation of hybrid zones, but also imply that, for such populations, a balance between dispersal and selection should make hybrid zones evolve in time. It is then an interesting mathematical and biological question to ask about the motion of the hybrid zones over time.

To capture this phenomenon, Etheridge, Freeman and Penington [EFP17] propose the following model. Consider a population in which selection acts on a single genetic locus, which can have two possible allelic types, a and A. Individuals are assumed to be diploid, and it is considered that heterozygotes are less fit than homozygotes. That is, individuals of type aA have a lower reproductive success. The starting point is to incorporate this selection mechanism into a simpler model without space, similar to the Wright-Fisher diffusion. For this, we suppose individuals reproduce by the following mechanism. Each individual produces germ cells, each carrying a copy of both of the alleles carried by the parent. Each germ cell then splits into two gametes, each carrying a single copy of each of the alleles of the parental individual. This produces a pool of gametes from all individuals in the parental generation. Offspring are then formed by picking two gametes at random. The offspring generation then replaces the parental generation. Call the proportion of a-alleles of the population $\overline{w}$. We assume Hardy-Weinberg equilibrium, meaning there are:

$$\begin{aligned} \overline{w}^2 &\text{ individuals of type } aa, \\ (1-\overline{w})^2 &\text{ individuals of type } AA, \\ 2\overline{w}(1-\overline{w}) &\text{ individuals of type } aA. \end{aligned}$$

While this may sound unreasonable at first, populations reach Hardy-Weinberg equilibrium rather easily [Har08].

To model the selection against heterozygotes we suppose that they have a lower *relative fitness*, meaning they produce fewer offspring than homozygotes in reproductive events. Specifically, during reproductive events heterozygotes produce $1-\bar{s}$ times as many germ cells as homozygotes, for a small $\bar{s} > 0$. If $\overline{w}$ is the proportion of type a-alleles before a reproduction event, then the proportion of type a gametes in the pool is:

$$\begin{aligned} \frac{\overline{w}^2 + \overline{w}(1-\overline{w})(1-\bar{s})}{1-2\bar{s}\overline{w}(1-\overline{w})} &= (1-\bar{s})\overline{w} + \bar{s}(3\overline{w}^2 - 2\overline{w}^3) + \mathcal{O}(\bar{s}^2), \\ &= \overline{w} + \bar{s}(1-\overline{w})\overline{w}(2\overline{w}-1) + \mathcal{O}(\bar{s}^2). \end{aligned}$$

Hence, if $\bar{s} \sim \frac{\boldsymbol{s}}{N}$ for large N and $\boldsymbol{s}$ constant, then, as N goes to infinity, the change in $\overline{w}$ in units of time N is approximated by the equation

$$\frac{d\overline{w}}{dt} = \boldsymbol{s}(1-\overline{w})(2\overline{w}-1).$$

Furthermore, supposing individuals disperse in a radially symmetric way with finite variance we arrive, after rescaling, at the following equation for the proportion of a-allele in the population across the Euclidean space:

$$\frac{\partial w}{\partial t} = \frac{m}{2}\Delta w + \boldsymbol{s}w(1-w)(2w-1). \tag{1.8}$$

This is a special case of the Allen-Cahn equation. From now on, for convenience, we take $m = 2$ and $\boldsymbol{s} = 1$. Since we expect hybrid zones to be a narrow region with respect to the population range, we *zoom out* and study the equation over large spatial and temporal scales. That is, we perform a diffusive re-scaling $t \to \varepsilon^{-2}t$, $x \to x\varepsilon^{-1}$ so the equation becomes

$$\frac{\partial w}{\partial t} = \Delta w + \frac{1}{\varepsilon^2}w(1-w)(2w-1). \tag{1.9}$$

It is a known result that, as $\varepsilon \to 0$, the solution to equation (1.9) converges to the indicator function of a set, whose boundary evolves according to mean curvature flow [Che92]. We proceed to explain what this statement means.

Let $\mathbb{S}^1$ denote the unit sphere in $\mathbb{R}^{\mathrm{d}}$. Let $\Gamma_t : \mathbb{S}^1 \to \mathbb{R}^{\mathrm{d}}$ be a family of smooth embeddings, defined up to some time $\mathscr{T}$, indexed by $t \in [0, \mathscr{T})$. We write $\widehat{n}_t(\Phi)$ for the unit (inward) normal vector to Γ_t at Φ. We also write $\kappa = \kappa_t(\Phi)$ for the mean curvature of Γ_t at Φ. Then we say Γ_t evolves according to mean curvature flow if

$$\frac{\partial_t \Gamma_t(\Phi)}{\partial t} = \kappa_t(\Phi)\widehat{n}_t(\Phi). \tag{1.10}$$

In dimension 2, mean curvature flow is just curvature flow. In that case, the evolution is almost completely understood. The surface in dimension 2 is a curve, and under mean curvature flow it becomes convex in a finite time [Gra87] and then shrinks to a point [GH86]. In dimensions 3 or higher, mean curvature flow can develop singularities if the initial surface is not convex. We refer to [Man11] for a detailed discussion of mean curvature flow.

It is worth mentioning that it was first observed by Allen and Cahn in [AC79] that the velocity of the interfacial motion described by equation (1.9) was proportional to the mean curvature. Bronsard and Kohn [Bro90], [BK91] and Mottoni and Schatzman [MS89], [MS90] provided a rigorous proof of this under a variety of dimensional and regularity restrictions (see [Che92] for a detailed account), and in 1992, Chen proved this result in all dimensions under the relatively weak regularity conditions we state below in Assumptions 1.16, [Che92].

We now specify what we mean by the convergence of w, the solution to equation (1.9), to an indicator function of a set whose boundary evolves according to mean curvature flow. We choose the initial condition of equation (1.9), p, so that the following regularity conditions hold.

Assumptions 1.16. Let $p : \mathbb{R}^{\mathrm{d}} \to [0,1]$ and Γ_0 be given by;

$$\Gamma_0 = \{x \in \mathbb{R}^{\mathrm{d}} : p(x) = 1/2\}$$

(A) Γ_0 is C^a for some $a > 3$.

(B) For x inside Γ_0, $p(x) > \frac{1}{2}$, and for x outside Γ_0, $p(x) < \frac{1}{2}$.

(C) There exist $\eta, \mu > 0$ such that, for all $x \in \mathbb{R}^{\mathrm{d}}$, $|p(x) - \frac{1}{2}| \geq \mu\left(\mathrm{dist}(x,\Gamma) \wedge \eta\right)$.

The last condition indicates that w_0 has a slope near Γ_0 that is not *too shallow*. Together, these conditions ensure that the mean curvature flow started from Γ_0, $(\Gamma_t(\cdot))_{t\geq 0}$, exists up until some finite positive time $\mathscr{T}$. Let $d(x,t)$ denote the signed distance from x to Γ_t, chosen to be positive inside of Γ_t. The convergence of w can be rigorously stated as follows.

Theorem 1.17 ([Che92]). *Assume Assumptions 1.16 hold. Fix $T^* \in (0, \mathscr{T})$ and $k \in \mathbb{N}$. Then there exist $\varepsilon(k) > 0$ and $a(k), c(k) \in (0,\infty)$ such that, for all $\varepsilon \in (0, \varepsilon(k))$ and for all t in the interval $[a\varepsilon^2|\log(\varepsilon)|, T^*]$,*

1. *for x with $d(x,t) \geq c\varepsilon|\log(\varepsilon)|$ we have $w(t,x) \geq 1 - \varepsilon^k$,*
2. *for x with $d(x,t) \leq -c\varepsilon|\log(\varepsilon)|$ we have $w(t,x) \leq \varepsilon^k$.*

Since the hybrid zone for this deterministic model can be understood as the region where, for some small δ, we have $\delta < w(t,x) < 1 - \delta$, Theorem 1.17 could be paraphrased as saying that hybrid zones evolve by mean curvature flow.

Populations in nature are not entirely deterministic. Individuals experience fluctuations that come from the randomness of reproduction, an effect referred to as genetic drift. Naively, one could argue that, if genetic drift is small, then Theorem 1.17 guarantees that hybrid zones evolve as mean curvature flow. Sadly, mathematics is more complicated than that.

In [HRW12] the authors studied an Allen-Cahn equation with a mollified space-time white noise, whose solution takes value in $[-1, 1]$. It was proved that, if the mollifier is removed by using an increasing cut-off, solutions converge weakly to 0. On the other hand, if the intensity of the white noise is simultaneously converging to 0 sufficiently quickly, then the deterministic equation is recovered. This roughly says that the strength of the noise could drastically change the behaviour of the limiting process. Of course, space-time white noise is not a good model for genetic drift, but the problem is relevant for any type of noise. Will hybrid zones still evolve as curvature flow if we include the effect of genetic drift?

To model genetic drift we will use the SLFV as a base model. We need to include a mechanism of selection against heterozygosity into reproduction events to have any hope of recovering hybrid zones. To justify how we incorporate selection we return to our simplest model from the start of this section. Recall that, in this model, if the proportion of a-alleles in the population is $\overline{w}$, then, under selection against heterozygotes, the proportion of a-alleles in the pool of gametes of a reproduction event is:

$$\begin{aligned}\frac{\overline{w}^2 + \overline{w}(1-\overline{w})(1-\boldsymbol{s})}{1-2\boldsymbol{s}\overline{w}(1-\overline{w})} &= (1-\boldsymbol{s})\overline{w} + \boldsymbol{s}(3\overline{w}^2 - 2\overline{w}^3) + \mathcal{O}(\boldsymbol{s}^2),\\ &= (1-\boldsymbol{s})\overline{w} + \boldsymbol{s}(\overline{w}^3 + 3\overline{w}^2(1-\overline{w})) + \mathcal{O}(\boldsymbol{s}^2).\end{aligned}$$

Assuming birth events arrive as a Poisson process in our model, by using Poisson thinning, the last expression suggests we can split the reproduction events into two terms. The term $(1-\boldsymbol{s})\overline{w}$ corresponds to a neutral event, in which we sample a single individual and the offspring inherits its type, which happens with probability $1-\boldsymbol{s}$. The term $\boldsymbol{s}(\overline{w}^3 + 3\overline{w}^2(1-\overline{w}))$ corresponds to a selective event, happening with probability $\boldsymbol{s}$, in which three *potential parents* are sampled and the offspring is of type a if the majority of the potential parents are. This is what motivates the following definition of the Spatial Λ-Fleming Viot process with selection. The first introduction of genic selection in the SLFV process was done in [EVY20], but we present the version for selection against homozygotes described in [EFP17]. To avoid overcharging the nomenclature with more letters, we will still abbreviate this to SLFVS.

Definition 1.18 (Spatial Λ-Fleming Viot process with selection against heterozygotes (SLFVS)). The Spatial Λ-Fleming Viot process $\{w_t(x); x \in \mathbb{R}^{\mathrm{d}}, t \geq 0\}$ specifies a random function. Its dynamics are driven by a Poisson point process Π on $\mathbb{R}_+ \times \mathbb{R}^d \times \mathbb{R}_+ \times (0, 1]$, with intensity $dt \otimes dx \otimes \mu(dr)\nu_r(du)$, where $\mu(dr)$ is a measure on $\mathbb{R}_+$, $\nu_r(du)$ is a probability measure on $(0, 1]$ for every $r \in (0, \infty)$ and μ, ν_r satisfy (1.7).

At a point $(t, x, r, u) \in \Pi$, a reproduction event takes place in the closed ball $B(x, r)$. With probability $1 - \boldsymbol{s}$ this event is neutral, meaning that the event proceeds as in Definition 1.10 With probability $\boldsymbol{s}$ the event is selective, in which case,

1. Choose three *potential parental locations* z_1, z_2, z_3 independently from $B(x, r)$ uniformly at random.

2. For each $i = 1, 2, 3$, select independently a type k_i, which can be a or A, according to the distribution of types at z_i. That is $1_{k_i=a} \sim Ber(w_{t^-}(z_i))$.

3. Select $\widehat{k}$ as the most common allelic type in $\{k_1, k_2, k_3\}$. For all $y \in B(x, r)$ replace a proportion u of individuals by individuals of type $\widehat{k}$. That is,

$$w_t(y) = (1 - u)w_{t^-}(y) + u1_{\widehat{k}=a}.$$

Sites outside $B(x, r)$ are not affected by the reproduction event. That is, for all $y \notin B(x, r)$,

$$w_t(y) = w_{t^-}(y).$$

This process gives a way to study the interplay between genetic drift, spatial structure and the particular form of selection we are considering. For this model, we present the scaling used in [EFP17] and then state the main result therein. Consider a SLFVS where $\mu(dr)$ is a measure on $(0, R]$, for some fixed $R \in [0, \infty)$, and $\nu_r = \delta_u$ for some fixed $u \in [0, 1]$. We call this fixed u the impact parameter. In the scaling, we modify the units of space x, time t, the impact parameter u and the selection parameter s. For a fixed β we scale time from t to nt and space from x to $n^\beta x$. To be explicit, this means in the scaled model, the SLFVS w^n is driven by a Poisson point process Π^n with intensity

$$ndt \otimes n^\beta dx \otimes \mu^n(dr),$$

where $\mu^n(A) = \mu(n^\beta A)$ for all Borel subsets A of $(0, \infty)$. For the impact and selection, we assume they are of the form:

$$u_n = \frac{u}{n^{1-2\beta}}, \qquad \boldsymbol{s}_n = \frac{1}{\varepsilon_n^2 n^{2\beta}}. \tag{1.11}$$

Assumption 1.19. We suppose $(\varepsilon_n)_{n\in\mathbb{N}}$ is a sequence of strictly positive numbers such that $\varepsilon_n \to 0$ and that $\log(n)^{1/2}\varepsilon_n \to \infty$ as $n \to \infty$.

It is convenient to define the following constant

$$\sigma^2 = \frac{u}{2\mathrm{d}} \int_0^R \int_{\mathbb{R}^\mathrm{d}} |z|^2 \frac{|B(0,r) \cap B(z,r)|}{|B(0,r)|} dz d\mu(r).$$

The following is a simplified version of the main result in [EFP17].

Theorem 1.20. *Suppose that $\beta \in (0, 1/4)$ and that $(\varepsilon_n)_{n\in\mathbb{N}}$ is a sequence of positive numbers satisfying Assumption 1.19. Let $(w_t^n)_{t\geq 0}$ be the SLFVS with u_n and $\boldsymbol{s}_n$ satisfying (1.11). Suppose $w_0^n(x) = w_0(x)$, where w_0 satisfies Assumptions 1.16. Consider $\mathscr{T}$ and $d(x,t)$ as in Theorem 1.17 and let $T^* < \mathscr{T}$. Then, for all $k \in \mathbb{N}$, there exist $n_*(k) \in \mathbb{N}$ and $a_*(k), d_*(k) > 0$ such that, for all $n \geq n_*(k)$ and all $t \in [a_*\varepsilon_n^2|\log(\varepsilon_n)|, T^*]$,*

1. *for almost every x with $d(x, \sigma^2 t) \geq d_*\varepsilon_n|\log(\varepsilon_n)|$ we have $\mathbb{E}[w_t^n(x)] \geq 1 - \varepsilon_n^k$.*
2. *for almost every x with $d(x, \sigma^2 t) \leq -d_*\varepsilon_n|\log(\varepsilon_n)|$ we have $\mathbb{E}[w_t^n(x)] \leq \varepsilon_n^k$.*

Comparing Theorem 1.17 with Theorem 1.20 we see that the latter result says hybrid zones move as curvature flow in the limit of the particular scaled version of the SLFVS we stated. This similarity is not a coincidence. To prove Theorem 1.20 the strategy is to give a probabilistic proof of Theorem 1.17, which is flexible enough to later incorporate the genetic drift present in the SLFVS. We sketch the probabilistic proof of Theorem 1.17 as it not only serves to understand the scaling we state for the SLFVS but also to understand some of the new contributions of this thesis.

We start with a ternary branching Brownian motion started from a single individual at the point $x \in \mathbb{R}^\mathrm{d}$. That individual has an exponential lifetime with parameter ε^{-2}, during which it follows a Brownian motion (running at speed 2, to avoid problems with constants).

At the end of its lifetime, it dies and is replaced by three offspring. These offspring start at the position where the parent died, and then behave in the same way as their parent, independently of each other. We write $\boldsymbol{W}(t)$ for the set of historical paths of the branching Brownian motion up to time t. This encodes all the paths of all the Brownian particles that have lived up to time t. For a fixed function $p : \mathbb{R}^{\mathrm{d}} \to [0, 1]$, we define a voting procedure on $\boldsymbol{W}(t)$ as follows:

1. each leaf i votes 1 with probability $p(W_i(t))$ and 0 otherwise. Here $W_i(t)$ denotes the spatial position of the i-th leaf at time t,
2. at each branch point in $\boldsymbol{W}(t)$, the vote of the parent particle is the majority vote of the children.

This defines a voting procedure that runs from the leaves of $\boldsymbol{W}(t)$ to the root. Let $\mathbb{V}_p(\boldsymbol{W}(t))$ be the vote associated with the root. Let $\mathbb{P}^{\varepsilon}_x$ denote the law under which the branching Brownian motion branches at rate ε^{-2} and starts from a single ancestor at position x.

Proposition 1.21 ([EFP17]). *Let $u(t, x) = \mathbb{P}^{\varepsilon}_x[\mathbb{V}_p(\boldsymbol{W}(t)) = 1]$. Then u solves (1.9) with initial condition $p(x)$.*

With this representation, Theorem 1.17 becomes

1. for x with $d(x, t) \geq c\varepsilon|\log(\varepsilon)|$, we have $\mathbb{P}^{\varepsilon}_x[\mathbb{V}_p(\boldsymbol{W}(t)) = 1] \geq 1 - \varepsilon^k$,
2. for x with $d(x, t) \leq -c\varepsilon|\log(\varepsilon)|$, we have $\mathbb{P}^{\varepsilon}_x[\mathbb{V}_p(\boldsymbol{W}(t)) = 1] \leq \varepsilon^k$.

This is achieved in [EFP17] by an analysis of how a small bias changes the result of the voting system and a coupling of the distance of a Brownian motion to a surface evolving according to mean curvature flow to a one dimensional Brownian motion. Rather than expand on this proof, we explain how this gives Theorem 1.20. The core idea is to write a duality relationship for the SLFVS, similar to the one given by Proposition 1.14 for the SLFV, and then argue that the dual is close to a branching Brownian motion. We begin by defining the dual to the SFLVS. This is an adaptation of Definition 1.13 to include the distinction between selective and neutral events.

Definition 1.22. (SLFVS dual particle system) Let Π^n be a Poisson point process on $\mathbb{R}_+ \times \mathbb{R}^{\mathrm{d}} \times \mathbb{R}_+$, with intensity $ndt \otimes n^{\beta}dx \otimes \mu^n(dr)$. The dual particle system of the scaled SLFVS $(\mathcal{P}_t^n)_{t\geq 0} = (\xi_t^{n,1}, \xi_t^{n,2}, \dots, \xi_t^{n,N(t)})_{t\geq 0}$ is the $\bigcup_{m\geq 1}(\mathbb{R}^{\mathrm{d}})^m$-valued process with dynamics given as follows. Each point $(t,x,r) \in \Pi^n$ has a probability $1 - \boldsymbol{s}_n$ of being neutral. In that case

1. each $\xi_t^i \in B(x,r)$ is independently marked with probability u_n,
2. if at least one particle is marked, all marked particles are replaced with a single particle at a position picked uniformly in $B(x,r)$. If there were no marked particles nothing happens.

If the event is not neutral it is selective. In that case:

1. each $\xi_t^i \in B(x,r)$ is independently marked with probability u_n,
2. if at least one particle was marked, all marked particles are replaced by three particles. Those particles have positions picked independently and uniformly at random in $B(x,r)$. If there were no marked particles nothing happens.

Let $(\Xi^n(t))_{t\geq 0}$ be the collection of historical paths of the process $(\mathcal{P}_s)_{0\leq s\leq t}$. We can define a voting system on $(\Xi^n(t))_{t\geq 0}$ in a similar way to what we did for the historical paths of Brownian motion $(\boldsymbol{W}(t))_{t\geq 0}$. Let $p : \mathbb{R}^{\mathrm{d}} \to [0,1]$. Each leaf votes 1 with probability $p(\xi_t^{n,j})$. We trace the votes backwards in time as follows:

1. at selective events all marked individuals adopt the majority vote of the three parental particles,
2. at neutral events all marked individuals adopt the vote of the (unique) parent.

This gives us a unique vote associated to the root, which we denote by $\mathbb{V}_p(\Xi^n(t))$. We have the following duality between the SLFVS and $\Xi^n(t)$.

Proposition 1.23 ([EFP17]). *The scaled SLFVS is dual to $(\Xi_t)_{t\geq 0}$ (both driven by the same measure Π^n) in the sense that, for every $\psi \in \mathcal{C}(\mathbb{R}^{\mathrm{d}}) \cap L^1(\mathbb{R}^{\mathrm{d}})$,*

$$\mathbb{E}_p \left[\int_{\mathbb{R}^{\mathrm{d}}} \psi(x) w_t^n(x) dx \right] = \int_{\mathbb{R}^{\mathrm{d}}} \psi(x) \mathbb{P}_x \left[\mathbb{V}_p(\Xi^n(t)) = 1 \right] dx. \tag{1.12}$$

It is by virtue of Proposition 1.23 that we can study $\mathbb{P}_x\left[\mathbb{V}_p(\Xi^n(t)) = 1\right]$ to understand the hybrid zones in the SLFVS. Furthermore, it is from this fact that the proof of Theorem 1.20 is an adaptation of Theorem 1.17. Although it is not entirely trivial to see, under the scaling we perform, the dual process $\Xi^n(t)$ has three asymptotic properties as n goes to infinity:

1. individual ancestral lineages $\xi_t^{n,j}$ move, approximately, as Brownian motions;

2. coalescence of particles in $\Xi^n(t)$ happens with *arbitrarily low* probability;

3. selective events in $\Xi^n(t)$ happen at rate ε_n^{-2}.

Since selective events translate into branching in the system of random walkers $(\Xi^n(t))_{t\geq 0}$, the three properties imply that

$$\mathbb{P}_x\left[\mathbb{V}_p(\Xi^n(t)) = 1\right] \approx \mathbb{P}_x\left[\mathbb{V}_p(\boldsymbol{W}(t)) = 1\right],$$

and this is why it is not surprising that Theorem 1.20 follows from adapting the proof of Theorem 1.17 and Proposition 1.21.

Although Theorem 1.20 gives a precise description of the hybrid zones of the SLFVS under a particular scaling regime, a lot of questions are left unanswered about the model. What about other scaling regimes? What if a drastically different measure drives the SLFVS? What about different selection mechanisms? What about fluctuations of hybrid zones? This thesis partially answers some of these questions. Before introducing these results we explain how to relax the symmetry in the selection mechanism, resulting in a model that plays an important role in the results we later present.

Suppose that homozygotes aa and AA are not equally fit. Without loss of generality we assume that aa has a higher relative fitness than AA. Assuming that heterozygotes are still less fit than both homozygotes, and proceeding as we did to obtain (1.8), we recover the equation

$$\frac{\partial w}{\partial t} = \Delta w + w(1-w)(2w - (1-\gamma)), \tag{1.13}$$

for some $\gamma > 0$, where we have again used w to describe the proportion of a-alleles in the population. We could just perform a diffusive re-scaling, but equation (1.13) is very sensitive to the asymmetry. Thus, in addition to the diffusive rescaling, we set $\gamma = \nu\varepsilon$, which could be interpreted as a weak asymmetric regime. In that case, the equation becomes

$$\frac{\partial w}{\partial t} = \Delta w + \frac{1}{\varepsilon^2} w(1-w)(2w-(1-\nu\varepsilon)). \tag{1.14}$$

In [Goo18] it was shown that the limiting behaviour of the hybrid zone for a population evolving according to equation (1.14) is no longer mean curvature flow, but rather a mix of mean curvature flow and *constant* flow. That is the next result.

Theorem 1.24 ([Goo18], Theorem 2.4). *Let w^ε solve equation (1.14) with initial condition p satisfying Assumptions 1.16, and let*

$$\partial_t \widetilde{\Gamma}_t(s) = (\kappa_t(s) - \nu_\varepsilon)\mathbf{n}_t(s), \tag{1.15}$$

until the time $\mathscr{T}$ at which $\widetilde{\Gamma}$ develops a singularity. Write d for the signed distance to Γ (chosen to be positive inside Γ). Fix $T^* \in (0, \mathscr{T})$. Let $k \in \mathbb{N}$. There exists $\varepsilon_{\mathbb{d}}(k) > 0$, and $a_{\mathbb{d}}(k), c_{\mathbb{d}}(k) \in (0,\infty)$ such that for all $\varepsilon \in (0, \varepsilon_{\mathbb{d}})$ and t satisfying $a_{\mathbb{d}}\varepsilon^2|\log\varepsilon| \le t \le T^*$,

1. *for x such that $d(x,t) \ge c_{\mathbb{d}}\varepsilon|\log\varepsilon|$, we have $v^\varepsilon(t,x) \ge 1 - \varepsilon^k$;*

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

Theorem 1.24 is a special case of Theorem 1.3 of [AHM08] who consider the generation and propagation of a sharp interface for more general versions of the Allen-Cahn equation with a slightly unbalanced bistable nonlinearity. However, [Goo18] proves also that an analogous result holds for a SLFVS with asymmetric selection, under a similar reescaling and conditions to Theorem 1.20.

1.6 Hybrid zones and how to stop them.

Although Theorem 1.20 and the preceding result give a precise description of the hybrid zones of the SLFVS under a particular scaling regime, there is much more we would like to understand about this model. In the case of asymmetric selection, equation (1.14) presents two opposing forces. On the one hand, the curvature flow pushes hybrid zones inward (with respect to the normal direction) on zones of positive curvature, shrinking the range of the hybrid zone. On the other hand the constant flow pushes outwards, always trying to expand the range of the fittest type of individual in the population. It is then possible to create an equilibrium: for example, in dimension 2 this can be achieved by considering the hybrid zone starting as a circle of radius $1/\nu$. By contrast, if the hybrid zone is a straight line, splitting the Euclidean plane in two, then there is no curvature and the constant flows push the hybrid zone indefinitely. In that case, the fittest allelic type, aa, will take over the population. In joint work with Alison Etheridge and Mitchell Gooding, we try to understand the possible equilibrium or lack thereof, created by other factors. In particular, we focus on the interplay between selection, the shape of the domain and the intensity of the genetic drift.

Populations in nature are not free to move in the whole space. Plants are surrounded by mountains that make the spreading of seeds impossible; fish face seawalls they cannot traverse; and so there are numerous examples one can think of where the habitat constrains the movement and reproduction of individuals. We call those *barriers.* We seek to understand how those barriers can disrupt the movement of hybrid zones.

As a starting point, we return to the deterministic model given by equation (1.14). We assume the habitat of the individuals is a subset of Euclidean space, given by some domain Ω. Its effect is only to impede the movement of individuals, reflecting them into the domain whenever they reach the boundary. Considering how we deduce (1.14), a deterministic model with which to study this problem is the system:

$$\begin{cases} \partial_t u^\varepsilon = \Delta u^\varepsilon + \frac{1}{\varepsilon^2} u^\varepsilon (1-u^\varepsilon)(2u^\varepsilon - (1-\nu\varepsilon)) & x \in \Omega,\, t>0, \\ \partial_n u^\varepsilon = 0 & x \in \partial\Omega,\, t>0, \\ u^\varepsilon(x,0) = \mathbf{1}_{x_1 \geq 0} & x \in \Omega, \end{cases} \tag{1.16}$$

with $\nu \in (0,\infty)$, ε a small parameter, Ω an unbounded domain in $\mathbb{R}^{\mathrm{d}}$, and $\partial_n u^\varepsilon$ the normal derivative at the boundary.

While we could try to understand how hybrid zones behave under the model given by equation (1.16), in a similar way to Theorem 1.20, this appears to be technically challenging. We expect hybrid zones to retain some of the flavour of mean curvature flow. However, mean curvature flow does not behave well with boundaries, and some care is needed [Buc05], [Hui89]. Consequently, we restrict ourselves to a subclass of domains for which we can overcome this issue.

Here we shall ask about propagation of solutions through other unbounded domains. We are particularly interested in the question of when the spread of a population may be halted, for example when passing through an isthmus, and whether this will change in the presence of noise. Theorems 1.26 and 1.27 below identify conditions under which such 'blocking' can and cannot occur for a simple domain illustrated in Figure 1.1 below. We later indicate how to extend these results to a multitude of other domains. Theorem 1.34 discusses a stochastic analogue of (1.16) based on the spatial Λ-Fleming-Viot process. For $\Omega = \mathbb{R}^d$, Theorem 1.20 and Theorem 1.24 both identified noisy regimes in which the behaviour is the same as for the deterministic equation; here we complement those results by also identifying conditions under which the noise is strong enough to break down the structure induced by the potential in (1.16).

[BBC16] consider a general bistable nonlinearity $f(u)$ in place of the particular cubic term that arises in our application, and they take the domain to be 'cylinder-like':

$$\Omega = \left\{(x_1, x'), x_1 \in \mathbb{R}, x' \in \phi(x_1) \subseteq \mathbb{R}^{\mathrm{d}-1}\right\}. \tag{1.17}$$

with $\phi(x_1)$ regular enough so Ω has a $\mathcal{C}^1(\mathbb{R}^{\mathrm{d}})$ boundary. For our equation, their results show that depending on the geometry of the domain, we can have different long-term behaviours of the solution of equation (1.16).

Theorem 1.25 ([BBC16], Theorems 1.4, 1.5, 1.6, 1.7, paraphrased). *Depending on the geometry of the domain Ω we have one of three possible asymptotic behaviours of the solution of equation (1.16).*

1. *there can be* complete invasion*, that is $u(x,t) \to 1$ as $t \to \infty$ for every $x \in \Omega$.*

2. *there can be* blocking *of the solution, meaning that $u(x,t) \to u_\infty(x)$ as $t \to \infty$, with $u_\infty(x) \to 0$ as $x_1 \to -\infty$.*

3. *there can be* axial partial propagation*, meaning that $u(x,t) \to u_\infty(x)$ as $t \to \infty$, with $1 > \inf_{x\in\mathbb{R}\times B_R} u_\infty(x) > c > 0$ for some $R > 0$, where B_R is the ball of radius R centred at* 0 *in* $\mathbb{R}^{d-1}$.

Which behaviour is observed depends on the geometry of the domain Ω. There will be complete invasion if Ω is decreasing as $x_1 \to -\infty$; there will be axial partial propagation if it contains a straight cylinder of sufficiently large cross-section; and there can be blocking if there is an abrupt change in the geometry.

We remark that in [BBC16] the convention is to consider invasions from $-\infty$ to $+\infty$; our choice, which is the opposite, makes it easier to borrow results from [Goo18]. Our results complement those of [BBC16], while being more quantitative in the conditions imposed on the geometry of the domain and allowing for the inclusion of noise, corresponding to 'genetic drift'. However, our results do not contain or imply the ones of [BBC16]. We introduce the new parameter ε to the equation and study the behaviour of the equation as ε gets close to 0, which prevents a direct comparison between the two sets of results.

Since hybrid zones are typically narrow compared to the range of the population, we take ε to be small, which necessitates taking γ to be small if the motion of the hybrid zone is not to be unreasonably fast. As we shall see, the motion of the hybrid zone can be blocked if the domain Ω has an abrupt wide opening. Although we shall consider more general domains later on, to introduce the main ideas we shall begin by focusing on a domain of the form shown in Figure 1.1. This means that, in the notation above, $\phi(x) \equiv 1_{\|x'\|<R_0}$ for $x_1 < 0$, $\phi(x) \equiv 1_{\|x'\|<r_0}$ for $x_1 > 0$, where $r_0 < R_0$. This domain has the advantage of notational simplicity, while allowing us to introduce all the key ideas required to understand conditions for blocking/invasion in much more general domains in the next section. While this domain does not have the regularity of the domains in [BBC16], results in [CG05] gives us existence and uniqueness of the solutions of (1.16) in this domain (after deciding that the normal in the corners points *in a diagonal direction*) and also provides that invasion or blocking can happen in this setting. Therefore, we do not lose tools by focusing on this domain first.

Our first result shows how the behaviour of (1.16) on the domain Ω can be very different from that on the whole of Euclidean space. As $\varepsilon \to 0$, the solution will look increasingly

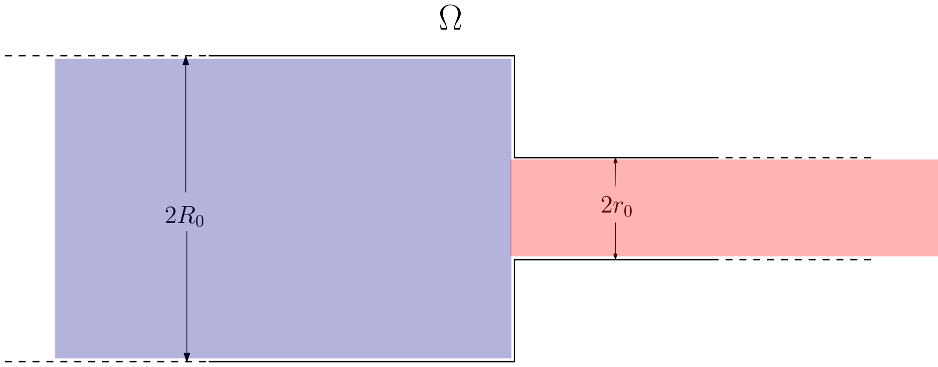

Figure 1.1: Representation of the domain Ω of Theorem 1.26 and the initial condition, with individuals carrying the fittest allele in red (on the right side of the domain) and the less fit in blue (on the left side of the domain). The two populations meet at the opening point $x_1 = 0$.

like the indicator function of a region whose boundary evolves according to (1.15), with the additional condition that it is perpendicular to the boundary $\partial\Omega$ of Ω where the two intersect. If $r_0 = R_0$ then the solution will simply propagate from right to left. Indeed it will converge to a travelling wave, moving in x_1, whose profile is given by

$$u(x,t) = \frac{1}{2}\left(1 - \tanh\left(\sqrt{\frac{1}{2\varepsilon^2}}(x - ct)\right)\right), \tag{1.18}$$

with $c = \gamma/\sqrt{2}$. If $R_0 \gg r_0$, then the solution will try to spread out from the opening at $x_1 = 0$. Approximating the solution as above, the interface between the region where u is close to 0 and where it is close to 1 is pushed to the left by the constant normal flow, while being pushed back by the mean curvature (a kind of 'surface tension'). In the limit, these two forces will balance if $\nu = \kappa$, that is precisely when the interface to the left of the origin is a hemispherical shell of radius $r = \frac{d-1}{\nu}$. This is, of course, the same as the radius at which the forces will balance for a solution on the whole of $\mathbb{R}^d$ started from the indicator function of a sphere. However, working on all of $\mathbb{R}^d$ does not give a satisfactory example of blocking, as any perturbation above or below this radius will result in complete invasion or extinction of the dominant phenotype respectively. The blocking that we see in domains such as Ω is robust to this type of perturbation. Making a rigorous proof for this behaviour is the first of our original results.

Theorem 1.26. *Let u^ε denote the solution to equation (1.16) with Ω as in Figure 1.1. Suppose $r_0 < \frac{\mathbb{d}-1}{\nu} \wedge R_0$. Define*

$$N_{\mathbb{r}} = \{x \in \Omega \,:\, \|x\| = \mathbb{r},\, x_1 < 0\},$$

where $R_0 > \mathbb{r} > r_0$, and let $\widehat{d}(x)$ be the signed (Euclidean) distance of any point $x \in \Omega$ to $N_{\mathbb{r}}$ (chosen to be negative as $x_1 \to -\infty$). Let $k \in \mathbb{N}$. Then there are $\widehat{\varepsilon}(k) > 0$ and $M(k) > 0$ such that, for all $\varepsilon \in (0, \widehat{\varepsilon})$, and all $t \geq 0$,

$$\text{if } x = (x_1, ..., x_{\mathbb{d}}) \in \Omega \text{ is such that } \widehat{d}(x) \leq -M(k)\varepsilon|\log(\varepsilon)| \text{ then } u^\varepsilon(x,t) \leq \varepsilon^k.$$

In other words, if the aperture r_0 is too small, and R_0 is large enough, then, for sufficiently small ε, blocking occurs. The converse is also true in the following sense.

Theorem 1.27. *Let u^ε denote the solution to equation (1.16) with Ω as in Figure 1.1. Suppose $R_0 > r_0 > \frac{\mathbb{d}-1}{\nu}$, then for all $x \in \Omega$ and $\delta > 0$ there are $\widehat{t} := \widehat{t}(x_1, R_0, r_0) > 0$ and $\widehat{\varepsilon}$ such that, for all $\varepsilon \in (0, \widehat{\varepsilon})$ and $t \geq \widehat{t}$ we have $u^\varepsilon(t, x) \geq 1 - \delta$.*

Together, Theorem 1.26 and Theorem 1.27 say that there is a sharp transition at the critical radius $(\mathbb{d} - 1)/\nu$. To the best of our knowledge, such an explicit condition for this phase transition is new.

The proof uses a voting system, similar to the one in Proposition 1.21, but where the motion is given by Brownian motion with reflection at the boundary. To prove Theorem 1.26 we use an upper solution argument to compare the system with initial condition $1_{x_1 \geq 0}$ to one with initial condition approximately given by $1_{\{x_1 \geq 0\} \cup B(0,r)}$ for some small $r > 0$. The advantage of this new initial condition is that the radial symmetry allows us to follow closely most of the calculations in the probabilistic proof of Theorem 1.17 with asymmetric selection found in [Goo18]. For Theorem 1.27 we use a lower solution argument: that is, we find subsolutions with a radial symmetry (which again allows one to use the calculations in [Goo18]) that are arbitrarily close to 1 in any point of Ω. A generalisation to a large family of domains can be found in Theorem 1.29 and Theorem 1.30 below.

Remark 1.28. While, for simplicity and clarity of the proofs, we state Theorem 1.26 as a one-sided inequality, as a by-product of our methods one can obtain a lower bound on the right hand side of the domain. That is, in the context of Theorem 1.26, if x is such that $x_1 \geq M(k)\varepsilon|\log(\varepsilon)|$, then $u^\varepsilon(x,t) \geq 1-\varepsilon^k$. The proof follows exactly the same arguments as Theorem 1.27; for more details see Theorem 2.4 in [Goo18]. Biologically, the function $u^\varepsilon(x,t)$ models the proportion of a particular type in the population at position x and time t. Hence Theorem 1.26 says that, if we have blocking, we have coexistence of the two alleles, with a narrow interface of width $\mathcal{O}(\varepsilon|\log(\varepsilon)|)$ between the aa and AA homozygotes near the opening of the domain. Note this is in sharp contrast to Theorem 1.27, where we have fixation of the type a-allele across the whole population.

Other domains

The domain Ω of Theorem 1.26 is very special. [MNL06] consider a plane curve evolving according to equation (1.15) in a two-dimensional cylinder with a periodic saw-toothed boundary. More precisely, they take

$$\Omega_\delta = \left\{(x,y) \in \mathbb{R}^2 : x \in \big(-H - h_\delta(y), H + h_\delta(y)\big)\right\},$$

where H is a positive constant and $h_\delta(y) = \delta h_1(y/\delta)$ with h_1 being a smooth, 1-periodic function, satisfying

$$h_1(0) = h_1(1) = 0, \quad h_1(y) \geq 0, \quad \forall y \in \mathbb{R}; \tag{1.19}$$

see Figure 1.2 for an illustration. They impose the additional condition that at the points where they meet, the curve and the boundary of the domain are perpendicular, and their convention is that the normal to the curve points into the 'bottom half' of the domain. This corresponds to the limit of the solution u^ε of (1.16) on their domain as $\varepsilon \to 0$. There will be no travelling wave solution to their equation, in the classical sense, unless the cylinder is flat, and so they define a solution to be a *periodic* travelling wave if $\Gamma_{t+T_\delta}(s) = \Gamma_t(s) + \delta$ for some $T_\delta > 0$. Its effective speed is then $c_\delta := \delta/T_\delta$. Recalling that $h_\delta(y) = \delta h_1(y/\delta)$ and letting $\delta \to 0$, they then investigate the homogenisation limit of the travelling wave, with corresponding speed $c_0 := \lim_{\delta\to 0} c_\delta$. They show, in particular, that $c_0 > 0$ if $\nu H > \sin\alpha$ where α is determined by $\tan\alpha = \max_y h'(y)$, but that the wave is blocked for small enough δ when $\nu H < \sin\alpha$. In Section 2.1.8 we shall sketch the proof of the following

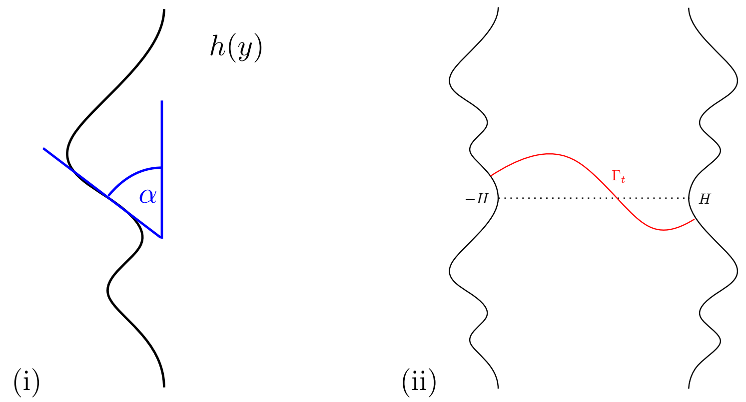

Figure 1.2: (i) An example of function h for the boundary of the domain (ii) An example of a periodic saw-toothed domain considered by [MNL06]

multidimensional analogue of this blocking result for solutions to (1.16) in cylindrical domains from our approach. See Remark 1.31 for an explanation of how the two results are connected.

Theorem 1.29. *Suppose that u^ε solves (1.16) where $\Omega \subseteq \mathbb{R}^{\mathbb{d}}$ is defined as in (1.17) with*

$$\phi(x_1) = \{x : \|x'\| \leq H + h(-x_1)\},$$

and h being a non-negative, $C^1(\mathbb{R})$ (not necessarily periodic) function. Suppose that

$$\inf_{z>0} \left\{ H + h(z) - \left(\frac{\mathbb{d}-1}{\nu}\right) \frac{h'(z)}{\sqrt{1+h'(z)^2}} \right\} < 0. \tag{1.20}$$

Fix $k \in \mathbb{N}$. There exist $x_0 < 0$, $\widehat{\varepsilon}(k) > 0$ and $M(k) > 0$ such that, for all $\varepsilon \in (0, \widehat{\varepsilon})$ and $t \geq 0$,

$$\textit{if } x = (x_1, ..., x_{\mathbb{d}}) \in \Omega \textit{ is such that } x_1 \leq x_0 - M(k)\varepsilon|\log(\varepsilon)| \textit{ then } u^\varepsilon(x,t) \leq \varepsilon^k.$$

In other words, the solution is blocked if the cylindrical domain Ω opens out too quickly. Indeed, in the proof of Theorem 1.29 we compute the angle at which the boundary *opens up*. We shall see that condition (1.20) ensures that this angle is 'big enough'. As a consequence, when (1.20) holds, we can insert a portion of a spherical shell of radius less than $(\mathbb{d}-1)/\nu$ into the domain in such a way that expanding the shell radially one stays within the domain, at least for a short time. With this we can adapt the analysis performed for Theorem 1.26 and conclude that there is blocking. Conversely, if the angle is not big enough, we have invasion.

Theorem 1.30. *Suppose that u^ε solves (1.16), where $\Omega \subseteq \mathbb{R}^{\mathbb{d}}$ is defined as in (1.17) with*

$$\phi(x_1) = \{x : \|x'\| \le H + h(-x_1)\},$$

and h is a non-negative, $C^1(\mathbb{R})$ function. Suppose that

$$\inf_{z>0}\left\{H + h(z) - \left(\frac{\mathbb{d}-1}{\nu}\right)\frac{h'(z)}{\sqrt{1+h'(z)^2}}\right\} > 0. \tag{1.21}$$

Then for all $x \in \Omega$ and $\delta > 0$, there are $\widehat{t} := \widehat{t}(x_1, R_0, r_0) > 0$ and $\widehat{\varepsilon}$ such that, for all $\varepsilon \in (0, \widehat{\varepsilon})$ and $t \ge \widehat{t}$ we have $u^\varepsilon(t,x) \ge 1-\delta$.

Remark 1.31. Note that $h'(z)/\sqrt{1+h'(z)^2} = \sin(\arctan(h'(z))$. So this fractional term is just the sine of the slope of the boundary when the first coordinate is $x_1 = z$. Using again $\alpha = \max_y h'(y)$ and connecting back to the results in [MNL06], one can check that for $\mathbb{d} = 2$, if $H\nu < \sin\alpha$, then, for small enough δ, equation (1.20) is satisfied and so we see blocking for the domain Ω_δ, whereas if $H\nu > \sin\alpha$ equation (1.21) is satisfied for any $\delta > 0$ and there is no blocking. Thus we have recovered a multi-dimensional analogue of Theorem 2.1 in [MNL06], with weaker conditions on the function h.

While conditions (1.20) and (1.21) are hard to verify in general, there are cases where it is trivial to determine if one of them holds. For example, narrowing domains, or those that only narrow beyond the point at which their diameter is smaller than $2(\mathbb{d}-1)/\nu$, clearly satisfy (1.21). See Figure 1.3 for examples of such domains.

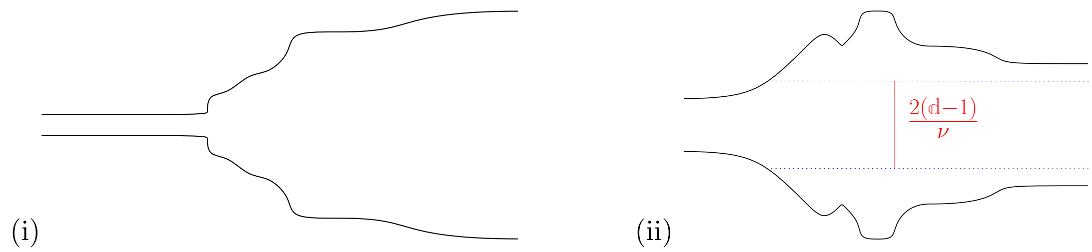

Figure 1.3: (i) Example of a domain for which the opening only gets narrower. (ii) Example of a domain for which the opening gets smaller once it is less than $(\mathbb{d}-1)/\nu$.

1.6.1 The effect of the strength of genetic drift

In the previous subsection, we justified using equation (1.16) to model the motion of a hybrid zone in population genetics. However, a deterministic equation like this rests on the assumption of infinite population density. We should like to understand the effects of the random fluctuations caused by reproduction in a finite population.

Several ways of introducing noise into the Allen-Cahn equation have been considered in the literature. [Fun99] considers (1.16) on a bounded domain in $\mathbb{R}^2$, and with an additional additive noise term of the form $\xi^\varepsilon(t)/\varepsilon$, where ξ^ε is centred and smooth in t, but behaves like white noise in the limit as $\varepsilon \to 0$. Once again the solution generates an interface as $\varepsilon \to 0$, which Funaki calls *randomly perturbed motion by curvature*. [Alf+18] consider (1.16) with the same form of additive noise as [Fun99], this time on a bounded domain in $\mathbb{R}^d$. Their results show that, just as in the deterministic case, an interface develops in a very short time, and that the law of motion of the interface is now given by mean curvature flow perturbed by white noise. The profile of the solution near the interface is not destroyed by the random noise, as long as the noise depends only on the time variable. [Lee18] considers a space-time noise, but the noise is smooth in space, and although it is shown that an interface is generated, the law of motion of the interface is not established. We already mentioned that [HRW12] consider the equation

$$\partial_t u = (\Delta u + u - u^3)dt + \sigma dW,$$

on $\mathbb{R}^2$, where W is a space-time white noise, mollified in space. Setting $v = (1+u)/2$, we recover (1.16) with $\varepsilon = 1$ and an additional mollified white noise. They show that, if the mollifier is removed, then the solution converges weakly to zero, but that if the intensity of the noise simultaneously converges to zero sufficiently quickly, they recover the solution to the deterministic equation. In other words, unless the noise is small, it can completely destroy the structure of the deterministic equation.

Additive white noise (with or without a spatial component) is not a good model for the randomness due to reproduction in a biological population, but the same challenges may arise for genetic drift. Our approach is to use the spatial Λ-Fleming-Viot process, for which the replacement mechanism produces a noise that better represents genetic drift, to approximate the solution of (1.16). In Section 2.2 we use an approach that mimics the proof of Theorem 1.20 to provide a stochastic analogue of Theorem 1.26. Furthermore, we prove a complementary result, in which we identify rather precisely the relative strength of genetic drift and selection that results in breakdown of the structure of the deterministic equation.

The key to understanding blocking in the presence of noise is to establish whether a stochastic analogue of Theorem 1.24 holds on the whole of Euclidean space, and so, for the purposes of this introduction, we shall take $\Omega = \mathbb{R}^{\mathrm{d}}$.

First, we define a version of the Spatial Λ-Fleming-Viot process with selection introduced in [Goo18]. This process generalises a bit the process in Definition 1.18 and provides a stochastic analogue of the solution to equation (1.16). With a slight abuse of nomenclature, we will still refer to this process as SLFVS.

Definition 1.32 (Spatial Λ-Fleming Viot process with asymmetric selection against heterozygotes (SLFVS)). Fix $u, \gamma \in (0, 1]$, $\boldsymbol{s} \in (0, 1/(1+\gamma))$, $\mathcal{R} > 0$. Let μ be a finite measure on $(0, \mathcal{R}]$. Let Π be a Poisson point process on $\mathbb{R}_+ \times \mathbb{R}^{\mathrm{d}} \times (0, \mathcal{R}]$ with intensity measure

$$dt \otimes dx \otimes \mu(dr). \tag{1.22}$$

The *spatial Λ-Fleming-Viot process with selection (SLFVS)* driven by Π, with *selection coefficient $\boldsymbol{s}$* and *impact parameter u*, 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 $B(x, r)$ of radius r centred on x. With probability $1 - (1+\gamma)\boldsymbol{s}$ the event is *neutral*, in which case the event proceeds as in Definition 1.10. With the complementary probability $(1+\gamma)\boldsymbol{s}$ the event is *selective*, in which case:

1. Choose three 'potential' parental locations $z_1, z_2, z_3 \in \mathbb{R}^{\mathrm{d}}$ independently and uniformly at random from $B(x, r)$. At each of these sites sample 'potential' parental types α_1, α_2, α_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 $\alpha_1, \alpha_2, \alpha_3$, except that if precisely one of α_1, α_2, α_3, is a, with probability

$\frac{2\gamma}{3+3\gamma}$ set $\widehat{\alpha} = a$.

2. For every $y \in B(x,r)$ set $w_t(y) = (1-u)w_{t-}(y) + u1_{\{\widehat{\alpha}=a\}}$.

Before going any further, we explain the origin of the reproduction rule in Definition 1.32. Comparing to our justification of equation (1.16), recalling that w is the proportion of a-alleles in the population, we first write

$$\begin{aligned} &\boldsymbol{s}w(1-w)(2w-(1-\gamma)) \\ &\quad = \left((1-(1+\gamma)\boldsymbol{s})w + (1+\gamma)\boldsymbol{s}\left(w^3 + 3w^2(1-w) + \left(\frac{2\gamma}{3(1+\gamma)}\right)3w(1-w)^2\right) - w\right). \end{aligned} \tag{1.23}$$

In the SLFVS framework, reproduction events arrive as a Poisson process (as opposed to the deterministic generations times in our justification of (1.16)). With probability $1-(1+\gamma)\boldsymbol{s}$ an event is neutral, so that the chance that offspring are of type a is simply the probability w that a randomly chosen parent is of type a, and we recognise the first term on the right of (1.23). With probability $(1+\gamma)\boldsymbol{s}$, the event is selective. If we sample three individuals from the population, the probability that the majority are type a is $w^3 + 3w^2(1-w)$; whereas the probability that exactly one is type a is $3w(1-w)^2$. In the latter case, we multiply further by $2\gamma/(1+\gamma)$ to recover the probability that the offspring are type a, and we recognise the second and third terms on the right of (1.23). Overall, equation (1.23) represents the change in proportion of a alleles in the portion of the population replaced during the event.

Remark 1.33. We could equally have taken two types of selective events, one corresponding to selection against heterozygosity, and one to genic selection. To see why, we rewrite the part of (1.23) corresponding to selective events as

$$\begin{aligned} &(1+\gamma)\boldsymbol{s}\left(w^3 + 3w^2(1-w) + \frac{2\gamma}{3(1+\gamma)}3w(1-w)^2 - w\right) \\ &\qquad = \boldsymbol{s}\left(w^3 + 3w^2(1-w) - w\right) + \gamma\boldsymbol{s}\left(w^3 + 3w^2(1-w) + 2w(1-w)^2 - w\right) \\ &\qquad = \boldsymbol{s}\left(w^3 + 3w^2(1-w) - w\right) + \gamma\boldsymbol{s}\left(w^2 + 2w(1-w) - w\right) \end{aligned} \tag{1.24}$$

This suggests that with probability $\boldsymbol{s}$ an event corresponds to selection against heterozygosity: three potential parents are sampled and offspring adopt the type of the majority of those individuals; with probability $\gamma \boldsymbol{s}$ an event corresponds to genic selection: *two* potential parents are sampled and if either of them is type a, then the offspring is of type a.

Although this leads to the same process of allele frequencies as the apparently more complex mechanism that we introduced in Definition 1.32, in our proof it will be convenient to have a single rule for selective events, based on three potential parents, which will be encoded in the function g of equation (2.1) below.

To study the relationship between genetic drift and selection we will introduce two possible scalings for the SLFVS. While the choice of the scaling parameters may seem obscure, once we work with a branching and coalescing dual for the SLFVS, the reason for these choices will become clear. If in both cases we fix the same values for the parameters that dictate selection (corresponding to $\boldsymbol{s}_n$ and γ_n below), the difference between the two scalings is entirely in the strength of the genetic drift, while the 'deterministic part' of the evolution (corresponding to dispersion and selection) is identical. We return to this in Remark 2.28.

Before presenting the scalings, we recall that Assumption 1.19 states that $(\varepsilon_n)_{n \geq \mathbb{N}}$ is a sequence of positive numbers converging to 0, for which $\varepsilon_n \log(n)^{1/2}$ goes to infinity.

Weak noise/selection ratio

Our first scaling is what we shall call the *weak noise/selection ratio* regime. In this regime selection overwhelms genetic drift. It mirrors that explored in [EFP17] and is also considered in [Goo18]. For each $n \in \mathbb{N}$, and for some $\beta \in (0, 1/4)$, we define the finite measure μ^n on $(0, \mathcal{R}_n]$, where $\mathcal{R}_n = n^{-\beta}\mathcal{R}$, by setting $\mu^n(B) = \mu(n^\beta B)$ for all Borel subsets B of $(0, \infty)$. In the weak noise/selection ratio regime the rescaled SLFVS is driven by the Poisson point process Π^n on $\mathbb{R}^+ \times \mathbb{R}^{\mathrm{d}} \times (0, \infty)$ with intensity measure

$$ndt \otimes n^\beta dx \otimes \mu^n(dr). \tag{1.25}$$

Here the linear dimension of the infinitesimal region dx is scaled by n^β (so that when we integrate, the volume of a region is scaled by $n^{\mathrm{d}\beta}$). Let $\nu > 0$. We denote by u_n the impact parameter and by $\boldsymbol{s}_n$ the selection parameter at the nth stage of the scaling. They will be given by

$$\gamma_n = \nu\varepsilon_n, \qquad u_n = \frac{u}{n^{1-2\beta}}, \qquad \boldsymbol{s}_n = \frac{1}{\varepsilon_n^2 n^{2\beta}}. \tag{1.26}$$

where we assume ε_n satisfies Assumption 1.19. Adapting the proof of Theorem 1.11 in [EVY20], and the arguments in Section 3 of [EFP17], one can show that under this scaling, for large n, the SLFVS will be close to the solution of problem (1.16).

Strong noise/selection ratio

We shall refer to our second scaling as the *strong noise/selection ratio* regime. In this regime genetic drift overcomes selection. In this scenario, we consider any sequence of impact parameters $(u_n)_{n\in\mathbb{N}} \subseteq (0,1)$. Consider $\beta \in (0, 1/2)$ and let $\widehat{u}_n := u_n n^{1-2\beta}$. We scale time by $n/\widehat{u}_n$ and space by n^β. At the nth stage of the rescaling, Π^n is a Poisson measure on $\mathbb{R}^+ \times \mathbb{R}^{\mathrm{d}} \times (0,\infty)$ with intensity measure

$$\frac{n}{\widehat{u}_n} dt \otimes n^\beta dx \otimes \mu^n(dr). \tag{1.27}$$

We consider a sequence of selection coefficients, $(\boldsymbol{s}_n)_{n\in\mathbb{N}} \subseteq (0,1)$, satisfying one of the following conditions:

$$\begin{cases} \boldsymbol{s}_n n^{2\beta} \to 0 & \liminf_{n\to\infty} u_n \log n < \infty \text{ or } \mathrm{d} \geq 3, \\ \frac{\boldsymbol{s}_n n^{2\beta}}{u_n \log n} \to 0 & \liminf u_n \log n = \infty \text{ and } \mathrm{d} = 2. \end{cases} \tag{1.28}$$

The strength of genetic drift (noise) is determined by the impact parameter. The first case includes some choices of impact that were allowed in the first (weak noise/selection ratio) regime; it is the strength of drift *relative to selection* that matters. In this regime, we can take the parameters $(\gamma_n)_{n\in\mathbb{N}}$ that dictate the asymmetry in our selection to be any sequence in $(0,1)$.

The stochastic result

With the two scaling regimes defined we can state a result. Recall that for this model we are working on the whole of Euclidean space.

Theorem 1.34. *Let $(w_t^n)_{t\geq 0}$ be the scaled SLFVS with initial condition $w_0^n(x) = \mathbf{1}_{B(0,(\mathrm{d}-1)/\nu)}(x)$.*

1. *Under the weak noise/selection ratio regime, for each* $k \in \mathbb{N}$*, there exist* $n_*(k) < \infty$*, and* $d_*(k) \in (0,\infty)$*, such that for all* $n \geq n_*$*, and* $t > 0$*,*

   (a) *for almost every* x *such that* $\|x\| \geq (\mathbb{d}-1)/\nu + d_*\varepsilon_n|\log \varepsilon_n|$*, we have* $\mathbb{E}[w_t^n(x)] \leq \varepsilon_n^k$*;*

   (b) *for almost every* x *such that* $\|x\| \leq (\mathbb{d}-1)/\nu - d_*\varepsilon_n|\log \varepsilon_n|$*, we have* $\mathbb{E}[w_t^n(x)] \geq 1 - \varepsilon_n^k$*.*

2. *Under the strong noise/selection ratio regime, for all* $t > 0$*, there is* $\sigma^2 > 0$ *(independent of* t*) and* $n_* := n_*(t) < \infty$*, such that for all* $n \geq n_*$*, and all* $x \in \mathbb{R}^{\mathbb{d}}$*,*

$$\left|\mathbb{E}_{w_0}[w_t^n(x)] - \mathbb{P}_x\left[\|W(\sigma^2 t)\| \leq (\mathbb{d}-1)/\nu\right]\right| \leq \varepsilon, \tag{1.29}$$

   where $(W_t)_{t\geq 0}$ *is a standard Brownian motion in* $\mathbb{R}^{\mathbb{d}}$ *and the subscript* x *on* $\mathbb{P}_x$ *indicates that* $W(0) = x$*.*

More generally, one can show that in the strong noise/selection ratio regime for $x \neq y$, $w_t(x)$ and $w_t(y)$ decorrelate as $n \to \infty$. In other words, in the weak noise/selection ratio regime the SLFVS behaves approximately as the deterministic equation (1.16), while in the strong noise/selection ratio regime the genetic drift is strong enough to overcome the effects of selection and it breaks down the interface. (The corresponding breakdown of the interface under the strong noise/selection ratio regime in one dimension requires us to replace $u_n \log n$ above by $u_n n^{1/2}$.)

The intuition for why the last phenomenon occurs is much easier to grasp if we compute the rate at which particles in the dual make an excursion of order 1, before coalescing. In dimension 2, if $u_n \log(n)$ goes to infinity, the rate at which particles are born and make an excursion of order 1 is given by $\frac{\boldsymbol{s}_n n^{2\beta}}{u_n \log(n)}$ (c.f. Remark 2.39). Our restriction of $\boldsymbol{s}_n$ in (1.28) is just imposing that particles in the dual never do a meaningful excursion, meaning that in the limit ancestral lineages are given by a single Brownian particle. It is then no surprise that we see the heat equation. Furthermore, in dimension 2, $u_n = (\log(n))^{-1/2}$ and $\boldsymbol{s}_n = (\log(n))^{1/3}/n^{2\beta}$ fall into such a scaling regime. In that case, the branching rate of an ancestral lineage is going to infinity, but the probability of seeing more than one particle in the dual in the limit is converging to 0. This illustrates that, even if dual particles are created arbitrarily fast, genetic drift is sufficiently strong that the coalescence happens over an even faster time scale.

Remark 1.35. As we discuss in a little more detail in Section 2.2.3, the conditions (1.28) are essentially optimal. In this regime, if the scaled limits of $\boldsymbol{s}_n$ tend to a positive constant, rather than 0, then we expect the limit to behave like the weak noise/selection regime, except with ρ^* replaced by a smaller value.

Everything in this subsection, and in particular Theorem 1.34, has concerned the SLFVS on the whole of $\mathbb{R}^{\mathrm{d}}$. In the context of the focus of this chapter, this result is enough to determine conditions under which genetic drift will break down the effect of the curvature flow and prevent blocking. A technical point that we have to address is that the SLFVS has only previously been studied on all of $\mathbb{R}^{\mathrm{d}}$ or on a torus. In Appendix A, we present an approach to defining the SLFVS on the domain Ω of Figure 1.1 with a natural analogue of the reflecting boundary condition in (1.16). We call this process the SLFVS on Ω.

Theorem 1.36. *Suppose $r_0 < (\mathrm{d}-1)/\nu$. Let $(w_t^n)_{t\geq 0}$ be the scaled SLFVS on Ω with initial condition $w_0^n(x) = 1_{x_1 \geq 0}$.*

1. *Under the weak noise/selection ratio regime, for any $k \in \mathbb{N}$, there exist $n_*(k) < \infty$ and $a_*(k), d_*(k) \in (0, \infty)$ such that, for all $n \geq n_*$ and all $t > 0$. we have that*

$$\textit{for almost every } x \textit{ such that } x_1 \leq -d_*\varepsilon_n|\log \varepsilon_n|, \qquad \mathbb{E}\left[w_t^n(x)\right] \leq \varepsilon_n^k.$$

2. *Under the strong noise/selection ratio regime, a sharp interface does not develop as n goes to infinity. Instead, there is $\sigma^2 > 0$ such that. for every $\varepsilon > 0$ and $t \geq 0$, there exists a reflected Brownian motion, $(W_t)_{t\geq 0}$, and $n_* < \infty$, such that, for all $n \geq n_*$*

$$\left|\mathbb{E}_{w_0}\left[w_t^n(x)\right] - \mathbb{P}_x[W(\sigma^2 t) \geq 0]\right| \leq \varepsilon. \tag{1.30}$$

Theorem 1.36 provides an adaptation of Theorem 1.34 to the domain Ω of Figure 1.1. In that case, under the weak noise/selection ratio regime, we still see blocking, but in the strong noise/selection ratio regime, the proportion of a-alleles spreads approximately according to heat flow in Ω (with a reflecting boundary condition), and, in particular, blocking no longer occurs.

Our results demonstrate that, even in cylindrical domains, it can be misleading to caricature allele frequencies by one-dimensional travelling waves, and that the strength of genetic drift plays an important role in determining the fate of a favoured allele.

1.7 Hybrid zones in populations with long-range dispersal

We return to Theorem 1.20 to examine the robustness of our description of hybrid zones to changes in the underlying population model. Different factors in nature could easily break the assumptions made, and so possibly change the conclusion of the aforementioned theorem. In joint work with Kimberly Becker and Alison Etheridge, we study populations in which individuals exhibit long-range dispersal. Long-range dispersal refers to the ability of some individuals to have offspring at a huge distance from their own location. Long-range dispersal is ubiquitous in nature and is a crucial survival mechanism for many organisms, particularly those insular species that must travel vast distances to populate new regions. Examples include the dispersal of plant seeds (which can travel by wind, by water, and internally or externally by animals) [CMS00], fungi [Bus07], and small insects such as flies, moths, and bees (which have the secondary effect of facilitating the hybridisation of the flora they pollinate) [Bar96], [Rea21]. We can then ask, will hybrid zones in populations that exhibit long-range dispersal still evolve as mean curvature flow? In this thesis, we prove that, under a family of scalings, the fractional Allen–Cahn equation, with index $\alpha \in (1,2)$, started from a suitable initial condition converges to the indicator function of a set whose boundary evolves according to mean curvature flow in dimension $\mathbb{d} \geq 2$. The main difficulty will be controlling the effect of jumps with an infinite variance that arise in studying the fractional Allen–Cahn equation. Our tool for dealing with this issue will be a procedure of *marking* big jumps to quantify how substantial is their effect on the hybrid zone.

We now describe the mathematical model we will use for studying hybrid zones with long-range dispersal. Analogous reasoning to that which yields equations (1.8) suggests that a good (deterministic) starting point to understand this type of population is the equation

$$\frac{\partial w}{\partial t} = (-\Delta)^{\alpha/2} w + w(1-w)(2w-1), \tag{1.31}$$

where $(-\Delta)^{\alpha/2}$ denotes the fractional Laplacian, see [Lis+20] for a comprehensive reference on the properties of this operator. From a probabilistic perspective this operator arises as the generator of an α-stable process [Sat99]. These processes are widely used in models for long-range dispersal, arising as an analogue of Brownian motion when we take scaling limit of population models with long-range jumps [EVY20].

In recent years, the fractional equation (1.31) has become a topic of interest, particularly in the PDE community. Important applications include dislocation dynamics and statistical mechanics [IS09]. Dislocation dynamics relate to the motion of crystalline defects which can sometimes be modelled by non-local diffusions, as in Peierls-Nabarro models [KCO02]. For a comprehensive list of references, we refer to [IS09].

When $\alpha = 2$, the fractional Laplacian and ordinary Laplacian coincide, so in view of Theorem 1.17, we expect mean curvature flow to be preserved for α sufficiently close to 2, but as α approaches 0, the severity and frequency of large jumps increases, so for small α it seems unlikely that motion by mean curvature flow will be seen in the limit. This intuition is supported by results in the PDE literature, with the threshold between the two behaviours occurring at $\alpha = 1$. It is known that threshold dynamics type algorithms (i.e. algorithms that track the evolution of sets in $\mathbb{R}^{\mathrm{d}}$ by repeated use of a threshold condition, see Remark 3.5 for more details) corresponding to the fractional Laplacian, when appropriately scaled, converge to motion by mean curvature for $\alpha \in (1, 2)$, and to motion by fractional curvature for $\alpha \in (0, 1)$. This was first proved using the 'level set approach' by Caffarelli and Souganidis [CS10]. It was later shown in [IS09] that equation (1.31), when correctly scaled, converges to motion by mean curvature flow when $\alpha \in (1, 2)$ (see Remark 3.8 for more details on this result). As we will see, the scaling of (1.31) taken in [IS09] is distinct from our family of scalings, so the deterministic convergence result we present here is new.

The central ideas of this work take inspiration again from [EFP17] and their proof of Theorem 1.17, our use of a long-range dispersal mechanism throughout necessitates new (often delicate) arguments. In the Brownian setting of [EFP17], Proposition 1.21 gives a dual to the Allen–Cahn equation through a ternary branching Brownian motion, endowed with a voting system. The corresponding duality for (1.31) follows by almost identical arguments, but replacing all Brownian motions therein by α-stable processes. However, the rare long jumps of the α-stable process render working with this representation directly very challenging. To overcome this, we couple a ternary branching α-stable motion to a ternary branching subordinated Brownian motion (with subordinator given by a *truncated* $\alpha/2$-stable motion) in such a way that their root votes are comparable. This coupling is crucial to our work, and in fact quantifies the degree to which large jumps of the α-stable process affect the error bounds in our final result.

Recall that $(-\Delta)^{\frac{\alpha}{2}}$ is the fractional Laplacian, defined on smooth functions u with sufficient decay by

$$(-\Delta)^{\frac{\alpha}{2}} u := C_\alpha \lim_{\delta \to 0} \int_{\mathbb{R}^{\mathbb{d}} \setminus B(x,\delta)} \frac{u(y) - u(x)}{\|y - x\|^{\mathbb{d}+\alpha}} dy \tag{1.32}$$

where $C_\alpha := \frac{2^\alpha \Gamma\left(\frac{\alpha}{2} + \frac{\mathbb{d}}{2}\right)}{\pi^{\mathbb{d}/2} \left|\Gamma\left(-\frac{\alpha}{2}\right)\right|}$. Set

$$\sigma^2 \equiv \sigma^2_{\mathbb{d}}(\alpha) := \frac{2C_\alpha}{2 - \alpha}. \tag{1.33}$$

Our aim is to show that, in dimension $\mathbb{d} \geq 2$, for suitable initial conditions, as $\varepsilon \to 0$, the solution of the scaled fractional Allen–Cahn equation

$$\partial_t v^\varepsilon = \frac{\sigma^{-2}}{I(\varepsilon)^{2-\alpha}} (-\Delta)^{\frac{\alpha}{2}} v^\varepsilon + \frac{1}{\varepsilon^2} v^\varepsilon (1 - v^\varepsilon)(2v^\varepsilon - 1), \qquad v^\varepsilon(0, x) = p(x), \tag{1.34}$$

converges to the indicator function of a set whose boundary evolves according to mean curvature flow. Here, the function $I : \mathbb{R}_+ \to \mathbb{R}_+$ can be any function satisfying the following.

Assumptions 1.37. Let $I : \mathbb{R}_+ \to \mathbb{R}_+$.

(A) $\lim_{\varepsilon \to 0} I(\varepsilon) |\log(\varepsilon)|^k = 0 \quad \forall k \in \mathbb{N}$.

(B) $\lim_{\varepsilon \to 0} \varepsilon^2 |\log(\varepsilon)| I(\varepsilon)^{-2} = 0$.

(C) $\lim_{\varepsilon \to 0} \left(I(\varepsilon)^2 |\log(\varepsilon)| \varepsilon^{\frac{2\mathbb{d}}{\alpha} - \mathbb{d} - 1} + I(\varepsilon)^{2\alpha} \varepsilon^{-2} |\log(\varepsilon)|^\alpha \right) = 0$.

The rate of convergence and width of the ‘solution interface’ will ultimately depend on this choice of I. Note that Assumption 1.37 (B) and Assumption 1.37 (C) are incompatible as soon as $\alpha \leq 1$. When $\alpha > 1$ an example that fulfills all the conditions is $I(\varepsilon) = \varepsilon|\log(\varepsilon)|$.

Remark 1.38. Equation (1.34) can be obtained from the unscaled Fractional Allen-Chan equation by scaling time and space by $t \mapsto \varepsilon^2 t$ and $x \mapsto \sigma^{-2/\alpha}\varepsilon^{2/\alpha}I(\varepsilon)^{1-2/\alpha}x$. When $\alpha = 2$, this is consistent with the diffusive scaling considered in the Brownian setting of equation (1.14).

Lastly, for $M > 0$, we define $F(\varepsilon) := F(\varepsilon, M, I, \alpha)$ by

$$F(\varepsilon) = M\left(I(\varepsilon)^2|\log(\varepsilon)|\varepsilon^{\frac{2\mathrm{d}}{\alpha}-\mathrm{d}-1} + \frac{I(\varepsilon)^{2\alpha}}{\varepsilon^2}|\log(\varepsilon)|^{\alpha} + I(\varepsilon)^{\alpha-1}\right). \tag{1.35}$$

Note that, for any $\alpha \in (1, 2)$ and for any function $I : \mathbb{R}_+ \to \mathbb{R}_+$ satisfying Assumptions 1.37, $F(\varepsilon)$ converges to 0 as ε goes to 0. As we have done with F, to ease notation, we will only state the parameters on which a new variable depends when it is first introduced. Our main theorem for this model is the following.

Theorem 1.39. *Let $\alpha \in (1, 2)$ and fix a function $I : \mathbb{R}_+ \to \mathbb{R}_+$ satisfying Assumptions 1.37. Suppose v^ε solves equation (1.34) with initial condition p satisfying Assumptions 1.16. Let $\mathscr{T}$ and $d(x,t)$ be as above, F be as in (1.35) and fix $T^* \in (0, \mathscr{T})$. Then there exist $\varepsilon_{\mathrm{d}}(\alpha, I)$, $a_{\mathrm{d}}(\alpha, I)$, $c_{\mathrm{d}}(\alpha, I)$, $M(\alpha, I) > 0$ such that, for $\varepsilon \in (0, \varepsilon_{\mathrm{d}})$ and $a_{\mathrm{d}}\varepsilon^2|\log\varepsilon| \leq t \leq T^*$,*

(1) for x with $d(x,t) \geq c_{\mathrm{d}}(\alpha)I(\varepsilon)|\log\varepsilon|$, we have $v^\varepsilon(t,x) \geq 1 - \frac{\varepsilon^2}{I(\varepsilon)^2} - F(\varepsilon, M, I, \alpha)$;

(2) for x with $d(x,t) \leq -c_{\mathrm{d}}(\alpha)I(\varepsilon)|\log\varepsilon|$, we have $v^\varepsilon(t,x) \leq \frac{\varepsilon^2}{I(\varepsilon)^2} + F(\varepsilon, M, I, \alpha)$.

Throughout this work, we often discuss the ‘solution interface’ and its corresponding width. The term solution interface refers to the spatial region outside of which the solution $v^\varepsilon(\cdot, t)$ is very close to zero or one. Explicitly, in Theorem 1.39 the solution interface is the set of $x \in \mathbb{R}^{\mathrm{d}}$ for which $|d(x,t)| \leq c_{\mathrm{d}}(\alpha)I(\varepsilon)|\log\varepsilon|$ for all $a_{\mathrm{d}}\varepsilon^2|\log\varepsilon| \leq t \leq T^*$, and we call $c_{\mathrm{d}}(\alpha)I(\varepsilon)|\log\varepsilon|$ the interface width. We will refer to the error bounds on v^ε in Theorem 1.39 as the *tolerance* of the interface.

Remark 1.40. We note two key differences between Theorem 1.39 and Theorem 1.17. Firstly, in Theorem 1.17, the width of the solution interface was $\mathcal{O}(\varepsilon|\log\varepsilon|)$, compared to the strictly larger width of $\mathcal{O}(I(\varepsilon)|\log\varepsilon|)$ in the fractional setting. Secondly, the tolerance of solutions to equation (1.9) in Theorem 1.17 is of polynomial order, ε^k, compared to the much larger tolerance in Theorem 1.39. Both of these differences are consistent with our intuition from considering the (soon to be formalised) probabilistic representation of solutions to the ordinary and fractional Allen–Cahn equations in terms of Brownian and α-stable motions, respectively. In the latter case, the rare large jumps of the α-stable motion act to 'fatten' the interface compared to the Brownian case. While the effect of these large jumps is not so great that mean curvature flow is impeded (as $\alpha > 1$), it does require much smaller ε for solutions to converge in the fractional setting. Indeed, the error term $\varepsilon^2/I(\varepsilon)^2$ will appear when we compute the effect of erasing the *big jumps* of the α-stable processes. On the other hand, we will need to use these processes without big jumps in a short-term duality that compares the multidimensional case to the one-dimensional one. The fact that our processes have a larger variance than the Brownian case makes this comparison have a new error term, which in turn results in $F(\varepsilon)$ appearing in our statements.

Examples 1.41.

1. It is easy to verify that $I(\varepsilon) = \varepsilon|\log\varepsilon|^p$ for any $p > \frac{1}{2}$ satisfies Assumptions 1.37, so for this choice of I the interface width is $\mathcal{O}\left(\varepsilon|\log\varepsilon|^{p+1}\right)$. One can also check that $F(\varepsilon) < \frac{\varepsilon^2}{I(\varepsilon)^2} = \frac{1}{|\log\varepsilon|^{2p}}$, so the tolerance of the interface given in Theorem 1.39 is $\mathcal{O}\left(|\log\varepsilon|^{-2p}\right)$.

2. It is not always the case that $\frac{\varepsilon^2}{I(\varepsilon)^2}$ dominates $F(\varepsilon, I)$. Set $I(\varepsilon) = \varepsilon^{q_\alpha}$ where $q_\alpha = \frac{3\alpha+1}{2\alpha(1+\alpha)}$. This choice of I satisfies Assumptions 1.37, and the resulting interface width and tolerance of the interface given by Theorem 1.39 are $\mathcal{O}\left(\varepsilon^{q_\alpha}|\log\varepsilon|\right)$ and $\mathcal{O}(\varepsilon^{\frac{\alpha-1}{\alpha+1}}|\log\varepsilon|^\alpha)$, respectively.

3. While the Brownian interface width $\mathcal{O}(\varepsilon|\log\varepsilon|)$ is not achievable in Theorem 1.39, we may approximate it by choosing $I(\varepsilon) = \varepsilon^\delta$ where $\alpha^{-1} < \delta < 1$, and consider $\delta \to 1$. The interface width and tolerance in this case are $\mathcal{O}(\varepsilon^\delta|\log\varepsilon|)$ and $\varepsilon^{2-2\delta}$, respectively.

Remark 1.42. At this point we can describe the origin of our scaling. For any $\mathbb{d}$-dimensional α-stable process, $(Y_t)_{t\geq 0}$, run at speed $I(\varepsilon)^{\alpha-2}$, we can write

$$Y_t \stackrel{D}{=} Y_t^\varepsilon + R_t, \tag{1.36}$$

(where $\stackrel{D}{=}$ denotes equality in distribution), for $(Y_t^\varepsilon)_{t\geq 0}$ an $I(\varepsilon)$-truncated α-stable process, that is, the Lévy process with Lévy measure $\nu^\alpha(dx)\mathbb{1}_{\{x\leq I(\varepsilon)\}}$, for ν^α the Lévy measure of the symmetric α-stable process run at speed $I(\varepsilon)^{\alpha-2}$, and some process $R_t \stackrel{D}{=} Y_t - Y_t^\varepsilon$. The truncated process Y_t^ε can then be approximated as a Brownian motion W_t, so that

$$Y_t^\varepsilon \stackrel{D}{=} W_t + \hat{R}_t \tag{1.37}$$

for some process $\hat{R}_t \stackrel{D}{=} Y_t^\varepsilon - W_t$. We emphasise that, by our choice of scaling, the Brownian motion W_t has not undergone a speed change, making it comparable to the Brownian motions in Proposition 1.21. This decomposition, which is often exploited in simulations of stable processes, appears, for example, in [AR01], [CR07], [KT10], [Dia13]. Equation (1.37) suggests that, by Theorem 1.17, if one obtains sufficient control on the remainders R_t and $\hat{R}_t$, Theorem 1.39 should hold. In reality, we choose to work with a subordinated Brownian motion, rather than the truncated stable process, throughout (more on this in Remark 1.43), but our motivation for the chosen scaling remains the same.

Remark 1.43. For any $\mathbb{d}$-dimensional α-stable process, $(Y_t)_{t\geq 0}$ (run at speed $I(\varepsilon)^{\alpha-2}$) it is a standard fact that for a standard $\mathbb{d}$-dimensional Brownian motion $(W_t)_{t\geq 0}$ and an $\frac{\alpha}{2}$-stable subordinator $(R_t)_{t\geq 0}$ (run at speed $I(\varepsilon)^{\alpha-2}$),

$$Y_t \stackrel{D}{=} W(R_t),$$

(see, for instance, [Boc55]). As we shall see, the heavy-tailed nature of Y_t does not allow us to easily adapt the techniques of [EFP17]. To overcome this, we consider subordinating the Brownian motion W by a *truncated* stable subordinator. More precisely, we decompose R_t as

$$R_t \stackrel{D}{=} R_t^\varepsilon + L_t,$$

where $(R_t^\varepsilon)_{t\geq 0}$ is an $I(\varepsilon)^2$-truncated $\frac{\alpha}{2}$-stable subordinator and L_t is the remaining 'large jump process'. We then prove an analogue of Theorem 1.39 using $W(R_t^\varepsilon)$, and Theorem 1.39

then holds by bounding the error resulting from our use of $W(R_t^\varepsilon)$ instead of $W(R_t)$, which manifests as the $F(\varepsilon)$ term in the tolerance of our interface. The subordinated process $W(R_t^\varepsilon)$ enjoys many of the same properties as the truncated stable motion Y_t^ε from Remark 1.42, but allows us to adapt the techniques of [EFP17] more easily.

At this point, we have mentioned two natural ways to control the large jumps of an α-stable Lévy process, namely, by truncation at level $I(\varepsilon)$ or by truncation of the $\frac{\alpha}{2}$-stable subordinator at level $I(\varepsilon)^2$. It is not difficult to show that the errors resulting from each of these approximations are of the same order. Indeed, the arrival rate of jumps larger than $I(\varepsilon)$ in an α-stable process is

$$\frac{C_\alpha \sigma^{-2}}{I(\varepsilon)^{2-\alpha}} \int_{I(\varepsilon)}^{\infty} \frac{1}{x^{\alpha+1}} dx = \mathcal{O}\left(I(\varepsilon)^{-2}\right)$$

and the arrival rate of jumps larger than $I(\varepsilon)^2$ in the $\frac{\alpha}{2}$-stable subordinator is

$$\frac{C_\alpha \sigma^{-2}}{I(\varepsilon)^{2-\alpha}} \int_{I(\varepsilon)^2}^{\infty} \frac{1}{x^{\frac{\alpha}{2}+1}} dx = \mathcal{O}\left(I(\varepsilon)^{-2}\right).$$

To prove Theorem 1.39, we follow the two-step approach used in [EFP17] to obtain the analogous Brownian result. First, we prove a one-dimensional version of the result for the initial condition $p(x) = \mathbb{1}_{\{x \geq 0\}}$. This, together with a coupling of one-dimensional and $\mathbb{d}$-dimensional α-stable motions (described in Theorem 3.25), is used to prove Theorem 1.39. Just as in [EFP17], probabilistic arguments are used throughout. Namely, we exploit the duality between solutions of (1.34) and branching (α-stable) motions. While we follow the broad format of [EFP17], we require new techniques to control the rare large jumps that may appear in the α-stable motion. Most crucial among these is our *marked voting system* (Definition 3.9), which will allow us to work with subordinated Brownian motions with *truncated* subordinators throughout.

These results prove that long-range jumps do not destroy the structure of hybrid zones over a scale such that small jumps look like Brownian motion and big jumps are sufficiently rare. Furthermore, we believe the mathematical techniques we use have a value in their own right. Mainly, we take advantage of voting systems in branching processes, which seems to offer great flexibility to deal with a population with the complicated quality that is long-range dispersal. We hope this work is a starting point for the use of the same devices to deal with other complicated populations.

1.8 Outline of the rest of this thesis

We prove the results of Section 1.6 in Chapter 2 and the results of Section 1.7 in Chapter 3. In Chapter 4 we discuss some points of further work and direction of research related to the results we have obtained.

2

Hybrid zones and how to stop them

This Chapter includes the proofs of the results stated in Section 1.6. This includes Theorem 1.26, Theorem 1.27, Theorem 1.29, Theorem 1.30 and Theorem 1.34.

To prove Theorem 1.26, we proceed as follows. First, in Section 2.1.1 we provide a stochastic representation for the solution of equation (1.16). This is entirely analogous to that in [Goo18] for the equation on $\mathbb{R}^{d}$. The solution at the point x at time t is the expected result of a 'voting procedure' defined on the tree of paths traced out by branching reflecting Brownian motion in Ω up to time t, starting from a single individual at the point x. Some immediate properties of the solution that follow from this representation are presented in Section 2.1.2. In particular, it allows us to compare the solution from different initial conditions, which in turn allows us to bound the solution to (1.16) from above by the solution of the same equation with a bigger initial condition. This will be convenient in formalising the idea of the mean curvature flow 'fighting against' the constant flow. This is developed in Section 2.1.3. In Section 2.1.4 we present a coupling result which is the key to establishing a concrete bound on the effect of the mean curvature flow. With this, we prove the blocking result for the equation with a larger initial condition and so, a fortiori, for (1.16). We first bound the solution over a small window of time in Section 2.1.5, and then use a bootstrapping argument in Section 2.1.6 to show that this bound is uniform in time. Theorem 1.27 is proved in Section 2.1.7. In Section 2.1.8 we sketch the extension of these results to more general domains. In particular, we present key elements of the proof of Theorem 1.29 and some of the arguments of Theorem 1.30.

In Section 2.2 we turn to the stochastic version of our problem. The results in this section will be on the whole of $\mathbb{R}^d$. As usual for models of this type, the key to the analysis will be a dual process of branching and coalescing lineages which we introduce in Section 2.2.1. The duality function relating this process to the SLFVS will involve the same voting procedure that we use for the branching reflecting Brownian motion in the deterministic setting. Theorem 1.34, is proved in Section 2.2. In the weak noise/selection ratio regime, the key will be to show that asymptotically we no longer see coalescence in the dual process, which is then well-approximated by branching Brownian motion. This tells us that the solution to the stochastic equation will be close to that of (1.16). In the strong noise/selection ratio regime, branching events in the dual process are quickly annulled by coalescence and, over long time scales, the dual is close to a single Brownian motion. This allows us to approximate the stochastic evolution by heat flow, in sharp contrast to the first regime.

Boundary conditions have not previously been explicitly treated in the SLFVS. As will become clear, the details of what happens at the boundary should not affect our results and so we relegate a description of how one can construct what can be reasonably called an SLFVS with reflecting boundary conditions (for some rather special domains) to the Appendix A.

2.1 Deterministic Model

2.1.1 A stochastic representation of the solution to the Allen-Cahn equation

Our first goal is to present a stochastic representation of the solution of equation (1.16). This is an essentially trivial adaptation of the representation of the solution to the corresponding equation on $\mathbb{R}^d$ presented in [Goo18], which builds on [EFP17], which in turn is closely related to results of [DFL86], but we include it here as it will be central to what follows.

The representation is based on ternary branching reflected Brownian motion in Ω. The dynamics of this process have three ingredients:

1. *Spatial motion:* during its lifetime each particle moves according to a reflected Brownian

motion in Ω.

2. *Branching rate, V:* each individual has an exponentially distributed lifetime with parameter V.

3. *Branching mechanism:* when a particle dies it leaves behind (at the location where it died) exactly three offspring. Conditional on the time and place of their birth, the offspring evolve independently of each other, and in the same manner as the parent.

Assumptions 2.1.

1. For consistency with the PDE literature **we adopt the convention that all Brownian motions run at speed** 2 (and so have infinitesimal generator Δ).

2. We shall write $\gamma_\varepsilon := \varepsilon\nu$ and set the branching rate $V = (1+\gamma_\varepsilon)/\varepsilon^2$.

The stochastic representation of (1.16) is reminiscent of the classical representation of the solution to the Fisher-KPP equation in terms of binary branching Brownian motion ([Sko64; McK67]), but now it will depend not just on the leaves of the tree swept out by the branching Brownian motion, but on the whole tree. We need some notation. We write $\boldsymbol{W}(t)$ for the *historical process* of the ternary branching reflected Brownian motion; that is, $\boldsymbol{W}(t)$ records the spatial position of all individuals alive at time s for all $s \in [0,t]$.

The set of Ulam-Harris labels $\mathcal{U} = \{\emptyset\} \cup \bigcup_{k=1}^{\infty}\{1,2,3\}^k$ will be used to label the vertices of the infinite ternary tree, which consists of a single vertex of degree 1, which we call the root, and every other vertex having degree 4. The unique path from a given vertex to the root distinguishes exactly one of its four neighbours, and we consider that neighbour to be its parent, and the remaining three neighbours to be its offspring. The root is given label $\emptyset$, and if a vertex has label $(i_1,...,i_n) \in \mathcal{U}$, its offspring have labels $(i_1,...,i_n,1)$, $(i_1,...,i_n,2)$ and $(i_1,...,i_n,3)$. For example, $(3,1)$ is the first child of the third child of the initial ancestor $\emptyset$.

Definition 2.2. 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 corresponding label of the parent vertex of v.

If we ignore the spatial positions of the individuals in our ternary branching reflected Brownian motion, then each realisation of the process traces out a time-labelled ternary tree which records the relatedness (*the genealogy*) between individuals, and associates a time to each branching event. We use $N(t) \subset \mathcal{U}$ to denote the set of individuals at time t. We abuse notation to write $(W_i(t))_{i \in N(t)}$ for the spatial location of the individuals at time t and $(W_i(s))_{0 \leq s \leq t}$ for the unique path that connects the leaf i to the root. We write $\mathcal{T}(\boldsymbol{W}(t))$ for the time-labelled ternary tree determined by the branching structure of $\boldsymbol{W}(t)$.

Now, for a fixed function $p : \mathbb{R}^{\mathrm{d}} \to [0,1]$ and parameter $\gamma_\varepsilon \in (0,1)$, we define a *voting procedure* on the tree $\mathcal{T}(\boldsymbol{W}(t))$:

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

2. At each branching point in $\mathcal{T}(\boldsymbol{W}(t))$ the vote of the parent particle j is the *majority vote* of the children $(j,1)$,$(j,2)$ and $(j,3)$, unless precisely one of the offspring votes is 1, in which case the parent votes 1 with probability $\frac{2\gamma_\varepsilon}{3+3\gamma_\varepsilon}$, otherwise the parent votes 0.

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

Definition 2.3. With the voting procedure above, we define $\mathbb{V}_p^\gamma(\boldsymbol{W}(t))$ to be the vote associated to the root $\emptyset$.

Recalling Assumptions 2.1, we use $\mathbb{P}_x^\varepsilon$ to denote the law of $\boldsymbol{W}$ when started from a single individual at the point $x \in \Omega$. We have the following representation.

Proposition 2.4. *Let $u_0^\varepsilon : \Omega \to [0,1]$. Then $u^\varepsilon(x,t) = \mathbb{P}_x^\varepsilon[\mathbb{V}_{u_0}^\gamma(\boldsymbol{W}(t)) = 1]$ is a solution to the problem (1.16) with initial condition u_0^ε.*

Proof. The proof follows a standard pattern (c.f. [Sko64; McK75]), which we now sketch. To ease notation, we write

$$\mathbb{P}_x[\mathbb{V}(t) = 1] := \mathbb{P}_x^\varepsilon[\mathbb{V}_{u_0^\varepsilon}^\gamma(\boldsymbol{W}(t)) = 1].$$

First we check that u^ε satisfies (1.16) in the interior of Ω. For this, define the function g by

$$\begin{aligned} g(p_1, p_2, p_3) &= p_1p_2p_3 + p_1p_2(1-p_3) + p_1(1-p_2)p_3 + (1-p_2)p_2p_3 \\ &\quad + \frac{2\gamma_\varepsilon}{3+3\gamma_\varepsilon}\Big[(1-p_1)(1-p_2)p_3 + (1-p_1)p_2(1-p_3) + p_1(1-p_2)(1-p_3)\Big]. \end{aligned}$$

This is the vote of a branching point given that the three offspring vote 1 with probabilities p_1, p_2, p_3 respectively. We abuse notation to write $g(p) := g(p,p,p)$. Note that we have the identity:

$$g(p) - p = \frac{1}{1+\gamma_\varepsilon} p(1-p)(2p - (1-\gamma_\varepsilon)), \tag{2.1}$$

in which we recognise the non-linearity in (1.16) divided by $(1+\gamma_\varepsilon)$ (c.f. the discussion below Definition 1.32).

Fix $t > 0$ and $x \in \Omega$. Denoting by S the first branching time of $\boldsymbol{W}$ and partitioning on the events $\{S \le h\}$, $\{S > h\}$, for small h we obtain

$$\begin{aligned} \mathbb{P}_x[\mathbb{V}(t+h) = 1] &= \mathbb{P}_x\big[\mathbb{V}(t+h) = 1 | S \le h\big]\mathbb{P}[S \le h] + \mathbb{P}_x\big[\mathbb{V}(t+h) = 1 | S > h\big]\mathbb{P}[S > h] \\ &= \mathbb{E}_x\big[g(\mathbb{P}_{W_S}[\mathbb{V}(t) = 1]) | S \le h\big](1 - e^{-Vh}) + \mathbb{E}_x\big[\mathbb{P}_{W_h}[\mathbb{V}(t) = 1]\big]e^{-Vh}, \end{aligned}$$

where we have used the notation $V := \varepsilon^{-2}(1+\gamma_\varepsilon)$ from Assumptions 2.1. Now, given the regularity of the heat semigroup, and the continuity of g we have that, if h is small,

$$\mathbb{E}_x[g(\mathbb{P}_{W_S}[\mathbb{V}(t) = 1]) | S \le h] = g(\mathbb{P}_x[\mathbb{V}(t) = 1]) + \mathcal{O}(h).$$

Using this we can compute:

$$\begin{aligned} \partial_t \mathbb{P}_x[\mathbb{V}(t) = 1] &= \lim_{h\to 0} \frac{\mathbb{E}_x[\mathbb{P}_{W_h}[\mathbb{V}(t) = 1] - \mathbb{P}_x[\mathbb{V}(t) = 1]]}{h} e^{-Vh} \\ &\quad + \lim_{h\to 0} \frac{1 - e^{-Vh}}{h} \Bigg\{ g(\mathbb{P}_x[\mathbb{V}(t) = 1]) - \mathbb{P}_x[\mathbb{V}(t) = 1] \Bigg\} \\ &= \Delta(\mathbb{P}_\cdot[\mathbb{V}(t) = 1])(x) + V\Big\{ g(\mathbb{P}_x[\mathbb{V}(t) = 1]) - \mathbb{P}_x[\mathbb{V}(t) = 1] \Big\}. \end{aligned}$$

Substituting for V and using identity (2.1) the equation follows.

The boundary condition is inherited from the reflecting Brownian motion in the Lipschitz domain Ω, see [BH91]. $\square$

2.1.2 Basic results

In this section we record some easy results from [Goo18], adapted to our setting. It will be convenient to be able to refer to these results in the proof of Theorem 1.26.

A one-dimensional travelling wave.

We will later approximate the profile of the solution to (1.16) in a neighbourhood of the region in which it takes the value $\frac{1-\nu\varepsilon}{2}$ by the one-dimensional function

$$p(x,t) = \left(\exp\left(-\frac{(x+\nu t)}{\varepsilon}\right)+1\right)^{-1}. \tag{2.2}$$

Note (c.f. equation (1.18)) that this is a travelling wave solution with speed $-\nu$ and connecting 0 at $-\infty$ to 1 at ∞, of the one-dimensional equation

$$\partial_t u = \Delta u + \frac{1}{\varepsilon^2}u(1-u)(2u-(1-\nu\varepsilon)).$$

We will need to control the 'width' of the one dimensional wavefront for small ε. This is readily obtained from the explicit form of (2.2).

Lemma 2.5 ([Goo18], special case of Theorem 2.11). *For all $k \in \mathbb{N}$, and for sufficiently small $\varepsilon = \varepsilon(k) > 0$,*

1. *for $z \geq k\varepsilon|\log(\varepsilon)| - \nu t$, we have $p(z,t) \geq 1-\varepsilon^k$;*
2. *for $z \leq -k\varepsilon|\log(\varepsilon)| - \nu t$, we have $p(z,t) \leq \varepsilon^k$.*

In our case, the proof of this result is a simple calculation. We also need control of the slope of $p(z,t)$.

Proposition 2.6 ([Goo18], special case of Proposition 2.12). *Let ε be such that*

$$\varepsilon < \min\left(\frac{1}{2\nu}, \exp\left(-\frac{36}{23}\right)\right).$$

Suppose that, for some $t \in (0,\infty), z \in \mathbb{R}$ we have

$$\left|p(z,t)-\frac{1}{2}\right| \leq \frac{5+\gamma_\varepsilon}{12}, \tag{2.3}$$

and let $w \in \mathbb{R}$ satisfy $|z-w| \leq \varepsilon$. Then

$$|p(z,t) - p(w,t)| \geq \frac{|z-w|}{48\varepsilon|\log(\varepsilon)|}.$$

This proof can be found in [Goo18], Proposition 2.12. Denoting by $\mathbf{B}(t)$ the historical process of a one-dimensional ternary branching Brownian motion with branching rate $V = (1+\gamma_\varepsilon)/\varepsilon^2$ and Brownian motions run at rate 2, the argument in the proof of Proposition 2.4 yields that, for $p(x,t)$ given by (2.2),

$$\mathbb{P}_x^\varepsilon[\mathbb{V}_p^\gamma(\mathbf{B}(t)) = 1] := \mathbb{P}_x^\varepsilon[\mathbb{V}_{p(\cdot,0)}^\gamma(\mathbf{B}(t)) = 1] = p(x,t). \tag{2.4}$$

The conclusion of Proposition 2.6 then becomes that for z with

$$\left|\mathbb{P}_z^\varepsilon\big[\mathbb{V}_p^\gamma(\mathbf{B}(t)) = 1\big] - \frac{1}{2}\right| \leq \frac{5+\gamma_\varepsilon}{12},$$

and $w \in \mathbb{R}$ satisfying $|z-w| \leq \varepsilon$,

$$\Big|\mathbb{P}_z^\varepsilon\big[\mathbb{V}_p^\gamma(\mathbf{B}(t)) = 1\big] - \mathbb{P}_w^\varepsilon\big[\mathbb{V}_p^\gamma(\mathbf{B}(t)) = 1\big]\Big| \geq \frac{|z-w|}{48\varepsilon|\log(\varepsilon)|}.$$

In what follows we will always use $\mathbf{B}(t)$ to denote the historical paths of a one dimensional ternary branching Brownian motion, whereas $\boldsymbol{W}(t)$ will denote the historical paths of the multidimensional (reflected) ternary branching Brownian motion.

Results on the voting system

From a probabilistic perspective, the effect of the potential in (1.16) is captured by the voting mechanism on our tree: the function $g(p)$ amplifies the difference between p and the unstable fixed point $(1-\gamma_\varepsilon)/2$. As ε decreases, we see more and more rounds of voting in the tree by time t, leading to a rapid transition in the solution u^ε to (1.16) from values close to 0 to values close to 1. We need to control the amplification of $|p-(1-\gamma_\varepsilon)/2|$ arising from multiple rounds of voting.

We assemble the results that we need related to our voting scheme. It is immediate from equation (2.1) that if $p < (1-\gamma_\varepsilon)/2$, then $g(p) < p$, and, iterating, $g^{(n)}(p)$ will be a decreasing sequence in n. Our first lemma (whose proof mimics that of Lemma 2.9 in [EFP17]) controls how rapidly it converges to zero as n increases.

Lemma 2.7 ([Goo18], Lemma 2.14). *For all $k \in \mathbb{N}$ there exists $A(k) < \infty$ such that the following hold:*

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

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

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

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

The proof can be found in [Goo18]. It is convenient to record the following easy bound on the function g. For $0 \leq p_1, p_2, p_3 \leq \frac{1-\gamma_\varepsilon}{2}$,

$$g(p_1, p_2, p_3) \leq \max(p_1, p_2, p_3). \tag{2.5}$$

In order to exploit Lemma 2.7, we will need to know how small ε must be for us to be able to find (with high probability) a regular n-generation ternary tree inside $\boldsymbol{W}(t)$. Let $\mathcal{T}_n^{reg} := \bigcup_{k \leq n} \{1,2,3\}^k \subseteq \mathcal{U}$ denote the n-level regular ternary tree, and for $l \in \mathbb{R}$ set $\mathcal{T}_l^{reg} := \mathcal{T}_{\lceil l \rceil}^{reg}$. For $\mathcal{T}$ a time-labelled ternary tree, we write $\mathcal{T}_l^{reg} \subseteq \mathcal{T}$ to mean that as subtrees of $\mathcal{U}$, $\mathcal{T}_l^{reg}$ is contained inside $\mathcal{T}$ (ignoring its time labels).

The proof of the following result is a simple modification of that of the corresponding result (Lemma 2.10) in [EFP17].

Lemma 2.8 ([Goo18], Lemma 2.16). *Let $k \in \mathbb{N}$ and $A(k)$ as in Lemma 2.7. Then there exist $a(k)$ and $\widehat{\varepsilon}(k)$ such that for all $\varepsilon \in (0, \widehat{\varepsilon}(k))$ and $t \geq a(k)\varepsilon^2|\log(\varepsilon)|$ we have*

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

Sketch of proof. The time-length of the path to any leaf in the regular tree of height $A(k)|\log \varepsilon|$ is the sum of this number of independent exponentially distributed random variables with parameter $(1+\gamma_\varepsilon)/\varepsilon^2$. Use a large deviation principle for the sum of these exponentials to estimate the probability that a particular leaf in the regular tree has *not* been born before time t (and therefore is not contained in $\mathcal{T}(\boldsymbol{W}(t))$), and then use a union bound to control the probability that there is a leaf of the regular tree not contained in $\mathcal{T}(\boldsymbol{W}(t))$. $\square$

Monotonicity in the initial condition

The following comparison result will simplify our analysis of the solution of (1.16).

Proposition 2.9. *Let $u_0, v_0 : \Omega \to [0,1]$ with $u_0(x) \leq v_0(x)$ for all $x \in \Omega$. Then*

$$\mathbb{P}_x^\varepsilon[\mathbb{V}_{u_0}^\gamma(\boldsymbol{W}(t)) = 1] \leq \mathbb{P}_x^\varepsilon[\mathbb{V}_{v_0}^\gamma(\boldsymbol{W}(t)) = 1] \qquad \forall x \in \Omega.$$

Proof. An analytic proof would use the maximum principle and monotonicity, but here we present a probabilistic proof based on a simple coupling argument. Since there are so many different sources of randomness in our stochastic representation of solutions, we spell the argument out. Let us write $u(x,t) := \mathbb{P}_x^\varepsilon[\mathbb{V}_{u_0}^\gamma(\mathbf{W}(t)) = 1]$ and $v(x,t) := \mathbb{P}_x^\varepsilon[\mathbb{V}_{v_0}^\gamma(\mathbf{W}(t)) = 1]$. We shall say 'the probability law defining u' (respectively v) to mean the law in which each of the leaves of $\boldsymbol{W}(t)$ votes 1 with probability u_0 (respectively v_0).

Consider a given realisation of $\boldsymbol{W}(t)$. To determine the vote of a leaf $W_i(t)$, we sample U_i uniformly on $[0,1]$. If $U_i \leq u_0(W_i(t))$ then the vote of $W_i(t)$ is 1 under the law defining u. Similarly, the vote of $W_i(t)$ under the law defining v is one if and only if $U_i \leq v_0(W_i(t))$. Notice that since, by assumption, $u_0(x) \leq v_0(x)$ for all $x \in \Omega$, coupling by using the same uniform random variables U_i for determining the leaf votes for u and v, all votes that are 1 under the law defining u are also 1 under the law defining v.

Now consider the votes on the interior of the tree. Suppose that the votes of the offspring of a branch are (i_1, i_2, i_3) under u and (j_1, j_2, j_3) under v, and $i_1+i_2+i_3 \leq j_1+j_2+j_3$. Recall that under our voting procedure, the vote is determined by majority voting unless exactly one vote is one. Evidently, if the majority vote under u is 1, then it is also 1 under v; and if $j_1 + j_2 + j_3 = 0$, then the vote will be zero under both u and v. The only case of interest is when $i_1 + i_2 + i_3 = 1 = j_1 + j_2 + j_3$. The vote is then determined by a Bernoulli random variable, resulting in a 1 with probability $2\gamma_\varepsilon/(3 + 3\gamma_\varepsilon)$. For the final stage of the coupling, we use the same Bernoulli random variable for determining the vote under u and under v.

This coupling guarantees that if the vote under u is one, then necessarily the vote under v must also be one. Proceeding inductively down the tree $\boldsymbol{W}$, which is a.s. finite, we find that, if under u the vote of the root is 1, then under v it must also be 1, and the result follows. □

2.1.3 The upper solution

Proposition 2.9 allows us to bound the solution u to (1.16) on the domain Ω of Figure 1.1 by that obtained by starting from a larger initial condition. It is convenient to start from an initial condition that is radially symmetric on the left side of our domain (i.e. when $x_1 < 0$). We set

$$N_{\mathbb{r}} = \{x \in \Omega \,:\, \|x\| = \mathbb{r},\, x_1 < 0\},$$

where $R_0 > \mathbb{r} > r_0$ (see Figure 2.1). We can then write Ω as the disjoint union

$$\Omega = \Omega_+ \sqcup \Omega_- \sqcup N_{\mathbb{r}},$$

where

$$\Omega_+ = \{x \in \Omega \,:\, \|x\| < \mathbb{r} \text{ or } x_1 > 0\},$$

and

$$\Omega_- = \{x \in \Omega \,:\, \|x\| > \mathbb{r} \text{ and } x_1 < 0\}.$$

The signed distance $\widehat{d}$ of any point in Ω to $N_{\mathbb{r}}$ is defined up to a change of sign and we impose $\widehat{d}(x) > 0$ for all x in Ω_+. Here the distance is that inherited from $\mathbb{R}^{\mathrm{d}}$. Note that for all $x \in \Omega$ such that $x_1 < 0$,

$$\widehat{d}(x) = \mathbb{r} - \|x\|. \tag{2.6}$$

Assumptions 2.10. Let the function $\widehat{p} := \widehat{p}(x)$ satisfy the following conditions:

(A) $\widehat{p}(x) = 1$ for all $x = (x_1, ..., x_{\mathrm{d}}) \in \Omega$ such that $x_1 \geq 0$;

(B) $\widehat{p}(x) = \frac{1-\gamma_\varepsilon}{2}$ for all $x \in N_{\mathbb{r}}$;

(C) $\widehat{p}(x) > \frac{1-\gamma_\varepsilon}{2}$ if $\widehat{d}(x) > 0$, and $\widehat{p} < \frac{1-\gamma_\varepsilon}{2}$ if $\widehat{d}(x) < 0$;

(D) $\widehat{p}(x)$ is continuous and there exists $\mu, \lambda > 0$ such that $|\widehat{p}(x) - \frac{1-\gamma_\varepsilon}{2}| \geq \mu\big(\mathrm{dist}(x, N_{\mathbb{r}}) \wedge \eta\big)$.

The first condition means that $\widehat{p}$ dominates the initial condition for (1.16); the second and third that $\widehat{p}(1-\widehat{p})\big(2\widehat{p}-(1-\gamma_\varepsilon)\big)$ vanishes only on $N_{\mathbb{r}}$; and the final one is a lower bound on the slope of $\widehat{p}$ in a neighbourhood of $N_{\mathbb{r}}$. Since $\mathbb{r} > r_0$, one can readily construct a function $\widehat{p}$ that is radially symmetric in $\Omega \cap \{x_1 < 0\}$ and satisfies Assumptions 2.10. (The first, second and fourth condition are incompatible when $\mathbb{r} < r_0$.)

We are going to show that for $\mathbb{r} < \frac{\mathbb{d}-1}{\nu}$, and small enough ε, the solution to

$$(AC^*_\varepsilon) = \begin{cases} \partial_t u = \Delta u + \frac{1}{\varepsilon^2} u(1-u)(2u-(1-\varepsilon\nu)), \\ u_0 = \widehat{p}(x), \\ \partial_n u = 0 \end{cases}$$

is blocked. More precisely:

Theorem 2.11. *Suppose that $r_0 < \mathbb{r} < R_0 \wedge \frac{\mathbb{d}-1}{\nu}$, and that $\widehat{p}(x)$ satisfies Assumptions 2.10. Fix $k \in \mathbb{N}$. Then there exist $\widehat{\varepsilon}(k) > 0$ and $c(k) > 0$ such that for all $\varepsilon \in (0, \widehat{\varepsilon})$, $t \in (0, \infty)$,*

$$\text{for } z \in \Omega \text{ such that } d(z) \leq -c(k)\varepsilon|\log(\varepsilon)|, \text{ we have } u(z,t) \leq \varepsilon^k,$$

where $u(z,t)$ solves (AC^*_ε).

Since $\widehat{p}(x) \geq \mathbf{1}_{x_1 \geq 0}$, Theorem 1.26 follows from Theorem 2.11 and Proposition 2.9. To prove Theorem 2.11 we will control $u(x,t)$ close to $N_{\mathbb{r}}$. The key to this will be a coupling argument which will allow us to control the change in the solution over small time intervals in terms of a one-dimensional problem.

2.1.4 Coupling around a hemispherical shell

Our next goal is to quantify the effect of the mean curvature flow in 'pushing back' the solution of (1.16). We will need a notation for the 'opening' in Ω, which we will denote $\mathcal{O}$ (see Figure 2.1). Recalling that we write $x \in \mathbb{R}^{\mathbb{d}}$ as (x_1, x') with $x' \in \mathbb{R}^{\mathbb{d}-1}$,

$$\mathcal{O} = \{x \in \Omega : x_1 = 0, \|x'\| < r_0\}.$$

The following elementary result will play the role of coupling results in [EFP17] (Proposition 2.14) and [Goo18] (Proposition 2.13), but is completely straightforward since $N_{\mathbb{r}}$ takes such a simple form.

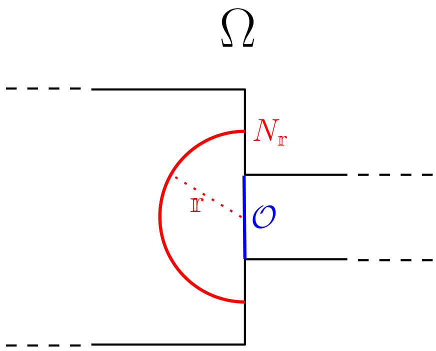

Figure 2.1: The domain Ω with $N_{\mathbb{r}}$ and $\mathcal{O}$ represented in it.

Lemma 2.12. *Suppose that $r_0 < \mathbb{r} < R_0$. Let W be a Brownian motion in Ω and $\beta > 0$ be such that*

$$\beta \leq \frac{\mathbb{r} - r_0}{2} \wedge \frac{R_0 - \mathbb{r}}{2}.$$

Define T_β by

$$T_\beta = \inf\{t \geq 0 : |\widehat{d}(W_t)| \geq \beta\}.$$

Then there exists a one-dimensional Brownian motion $\widehat{B}$, started from $\widehat{d}(W_0)$ such that, for $0 \leq s \leq T_\beta$,

$$\widehat{B}_s - \frac{s(\mathbb{d}-1)}{\mathbb{r}+\beta} \geq \widehat{d}(W_s) \geq \widehat{B}_s - \frac{s(\mathbb{d}-1)}{\mathbb{r}-\beta}.$$

Proof. Note that for $s \leq T_\beta$, the first coordinate of W_s, satisfies $(W_1)_s < 0$. Indeed, to reach the region in which $(W_1)_s \geq 0$, we must pass through $\mathcal{O}$, which requires $|\widehat{d}(W_s)| \geq \mathbb{r} - r_0 > \beta$. In particular, by (2.6), we have that $\widehat{d}(W_s) = \mathbb{r} - \|W_s\|$ for all $s \leq T_\beta$. Moreover, for $0 \leq s < T_\beta$, $\|W_s\|$ is a (time changed since our Brownian motion runs at rate 2) Bessel process, whose drift (by definition of T_β) lies in $(\frac{\mathbb{d}-1}{\mathbb{r}+\beta}, \frac{\mathbb{d}-1}{\mathbb{r}-\beta})$, and the result follows. □

2.1.5 Generation of an interface

The next step is to check that the solution to the equation with initial condition $\widehat{p}$ decays rapidly for small times. The proof mimics those that guarantee the formation of an interface in [EFP17] (Proposition 2.16) and [Goo18] (Lemma 2.17). The key is to control the maximal displacement of particles of the ternary branching reflected Brownian motion, with the only twist now being to handle the fact that particles are reflected at the boundary of the domain.

Proposition 2.13. *Let $k \in \mathbb{N}$ and $a(k)$ be given by Lemma 2.8. Then there exist $e(k)$ and $\widehat{\varepsilon}(k)$ such that, for all $\varepsilon \in (0, \widehat{\varepsilon}(k))$, all $s \leq (a(k)+k+1)\varepsilon^2|\log(\varepsilon)|$, and all $x \in \Omega$,*

$$\mathbb{P}_x^\varepsilon\Big[\exists i \in N(s) : \|W_i(s) - x\| \geq e(k)\varepsilon|\log(\varepsilon)|\Big] \leq \varepsilon^k.$$

Proof. Note that

$$\begin{aligned}
\mathbb{P}_x^\varepsilon\Big[\exists i \in N(s) :&\|W_i(s) - x\| \geq e(k)\varepsilon|\log(\varepsilon)|\Big] \\
&\leq \mathbb{E}^\varepsilon\big[N(s)\big]\mathbb{P}_x^\varepsilon\Big[\|W(s) - x\| \geq e(k)\varepsilon|\log(\varepsilon)|\Big] \\
&= e^{2\frac{s(1+\varepsilon\nu)}{\varepsilon^2}}\mathbb{P}_x^\varepsilon\Big[\|W(s) - x\| \geq e(k)\varepsilon|\log(\varepsilon)|\Big] \\
&\leq e^{2(a(k)+k+1)|\log(\varepsilon)|(1+\varepsilon\nu)}\mathbb{P}_x^\varepsilon\Big[\|W(s) - x\| \geq e(k)\varepsilon|\log(\varepsilon)|\Big] \\
&\leq e^{2(a(k)+k+1)|\log(\varepsilon)|(1+\varepsilon\nu)} \\
&\quad\times \mathbb{P}_x^\varepsilon\Big[\sup_{0\leq s\leq(a(k)+k+1)\varepsilon^2|\log(\varepsilon)|}\|W(s) - x\| \geq e(k)\varepsilon|\log(\varepsilon)|\Big].
\end{aligned} \tag{2.7}$$

To bound the last term, consider first $(\widehat{W}_t)_{t\geq 0}$, defined to be a Brownian motion reflected on the boundary of $\Omega \cap B(x,1)$, which we take to coincide with $(W_t)_{t\geq 0}$ until the first hitting time of the boundary of $\Omega \cap B(x,1)$. Since $\Omega \cap B(x,1)$ is a bounded domain with Lipschitz boundary, Theorem 2.7 of [BH91] implies the existence of constants $c_1, c_2 > 0$ (independent of x) such that,

$$\begin{aligned}
\mathbb{P}_x^\varepsilon\Big[\sup_{0\leq s\leq(a(k)+k+1)\varepsilon^2|\log(\varepsilon)|}\|\widehat{W}(s) - x\| \geq e(k)\varepsilon|\log(\varepsilon)|\Big] &\leq c_1\exp\left(\frac{-e(k)^2|\log(\varepsilon)|}{c_2(a(k)+k+1)}\right) \\
&= \varepsilon^{e(k)^2(c_2(a(k)+k+1))^{-1}}.
\end{aligned} \tag{2.8}$$

Choose $e(k)$ sufficiently large that

$$-2\big(a(k)+k+1\big)(1+\varepsilon\nu) + \frac{e(k)^2}{c_2\big(a(k)+k+1\big)} > k. \tag{2.9}$$

Finally, choose $\widehat{\varepsilon}(k)$ sufficiently small that $e(k)\varepsilon|\log(\varepsilon)| < 1$ for all $\varepsilon < \widehat{\varepsilon}$. Since $\widehat{W}(s) = W(s)$ until $\widehat{W}(s)$ hits the boundary of $B(x,1)$, this choice of $\widehat{\varepsilon}$ ensures that

$$\begin{aligned}
&\left\{\sup_{0\leq s\leq(a(k)+k+1)\varepsilon^2|\log(\varepsilon)|}\|\widehat{W}(s) - x\| \geq e(k)\varepsilon|\log(\varepsilon)|\right\} \\
&\qquad\qquad = \left\{\sup_{0\leq s\leq(a(k)+k+1)\varepsilon^2|\log(\varepsilon)|}\|W(s) - x\| \geq e(k)\varepsilon|\log(\varepsilon)|\right\}.
\end{aligned}$$

Substituting (2.8) in (2.7) and using (2.9) yields the desired bound. □

We are now in a position to prove the rapid decay of the solution to (AC_ε^*) to the left of $N_{\mathbb{r}}$.

Proposition 2.14. *Let $k \in \mathbb{N}$. There exist $\widehat{\varepsilon}(k)$, $a(k)$, $b(k) > 0$, such that, for all $\varepsilon \in (0, \widehat{\varepsilon}(k))$, setting*

$$\delta(k,\varepsilon) := a(k)\varepsilon^2|\log(\varepsilon)| \text{ and } \delta'(k,\varepsilon) := (a(k)+k+1)\varepsilon^2|\log(\varepsilon)|,$$

for $t \in [\delta(k,\varepsilon), \delta'(k,\varepsilon)]$, and x with $\widehat{d}(x) \leq -b(k)\varepsilon|\log(\varepsilon)|$, $\mathbb{P}_x^\varepsilon[\mathbb{V}_{\widehat{p}}^\gamma(\boldsymbol{W}(t)) = 1] \leq \varepsilon^k$.

Proof. By Lemma 2.8, given $k \in \mathbb{N}$ there exist $a(k)$ and $\widehat{\varepsilon}(k)$ such that, for all $\varepsilon \in (0, \widehat{\varepsilon}(k))$ and $t \geq a(k)\varepsilon^2|\log(\varepsilon)|$,

$$\mathbb{P}^\varepsilon\left[\mathcal{T}^{reg}_{A(k)|\log(\varepsilon)|} \subseteq \mathcal{T}(\boldsymbol{W}(t))\right] \geq 1 - \varepsilon^k. \tag{2.10}$$

Similarly by Proposition 2.13 there exists $e(k)$ such that, taking a smaller $\widehat{\varepsilon}(k)$ if necessary, for $\varepsilon \in (0, \widehat{\varepsilon}(k))$ and $t \in [\delta(k,\varepsilon), \delta'(k,\varepsilon)]$, we have,

$$\mathbb{P}_x^\varepsilon\Big[\exists i \in N(t) : \|W_i(t) - x\| \geq e(k)\varepsilon|\log(\varepsilon)|\Big] \leq \varepsilon^k. \tag{2.11}$$

Set $b(k) = 2e(k)$. Then if x is such that $\widehat{d}(x) \leq -b(k)\varepsilon|\log(\varepsilon)|$ and $\|W_i(t) - x\| \leq e(k)\varepsilon|\log(\varepsilon)|$ (noting that this implies that $W_1(t) \leq 0$), we obtain (using (2.6) and the triangle inequality)

$$\begin{aligned}\widehat{d}(W_i(t)) = \widehat{d}(x) + \|x\| - \|W_i\| &\leq \widehat{d}(x) + \|W_i(t) - x\| \\ &\leq -b(k)\varepsilon|\log(\varepsilon)| + e(k)\varepsilon|\log(\varepsilon)| = -e(k)\varepsilon|\log(\varepsilon)|,\end{aligned}$$

Reducing $\widehat{\varepsilon}(k)$ if necessary, we can ensure that $\varepsilon < \mu\eta$ with μ, η as in Assumption 1.16 (D), and $\varepsilon < \mu e(k)\varepsilon|\log(\varepsilon)|$ for $\varepsilon \in (0, \widehat{\varepsilon}(k))$. Then since $\widehat{p}$ satisfies Assumption 1.16 (D), we have that

$$\widehat{p}(W_i(t)) \leq \frac{1-\gamma_\varepsilon}{2} - \mu\big(|\widehat{d}(W_i(t))| \wedge \eta\big) \leq \frac{1-\gamma_\varepsilon}{2} - \varepsilon.$$

Using (2.10) and (2.11), with probability at least $1 - 2\varepsilon^k$, $\mathcal{T}(\boldsymbol{W}(t))$ contains a regular tree with at least $A(k)|\log(\varepsilon)|$ generations, and each leaf of $\mathcal{T}(\boldsymbol{W}(t))$ votes 1 with probability at most $\frac{1-\gamma_\varepsilon}{2} - \varepsilon$. Recalling from equation (2.5), that if $p_i \leq \frac{1-\gamma_\varepsilon}{2}$ for $i = 1, 2, 3$ then $g_{\gamma_\varepsilon}(p_1, p_2, p_3) \leq \max(p_1, p_2, p_3)$, we deduce that in this case, each vertex of $\mathcal{T}$ that corresponds to a leaf of the regular tree of height $A(k)|\log(\varepsilon)|$ votes 1 with probability at most $\frac{1-\gamma_\varepsilon}{2} - \varepsilon$. Lemma 2.7 then gives

$$\mathbb{P}_x^\varepsilon[\mathbb{V}_{\widehat{p}}^\gamma(\boldsymbol{W}(t)) = 1] \leq g_{\gamma_\varepsilon}^{(\lceil A(k)|\log(\varepsilon)|\rceil)}\left(\frac{1-\gamma_\varepsilon}{2} - \varepsilon\right) + 2\varepsilon^k \leq 3\varepsilon^k.$$

□

2.1.6 Blocking (proof of Theorem 2.11)

To prove Theorem 2.11, we establish a comparison between the solution to the multidimensional problem and that of the one-dimensional travelling wavefront (2.2). This is the content of Proposition 2.15 (corresponding to Proposition 2.17 in [EFP17], and Proposition 2.19 in [Goo18]), whose proof relies on a bootstrapping argument, the key ingredient for which is Lemma 2.16. The radial symmetry simplifies some of the expressions in our setting.

Proposition 2.15. *Let $l \in \mathbb{N}$ with $l \geq 4$ and let $a(l)$ and $\delta(l,\varepsilon)$ be given by Proposition 2.14. Then there exist $K_1(l)$ and $\widehat{\varepsilon}(l,K_1) > 0$ such that, for all $\varepsilon \in (0,\widehat{\varepsilon}(l,K_1))$, and $t \in [\delta(l,\varepsilon),\infty)$ we have*

$$\sup_{x\in\Omega}\left(\mathbb{P}_x^\varepsilon[\mathbb{V}_{\widehat{p}}^\gamma(\boldsymbol{W}(t)) = 1)] - \mathbb{P}^\varepsilon_{\widehat{d}(x)-\nu t+K_1\varepsilon|\log(\varepsilon)|}[\mathbb{V}_p^\gamma(\mathbf{B}(t)) = 1]\right) \leq \varepsilon^l,$$

with $\mathbb{V}_p^\gamma(\mathbf{B}(t))$ as in (2.4).

The key point here is that after the time δ that it takes for the interface to develop, we are comparing the multidimensional solution to the stationary one-dimensional interface.

We extend the domain of the function g from $[0,1]$ to $\mathbb{R}$ by setting $g(p) = 0$ for $p < 0$ and $g(p) = 1$ for $p > 1$. The following result is needed for Proposition 2.15, and will be proved afterwards.

Lemma 2.16. *Let $l \in \mathbb{N}$ with $l \geq 4$, $K_1 > 0$. There exists $\widehat{\varepsilon}(l,K_1) > 0$ such that for all $\varepsilon \in (0,\widehat{\varepsilon}(l,K_1))$, $x \in \Omega$, $s \in [0,(l+1)\varepsilon^2|\log(\varepsilon)|]$ and $t \in [s,\infty)$,*

$$\begin{aligned}\mathbb{E}_x^\epsilon&\left[g\left(\mathbb{P}^\varepsilon_{\widehat{d}(W_s)-\nu(t-s)+K_1\varepsilon|\log(\varepsilon)|}[\mathbb{V}_p^\gamma(\mathbf{B}(t-s) = 1] + \varepsilon^l\right)\right]\\ &\leq \frac{3+5\gamma_\varepsilon}{4(1+\gamma_\varepsilon)}\varepsilon^l + \mathbb{E}^\varepsilon_{\widehat{d}(x)}\left[g\left(\mathbb{P}^\varepsilon_{B_s-\nu t+K_1\varepsilon|\log(\varepsilon)|}[\mathbb{V}_p^\gamma(\mathbf{B}(t-s)) = 1]\right)\right] + 1_{s\leq\varepsilon^3}\varepsilon^l. \quad (2.12)\end{aligned}$$

Remark 2.17. In the proof of Proposition 2.15, the time s in Lemma 2.16 will be the time of the first branching event in our ternary branching process. In the symmetric case considered in [EFP17], the signed distance between a multi-dimensional Brownian motion at time s and an interface driven by mean curvature flow at time $t-s$ is coupled to a one-dimensional Brownian motion (for small times). Here, $\widehat{d}$ is the signed distance to N_{r}, which is subject to a mixture of constant flow and curvature flow and so we consider $\widehat{d}(W_s) - \nu(t-s)$. The inequality (2.12) will be used to control one round of branching in $\boldsymbol{W}$.

The proof of Proposition 2.15 first compares the multi-dimensional solution to the one-dimensional one over a small time window. It then bootstraps to longer times by conditioning on the first branch time in $\boldsymbol{W}$. The point of Lemma 2.16 is that at the branch time we have g evaluated on the one-dimensional solution plus a small error. If the solution is far from the interface $(1-\gamma_\varepsilon)/2$, then it is easy to see that the vote will not be changed by the small error. The difficult case is close to the interface, where we make use of estimates of the slope of the (one-dimensional) interface from Proposition 2.6.

Proof of Proposition 2.15. We follow the proof of Proposition 2.19 of [Goo18] which, in turn, adapts that of Proposition 2.17 of [EFP17].

Take $K_1 = b(l) + l$, where $b(l)$ is given by Proposition 2.14. Take $\widehat{\varepsilon}$ less than the corresponding $\widehat{\varepsilon}(l)$ in each of Lemma 2.5, Proposition 2.14, and Lemma 2.16.

Define δ and δ' as in Proposition 2.14. We first show that for $\varepsilon \in (0, \widehat{\varepsilon})$, $t \in [\delta, \delta']$, and $x \in \Omega$ we have

$$\mathbb{P}_x^\varepsilon[\mathbb{V}_{\widehat{p}}^\gamma(\boldsymbol{W}(t)) = 1] \leq \mathbb{P}^\varepsilon_{\widehat{d}(x)-\nu t+K_1\varepsilon|\log(\varepsilon)|}[\mathbb{V}_p^\gamma(\mathbf{B}(t)) = 1] + \varepsilon^l. \tag{2.13}$$

Indeed, if $\widehat{d}(x) \leq -b(l)\varepsilon|\log(\varepsilon)|$ then, by Proposition 2.14, we have $\mathbb{P}_x^\varepsilon[\mathbb{V}_{\widehat{p}}(\boldsymbol{W}(t)) = 1] \leq \varepsilon^l$. On the other hand if $\widehat{d}(x) \geq -b(l)\varepsilon|\log(\varepsilon)|$ then $\widehat{d}(x) - \nu t + K_1\varepsilon|\log(\varepsilon)| \geq -\nu t + l\varepsilon|\log(\varepsilon)|$ (since $K_1 = b(l) + l$), and so, by Lemma 2.5, the right hand side of (2.13) is greater than 1 and hence the inequality obviously holds.

Now consider $t \geq \delta'$. Suppose that there exists $t \in [\delta', \infty)$ for which the inequality (2.13) does not hold, that is for which there is $x \in \Omega$ such that

$$\mathbb{P}_x^\varepsilon[\mathbb{V}_{\widehat{p}}^\gamma(\boldsymbol{W}(t)) = 1] - \mathbb{P}^\varepsilon_{\widehat{d}(x)-\nu t+K_1\varepsilon|\log(\varepsilon)|}[\mathbb{V}_p^\gamma(\mathbf{B}(t)) = 1] > \varepsilon^l. \tag{2.14}$$

Let T' be the infimum of such t, and choose $T \in [T', T' + \varepsilon^{l+3}]$ for which (2.14) holds. By definition there is $x = x(l, \varepsilon) \in \Omega$ such that

$$\mathbb{P}^{\varepsilon}_x[\mathbb{V}^{\gamma}_{\widehat{p}}(\boldsymbol{W}(T)) = 1] - \mathbb{P}^{\varepsilon}_{\widehat{d}(x)-\nu T+K_1\varepsilon|\log(\varepsilon)|}[\mathbb{V}^{\gamma}_{p}(\mathbf{B}(T)) = 1] > \varepsilon^l.$$

We now seek to show that under this assumption

$$\mathbb{P}^{\varepsilon}_x[\mathbb{V}^{\gamma}_{\widehat{p}}(\boldsymbol{W}(T)) = 1] \leq \frac{7 + 9\gamma_\varepsilon}{8(1+\gamma_\varepsilon)}\varepsilon^l + \mathbb{P}^{\varepsilon}_{\widehat{d}(x)-\nu T+K_1\varepsilon|\log(\varepsilon)|}[\mathbb{V}^{\gamma}_{p}(\mathbf{B}(T)) = 1]. \tag{2.15}$$

To establish a contradiction we then choose ε sufficiently small that $\frac{7+9\gamma_\varepsilon}{8(1+\gamma_\varepsilon)}\varepsilon^l < \varepsilon^l$, so that (2.15) contradicts the assumption that (2.14) is satisfied at time T.

It remains to prove (2.15) under the assumption that there exist $t \in [\delta', \infty)$ for which (2.14) holds. We write S for the time of the first branching event in $\boldsymbol{W}(t)$, and W_S for the location of the individual that branches at that time. Using the strong Markov property,

$$\begin{aligned}\mathbb{P}^{\varepsilon}_x[\mathbb{V}^{\gamma}_{\widehat{p}}(\boldsymbol{W}(T)) = 1] = \mathbb{E}^{\varepsilon}_x\left[g\left(\mathbb{P}^{\varepsilon}_{W_S}[\mathbb{V}^{\gamma}_{\widehat{p}}(\boldsymbol{W}(T-S)) = 1]\right)\mathbf{1}_{S\leq T-\delta}\right]&\\ + \mathbb{E}^{\varepsilon}_x[\mathbb{P}^{\varepsilon}_{W_{T-\delta}}[\mathbb{V}^{\gamma}_{\widehat{p}}(\boldsymbol{W}(\delta)) = 1]\mathbf{1}_{S\geq T-\delta}].&\end{aligned} \tag{2.16}$$

As $T - \delta \geq \delta' - \delta = (l+1)\varepsilon^2|\log(\varepsilon)|$ and $S \sim \texttt{Exp}((1+\gamma_\varepsilon)\varepsilon^{-2})$ we can bound the second term in (2.16) by $\mathbb{P}^{\varepsilon}[S \geq (l+1)\varepsilon^2|\log(\varepsilon)|] \leq \varepsilon^{l+1}$. To bound the other term we partition over the event $\{S \leq \varepsilon^{l+3}\}$ (which we note has probability at most $(1+\gamma_\varepsilon)\varepsilon^{l+1}$) and its complement:

$$\begin{aligned}&\mathbb{E}^{\varepsilon}_x\left[g\left(\mathbb{P}^{\varepsilon}_{W_S}[\mathbb{V}^{\gamma}_{\widehat{p}}(\boldsymbol{W}(T-S)) = 1]\right)\mathbf{1}_{S\leq T-\delta}\right]\\ &\leq \mathbb{E}^{\varepsilon}_x\left[g\left(\mathbb{P}^{\varepsilon}_{W_S}[\mathbb{V}^{\gamma}_{\widehat{p}}(\boldsymbol{W}(T-S)) = 1]\right)\mathbf{1}_{S\leq T-\delta}\mathbf{1}_{S\geq\varepsilon^{l+3}}\right] + \mathbb{P}^{\varepsilon}[S \leq \varepsilon^{l+3}]\\ &\leq \mathbb{E}^{\varepsilon}_x\left[g\left(\mathbb{P}^{\varepsilon}_{\widehat{d}(W_S)-\nu(T-s)+K_1\varepsilon|\log(\varepsilon)|}[\mathbb{V}^{\gamma}_{p}(\mathbf{B}(T-S)) = 1] + \varepsilon^l\right)\mathbf{1}_{S\leq T-\delta}\right] + (1+\gamma_\varepsilon)\varepsilon^{l+1},\end{aligned}$$

where the last line follows from the minimality of T' (with the observation that, if $\varepsilon^{l+3} \leq S \leq T - \delta$, then $T - S \in [\delta, T')$), and the monotonicity of g.

As the path of a particle is conditionally independent of the time at which it branches, we can condition further on S to obtain

$$\begin{aligned}
&\mathbb{E}_x^\varepsilon\left[g\left(\mathbb{P}^\varepsilon_{\widehat{d}(W_S)-\nu(T-s)+K_1\varepsilon|\log(\varepsilon)|}[\mathbb{V}_p^\gamma(\mathbf{B}(T-S))=1]+\varepsilon^l\right)\mathbf{1}_{S\le T-\delta}\right]\\
&\quad\le\int_0^{(l+1)\varepsilon^2|\log(\varepsilon)|}(1+\gamma_\varepsilon)\varepsilon^{-2}e^{-(1+\gamma_\varepsilon)\varepsilon^{-2}s}\\
&\qquad\qquad\times\mathbb{E}_x\left[g\left(\mathbb{P}^\varepsilon_{\widehat{d}(W_s)-\nu(T-s)+K_1\varepsilon|\log(\varepsilon)|}[\mathbb{V}_p^\gamma(\mathbf{B}(T-s))=1]+\varepsilon^l\right)\right]ds\\
&\qquad\qquad+\mathbb{P}^\varepsilon[S\ge(l+1)\varepsilon^2|\log(\varepsilon)|]\\
&\quad\le\int_0^{(l+1)\varepsilon^2|\log(\varepsilon)|}(1+\gamma_\varepsilon)\varepsilon^{-2}e^{-\varepsilon^{-2}(1+\gamma_\varepsilon)s}\\
&\qquad\qquad\times\mathbb{E}_{\widehat{d}(x)}\left[g\left(\mathbb{P}^\varepsilon_{B_s-\nu T+K_1\varepsilon|\log(\varepsilon)|}[\mathbb{V}_p^\gamma(\mathbf{B}(T-s))=1]+\varepsilon^l\right)\right]ds\\
&\qquad\qquad\quad+\frac{3+5\gamma_\varepsilon}{4(1+\gamma_\varepsilon)}\varepsilon^l+\mathbb{P}^\varepsilon[S\le\varepsilon^3]\varepsilon^l+\varepsilon^{l+1}\\
&\quad\le\mathbb{E}^\varepsilon_{\widehat{d}(x)}\left[g\left(\mathbb{P}^\varepsilon_{B_{S'}-\nu T+K_1\varepsilon|\log(\varepsilon)|}[\mathbb{V}_p^\gamma(\mathbf{B}(T-S')=1]\right)\mathbf{1}_{S'\le T-\delta}\right]\\
&\qquad\qquad\quad+\frac{3+5\gamma_\varepsilon}{4(1+\gamma_\varepsilon)}\varepsilon^l+\varepsilon^{l+1}+\varepsilon^{l+1}(1+\gamma_\varepsilon),
\end{aligned}$$

where the second inequality follows by Lemma 2.16.

For the final inequality we have written S' for the time of the first branching event in $(\mathbf{B}(s))_{s\ge0}$ and $B_{S'}$ for the position of the ancestor particle at that time. Of course S' has the same distribution as S, and the inequality follows since $T\ge\delta'$ and so $T-\delta\ge(l+1)\varepsilon^2|\log(\varepsilon)|$.

Combining these computations, (2.16) becomes

$$\begin{aligned}
&\mathbb{P}_x^\varepsilon[\mathbb{V}_p^\gamma(\boldsymbol{W}(T))=1]\\
&\quad\le\mathbb{E}^\varepsilon_{\widehat{d}(x)}\left[g\left(\mathbb{P}^\varepsilon_{B_{S'}-\nu T+K_1\varepsilon|\log(\varepsilon)|}[\mathbb{V}_p^\gamma(\mathbf{B}(T-S')=1]\right)\mathbf{1}_{S'\le T-\delta}\right]\\
&\qquad\qquad\qquad\qquad\qquad+\frac{3+5\gamma_\varepsilon}{4(1+\gamma_\varepsilon)}\varepsilon^l+\varepsilon^{l+1}\big(2+\gamma_\varepsilon\big)\\
&\quad\le\mathbb{E}^\varepsilon_{\widehat{d}(x)-\nu T+K_1\varepsilon|\log(\varepsilon)|}[\mathbb{V}_p^\gamma(\mathbf{B}(T))=1]+\frac{3+5\gamma_\varepsilon}{4(1+\gamma_\varepsilon)}\varepsilon^l+\varepsilon^{l+1}\big(3+2\gamma_\varepsilon\big),
\end{aligned}$$

where, analogously to (2.16), in the last line we have applied the strong Markov property at time $S'\wedge(T-\delta)$. For sufficiently small ε we have that $\frac{3+5\gamma_\varepsilon}{4(1+\gamma_\varepsilon)}\varepsilon^l+2\varepsilon^{l+1}(1+\varepsilon+\gamma_\varepsilon\varepsilon)\le\frac{7+9\gamma_\varepsilon}{8(1+\gamma_\varepsilon)}\varepsilon^l<\varepsilon^l$, which completes the proof. □

Proof of Lemma 2.16. Our approach mirrors the proof of Lemma 2.20 of [Goo18], which builds on that of Lemma 2.18 of [EFP17]. The ideas are simple, but they are easily lost in the notation.

Let us write, for $u \geq 0$ and $z \in \mathbb{R}$,

$$\mathbb{Q}_z^{\varepsilon,u} = \mathbb{P}_z^{\varepsilon}[\mathbb{V}_p^{\gamma}(\mathbf{B}(u)) = 1], \tag{2.17}$$

and in what follows we use $\mathbb{Q}_z^u := \mathbb{Q}_z^{u,\varepsilon}$ when there is no confusion in doing so.

We consider three cases:

1. $\widehat{d}(x) \leq -(l + 2\mathbb{d}(l+1) + K_1)\varepsilon|\log(\varepsilon)|$.

2. $\widehat{d}(x) \geq (l + 2\mathbb{d}(l+1) + K_1)\varepsilon|\log(\varepsilon)|$.

3. $|\widehat{d}(x)| \leq (l + 2\mathbb{d}(l+1) + K_1)\varepsilon|\log(\varepsilon)|$.

Case 1: Let us define

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

We estimate the probability of A_x^c exactly as in the proof of Proposition 2.13. Indeed reading off from equation (2.8) with $l = k$, $e(k) = 2\mathbb{d}(k+1)$, $a(k) = 0$, we obtain that, for $\widehat{\varepsilon}$ small enough,

$$\mathbb{P}_x^{\varepsilon}\left[A_x^c\right] \leq c_1 \varepsilon^{\mathbb{d}(l+1)}.$$

Now suppose that A_x occurs. Then the first component of W_s is negative, and so, using equation (2.6),

$$\begin{aligned}
\widehat{d}(W_s) + K_1\varepsilon|\log(\varepsilon)| &= \widehat{d}(x) + K_1\varepsilon|\log(\varepsilon)| + \widehat{d}(W_s) - \widehat{d}(x) \\
&\leq -(l + 2\mathbb{d}(l+1))\varepsilon|\log(\varepsilon)| + |\widehat{d}(W_s) - \widehat{d}(x)| \\
&= -(l + 2\mathbb{d}(l+1))\varepsilon|\log(\varepsilon)| + \big|\|W_s\| - \|x\|\big| \leq -l\varepsilon|\log(\varepsilon)|.
\end{aligned} \tag{2.18}$$

Therefore, reducing $\widehat{\varepsilon}$ if necessary to ensure that $\gamma_{\widehat{\varepsilon}} \in (0,1)$, applying Lemma 2.5, and using the definition of g and the notation (2.17),

$$\begin{aligned}
\mathbb{E}_x\left[g\left(\mathbb{Q}^{t-s}_{\widehat{d}(W_s)-\nu(t-s)+K_1\varepsilon|\log(\varepsilon)|} + \varepsilon^l\right)\right] &\leq \mathbb{E}_x[g(\varepsilon^l + \varepsilon^l)\mathbf{1}_{A_x}] + \mathbb{P}_x[A_x^c] \\
&\leq 12\varepsilon^{2l} + \frac{4\gamma_\varepsilon}{1+\gamma_\varepsilon}\varepsilon^l + 4\mathbb{d}\varepsilon^{\mathbb{d}(l+1)}.
\end{aligned}$$

Reducing $\widehat{\varepsilon}$ still further if necessary, we can certainly arrange that for $\varepsilon \in (0, \widehat{\varepsilon})$ the last line is bounded above by $\frac{3+5\gamma_\varepsilon}{4(1+\gamma_\varepsilon)}\varepsilon^l$.

Case 2*:* First observe that, using the reflection principle and the standard bound, $\mathbb{P}[Z \geq x] \leq \exp\left(-x^2/2\right)$, on the tail of the standard normal distribution,

$$\mathbb{P}_{\widehat{d}(x)}\left[B_s \leq l\varepsilon|\log(\varepsilon)|\right] \leq 2\mathbb{P}_0\left[B_{(l+1)\varepsilon^2|\log\varepsilon|} < -2\mathbb{d}(l+1)\varepsilon|\log\varepsilon|\right] \leq 2\varepsilon^{\mathbb{d}^2(l+1)}.$$

Applying Lemma 2.5 again, we obtain

$$\begin{aligned}\mathbb{E}_{\widehat{d}(x)}\left[g\left(\mathbb{Q}^{t-s}_{B_s+K_1\varepsilon|\log(\varepsilon)|-\nu t}\right)\right] &\geq \mathbb{E}_{\widehat{d}(x)}\left[g\left(\mathbb{Q}^{t-s}_{B_s+K_1\varepsilon|\log(\varepsilon)|-\nu t}\right)\mathbf{1}_{B_s \geq l\varepsilon|\log(\varepsilon)|}\right] - 2\varepsilon^{\mathbb{d}^2(l+1)} \\ &\geq g(1-\varepsilon^l) - 2\varepsilon^{\mathbb{d}^2(l+1)} \\ &= 1 - \frac{(3+\gamma_\varepsilon)}{1+\gamma_\varepsilon}\varepsilon^{2l} + \frac{2}{1+\gamma_\varepsilon}\varepsilon^{3l} - 2\varepsilon^{\mathbb{d}^2(l+1)},\end{aligned}$$

where the final equality follows from the definition of g. Reducing $\widehat{\varepsilon}$ if necessary, we have that for $\varepsilon \in (0, \widehat{\varepsilon})$,

$$\mathbb{E}_{\widehat{d}(x)}\left[g\left(\mathbb{Q}^{t-s}_{B_s+K_1\varepsilon|\log(\varepsilon)|}\right)\right] \geq 1 - \frac{3+5\gamma_\varepsilon}{4(1+\gamma_\varepsilon)}\varepsilon^l,$$

and in this case the right hand side of (2.12) is greater than or equal to one, while the left hand side is less than or equal to one by definition.

Case 3*:* This is the most difficult case. We combine the coupling of Lemma 2.12 with our lower bound on the slope of the one-dimensional interface.

Again, suppose that A_x holds. Since we have chosen $\widehat{\varepsilon}$ small enough that $\mathbb{r} - r_0 \geq (l + K_1 + 4\mathbb{d}(l+1))\varepsilon|\log(\varepsilon)|$ for all $\varepsilon \in (0, \widehat{\varepsilon})$, on the event A_x the first component of W_s is negative, and, arguing as in (2.18), using equation (2.6) we obtain

$$|\widehat{d}(W_s)| \leq \varepsilon|\log(\varepsilon)|(4\mathbb{d}(l+1) + l + K_1).$$

Choosing $\widehat{\varepsilon}$ still smaller if necessary, for all $\varepsilon \in (0, \widehat{\varepsilon})$,

$$\left(4\mathbb{d}(l+1) + l + K_1\right)\varepsilon|\log(\varepsilon)| \leq \frac{\mathbb{r} - r_0}{2} \wedge \frac{R_0 - \mathbb{r}}{2}. \tag{2.19}$$

We write β for the quantity on the left hand side of (2.19) and use Lemma 2.12 to couple the reflected Brownian motion W started at $x \in \Omega$, with a one-dimensional Brownian motion B

started at $\widehat{d}(x)$ such that, up to time T_β, we have

$$\widehat{d}(W_s) \leq B_s - \frac{s(\mathbb{d}-1)}{\mathbb{r}+\beta}.$$

Combining this with the monotonicity of g and the fact that $\{T_\beta \geq s\} \subseteq A_x^c$ yields

$$\begin{aligned}
&\mathbb{E}_x\left[g\left(\mathbb{Q}^{t-s}_{\widehat{d}(W_s)-\nu(t-s)+K_1\varepsilon|\log(\varepsilon)|}+\varepsilon^l\right)\right] \\
&\qquad \leq \mathbb{E}_{\widehat{d}(x)}\left[g\left(\mathbb{Q}^{t-s}_{B_s-\nu(t-s)-\frac{s(\mathbb{d}-1)}{\mathbb{r}+\beta}+K_1\varepsilon|\log(\varepsilon)|}+\varepsilon^l\right)\right]+\mathbb{P}_x[T_\beta \geq s] \\
&\qquad \leq \mathbb{E}_{\widehat{d}(x)}\left[g\left(\mathbb{Q}^{t-s}_{B_s-\nu(t-s)-\frac{s(\mathbb{d}-1)}{\mathbb{r}+\beta}+K_1\varepsilon|\log(\varepsilon)|}+\varepsilon^l\right)\right]+4\mathbb{d}\varepsilon^{\mathbb{d}(l+1)}. \qquad (2.20)
\end{aligned}$$

Reducing $\widehat{\varepsilon}$ if necessary, we have that $\mathbb{r}^{-1}\beta < 1$ for all $\varepsilon \in (0,\widehat{\varepsilon})$. Thus

$$(\mathbb{d}-1)\frac{s}{\mathbb{r}}\left(\frac{1}{1+\mathbb{r}^{-1}\beta}\right) \geq \frac{s(\mathbb{d}-1)}{\mathbb{r}}(1-\mathbb{r}^{-1}\beta),$$

and so

$$\nu s - \frac{(\mathbb{d}-1)s}{\mathbb{r}+\beta} \leq s\left[\left(\nu - \frac{1}{\mathbb{r}}(\mathbb{d}-1)\right)+\frac{\mathbb{d}-1}{\mathbb{r}^2}\beta\right]. \qquad (2.21)$$

Recall that, $\mathbb{r} < (\mathbb{d}-1)/\nu$ so that $\nu - (\mathbb{d}-1)/\mathbb{r} < 0$, and so if $\widehat{\varepsilon}$ is small enough, we have that for $\varepsilon \in (0,\widehat{\varepsilon})$

$$\frac{\mathbb{d}-1}{\mathbb{r}^2}\beta + \varepsilon|\log(\varepsilon)| < \frac{\mathbb{d}-1}{\mathbb{r}} - \nu. \qquad (2.22)$$

In particular, the right hand side of (2.21) is negative.

Define

$$z = B_s - \nu t + s\left[(\nu - \frac{1}{\mathbb{r}}(\mathbb{d}-1)+\beta\frac{\mathbb{d}-1}{\mathbb{r}^2})\right] + K_1\varepsilon|(\log(\varepsilon)|.$$

Observe that, using (2.22), we have

$$\begin{aligned}
-s\left[\left(\nu-\frac{1}{\mathbb{r}}(\mathbb{d}-1)\right)+\frac{\mathbb{d}-1}{\mathbb{r}^2}\beta\right] &= s\left[\left(\frac{1}{\mathbb{r}}(\mathbb{d}-1)-\nu\right)-\frac{\mathbb{d}-1}{\mathbb{r}^2}\beta\right] \\
&\geq s\varepsilon|\log(\varepsilon)|,
\end{aligned}$$

so that

$$z \leq B_s - \nu t + (K_1 - s)\varepsilon|\log\varepsilon|. \qquad (2.23)$$

Consider the event:

$$E = \left\{\left|\mathbb{Q}^{t-s}_{B_s-\nu t+s[(\nu-\frac{1}{\mathbb{r}}(\mathbb{d}-1))+\frac{\mathbb{d}-1}{\mathbb{r}^2}\beta]+K_1\varepsilon|\log(\varepsilon)|}-\frac{1}{2}\right| \leq \frac{5+\gamma_\varepsilon}{12}\right\}.$$

As explained in [Goo18], although it looks slightly unnatural to take a set centred on the value 1/2 (about which g is symmetric only in the case when $\gamma = 0$), the importance of E is that it spans the interface (where $\mathbb{Q}$ takes the value $(1+\gamma_\varepsilon)/2$), and on E^c, $g'(\mathbb{Q}) < 1$.

Suppose first that E occurs. We apply Proposition 2.6 with z as above, and

$$w = z + s\varepsilon|\log(\varepsilon)| \leq B_s - \nu t + K_1\varepsilon|\log(\varepsilon)|.$$

Note that $|z-w| = s\varepsilon|\log(\varepsilon)| \leq (l+1)\varepsilon^3|\log(\varepsilon)|^2$, so reducing $\widehat{\varepsilon}$ if necessary so that $(l+1)\varepsilon^3|\log(\varepsilon)|^2 \leq \varepsilon$ for $\varepsilon \in (0,\widehat{\varepsilon})$, Proposition 2.6 implies

$$\mathbb{Q}^{t-s}_{B_s-\nu t+K_1\varepsilon|\log(\varepsilon)|+s[(\nu-\frac{1}{\mathbb{r}}(\mathbb{d}-1))+\frac{\mathbb{d}-1}{\mathbb{r}^2}\beta]}\mathbf{1}_E \leq \left(\mathbb{Q}^{t-s}_{B_s-\nu t+K_1\varepsilon|\log(\varepsilon)|} - \frac{s}{48}\right)\mathbf{1}_E. \tag{2.24}$$

Now suppose that E^c occurs. Since $g'(p) = \frac{2}{(1+\gamma_\varepsilon)}(1-p)(3p+\gamma_\varepsilon)$, for $p,\delta \geq 0$ with $p+\delta \leq \frac{1-\gamma_\varepsilon}{9}$ or $p \geq \frac{8+\gamma_\varepsilon}{9}$ it is easy to check that

$$g(p+\delta) \leq g(p) + \frac{2(1+2\gamma_\varepsilon)}{3(1+\gamma_\varepsilon)}\delta. \tag{2.25}$$

Let $C_{\gamma_\varepsilon} = \frac{2(1+2\gamma_\varepsilon)}{3(1+\gamma_\varepsilon)}$. Then, reducing $\widehat{\varepsilon}$ if necessary so that $\frac{1-\gamma_\varepsilon}{12} + \varepsilon^l \leq \frac{1-\gamma_\varepsilon}{9}$, for $\varepsilon \in (0,\widehat{\varepsilon})$, we obtain

$$\begin{aligned} g\left(\mathbb{Q}^{t-s}_{B_s-\nu t+K_1\varepsilon|\log(\varepsilon)|+s[(\nu-\frac{1}{\mathbb{r}}(\mathbb{d}-1))+\frac{\mathbb{d}-1}{\mathbb{r}^2}\beta]} + \varepsilon^l\right)\mathbf{1}_{E^c} \\ \leq \left(g\left(\mathbb{Q}^{t-s}_{B_s-\nu t+K_1\varepsilon|\log(\varepsilon)|}\right) + C_{\gamma_\varepsilon}\varepsilon^l\right)\mathbf{1}_{E^c}, \end{aligned} \tag{2.26}$$

where we have used (2.25), (2.23) and the monotonicity of g. Using (2.20), (2.24), and (2.26) we obtain

$$\begin{aligned} &\mathbb{E}_x\left[g\left(\mathbb{Q}^{t-s}_{\widehat{d}(W_s)-\nu(t-s)+K_1\varepsilon|\log(\varepsilon)|} + \varepsilon^l\right)\right] \\ &\leq \mathbb{E}_{\widehat{d}(x)}\left[g\left(\mathbb{Q}^{t-s}_{B_s-\nu t+K_1\varepsilon|\log(\varepsilon)|} - \frac{1}{48}s + \varepsilon^l\right)\mathbf{1}_E\right] \\ &\qquad + \mathbb{E}_{\widehat{d}(x)}\left[\left(g\left(\mathbb{Q}^{t-s}_{B_s-\nu t+K_1\varepsilon|\log(\varepsilon)|}\right) + C_{\gamma_\varepsilon}\varepsilon^l\right)\mathbf{1}_{E^c}\right] + 4\mathbb{d}\varepsilon^{\mathbb{d}(l+1)} \\ &\leq \mathbb{E}_{\widehat{d}(x)}\left[g\left(\mathbb{Q}^{t-s}_{B_s-\nu t+K_1\varepsilon|\log(\varepsilon)|}\right)\right] + C_{\gamma_\varepsilon}\varepsilon^l + \varepsilon^l 1_{s\leq 48\varepsilon^l} + 4\mathbb{d}\varepsilon^{\mathbb{d}(l+1)}, \end{aligned}$$

where we have used that $g'(p) \leq 1 + C_{\gamma_\varepsilon}$ for all $p \in [0,1]$. Finally notice that $C_{\gamma_\varepsilon} + \frac{1-\gamma_\varepsilon}{12(1+\gamma_\varepsilon)} = \frac{3+5\gamma_\varepsilon}{4(1+\gamma_\varepsilon)}$. Reducing $\widehat{\varepsilon}$ if necessary so that $4\mathbb{d}\varepsilon^{\mathbb{d}(l+1)} \leq \frac{1-\gamma_\varepsilon}{12(1+\gamma_\varepsilon)}\varepsilon^l$ and $48\varepsilon^l \leq \varepsilon^3$ for $\varepsilon \in (0,\widehat{\varepsilon})$ (which is possible since $l \geq 4$) completes the proof. $\square$

Proof of Theorem 2.11. Since we are concerned with small ε, it will be sufficient to prove the result for $k \geq 4$.

Let K_1 be given by Proposition 2.15, and let $\widehat{\varepsilon}$ be small enough that Proposition 2.15 and Lemma 2.16 hold. Set $c(k) = k + K_1$ so that, for any $\varepsilon \in (0, \widehat{\varepsilon})$ and $x \in \Omega$ such that $\widehat{d}(x) \leq -c(k)\varepsilon|\log(\varepsilon)|$, we have that $\widehat{d}(x) + K_1\varepsilon|\log(\varepsilon)| \leq -k\varepsilon|\log(\varepsilon)|$. Now choose $a(k)$ as in Proposition 2.15.

For $t \leq a(k)\varepsilon^2|\log(\varepsilon)|$ the result holds by Proposition 2.13. Indeed, since by definition $K_1 = b(k) + k = 2e(k) + k$, it holds that $K_1 \geq e(k)$. From this it follows that, if $\hat{d}(x) \leq c(k)\varepsilon|\log(\varepsilon)|$,

$$\mathbb{P}^{\varepsilon}_x\big[\mathbb{V}^{\gamma}_{\widehat{p}}(\boldsymbol{W}(t)) = 1)\big] \leq \mathbb{P}^{\varepsilon}_x\Big[\exists i \in N(s) : \|W_i(s) - x\| \geq e(k)\varepsilon|\log(\varepsilon)|\Big] \leq \varepsilon^k.$$

On the other hand, for any $t \in [a(k)\varepsilon^2|\log(\varepsilon)|, \infty)$,

$$\begin{aligned}\mathbb{P}^{\varepsilon}_x[\mathbb{V}^{\gamma}_{\widehat{p}}(\boldsymbol{W}(t)) = 1)] &\leq \varepsilon^k + \mathbb{P}^{\varepsilon}_{\widehat{d}(x)-\nu t+K_1|\log(\varepsilon)|}[\mathbb{V}^{\gamma}_{p}(\mathbf{B}(t)) = 1] \\ &\leq \varepsilon^k + \mathbb{P}^{\varepsilon}_{-k\varepsilon|\log(\varepsilon)|-\nu t}[\mathbb{V}^{\gamma}_{p}(\mathbf{B}(t)) = 1] \leq 2\varepsilon^k,\end{aligned}$$

where the last line is a consequence of Lemma 2.5. This completes the proof. □

2.1.7 Invasion (proof of Theorem 1.27)

We now turn to the proof of Theorem 1.27. Recall that we are now supposing that the narrower cylinder in the domain Ω of Figure 1.1 has radius $r_0 > (\mathbb{d} - 1)/\nu$. Our proof will mirror the 'sliding' technique used in the proof of complete propagation in [BBC16]. The key step is the following proposition, which establishes a lower bound on the solution started from $(1-\varepsilon)$ multiplied by the indicator function of a ball, whose radius is greater than $(\mathbb{d} - 1)/\nu$, sitting within $\Omega \cap \{x_1 > 0\}$.

Proposition 2.18. *Suppose that $r_0 > (\mathbb{d} - 1)/\nu$ and set*

$$r^* = \frac{r_0 + (\mathbb{d} - 1)/\nu}{2}.$$

Consider the solution $\widetilde{u}^\varepsilon$ to (1.16) with the initial condition replaced by $\widetilde{u}^\varepsilon(x, 0) = (1 - \varepsilon)\mathbf{1}_{B(x^0,r^*)}(x)$, where $x^0 = (x^0_1, \mathbf{0})$ and $\mathbf{0}$ denotes the origin in $\mathbb{R}^{\mathbb{d}-1}$. Let $\widetilde{\Gamma}$ be the solution to (1.15) started from the boundary of $B(x^0, r^*)$, and T_{r^*} be the time at which it is is *equal to the boundary of* $B(x^0,\rho^*)$ *with* ρ^* *defined by*

$$\rho^* = r^* + \frac{r_0 - r^*}{4}.$$

Then there exist constants $a \in \mathbb{R}$ *and* $\widehat{\varepsilon} > 0$ *such that for all* $\varepsilon \in (0, \widehat{\varepsilon})$,

$$\widetilde{u}^\varepsilon(x, T_{r^*}) > 1 - \varepsilon, \qquad \textit{for all } x \in B(x^0, \rho^* - a\varepsilon|\log\varepsilon|).$$

Outline of proof. Choose $\widehat{\varepsilon}$ so that $r_0 - r^* > \widehat{\varepsilon}|\log(\widehat{\varepsilon})|^2$.

By reducing $\widehat{\varepsilon}$ if necessary, assume that it is less than $\widehat{\varepsilon}(4)$ in each of Lemma 2.5, and Lemma 2.8; and small enough that the conditions of Proposition 2.6 are satisfied.

Denote the signed distance of the point $y \in \Omega$ to $\widetilde{\Gamma}_t$ by $\widetilde{d}_t(y)$ (with the convention that $\widetilde{d}_0(x^0) > 0$). Set

$$\beta = \frac{r^*}{2} \wedge \frac{r_0 - r^*}{2},$$

and

$$\widetilde{T}_\beta = \inf\{t \geq 0 : \widetilde{d}_t(W_t) \geq \beta\} \wedge T_{r^*}.$$

Exactly as in Lemma 2.12, for $0 \leq s \leq \widetilde{T}_\beta$, there exists a one-dimensional Brownian motion $\widetilde{B}$, started from $\widetilde{d}(W_0)$ such that

$$\widetilde{B}_s - \frac{s(\mathrm{d}-1)}{\rho^* + \beta} \geq \widetilde{d}(W_s) \geq \widetilde{B}_s - \frac{s(\mathrm{d}-1)}{r^* - \beta}. \tag{2.27}$$

An argument entirely analogous to the proof of Proposition 2.14 (with $k = 4$), where now we bound above the probability that a leaf votes 0 by ε (corresponding to it falling within $B(x^0, r^*)$) plus the probability that it lies outside $B(x^0, r^*)$, yields that by choosing $\widehat{\varepsilon}$ smaller still if necessary, there exist $a, b > 0$ such that for all $\varepsilon \in (0, \widehat{\varepsilon})$, setting $\delta = a\varepsilon^2|\log(\varepsilon)|$ and $\delta' = (a+5)\varepsilon^2|\log(\varepsilon)|$, for $t \in [\delta(\varepsilon), \delta'(\varepsilon)]$, and x with $\widetilde{d}(x) > b\varepsilon|\log(\varepsilon)|$,

$$\mathbb{P}_x^\varepsilon[\mathbb{V}_{\widetilde{u}^\varepsilon(\cdot,0)}(\boldsymbol{W}(t)) = 1] \geq 1 - \varepsilon^4.$$

Arguing as in the proof of Lemma 2.16, we can show that there exists $K_1 > 0$ such that, choosing $\widehat{\varepsilon}$ still smaller if necessary, for all $\varepsilon \in (0, \widehat{\varepsilon})$, $x \in \Omega$, $s \in [0, 5\varepsilon^2|\log(\varepsilon)|]$ and $t \in [s, T^*)$,

$$\begin{aligned}
&\mathbb{E}_x\left[g\left(\mathbb{P}^\varepsilon_{\widetilde{d}(W_s)-\nu(t-s)-K_1\varepsilon|\log(\varepsilon)|}[\mathbb{V}_p^\gamma(\mathbf{B}(t-s) = 0] + \varepsilon^4\right)\right] \\
&\qquad \leq \frac{3+5\gamma_\varepsilon}{4(1+\gamma_\varepsilon)}\varepsilon^4 + \mathbb{E}^\varepsilon_{\widetilde{d}(x)}\left[g\left(\mathbb{P}^\varepsilon_{B_s-\nu t-K_1\varepsilon|\log(\varepsilon)|}[\mathbb{V}_p^\gamma(\mathbf{B}(t-s)) = 0]\right)\right] + 1_{s \leq \varepsilon^3}\varepsilon^4.
\end{aligned} \tag{2.28}$$

The argument is once again simpler than the general case considered in [Goo18], since in this setting, over the time interval in which we are interested, $\widetilde{d}$ is simply the distance to the boundary of a ball.

From this we can proceed as in the proof of Proposition 2.15 to show that there exist K_1 and $\widehat{\varepsilon} > 0$ such that, for all $\varepsilon \in (0, \widehat{\varepsilon}$, $t \in [\delta(l, \varepsilon), T_{r^*}]$ we have

$$\sup_{x\in\Omega}\Big(\mathbb{P}_x^\varepsilon[\mathbb{V}^\gamma_{\widetilde{u}^\varepsilon(\cdot,0)}(\boldsymbol{W}(t))=0)]-\mathbb{P}^\varepsilon_{\widetilde{d}(x)-\nu t-K_1\varepsilon|\log(\varepsilon)|}[\mathbb{V}^\gamma_p(\mathbf{B}(t))=0]\Big)\leq\varepsilon^4,$$

with $\mathbb{V}^\gamma_p(\mathbf{B}(t))$ as in (2.4).

Finally, we can complete the proof with an argument that mirrors that of Theorem 2.11.

□

We are now in a position to prove Theorem 1.27.

Proof of Theorem 1.27. We use Proposition 2.18 to provide a series of lower solutions to (1.16). First take $x^0 = (r^*, \mathbf{0})$. By Proposition 2.9, for $\varepsilon < \widehat{\varepsilon}$ the solution to (1.16) at time T_{r^*} dominates $\widetilde{u}^\varepsilon(\cdot, T_{r^*})$. In particular, it is at least $1-\varepsilon$ on $B(x^1, r^*)$, where $x^1 = (r^* - (\rho^* - r^* - a\varepsilon|\log(\varepsilon)|), \mathbf{0})$. By choosing $\widehat{\varepsilon}$ smaller if necessary, we can certainly arrange that $\|x^0 - x^1\| \geq (r_0 - r^*)/8$. Notice also that, since $r^* > (\mathbb{d}-1)/\nu$, the time T_{r^*} is finite.

We can now simply iterate. The solution to (1.16) at time $2T_{r^*}$ dominates that started from $(1-\varepsilon)$ times the indicator function of the ball of radius r^* centred on x^1, which is at least $(1-\varepsilon)$ on the ball radius r^* centred on x^2, where $\|x^1 - x^2\| \geq (r_0 - r^*)/8$ and so on. As illustrated in Figure 2.2, any point $x \in \Omega$ can be connected to the right half-space by a finite chain of balls in this way, and the result follows. □

2.1.8 Other domains

The crux of the proof of Theorem 1.26 was the detailed analysis, close to $N_{\mathbb{r}}$, of the supersolution with initial condition $\widehat{p}$. The key to defining the supersolution was to be able to completely cover the opening $\mathcal{O}$ with a hemispherical shell of radius at most $(\mathbb{d}-1)/\nu$, that is completely contained in $\Omega \cap \{x_1 < 0\}$ (i.e. the portion of Ω to the left of the origin) and intersects the boundary $\partial\Omega$ at right angles. Evidently, we should be able to prove an entirely

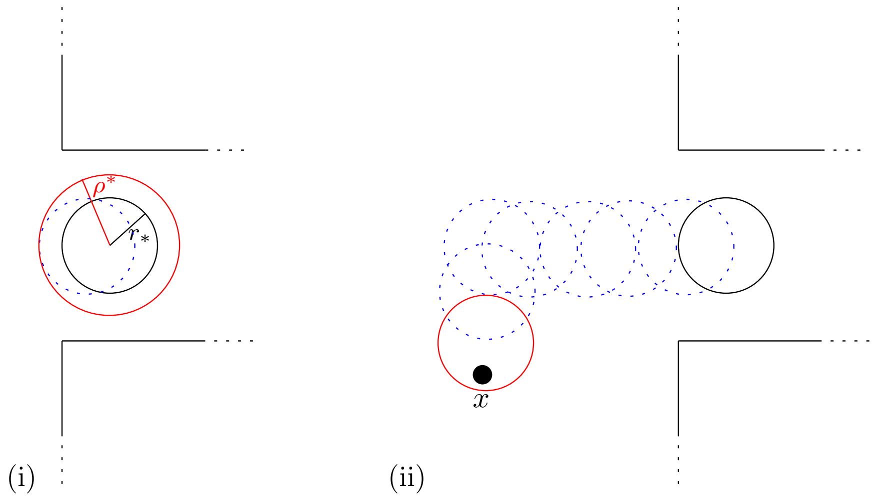

Figure 2.2: (i) Illustration of Lemma 2.18: started from $(1-\varepsilon)$ times the indicator function of the ball of radius r^*, at time T_{r^*} the solution exceeds $(1-\varepsilon)$ on the larger ball, and a fortiori on the dashed ball of radius r^* with centre shifted to the left. (ii) Illustration of how a chain of balls constructed in this way can link any point in Ω to the $\Omega \cap \{x : x_1 > 0\}$.

analogous result for any domain in which we can identify an appropriate analogue of N_{r} and control the solution around it.

The proof would pass through completely unchanged for domains of the form $\widehat{\Omega}$ of Figure 2.3, for example, provided that we could cover the disjoint union of openings by a single hemispherical shell of radius strictly less than $(\mathrm{d}-1)/\nu$. To see how we can recover an analogue of Theorem 1.26, for more general domains of the form (1.17), we first consider another special case.

Proposition 2.19. *Let*

$$\widetilde{\Omega} = \left\{(x_1, x') \subseteq \mathbb{R}^{\mathrm{d}} : x' \in \phi(x_1) \subseteq \mathbb{R}^d\right\},$$

with

$$\phi(x_1) = \begin{cases} \|x'\| < r_0 & x_1 \geq 0, \\ \|x'\| < r_0 - x_1 \tan(\alpha) & x_1 < 0, \end{cases}$$

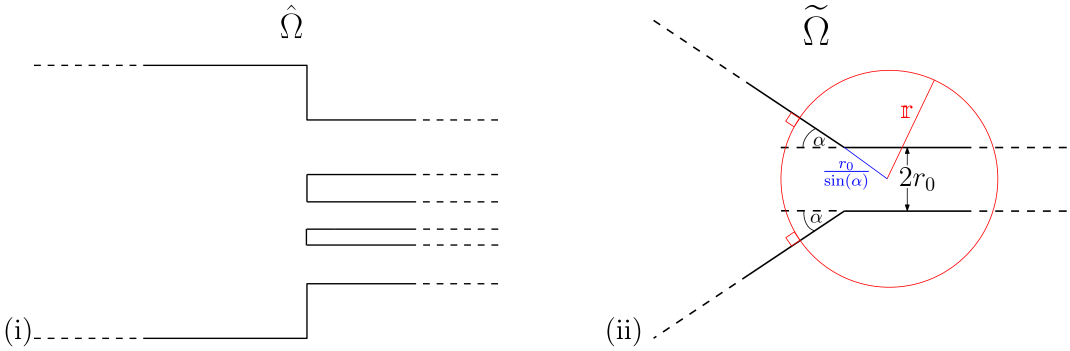

Figure 2.3: (i) Two dimensional representation of the domain $\widehat{\Omega}$. For $x_1 > 0$, we have multiple cylindrical domains that open into a single cylinder for $x_1 < 0$. (ii) Two dimensional representation of the domain $\widetilde{\Omega}$. It is composed of a cylindrical component on $x_1 > 0$ and a (truncated) cone for $x_1 < 0$.

for some $\alpha \in (0, \pi/2]$ *and suppose that*

$$r_0 < \frac{\mathbb{d}-1}{\nu} \sin\alpha. \tag{2.29}$$

Let $\mathbb{r}$ *satisfy*

$$\frac{r_0}{\sin\alpha} < \mathbb{r} < \frac{\mathbb{d}-1}{\nu}, \tag{2.30}$$

and define

$$\widetilde{N}_{\mathbb{r}} = \left\{ x = (x_1, x') \in \widetilde{\Omega} : x_1 < 0, \left\| \left(x_1 - \frac{r_0}{\tan\alpha}, x' \right) \right\| = \mathbb{r} \right\}.$$

We write $\widetilde{d}$ *for the signed distance to* $\widetilde{N}_{\mathbb{r}}$ *(chosen to be negative as* $x_1 \to -\infty$*). Let* $k \in \mathbb{N}$*. Then there exist* $\widehat{\varepsilon}(k) > 0$ *and* $a(k)$*,* $M(k) > 0$ *such that, for all* $\varepsilon \in (0, \widehat{\varepsilon})$*,* $t \in (a(k)\varepsilon^2 |\log(\varepsilon)|, \infty)$ *we have that:*

$$\textit{if} \;\; x = (x_1, \ldots, x_{\mathbb{d}}) \;\; \textit{is such that} \;\; \widetilde{d}(x) \leq -M\varepsilon|\log(\varepsilon)|, \;\; \textit{then} \;\; w(x,t) \leq \varepsilon^k.$$

Remark 2.20. The condition (2.29) becomes natural upon observing that any spherical shell intersecting the boundary of $\widetilde{\Omega}$ at right angles must have radius of size at least $r_0/\sin\alpha$,

Sketch of proof. The proof follows the same pattern as that of Theorem 1.26; first we dominate the solution by one with a larger initial condition, $\widetilde{p}$, then we put an interface with width of order $\varepsilon|\log(\varepsilon)|$ around the set on which the initial condition takes the value $(1-\varepsilon\nu)/2$ and we reproduce the proof of Lemma 2.16 to see how this interface moves.

The initial condition that we take satisfies

1. $\widetilde{p}(x) = 1$ for all $x \in \widetilde{\Omega}$ such that $x_1 \geq 0$;
2. $\widetilde{p}(x) = \frac{1-\gamma_\varepsilon}{2}$ for all $x \in \widetilde{N}_{\mathbb{r}}$.
3. $\widetilde{p}(x) > \frac{1-\gamma_\varepsilon}{2}$ if $\widetilde{d}(x) > 0$, and $\widetilde{p} < \frac{1-\gamma_\varepsilon}{2}$ if $\widetilde{d}(x) < 0$.
4. $\widetilde{p}(x)$ is continuous and there exist $\mu, \eta > 0$ such that $|\widetilde{p}(x) - \frac{1-\gamma_\varepsilon}{2}| \geq \mu(\mathrm{dist}(x, \widetilde{N}_{\mathbb{r}}) \wedge \eta)$.

The conditions of Proposition 2.19 are precisely what is required for these conditions to be compatible.

Our choice of $\widetilde{N}_{\mathbb{r}}$ enables us to prove the analogue of Theorem 2.11 for $\mathbb{P}_x^\varepsilon[\mathbb{V}_{\widetilde{p}}^\gamma(\boldsymbol{W}(t)) = 1]$ (using the same method). The obvious modification of Lemma 2.12 can then be used to obtain analogues of Lemma 2.16 and 2.15. $\square$

Armed with the proof of Proposition 2.19 the main argument behind Theorem 1.29 is easily understood. We recall that the main condition we impose is (1.20), that is,

$$\inf_{z>0} \left\{ H + h(z) - \left(\frac{\mathbb{d}-1}{\nu}\right) \frac{h'(z)}{\sqrt{1+h'(z)^2}} \right\} < 0. \tag{2.31}$$

In what follows, we recall the notation $\rho_* := (\mathbb{d}-1)/\nu$. Let

$$N_{\mathbb{r},a}^* = \left\{ x = (x_1, x') \in \Omega : x_1 < 0, \|x - (a, 0, ..., 0)\| = \mathbb{r} \right\},$$

and define $d_{\mathbb{r},a}^*$ to be the distance function to $N_{\mathbb{r},a}^*$ (chosen to be negative as $x_1 \to -\infty$).

Proposition 2.21. *Let W be Brownian motion and set*

$$T_\beta = \inf \left\{ t \geq 0 : |d_{\mathbb{r},a}^*(W_t)| \geq \beta \right\}.$$

Then, if (2.31) holds for some $z > 0$, there exist $C, \beta^* > 0$, $0 < \mathbb{r} < \rho_*$, $a < 0$ and a Brownian motion $\widehat{B}$, started from $d_{\mathbb{r},a}^*(W_0)$ such that, for all $0 < \beta \leq \beta^*$ and $0 \leq s \leq T_\beta$,

$$d_{\mathbb{r},a}^*(W_s) \leq \widehat{B}_s - \frac{s(\mathbb{d}-1)}{\mathbb{r}+\beta} + C\beta s. \tag{2.32}$$

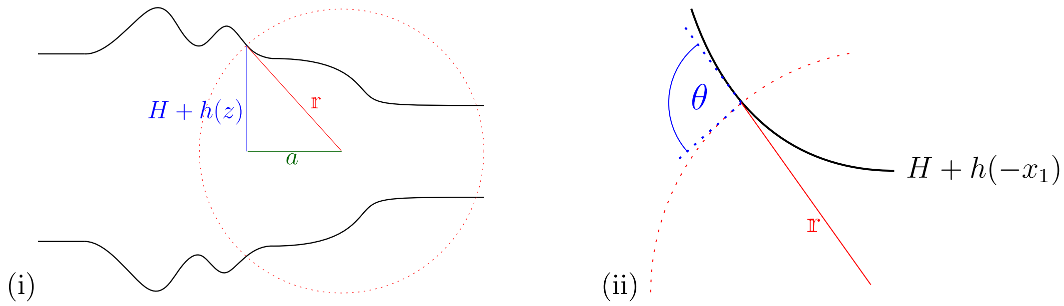

Figure 2.4: (i) Illustration of the set-up used in Proposition 2.21. The chosen value of a should be obvious from this picture. (ii) Illustration of θ, the angle at which the spherical shell intersects the boundary of Ω

Sketch of proof. Let z be such that

$$H + h(z) - \left(\frac{\mathbb{d}-1}{\nu}\right)\frac{h'(z)}{\sqrt{1+h'(z)^2}} < 0,$$

and set

$$\mathbb{r} = \frac{1}{2}\left(\frac{(H+h(z))\sqrt{1+h'(z)^2}}{h'(z)} + \rho_*\right). \tag{2.33}$$

Note that, by our choice of z, we have $\mathbb{r} < \rho_*$. Let $a = -z + \sqrt{\mathbb{r}^2 - (H+h(z))^2}$, and note that $N^*_{\mathbb{r},a}$ intersects the boundary of the domain exactly at the point where the first coordinate is $-z$. See Figure 2.4(i) for an illustration of this situation.

Choose β^* small enough that, for all $\beta < \beta^*$, the distance function $d^*_{\mathbb{r},a}$ satisfies the Assumptions 1.16 in the set $U_\beta = \{x : |d^*_{\mathbb{r},a}(x)| \leq \beta\}$. We will now reduce β further if necessary to ensure that, when $N^*_{\mathbb{r},a}$ intersects the boundary of the domain, the domain is still 'opening out' sufficiently fast. More precisely, let $x = (x_1, x') \in \partial\Omega$. Write $v_1(x)$ for the vector pointing from x to $(a, 0, ...0)$, and $v_2(x)$ for the normal vector to $\partial\Omega$ at x. We require that the angle θ between these two vectors is at least $\pi/2$. See Figure 2.4(ii) for an illustration. Some cumbersome computations that we defer to Appendix A give that

$$\langle v_1(x), v_2(x)\rangle = H + h(-x_1) + h'(-x_1)(-z - x_1 - \sqrt{\mathbb{r}^2 - (H+h(z))^2}). \tag{2.34}$$

Note that, if $x_1 = -z$, this gives,

$$H + h(z) - h'(z)\sqrt{\mathbb{r}^2 - (H + h(z))^2}$$
$$< H + h(z) - h'(z)\sqrt{\frac{(H + h(z))^2\big(1 + h'(z)^2\big)}{(h'(z))^2} - (H + h(z))^2} = 0,$$

which gives that θ is bigger than $\pi/2$. In particular, by smoothness of h, we can reduce β^*, so that for all $\beta < \beta^*$ we have that, for $x = (x_1, x') \in U_\beta \cap \partial\Omega$,

$$H + h(-x_1) + h'(-x_1)(-z - x_1 - \sqrt{\mathbb{r}^2 - (H + h(z))^2}) \leq 0. \tag{2.35}$$

Equation (2.35) encapsulates that the surface hits the boundary of the domain in such a way that θ is at least $\pi/2$.

Let $d(x) = \|x - a\| - \mathbb{r}$ be the distance of x from the circle of radius $\mathbb{r}$ and centre a in $\mathbb{R}^{\mathbb{d}}$. We can deduce the following fact: there exists a smooth function $f : \Omega \to \mathbb{R}$ such that,

$$d^*_{\mathbb{r},a}(x) - d(x) = \beta f(x) \quad \forall x \in U_\beta.$$

Indeed, as $d^*_{\mathbb{r},a}$ is a perturbation of at most β of $d(x)$, the last equation must hold. In particular, for $x \in U_\beta$, there exists a constant $C > 0$ such that,

$$\Delta d^*_{\mathbb{r},a}(x) \leq \Delta d(x) + \beta C = \frac{\mathbb{d} - 1}{\|x\|} + \beta C.$$

Then, from Itô's formula, we obtain that, for all $t \leq T_\beta$,

$$d^*_{\mathbb{r},a}(W_t) \leq d^*_{\mathbb{r},a}(W_0) + \int_0^t \nabla d^*_{\mathbb{r},a}(W_s) \cdot dW_s + \int_0^t \frac{(\mathbb{d} - 1)}{\|W_s\|} ds + \beta C t$$
$$+ \int_0^t \langle \nabla d^*_{\mathbb{r},a}(W_s), \hat{n} \rangle dL^{\partial\Omega}_s(W_s), \tag{2.36}$$

where $\hat{n}$ is the inward pointing normal.

The first two terms of the right hand side of (2.36) are a Brownian motion (by Lévy's characterisation). Since $t < T_\beta$, the third term on the right hand side of (2.36) can be bounded by the drift term on the right hand side of equation (2.32). To deal with the integral against the local time in (2.36), we just need to show that $\langle \nabla d^*_{\mathbb{r},a}(W_s), \hat{n} \rangle \leq 0$. Indeed, if the line segment from W_s to $(a, 0..., 0)$ intersects $\partial\Omega$, then $\nabla d^*_{\mathbb{r},a}(W_s)$ is a vector parallel to the surface, in which case the inequality holds trivially. On the other hand, if the said line segment does not intersect $\partial\Omega$, then $\nabla d^*_{\mathbb{r},a}(W_s)$ is a vector going from W_s to $(a, 0, ..., 0)$, in which case $\langle \nabla d^*_{\mathbb{r},a}(W_s), \hat{n} \rangle \leq 0$, since (2.35) holds from our choice of β^*. Combining these bounds gives (2.32) and completes the proof. □

From the last result the approach to prove Theorem 1.29 should be obvious. We proceed as in Proposition 2.19 by taking an initial condition similar to $\widetilde{p}$, but now it will be $(1-\gamma_\varepsilon)/2$ for $x \in N^*_{\mathbb{r},a}$. By Proposition 2.21 we can choose $\mathbb{r}$ and a such that (2.32) holds. From this the proof of the result is analogous to the one used to prove Proposition 2.19, giving Theorem 1.29.

Invasion under condition (1.21) can then be handled in a very similar way to the proof of Theorem 1.29.

Sketch of proof of Theorem 1.30. If (1.21) holds, then, for all a, we can choose $\mathbb{r}$ as in (2.33), so that (2.35) holds with the reverse inequality for points on the boundary that are close enough to $N^*_{\mathbb{r},a}$. That is, for any $x = (x_1, x') \in \partial\Omega$ that is close enough to $N^*_{\mathbb{r},a}$ we can prove,

$$H + h(-x_1) + h'(-x_1)(-z - x_1 - \sqrt{\mathbb{r}^2 - (H + h(z))^2}) \geq 0. \tag{2.37}$$

Proceeding as we did for equation (2.36), we obtain the existence of a Brownian motion $\widehat{B}_s$ and a constant $C > 0$, such that for small enough β,

$$d^*_{\mathbb{r},a}(W_s) \geq \widehat{B}_s - \frac{s(\mathbb{d}-1)}{\mathbb{r}-\beta} \geq \widehat{B}_s - s\nu - sC, \tag{2.38}$$

where the last inequality uses $\mathbb{r} > \rho_*$. Therefore, if we consider u^ε, the solution to (1.16) with initial condition $u^\varepsilon(x, 0) = (1-\varepsilon)1_{B(a,\mathbb{r})}$, we can prove an analogue of Proposition 2.18, by using equation (2.38) in place of equation (2.27). Iterating, we can then deduce that under condition (1.21) an analogue of Theorem 1.27 holds, thus giving the invasion result. □

2.2 Stochastics

We now turn to the SLFVS. As we have seen in the deterministic setting, the crucial step in determining whether or not there will be blocking is to understand the interplay between the selection against heterozygosity and the selective advantage of aa-homozygotes over AA-homozygotes in a small neighbourhood of a critical radius in which the influence of the boundary of the domain is unimportant. In this section we therefore focus on the SLFVS on the whole of Euclidean space, relegating a discussion of other domains (and, in particular, a suitable definition of the SLFVS on domains with boundary) to Appendix A.

2.2.1 The dual process

At the core of the proof of Theorem 1.15 was the duality between the deterministic equation (1.16) and ternary branching (reflected) Brownian motion endowed with a voting mechanism. In an entirely analogous way, we wish to exploit the duality between the SLFVS and a system of branching and coalescing lineages endowed with a similar voting mechanism.

The process of branching and coalescing lineages is driven by (the time-reversal of) the Poisson point process of events that determined the dynamics in the SLFVS. Recall that Π^n is a Poisson point process on $\mathbb{R}_+ \times \mathbb{R}^{\mathrm{d}} \times (0, \mathcal{R}_n)$, whose intensity is given by (1.25) in the weak noise/selection ratio regime, and by (1.27) in the strong noise/selection ratio regime. To emphasize that the dual process runs 'backwards in time', we shall write $\overleftarrow{\Pi}^n$ for the time-reversal of Π^n, which of course has the same intensity as Π^n. The impact, asymmetry and selection parameters $(u_n)_{n\in\mathbb{N}}$, $(\gamma_n)_{n\in\mathbb{N}}$ and $(\boldsymbol{s}_n)_{n\in\mathbb{N}}$ are respectively, given in (1.26) in the weak noise/selection ratio regime, or fulfil the condition (1.28) in the strong noise/selection ratio regime.

Definition 2.22. (SLFVS dual) For $n \in \mathbb{N}$, the process $(\mathcal{P}^n_t)_{t\geq 0}$ is the $\bigcup_{l\geq 1}(\mathbb{R}^{\mathrm{d}})^l$-valued Markov process with dynamics defined as follows.

The process starts from a single individual at the point $x \in \mathbb{R}^{\mathrm{d}}$. We write $\mathcal{P}^n_t = (\xi^n_1(t), \ldots, \xi^n_{N(t)}(t))$, where the random number $N(t) \in \mathbb{N}$ is the number of individuals alive at time t, and $\{\xi^n_i(t)\}_{i=1}^{N(t)}$ are the individuals' locations. For each $(t, x, r) \in \overleftarrow{\Pi}^n$, the corresponding event is neutral with probability $1 - (1 + \gamma_n)\boldsymbol{s}_n$, in which case:

1. for each $\xi^n_i(t-) \in B(x, r)$, 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, whose location is chosen uniformly in $B(x, r)$.

With the complementary probability $(1 + \gamma_n)\boldsymbol{s}_n$, the event is selective, in which case:

1. for each $\xi_i^n(t-) \in B(x, r)$, 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 a total of *three* offspring, whose locations are drawn independently and uniformly in $B(x, r)$.

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

Remark 2.23. From the perspective of the SLFVS, it would be more natural to call the individuals created during a reproduction event in the dual process 'parents' (or 'potential parents'), as they are situated at the locations from which the parental alleles are sampled. We choose to call them offspring in order to emphasize that the dual process plays the same role as ternary branching Brownian motion in the deterministic setting, and, indeed, much of the proof of Theorem 2.11 carries over with minimal changes to the SLFVS setting.

Just as for the deterministic setting, the duality relation that we exploit is between the SLFVS and the *historical process* of branching and coalescing lineages,

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

We write $\mathbb{P}_x$ for the law of Ξ^n when $\mathcal{P}_0^n$ is the single point x, and $\mathbb{E}_x$ for the corresponding expectation.

Just as for the branching process, we can use Ulam-Harris labels to define lines of descent from the root individual at x through $\mathcal{P}^n$. More precisely, for $\boldsymbol{i} = (i_1, i_2, \ldots) \in \{1, 2, 3\}^{\mathbb{N}}$, we write $(\xi_{\boldsymbol{i}}^n(\cdot))_{0 \leq s \leq t} \subseteq \Xi^n(t)$ for the $\mathbb{R}^{\mathrm{d}}$-valued path which jumps to the location of the (unique) offspring when the individual in $\mathcal{P}_s^n$ at its location is affected by a neutral event, and to the location of the i_kth offspring, the kth time that it is affected by a selective event. From the perspective of the SLFVS, $\xi_{\boldsymbol{i}}^n$ is an ancestral *lineage*, and we shall use the terminology lineage below.

Let $p : \mathbb{R}^{\mathrm{d}} \to [0,1]$. The voting procedure on $\Xi^n(t)$ is a natural modification of the one that we defined for the ternary branching Brownian motion:

1. Each leaf of $\Xi^n(t)$ independently votes 1 with probability $p(\xi_i(t))$, and 0 otherwise;
2. at each neutral event in $\overleftarrow{\Pi}^n$, all marked individuals adopt the vote of the offspring;
3. at each selective event in $\overleftarrow{\Pi}^n$, all marked individuals adopt the majority vote of the three offspring, unless precisely one vote is 1, in which case they all vote 1 with probability $\frac{2\gamma_n}{3+3\gamma_n}$, otherwise they vote 0.

This defines an iterative voting procedure, which runs inwards from the 'leaves' of $\Xi^n(t)$ to the ancestral individual $\emptyset$ situated at the point x.

Definition 2.24. With the voting procedure described above, we define $\mathbb{V}_p(\Xi^n(t))$ to be the vote associated to the root $\emptyset$.

We should like to have an analogue of the stochastic representation of the solution to (1.16) of Proposition 2.4 for the SLFVS, but recall from Remark 1.11 that the SLFVS is only defined up to a Lebesgue null set, and so we cannot expect such a representation at every point of $\mathbb{R}^{\mathrm{d}}$. However, a weak version of the representation is valid. The following result is easily proved using the approach introduced to prove Proposition 1.7 in [EVY20] (the corresponding result in the case of genic selection).

Theorem 2.25. *The SLFVS driven by Π^n, $(w_t^n(x), x \in \mathbb{R}^{\mathrm{d}})_{t\geq 0}$, is dual to the process $(\Xi^n(t))_{t\geq 0}$ of Definition 2.22 in the sense that for every $\psi \in C(\mathbb{R}^{\mathrm{d}}) \cap L^1(\mathbb{R}^{\mathrm{d}})$, we have*

$$\mathbb{E}_p\left[\int_\Omega \psi(x) w_t^n(x)\,dx\right] = \int_\Omega \psi(x)\mathbb{E}_x\left[\mathbb{V}_p\left(\Xi^n(t)\right)\right]\,dx = \int_\Omega \psi(x)\mathbb{P}_x\left[\mathbb{V}_p\big(\Xi^n(t)\big) = 1\right]\,dx. \quad (2.39)$$

Remark 2.26. Note that the expectations on the left and right of equation (2.39) are taken with respect to different measures. The subscripts on the expectations are the initial values for the processes on each side.

The duality reduces the proof of Theorem 1.34 to the following analogue of Theorem 2.11.

Theorem 2.27. *Define $\rho_* = (\mathrm{d}-1)/\nu$ and let $\widehat{p}(x) = 1_{B(0,\rho_*)}(x)$.*

1. *Under the weak noise/selection ratio regime, for any $k \in \mathbb{N}$, there exist $n_*(k) < \infty$ and $d_*(k) \in (0, \infty)$, such that, for all $n \geq n_*$ and all $t > 0$,*

$$\text{for all } x \in \mathbb{R}^{\mathrm{d}} \text{ with } \|x\| \geq \rho_* + d_* \varepsilon_n |\log \varepsilon_n|, \text{ we have } \mathbb{P}_x \left[\mathbb{V}_{\widehat{p}}(\Xi^n(t)) = 1 \right] \leq \varepsilon_n^k.$$
$$\text{for all } x \in \mathbb{R}^{\mathrm{d}} \text{ with } \|x\| \leq \rho_* - d_* \varepsilon_n |\log \varepsilon_n|, \text{ we have } \mathbb{P}_x \left[\mathbb{V}_{\widehat{p}}(\Xi^n(t)) = 1 \right] \geq 1 - \varepsilon_n^k.$$

2. *Under the strong noise/selection ratio regime, there is a constant $\sigma^2 > 0$ such that, for every fixed $t > 0$ and $\varepsilon > 0$, there is a Brownian motion, $(W_s)_{s \geq 0}$, and $n_* \in \mathbb{N}$, such that, for all $n \geq n_*$*

$$\left| \mathbb{P}_x \left[\mathbb{V}_{\widehat{p}}(\Xi^n(t)) = 1 \right] - \mathbb{P}_x \left[\|W(\sigma^2 t)\| \leq \rho_* \right] \right| \leq \varepsilon. \tag{2.40}$$

Motion of a single lineage

Our first task is to investigate the motion of a single lineage. It is a pure jump process. By spatial homogeneity, to establish the transition rates, it suffices to calculate the rate at which a lineage currently at the origin will jump to the point $z \in \mathbb{R}^{\mathrm{d}}$. This is given by

$$m_n(dz) = n u_n \chi_n n^{\mathrm{d}\beta} \int_0^{\mathcal{R}_n} \frac{V_r(0,z)}{V_r} d\mu^n(r) dz, \tag{2.41}$$

where $\chi_n = 1/\widehat{u}_n$ in the strong noise regime and 1 otherwise, V_r is the volume of $B(0,r)$, and $V_r(0,z)$ is the volume of $B(0,r) \cap B(z,r)$. To understand the expression (2.41), first observe that, in order for the lineage to jump from 0 to z, it must be affected by an event that covers both 0 and z. The region of possible centres for an event of radius r covering both 0 and z is the region given by the intersection of $B(0,r)$ and $B(z,r)$, which has volume $V_r(0,z)$, and so events fall in this region with rate $n\chi_n V_r(0,z)$. If an event occurs with centre x in the region, then the lineage jumps with probability u_n, and if it does so, then it jumps to a point chosen uniformly from $B(x,r)$; the chance of that point being z is dz/V_r.

Remark 2.28. We can now explain why our choices of scalings in the weak and strong regimes are appropriate for investigating the effect of increasing genetic drift. Notice that $n u_n \chi_n = u n^{2\beta}$ in the weak scaling and $n^{2\beta}$ in the strong scaling, so setting $u = 1$ in (1.26), the transition probabilities for a lineage will coincide under our two regimes. If the coefficients

$\boldsymbol{s}_n$ are the same, then so too will be the rate at which a lineage is affected by a selective event. From the perspective of lineages, the difference between the scalings is the probability of coalescence, driven by u_n. The parameter u_n determines the strength of the genetic drift (it can be thought of as proportional to the inverse of the population density). In particular, in two dimensions, Theorem 1.34 says that increasing the strength of the noise can indeed break down the structure of the solution that results from the selection term in the Allen-Cahn equation.

Integrating out over z in (2.41) and 'undoing' the scaling of $\mu^n(dr)$, we find that the total jump rate of the lineage is

$$\int_{\mathbb{R}^{\mathrm{d}}} m_n(dz) = nu_n\chi_n V_1 \int_0^{\mathcal{R}} r^d \mu(dr),$$

where, as noted above, the prefactor is $n^{2\beta}$ (up to a factor of u coming from (1.26)). Since each jump is of size $\Theta(n^{-\beta})$, we recognise the diffusive scaling. We can identify the diffusion constant for the limiting Brownian motion, by noting that

$$\sigma^2 = \int_{\mathbb{R}^{\mathrm{d}}} \|z\|^2 m_n(dz) = \frac{u_n n\chi_n}{n^{2\beta}2\mathrm{d}} \int_0^{\mathcal{R}} \int_{\mathbb{R}^{\mathrm{d}}} \|z\|^2 \frac{V_r(0,z)}{V_r(0)} dz\mu(dr). \tag{2.42}$$

The following coupling is Lemma 3.8 in [EFP17].

Lemma 2.29. *Let $(\xi^n(t))$ be a pure jump process with transition rates given by m_n and suppose that σ^2 is defined by (2.42). For fixed $t > 0$ there is a coupling of a Brownian motion W and ξ^n under which*

$$\mathbb{P}_x[|\xi^n(t) - W(\sigma^2 t)| \geq n^{-\beta/6}] = \mathcal{O}(n^{-\beta}(t \vee 1))$$

(where ξ^n and W both start from the point x).

We need to be able to couple the lineage at the time when it is first affected by a selective event (and so branches) with a Brownian motion. This is where we first use Assumption 1.19, $(\log n)^{1/2}\varepsilon_n \to \infty$, which guarantees that $\boldsymbol{s}_n = o\big(\log n/n^{2\beta}\big)$. We abuse our Ulam-Harris based notation $\xi_{\boldsymbol{i}}$, and write simply $\xi_1(\tau)$, $\xi_2(\tau)$, and $\xi_3(\tau)$ for the three possible positions of lineages at the first branching time of Ξ^n (corresponding to the positions of the three offspring of the first selective event to affect the lineage).

Corollary 2.30. *Let τ be the first branching of Ξ^n. Then there is a Brownian motion W and a coupling of Ξ^n and W under which the following holds. The branching time τ and W are independent, with $\tau \sim \mathtt{Exp}(\eta(1+\gamma_n)\varepsilon_n^{-2})$, where $\eta = uV_1 \int_0^{\mathcal{R}} r^{\mathrm{d}} \mu(dr)$. Furthermore, for $i = 1, 2, 3$,*

$$\mathbb{P}_x \left[|\xi_i^n(\tau) - W(\sigma^2 \tau)| \geq n^{-\beta/6} \right] = \mathcal{O}(n^{-\beta}),$$

where ξ^n and W both have the same starting point x.

Sketch of proof. By Poisson thinning, we can express the position of the lineage at time t as the sum of the jumps due to neutral events and those due to selective events. Lemma 2.29 allows us to couple the part corresponding to neutral events with a Brownian motion at time $\sigma^2(1-\boldsymbol{s}_n)t$. Since $\boldsymbol{s}_n = o\big(\log n / n^{2\beta}\big)$, Chebyshev's inequality gives

$$\mathbb{P}\left[\|W(\sigma^2 t) - W(\sigma^2(1-\boldsymbol{s}_n)t)\| \geq n^{-\beta/6} \right] = o\left(\frac{\log n}{n^{2\beta}} n^{\beta/3} (t \vee 1) \right).$$

The proof is completed by an application of the triangle inequality, using Poisson thinning to partition over the value of τ, and using that $|\xi_i^n(\tau) - \xi_1^n(\tau-)| \leq 2\mathcal{R}_n = 2n^{-\beta}\mathcal{R}$. $\square$

2.2.2 Proof of Theorem 2.27, weak noise/selection ratio regime

We outline the main steps of the proof of Theorem 2.27 in the weak scaling regime, that is one in which selection overwhelms noise.

There are two ways in which the dual to the SLFVS differs from our ternary branching Brownian motion. The first is that lineages can coalesce during reproduction events; the second is that lineages follow a continuous time and space random walk which only converges to Brownian motion in the scaling limit.

Following [EFP17], the first step in the proof of the limiting result in the weak scaling regime is to show that with high probability $\Xi^n(t)$ can be coupled to a branching jump process.

Definition 2.31 (Branching Jump Process). For a given $n \in \mathbb{N}$ and starting point $x \in \mathbb{R}^{\mathrm{d}}$, we define $(\Psi^n(t), t \geq 0)$ to be the historical process of the branching random walk described as follows.

1. Each individual has an independent lifetime, which is exponentially distributed with parameter $\eta(1+\gamma_n)\varepsilon_n^{-2}$.

2. During its lifetime, each individual, independently, evolves according to a pure jump process with jump rates given by $(1-(1+\gamma_n)\boldsymbol{s}_n)m_n$.

3. At the end of its lifetime an individual branches into three offspring. The locations of these offspring are determined as follows. First choose $r \in (0, \mathcal{R}_n]$ according to

$$\frac{r^{\mathrm{d}}\mu^n(dr)}{\int_0^{\mathcal{R}_n} \widetilde{r}^{\mathrm{d}}\mu^n(d\widetilde{r})}.$$

If the individual is at point z then the location of each offspring is sampled independently and uniformly from $B(z, r)$.

The process $\Psi^n(t)$ differs from $\Xi^n(t)$ only in that we have suppressed the coalescence events. Since, in this weak noise regime, coalescence events are extremely rare, we can couple the two processes in such a way that they coincide with high probability. The following is Lemma 3.12 of [EFP17].

Lemma 2.32. *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 Ξ^n and Ψ^n, both started with one particle at z such that, with probability at least $1-\varepsilon_n^k$, we have*

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

Sketch of proof. The key idea is to modify the dual process of Definition 2.22 in such a way that individuals are 'preemptively' marked. More precisely, at time zero each individual is marked with probability u_n. When an individual is in the region covered by a reproduction event, it will be affected only if it is marked. After the event, the marks of individuals within the region covered are removed and new marks are assigned (including to the offspring, if any) independently with probability u_n. Once marked, individuals remain marked until they are covered by an event. Unless both are marked, any two lineages evolve independently. The probability that two individuals are marked at time zero is u_n^2. We must control the probability that for a pair of 'root to leaf rays' in $\Xi^n(T^*)$ a reproduction event occurs during

$[0, T^*]$ after which both are marked. If they were not both marked at time zero, then in order for both to be marked, one of them must be affected by an event, after which the probability that they are both marked is u_n^2. Since order nT^* events affect a lineage over $[0, T^*]$, for any pair of root to leaf rays, the probability that there is a reproduction event after which both are marked is $\mathcal{O}(nu_n^2) = \mathcal{O}(n^{4\beta-1})$. (This is why we restrict to $\beta < 1/4$ in the weak noise/selection regime.)

We then use Assumption 1.19 for the second time. This time, it allows one to control the total number of such root to leaf rays in Ξ^n (by the number in a regular ternary tree of height $b \log n$ for a suitable constant b). A union bound over pairs of such rays shows that the probability that there is any time in $[0, T^*]$ when at least two rays are marked (which is required for a coalescence event to take place) is $\mathcal{O}(n^{-\alpha})$, for some $\alpha > 0$. $\square$

Since we have already checked that the motion of a lineage in Ξ^n, and therefore in Ψ^n, is close to a Brownian motion at the first branch time, we can already see why Theorem 2.27 should hold in the weak regime. Of course, there is still some work to do; as $\varepsilon_n \to 0$, the number of branching events in Ψ^n is very large, and it is not obvious that the convergence to Brownian motion along a single lineage will translate into sufficiently rapid convergence on the whole tree. The proof follows the same pattern as the deterministic case.

Generation of interface for the SLFVS

The first step is to show that, analogously to Proposition 2.14, the SLFVS generates an interface in a time window that is of order $\varepsilon_n^2 |\log(\varepsilon_n)|$. Evidently it suffices to work with Ψ^n.

Proposition 2.33. *Let $k \in \mathbb{N}$. Then there exist $n_*(k), a_*(k), b_*(k) > 0$ such that, for all $n \geq n_*$, if we set*

$$\delta_*(k,n) := a(k)\varepsilon_n^2 |\log(\varepsilon_n)| \quad \text{and} \quad \delta_*'(k,n) := (a(k) + \eta^{-1}(k+1))\varepsilon_n^2 |\log(\varepsilon_n)|, \tag{2.43}$$

then, for $t \in [\delta, \delta']$, we have that:

for any x such that $\|x\| \geq \rho_* + d_* \varepsilon_n |\log \varepsilon_n|$, we have $\mathbb{P}_x^{\varepsilon}[\mathbb{V}_{\widehat{p}}^{\gamma}(\Psi^n(t)) = 1] \leq \varepsilon_n^k$.

for any x such that $\|x\| \leq \rho_* - d_* \varepsilon_n |\log \varepsilon_n|$, we have $\mathbb{P}_x^{\varepsilon}[\mathbb{V}_{\widehat{p}}^{\gamma}(\Psi^n(t)) = 1] \geq 1 - \varepsilon_n^k$.

Our proof in the deterministic setting required that we could find a large ternary tree sitting within $\mathcal{T}(\boldsymbol{W})$. Here we also need the converse to prove an analogue of (2.7).

Lemma 2.34 ([EFP17], Lemma 3.16). *Let $k \in \mathbb{N}$ and let $A(k)$ be as in Lemma 2.7. There exist $a_*(k), B_*(k) \in (0,\infty)$ and $n_*(k) < \infty$ such that, for all $n \geq n_*$ and δ_*, δ_*' as defined in (2.43),*

$$\mathbb{P}\left[\mathcal{T}(\Psi^n(\delta_*)) \supseteq \mathcal{T}^{reg}_{A(k)|\log(\varepsilon_n)|}\right] \geq 1 - \varepsilon_n^k, \tag{2.44}$$

$$\mathbb{P}\left[\mathcal{T}(\Psi^n(\delta_*')) \subseteq \mathcal{T}^{reg}_{B(k)|\log(\varepsilon_n)|}\right] \geq 1 - \varepsilon_n^k. \tag{2.45}$$

Sketch of Proof of Proposition 2.33. The proof is based on the proof of Proposition 2.14, where the new ingredient is that now we use Lemma 2.29 to control the distance between a Brownian motion and the jump process with transitions given by m_n (which governs the lineage motion), over the time interval $[\delta_*, \delta_*']$. Since $\varepsilon_n^{-2} = o(\log n)$, for any constant d we can arrange that for large enough n, $d\varepsilon_n|\log(\varepsilon_n)| \geq 2n^{-\beta/6}$. Combining with our previous estimates for the Brownian motion (Proposition 2.13), for any root to leaf ray in $\Psi^n(\delta_*')$, we can use this to control the probability that the leaf is more than $\frac{1}{2}d_*\varepsilon_n|\log(\varepsilon_n)|$ from its starting point at any time in $[\delta_*, \delta_*']$. We can extend this to the whole of Ψ^n using equation (2.45) and a union bound.

The proof of the first inequality of Proposition 2.33 now mirrors that of Proposition 2.14. A symmetric computation gives the second inequality. □

Final steps for the weak noise/selection ratio regime of Theorem 2.27

To conclude the proof of the weak noise/selection ratio regime of Theorem 2.27. We need the following modification of Lemma 2.16. For ease of notation, we write $\|x\|_{\rho_*} := \|x\| - \rho_*$.

Lemma 2.35. *Let $l \in \mathbb{N}$ with $l \geq 4$, $K_1 > 0$. There exists n_* such that for all $n \geq n_*$,*

$x \in \mathbb{R}^{\mathrm{d}}$, $s \in [\sigma^2 \varepsilon_n^{l+3}, \sigma^2(l+1)\eta^{-1}\varepsilon_n^2|\log(\varepsilon_n)|]$ *and* $t \in [s, \infty)$,

$$\mathbb{E}_x \left[g\left(\mathbb{P}_{\|W_s\|_{\rho*} - \nu(t-s) + K_1\varepsilon_n|\log(\varepsilon_n)| + 3n^{-\beta/6}} [\mathbb{V}^\gamma(\mathbf{B}(t-s) = 1] + \varepsilon_n^l \right) \right]$$
$$\leq \frac{3+5\gamma_n}{4(1+\gamma_n)}\varepsilon_n^l + \mathbb{E}_{\|x\|_{\rho*}} \left[g\left(\mathbb{P}_{B_s - \nu t + K_1\varepsilon_n|\log(\varepsilon_n)|} [\mathbb{V}^\gamma(\mathbf{B}(t-s)) = 1] \right) \right] + 1_{s \leq \varepsilon_n^3}\varepsilon_n^l. \qquad (2.46)$$
$$\mathbb{E}_x \left[g\left(\mathbb{P}_{\|W_s\|_{\rho*} - \nu(t-s) - K_1\varepsilon_n|\log(\varepsilon_n)| + 3n^{-\beta/6}} [\mathbb{V}^\gamma(\mathbf{B}(t-s) = 0] + \varepsilon_n^l \right) \right]$$
$$\leq \frac{3+5\gamma_n}{4(1+\gamma_n)}\varepsilon_n^l + \mathbb{E}_{\|x\|_{\rho*}} \left[g\left(\mathbb{P}_{B_s - \nu t - K_1\varepsilon_n|\log(\varepsilon_n)|} [\mathbb{V}^\gamma(\mathbf{B}(t-s)) = 0] \right) \right] + 1_{s \leq \varepsilon_n^3}\varepsilon_n^l. \qquad (2.47)$$

Proof. The proof is identical to that of Lemma 2.16 except that we also approximate the jump process of a lineage by Brownian motion. Since $n^{-\beta/6} = o(s\varepsilon_n|\log(\varepsilon_n)|)$, the additional term $3n^{-\beta/6}$ is negligible for large n. For the case in which $\|W_s\|_{\rho_*}$ is close to 0, we use Lemma 2.12, where the distance is now to a sphere of fixed radius ρ_* (but the statement of the result does not change). □

The equivalent of Proposition 2.15 for Ψ^n is:

Proposition 2.36. *Let $l \in \mathbb{N}$ with $l \geq 4$. Let $a_*(l)$ and $\delta_*(l, \varepsilon)$ given by Proposition 2.33. There exists $K_1(l)$ and $n_*(l, K_1) > 0$ such that, for all $n \geq n_*$, $t \in [\delta_*(l, n), \infty)$ we have:*

$$\sup_{x \in \mathbb{R}^{\mathrm{d}}} \left(\mathbb{P}_x^\varepsilon[\mathbb{V}_{\widehat{p}}^\gamma(\Psi^n(t)) = 1] - \mathbb{P}^{\varepsilon_n}_{\|x\|_{\rho*} - \nu t + K_1\varepsilon_n|\log(\varepsilon_n)|}[\mathbb{V}^\gamma(\mathbf{B}(t)) = 1] \right) \leq \varepsilon_n^l.$$
$$\sup_{x \in \mathbb{R}^{\mathrm{d}}} \left(\mathbb{P}_x^\varepsilon[\mathbb{V}_{\widehat{p}}^\gamma(\Psi^n(t)) = 0] - \mathbb{P}^{\varepsilon_n}_{\|x\|_{\rho*} - \nu t - K_1\varepsilon_n|\log(\varepsilon_n)|}[\mathbb{V}^\gamma(\mathbf{B}(t)) = 0] \right) \leq \varepsilon_n^l.$$

Proof. The proof is essentially identical to that of Proposition 2.15, except that we use Lemma 2.35 in place of Lemma 2.16, and Proposition 2.33 in place of Proposition 2.14. □

Proof of Theorem 2.27, weak noise regime. As in the deterministic setting, it suffices to prove the result for sufficiently large $k \in \mathbb{N}$. By Lemma 2.32, it suffices to work with Ψ^n, and the result then follows from Proposition 2.36, in the same vein as the proof of Theorem 2.11. □

2.2.3 Proof of Theorem 2.27, strong noise/selection ratio regime

We now turn to the strong noise/selection ratio regime.

The total rate at which a lineage is affected by a selective event, and thus at which a new particle is created in the dual, is proportional to

$$\frac{n}{\widehat{u}_n}u_n\boldsymbol{s}_n = \boldsymbol{s}_n n^{2\beta}.$$

The first case of (1.28) corresponds to not seeing any creation of particles in the dual process in the limit, and we include it only for completeness.

The second condition in (1.28) is more interesting, and more complex. In this case, the parameter u_n is sufficiently large relative to the rate of creation of lineages that even though, asymptotically, new lineages may be created infinitely fast in the dual, they are effectively instantly annulled by coalescence. This is, for example, the case if $u_n = (\log n)^{-1/2}$ and $\boldsymbol{s}_n = (\log n)^{1/3}/n^{2\beta}$.

We remark that this cannot be achieved in $\mathbb{d} \geq 3$: the distance between two lineages in the dual is itself a homogeneous jump process with bounded jump size, and this will be transient in $\mathbb{d} \geq 3$, so that there is always a positive probability of the two lineages 'escaping' from one another and never coalescing. In $\mathbb{d} = 2$, even if two lineages initially move apart, there will, with probability one, be a later time at which they are close enough together to be covered by the same event, and therefore have a chance to coalesce. Indeed, in the strong noise/selection ratio regime, most lineages will coalesce with their 'siblings' very soon after being created, with the result that, in a way that we shall make precise below, the 'effective' rate of creation of new lineages in the dual is of order

$$\frac{n}{\widehat{u}_n}\frac{\boldsymbol{s}_n u_n}{u_n \log n} = \frac{\boldsymbol{s}_n n^{2\beta}}{u_n \log n}.$$

Here is a rigorous statement of the key result that we must prove.

Lemma 2.37. *Let $\left(\mathcal{P}^n(s)\right)_{s\geq 0} := \left(\xi_1^n(s), ..., \xi_{N(s)}^n(s)\right)_{s\geq 0}$ be the dual process to the SLFVS introduced in Definition 2.22. Let $x \in \mathbb{R}^{\mathbb{d}}$ and $t > 0$ be fixed. In the strong noise/selection ratio regime there is a Brownian motion $(W_s)_{s\geq 0}$ and a coupling of $\mathcal{P}^n$ and W such that, for any $\varepsilon > 0$, there is n_* such that, for all $n \geq n_*$*

$$\mathbb{P}_x\left[\left\{N(t) = 1\right\} \cap \left\{\|\xi_1^n(t) - W(\sigma^2 t)\| < n^{-\beta/6}\right\}\right] \geq 1 - \varepsilon, \tag{2.48}$$

where $\mathcal{P}^n$ starts from a single particle at x, which is also the starting point of W.

Armed with Lemma 2.37, the strong noise/ selection regime of Theorem 2.27 follows easily.

Proof of Theorem 2.27, strong noise/selection regime. From Lemma 2.37, there is a Brownian motion W such that, for n sufficiently large,

$$\begin{aligned}&\left|\mathbb{P}_x\left[\mathbb{V}_{\widehat{p}}(\Xi^n(t))=1\right]-\mathbb{P}_x[\|W(\sigma^2 t)\|\leq\rho_*]\right|\\&\quad\leq\left|\mathbb{P}_x\left[\|\xi_1^n(t)-W(\sigma^2 t)\|\leq n^{-\beta/6},\|\xi_1^n(t)\|\leq\rho_*\right]-\mathbb{P}_x[\|W(\sigma^2 t)\|\leq\rho_*]\right|+\varepsilon.\end{aligned}$$

Note that

$$\begin{aligned}&\mathbb{P}_x[\|W(\sigma^2 t)\|\leq\rho_*]-\mathbb{P}_x\Big[|\|W(\sigma^2 t)\|-\rho_*|\leq n^{-\beta/6}\Big]\\&\leq\mathbb{P}_x\left[|\xi_1^n(t)-W(\sigma^2 t)|\leq n^{-\beta/6},\|\xi_1^n(t)\|\leq\rho_*\right]\\&\leq\mathbb{P}_x\left[\|W(\sigma^2 t)\|\leq\rho_*\right]+\mathbb{P}\Big[|\|W(\sigma^2 t)\|-\rho_*|\leq n^{-\beta/6}\Big],\end{aligned}\tag{2.49}$$

To conclude, observe that

$$\mathbb{P}_x[|\|W(\sigma^2 t)\|-\rho_*|\leq n^{-\beta/6}]\leq C(\mathbb{d},\rho_*,x)n^{-\beta/6}\leq\varepsilon,\tag{2.50}$$

where the last inequality is valid for sufficiently large values of n_*. Using (2.50) in (2.49) gives the result. □

Proof of Lemma 2.37

Lemma 2.37 is a consequence of Lemma 2.29 and the following proposition.

Proposition 2.38. *Let $t>0$ and $\varepsilon>0$. Then in the strong noise/selection scaling ratio regime there is $n_*\in\mathbb{N}$, such that, for all $n\geq n_*$*

$$\mathbb{P}_x\big[N(t)>1\big]\leq\varepsilon,\tag{2.51}$$

where $N(s)$ is the number of particles alive at time s in the dual process $(\mathcal{P}_s^n)_{0\leq s\leq t}$ of Lemma 2.22.

We focus on the case where $\mathbb{d}=2$, $u_n\log n\to\infty$ and $\boldsymbol{s}_n n^{2\beta}/(u_n\log n)\to 0$ (note that $\liminf u_n\log n$ diverging implies that $\lim u_n\log n\to\infty$, and so we will use the latter condition in what follows). The proof is heavily inspired by [Eth+17]. In that work, it was shown that, in the case of genic selection (in which exactly two offspring are produced in the dual at a selective event), if $\beta = 1/2$, $u_n \equiv u$ and $\boldsymbol{s}_n = \log(n)/n$, there is an equilibrium of branching and coalescence in the dual, so that in the limit as $n \to \infty$ we see a branching Brownian motion (with an 'effective' branching rate). In our notation, this condition on $\boldsymbol{s}_n$ corresponds to $\boldsymbol{s}_n n^{2\beta}/(u_n \log n)$ being $\mathcal{O}(1)$.

Remark 2.39. In the case $u_n \equiv u$, if a pair of lineages is covered by a reproduction event, then there is a probability of order one that they will coalesce. As a result, even for large n, they coalesce having been hit by a finite number of events. In our setting, if one lineage is affected by an event, the probability that the second lineage is also affected is u_n and, as a consequence, a pair of lineages must come together in the order of $1/u_n$ times before they coalesce. The factor $\log n$ corresponds to the number of times that two lineages will come close enough together to be covered by the same event, before they escape to a separation of order one; once they do that, we can expect to wait a long time before they next come close together. This also explains why the condition (1.28) is to be expected to avoid seeing particles making an excursion of order one.

Sketch of proof of Proposition 2.38. Most of the work was done in [Eth+17], and indeed we do not require the precise estimates obtained there. First suppose that two lineages, $\xi^{n,1}$ and $\xi^{n,2}$, are created in a selective event occurring at time 0. Let $\eta^n = \xi^{n,1} - \xi^{n,2}$. The key step is to show that with high probability the two lineages will coalesce before time $1/(\log n)^c$ where c can be chosen to be at least 3. (We shall use a union bound on the complementary event to estimate the probability that all three lineages created in a selection event coalesce on this timescale.)

Much of the work arises from the fact that $\xi^{n,1}$ and $\xi^{n,2}$ only evolve independently when their separation is more than $2\mathcal{R}_n$.

Following [Eth+17], we consider three possible scenarios:

1. The separation $\|\eta^n\|$ exceeds $(\log n)^{-c}$ at some time before $(\log n)^{-c}$. This occurs with probability $\mathcal{O}\left(\frac{1}{u_n \log n}\right)$, and we say that η^n diverges.
2. The quantity $\|\eta^n\|$ does not exceed $(\log n)^{-c}$, and neither do the two lineages coalesce

before time $(\log n)^{-c}$. This happens with probability $\mathcal{O}\left(\frac{1}{(\log n)^{c-3/2}u_n}\right)$, and we say that η^n overshoots.

3. The lineages coalesce within time $(\log n)^{-c}$.

The idea, which can be traced to Lemma 4.2 of [EV12], is to characterise the behaviour of η^n in terms of 'inner' and 'outer' excursions, defined through sequences of stopping times. Set $\tau_0^{\text{out}} = 0$ and define

$$\tau_i^{\text{in}} = \inf\{s > \tau_i^{\text{out}} : |\eta_s^n| \geq 5\mathcal{R}_n\}, \tag{2.52}$$

$$\tau_{i+1}^{\text{out}} = \inf\{s > \tau_i^{\text{in}} : |\eta_s^n| \leq 4\mathcal{R}_n\}. \tag{2.53}$$

The interval $[\tau_i^{\text{out}}, \tau_i^{\text{in}})$ (and the path of η^n during it) is the ith inner excursion, and similarly $[\tau_{i-1}^{\text{in}}, \tau_i^{\text{out}})$ (and the corresponding path) is the ith outer excursion.

We use $\mathbb{P}_{[r_1,r_2]}$ to denote results that hold for any $|\eta_0^n| \in [r_1, r_2]$ and we set

$$L_n = (\log n)^{-c},$$

for a fixed $c \geq 3$. We then define

$$\tau^{\text{coal}} = \inf\{s > 0 : |\eta_s^n| = 0\},$$
$$\tau^{\text{div}} = \inf\{s > 0 : |\eta_s^n| \geq L_n\},$$
$$\tau^{\text{type}} = \tau^{\text{coal}} \wedge \tau^{\text{div}} \wedge L_n.$$

We say η^n coalesces if $\tau^{type} = \tau^{coal}$, diverges if $\tau^{type} = \tau^{div}$, and overshoots otherwise. We will also need the stopping times:

$$\tau_r = \inf\{s > 0 : |\eta_s^n| \leq r\}, \tag{2.54}$$

$$\tau^r = \inf\{s > 0 : |\eta_s^n| \geq r\}. \tag{2.55}$$

Lemma 4.7 in [Eth+17] yields that, as $n \to \infty$,

$$\mathbb{P}_{[5\mathcal{R}_n, 7\mathcal{R}_n]}[\tau^{L_n} < \tau_{4\mathcal{R}_n}] = \mathcal{O}\left(\frac{1}{\log n}\right). \tag{2.56}$$

The proof is unchanged in our setting; during outer excursions, η^n behaves like the difference between two independent jump processes, and it takes $\mathcal{O}(\log n)$ such excursions to achieve separation L_n.

In order to control $\mathbb{P}[\eta^n$ diverges], we control the number of inner excursions before we see a coalescence. We first note there exists $M > 0$ such that

$$\mathbb{P}_{[0,5\mathcal{R}_n]}[\tau_0 \leq \tau^{5\mathcal{R}_n}] \geq Mu_n \tag{2.57}$$

Indeed, from a separation of $5\mathcal{R}_n$, there is a strictly positive probability that within 3 jumps $\|\eta^n\| \leq \mathcal{R}_n$, and then a strictly positive probability that the next event will cover both lineages, in which they coalesce with probability u_n. This allows us to bound the number of inner excursions before coalescence from above by a geometric random variable with success probability Mu_n, and hence we have that the probability of divergence has order at most $1/(u_n \log n)$. Again the details follow exactly as in [Eth+17].

The proof that the probability that η^n overshoots is $\mathcal{O}\left(\frac{1}{(\log n)^{c-3/2}}\right)$ is a minor modification of that of [Eth+17], Lemma 4.4. The idea is that if $\|\eta^n_t\| = x$, with $x \in (0, 5\mathcal{R}_n)$, then with strictly positive probability θ, η^n will either coalesce or exit $B(0, 5\mathcal{R}_n)$ within three jumps. The number of jumps that η^n makes before either exiting $B(0, 5\mathcal{R}_n)$ or coalescing is therefore stochastically bounded by three times a geometric random variable with success probability θ. Since η^n jumps at least as fast as $\xi^{n,1}$, which jumps at rate m_n given by (2.41), this allows us to estimate

$$\mathbb{P}_{(0,5\mathcal{R}_n)}\left[\tau_0 \wedge \tau^{5\mathcal{R}_n} > n^{-\beta}\right] = \mathcal{O}\left(n^{-\beta/3} + (1-\theta)^{n^{2\beta/3}}\right),$$

where the first term is an estimate of the probability that the sum of $n^{2\beta/3}$ independent exponential random variables with parameter m_n exceeds $n^{-\beta}$ and the second is the probability that a Geometric random variable with parameter θ exceeds $n^{2\beta/3}$.

Now let i^* denote the index of the excursion during which τ^{type} occurs. From the argument that we used to control the probability of divergence, we can bound the number of inner excursions before coalescence above by a geometric random variable with success probability Mu_n. Let n be large enough that

$$\frac{(\log n)^{1/2}}{u_n}\left(n^{-\beta} + (\log n)^{-c-2}\right) \leq (\log n)^{-c}.$$

Note that this is possible since $u_n \log n \to \infty$ as $n \to \infty$, by assumption. If η^n overshoots, and $i^* < (\log n)^{1/2}/u_n$, then at least one inner excursion must have lasted at least $n^{-\beta}$ or at least one outer excursion must have lasted at least $(\log n)^{-c-1}$. Again using that when $\|\eta^n\| > 2\mathcal{R}_n$, η^n behaves as the difference of two independent walks (and Skorohod embedding), Lemma 4.7 of [Eth+17] shows that

$$\mathbb{P}_{[5\mathcal{R}_n, 7\mathcal{R}_n]}\left[\tau^{L_n} \wedge \tau_{4\mathcal{R}_n} > (\log n)^{-c-2}\right] = \mathcal{O}\left(\frac{1}{(\log n)^c}\right).$$

Combining the above,

$$\begin{aligned}
&\mathbb{P}\left[\eta^n \text{ overshoots}\right] \\
&\quad\leq \frac{(\log n)^{1/2}}{u_n}\left\{\mathbb{P}_{(0,5\mathcal{R}_n)}\left[\tau_0 \wedge \tau^{5\mathcal{R}_n} > n^{-\beta}\right] + \mathbb{P}_{[5\mathcal{R}_n,7\mathcal{R}_n]}\left[\tau^{L_n} \wedge \tau_{4\mathcal{R}_n} > (\log n)^{-c-2}\right]\right\} \\
&\qquad\qquad + \mathbb{P}\left[i^* > \frac{(\log n)^{1/2}}{u_n}\right] \\
&\quad\leq \frac{(\log n)^{1/2}}{u_n}\left(n^{-\beta} + (\log n)^{-c}\right) + (1 - Mu_n)^{(\log n)^{1/2}/u_n} = \mathcal{O}\Big(\frac{1}{(\log n)^{c-3/2} u_n}\Big).
\end{aligned}$$

To conclude the proof of Proposition 2.38, let us write ξ for a root to leaf ray connecting $\xi_0 = x$ to $\xi_1^n(t)$ (the first individual in $\mathcal{P}^n(t)$). Let S_m denote the time of the m-th selective event to affect the lineage, creating offspring $\xi^{n,1}, \xi^{n,2}, \xi^{n,3}$. Then, by the Markov property,

$$\begin{aligned}
&\mathbb{P}\left[\text{all individuals created at time } S_m \text{ have not coalesced by time } S_m + \frac{1}{(\log n)^c}\right] \\
&\quad\leq 3\mathbb{P}[\eta^n = (\xi^{n,1} - \xi^{n,2}) \text{ diverges or overshoots}] \\
&\quad= \mathcal{O}\left(\frac{1}{u_n \log n}\right) + \mathcal{O}\left(\frac{1}{u_n (\log n)^{c-3/2}}\right) = \mathcal{O}\left(\frac{1}{u_n \log n}\right).
\end{aligned}$$

Since selective events affect ξ according to a Poisson process of rate proportional to $\boldsymbol{s}_n u_n n/\widehat{u}_n = \boldsymbol{s}_n n^{2\beta}$,

$$\begin{aligned}
\mathbb{P}[N_t^n \neq 1] &\leq \mathbb{P}\Big[\text{a particle created in } [0, t - (\log n)^{-c}] \text{ did not coalesce}\Big] \\
&\quad + \mathbb{P}\Big[\text{a selective event occurred in } [t - (\log n)^{-c}, t]\Big] \\
&= \mathcal{O}\left(\frac{\boldsymbol{s}_n n^{2\beta} t}{u_n \log n}\right) + \mathcal{O}\left(\frac{\boldsymbol{s}_n n^{2\beta}}{(\log n)^{-c}}\right) \leq \mathcal{O}\left(\frac{\boldsymbol{s}_n n^{2\beta}(t \vee 1)}{u_n \log n}\right),
\end{aligned}$$

where our conditions on $\boldsymbol{s}_n$ guarantee that all the terms on the right hand side tend to 0 as $n \to \infty$, and the proof is complete. $\square$

3

Hybrid zones in populations with long-range dispersal

In this Chapter we present the proof of the main result stated in Section 1.7, Theorem 1.39.

To prove Theorem 1.39, we proceed as follows. First, in Section 3.1 we provide a stochastic representation for the solution of equation (1.34) in terms of branching α-stable processes. This is a long-range analogue of Proposition 1.21, where the solution at the point x at time t is the expected result of a *voting procedure* defined on the tree of paths traced out by branching α-stable processes, run up to time t, starting from a single individual at the point x.

In Section 3.2 we prove a one-dimensional version of Theorem 1.39. Two important tools are exploited here. First, that the equation (1.34), is invariant by transformation $v^\varepsilon(t,x) \to 1 - v^\varepsilon(t,x)$. Secondly, we have Theorem 3.10 and Theorem 3.12, which together roughly state that we can erase the *big jumps* of the α-stable process and get an error that goes to 0 with ε. Using this, the resulting processes behave, basically, as Brownian motions. This streamlines the computations and make the techniques in [EFP17] easy to apply.

In Section 3.3 we proceed to prove Theorem 1.39. Of particular importance is Section 3.3.1, in which we prove a coupling which allows us to compare our multidimensional α-stable processes with one-dimensional α-stable processes, given that we can get rid of the big jumps. In Section 3.3.1 we prove that we can indeed forget those big jumps in the α-stable process in higher dimensions and get a negligible error. Exploiting the two last facts in Section 3.3.3 and Section 3.3.4 we conclude Theorem 1.39.

Some proofs with longer computations are relegated to Appendix B to avoid disrupting the flow of the main results.

3.1 A probabilistic dual

In this section, we define a probabilistic dual to the scaled fractional Allen-Cahn equation (1.34), following closely the work of [EFP17].

To do this, we consider ternary branching α-stable motions in which each individual, independently, follows an α-stable motion, until the end of its exponentially distributed lifetime (with mean ε^2) at which point it splits into three particles. Let $Y(t)$ denote a $\mathbb{d}$-dimensional α-stable process and $\boldsymbol{Y}(t)$ denote a $\mathbb{d}$-dimensional historical ternary branching α-stable motion. That is, $\boldsymbol{Y}(t)$ traces out the space-time trees that record the position of all particles alive at time s for $s \in [0,t]$. Throughout this work, we adopt the convention that, for fixed $\varepsilon > 0$,

Assumption 3.1. All stable motions have generator $\sigma^{-2}I(\varepsilon)^{\alpha-2}(-\Delta)^{\frac{\alpha}{2}}$, for $\sigma^2 \equiv \sigma^2_{\mathbb{d}}(\alpha)$ as in (1.33).

That is, all $\mathbb{d}$-dimensional α-stable processes are assumed to be of the form $X_{\sigma^{-2}I(\varepsilon)^{\alpha-2}t}$, where the speed $\sigma^{-2}I(\varepsilon)^{\alpha-2}$ will be suppressed in our notation. As in Chapter 2, we employ the Ulam-Harris notation to record the genealogy of the process. We denote the time-labelled ternary tree traced out by $\boldsymbol{Y}(t)$ by $\mathcal{T}(\boldsymbol{Y}(t))$. For a fixed function $p : \mathbb{R}^{\mathbb{d}} \to [0,1]$ we define a voting procedure on $\mathcal{T}$ as follows:

(1) each leaf i of $\mathcal{T}(\boldsymbol{Y}(t))$ independently votes 1 with probability $p(Y_i(t))$ and otherwise votes 0;

(2) at each branching event in $\mathcal{T}(\boldsymbol{Y}(t))$, the vote of the parent particle j is given by the majority vote of its offspring $(j,1),(j,2),(j,3)$.

This voting procedure runs inward from the leaves of $\mathcal{T}(\boldsymbol{Y}(t))$ to the root $\emptyset$.

Definition 3.2 ($\mathbb{V}_p$). With the majority voting procedure described above, define $\mathbb{V}_p(\boldsymbol{Y}(t))$ to be the vote associated to the root $\emptyset$ of the ternary branching stable tree.

For $x \in \mathbb{R}^{\mathrm{d}}$, we write $\mathbb{P}_x^\varepsilon$ for the probability measure under which $(\boldsymbol{Y}(t), t \geq 0)$ has the law of the historical process of a ternary branching α-stable motion in $\mathbb{R}^{\mathrm{d}}$ with branching rate $1/\varepsilon^2$ started from a single particle at location x at time 0. We write $\mathbb{E}_x^\varepsilon$ for the corresponding expectation. Then the root vote $\mathbb{V}_p(\boldsymbol{Y}(t))$ provides us with a dual to equation (1.34) in the following sense.

Theorem 3.3. *Let $p : \mathbb{R}^{\mathrm{d}} \to [0,1]$. Then*

$$v^\varepsilon(t,x) = \mathbb{P}_x^\varepsilon[\mathbb{V}_p(\boldsymbol{Y}(t)) = 1] \tag{3.1}$$

is a solution to equation (1.34) with initial condition $v^\varepsilon(x,0) = p(x)$.

We emphasise that the variable ε used to define the speed of the stable process, $I(\varepsilon)^{\alpha-2}$, is the same variable ε that defines the branching rate, ε^{-2}. The proof of Theorem 3.3 is an analogue of the proof of Proposition 2.4 in Chapter 2, and so we omit it.

With this in mind, we can restate Theorem 1.39 as follows.

Theorem 3.4. *Let $\alpha \in (1,2)$ and fix a function $I : \mathbb{R}_+ \to \mathbb{R}_+$ satisfying Assumptions 1.37. Suppose the initial condition p satisfies Assumptions 1.16. Let $\mathscr{T}$ and $d(x,t)$ be as in Section 1.5, let F be as in (1.35), and fix $T^* \in (0, \mathscr{T})$. Then there exist $\varepsilon_{\mathrm{d}}(\alpha,I), a_{\mathrm{d}}(\alpha,I), c_{\mathrm{d}}(\alpha,I), M(\alpha,I) > 0$ such that, for $\varepsilon \in (0,\varepsilon_{\mathrm{d}})$ and $a_{\mathrm{d}}\varepsilon^2|\log\varepsilon| \leq t \leq T^*$,*

(1) for x with $d(x,t) \geq c_{\mathrm{d}}(\alpha)I(\varepsilon)|\log\varepsilon|$, we have $\mathbb{P}_x\left[\mathbb{V}_p(\boldsymbol{Y}(t)) = 1\right] \geq 1 - \frac{\varepsilon^2}{I(\varepsilon)^2} - F(\varepsilon, M)$;

(2) for x with $d(x,t) \leq -c_{\mathrm{d}}(\alpha)I(\varepsilon)|\log\varepsilon|$, we have $\mathbb{P}_x\left[\mathbb{V}_p(\boldsymbol{Y}(t)) = 1\right] \leq \frac{\varepsilon^2}{I(\varepsilon)^2} + F(\varepsilon, M)$.

Notation. *It will be convenient to distinguish between the one-dimensional and multi-dimensional α-stable processes. We adopt the convention that $X(t)$ will denote the one-dimensional α-stable process, with the corresponding historical branching stable process denoted by $\boldsymbol{X}(t)$. When $\mathrm{d} > 1$, we denote the α-stable process by $Y(t)$ and denote the corresponding historical branching stable process by $\boldsymbol{Y}(t)$.*

Remark 3.5. As we mentioned briefly, in [CS10], the authors construct an algorithm similar to the MBO algorithm but for the fractional heat equation. In the MBO algorithm we run the fractional heat equation at rate $h^{\frac{\alpha}{2}-1}$ for time h (for some small h), and then *snap* every value of the solution above $1/2$ to 1, and every value below $1/2$ to 0. We then iterate this process. The result in [CS10] states this procedure gives an indicator function the boundaries of whose support evolve as mean curvature flow.

We can translate this procedure into α-stable processes running at speed $h^{\frac{\alpha}{2}-1}$ for a time window h. After that time window, we would *snap* using the voting procedure in Definition 3.2. Note the time steps in the MBO algorithm are analogous to the $\texttt{Exp}(\varepsilon^{-2})$ branch times in our setting. If we approximate these branch times by their mean and set $h = \varepsilon^2$, [CS10] suggests that all stable processes should run at speed $\varepsilon^{\alpha-2}$. Notably, we do not believe Theorem 1.39 will hold under this predicted scaling. This is because we are concerned with interfaces that have width converging to zero, while [CS10] works with constant order interface widths (more on this in Remark 3.8).

They prove that the resulting interface converges to motion by mean curvature flow for $\alpha \geq 1$. We can compare the scaling taken in [CS10] to ours as follows. In [CS10], the authors consider time steps of length h and run the diffusive step for a time $h^{\frac{\alpha}{2}}$. In the continuous-time setting, this would be equivalent to constructing sets Ω_t for $t \geq 0$ after running the fractional heat equation for a time $h^{\frac{\alpha}{2}} h^{-1} t$ (i.e. running stable processes at speed $h^{\frac{\alpha}{2}-1}$).

Notably, this work requires all stable processes to run at speed $I(\varepsilon)^{\alpha-2}$ for $I(\varepsilon)$ *strictly greater* than ε. More precisely, by Assumption 1.37 (B), we require

$$\lim_{\varepsilon \to 0} \frac{\varepsilon^2}{I(\varepsilon)^2} |\log(\varepsilon)| = 0.$$

Heuristically, this condition arises because the arrival rate of a 'large' jump in all α-stable processes is $I(\varepsilon)^{-2}$. Bounding above each branch time S by its $(1-\varepsilon^k)$−quantile, $S \leq k\varepsilon^2|\log \varepsilon|$, and we find that each stable process is expected to make, at most,

$$k\frac{\varepsilon^2}{I(\varepsilon)^2}|\log \varepsilon|$$

'large' jumps in it's lifetime. To obtain a reasonable convergence result, we expect this quantity should go to zero with ε, and this is precisely Assumption 1.37 (B).

3.2 Majority voting in one dimension

In this section we consider only ternary branching α-stable motions in dimension one. As in Section 3.1, for each $x \in \mathbb{R}$, we write $\mathbb{P}_x^\varepsilon$ for the probability measure under which $(\boldsymbol{X}(t), t \geq 0)$ has the law of a one-dimensional historical ternary branching α-stable motion in $\mathbb{R}$, with branching rate ε^{-2}, started from a single particle at location x at time 0. In accordance with our earlier convention, each particle in a ternary branching α-stable motion is assumed to run at speed $\sigma^{-2} I(\varepsilon)^{\alpha-2}$. Write $\mathbb{E}_x^\varepsilon$ for the corresponding expectation. Let P_x be the law of an α-stable motion started at x, with corresponding expectation E_x. Denote the vote associated with the root of $\mathcal{T}(\boldsymbol{X}(t))$ under majority voting by $\mathbb{V}_p(\boldsymbol{X}(t))$.

We now aim to prove the analogous statement to Theorem 3.4 in one-dimension with initial condition

$$p(x) = \mathbb{1}_{\{x \geq 0\}},$$

i.e. when the leaves of $\mathcal{T}(\boldsymbol{X}(t))$ vote one if and only if they are on the right half line. For our choice of $p = \mathbb{1}_{\{x\geq 0\}}$, we write $\mathbb{V} := \mathbb{V}_p$. We will prove the following version of Theorem 1.39 in one dimension.

Theorem 3.6. *Let $\alpha \in (1,2)$ and $T^* \in (0,\infty)$. Suppose $I : \mathbb{R}_+ \to \mathbb{R}_+$ satisfies Assumptions 1.37 (A) and 1.37 (B). Then there exist $c_1(\alpha, I), \varepsilon_1(\alpha, I) > 0$ such that, for all $t \in [0, T^*]$ and all $\varepsilon \in (0, \varepsilon_1)$,*

1. *for $x \geq c_1(\alpha) I(\varepsilon) |\log \varepsilon|$, we have $\mathbb{P}_x^\varepsilon[\mathbb{V}(\boldsymbol{X}(t)) = 1] \geq 1 - \frac{\varepsilon^2}{I(\varepsilon)^2}$,*
2. *for $x \leq -c_1(\alpha) I(\varepsilon)|\log\varepsilon|$, we have $\mathbb{P}_x^\varepsilon[\mathbb{V}(\boldsymbol{X}(t)) = 1] \leq \frac{\varepsilon^2}{I(\varepsilon)^2}$.*

Note that, for positive x, 'typical' leaves of the branching process started at x are more likely to vote one than zero.

Remark 3.7. Let f be any bistable nonlinearity and consider the one-dimensional equation

$$\begin{cases} (-\Delta)^{\frac{\alpha}{2}} u(y) = f(u(y)), \quad \forall y \in \mathbb{R}, \\ \lim_{y\to\infty} u(y) = 1, \\ \lim_{y\to-\infty} u(y) = 0. \end{cases} \tag{3.2}$$

Then, by [GZ15, Proposition 3.2], a solution $u \in C^2(\mathbb{R})$ to (3.2) satisfies

$$\frac{A}{y^\alpha} \le 1 - u(y) \le \frac{B}{y^\alpha} \qquad \forall y > 1$$

for some $A, B > 0$ (and similarly for $u(y)$ when $y \le -1$). We rewrite this equation under our scaling by setting $z = \varepsilon^{\frac{2}{\alpha}} I(\varepsilon)^{1-\frac{2}{\alpha}} y$, and obtain

$$\frac{A\varepsilon^2 I(\varepsilon)^{\alpha-2}}{z^\alpha} \le 1 - u(z) \le \frac{B\varepsilon^2 I(\varepsilon)^{\alpha-2}}{z^\alpha}. \tag{3.3}$$

Hence, if $z \ge I(\varepsilon)$, this tells us that $1 - u(z)$ is exactly of order $\frac{\varepsilon^2}{I(\varepsilon)^2}$, the interface tolerance from Theorem 3.6. This suggests that Theorem 3.6 should hold for an interface of width $I(\varepsilon)$ (and this is the narrowest width possible for the given interface tolerance), but we were only able to prove this result for the larger interface width $I(\varepsilon)|\log \varepsilon|$. The logarithmic term in our interface width appears to be an unavoidable consequence of the need to control the deviations of the exponential branch times used to define $\boldsymbol{X}(t)$.

Remark 3.8. We can further consider [GZ15, Proposition 3..2] in the setting of the fractional MBO-like algorithm constructed in [CS10] (see Remark 3.5 for a description of the algorithm). Using the scaling from [CS10], we set $I(\varepsilon) = \varepsilon$. With this choice of $I(\varepsilon)$, [GZ15] tells us that, in order for the right and left bounds in (3.3) to tend to zero as ε goes to zero, the interface width would need to be of constant order. This also suggests that the definitions of interface used in this work and in [IS09] are not the same.

Theorem 3.6 tells us that even a small voting bias at the leaves is magnified under the majority voting procedure to a large voting bias at the root (and similarly when $x < 0$). We remark that, as in [EFP17], the choice of initial condition $p(x) = \mathbb{1}_{\{x \ge 0\}}$ affords us several useful inequalities. First, for all $x_1, x_2 \in \mathbb{R}$ with $x_1 \le x_2$, we have

$$\mathbb{P}^\varepsilon_{x_1}[\mathbb{V}(\boldsymbol{X}(t)) = 1] \le \mathbb{P}^\varepsilon_{x_2}[\mathbb{V}(\boldsymbol{X}(t)) = 1]. \tag{3.4}$$

As before, let $\mathcal{T}(\boldsymbol{X}(t))$ denote the time-labelled ternary tree traced out by the ternary branching α-stable motion $\boldsymbol{X}(t)$, and for any time-labelled ternary tree $\mathcal{T}$, write

$$\mathbb{P}^t_x(\mathcal{T}) := \mathbb{P}^\varepsilon_x[\mathbb{V}(\boldsymbol{X}(t)) = 1 | \mathcal{T}(\boldsymbol{X}(t)) = \mathcal{T}].$$

It then follows by symmetry of α-stable motions and the definition of $p(x)$ that, for any $x \in \mathbb{R}$ and time $t > 0$,

$$\mathbb{P}_x^t(\mathcal{T}) = 1 - \mathbb{P}_{-x}^t(\mathcal{T}). \tag{3.5}$$

Setting $x = 0$ in equation 3.5 yields $\mathbb{P}_0^t(\mathcal{T}) = \frac{1}{2}$, so by monotonicity (3.4).

$$\mathbb{P}_x^t(\mathcal{T}) \geq \frac{1}{2} \ \text{ for } \ x > 0, \ \text{ and } \ \mathbb{P}_x^t(\mathcal{T}) \leq \frac{1}{2} \ \text{ for } \ x < 0.$$

It will be convenient in our later calculations to introduce notation for the majority voting system. Mimicking [EFP17], define the function $g : [0,1]^3 \to [0,1]$ by

$$g(p_1, p_2, p_3) = p_1 p_2 p_3 + p_1 p_2 (1 - p_3) + p_2 p_3 (1 - p_1) + p_3 p_1 (1 - p_2). \tag{3.6}$$

This is the probability that a majority vote gives the result 1, in the special case when the three voters are independent and have probabilities p_1, p_2 and p_3 of voting 1. We will abuse notation slightly and write $g(p) := g(p, p, p)$.

3.2.1 A coupling of voting systems

In this section, we aim to couple the root vote of the branching α-stable motion defined above to the root vote of a branching subordinated Brownian motion, under a different voting system. This subordinated Brownian motion will have similar distributional properties to a truncated α-stable motion, but is advantageous in that it allows us to adapt the Brownian setting of [EFP17] more readily. As we shall see, the leaves of the subordinated Brownian tree rooted at $x > 0$, under this new voting system, will have smaller probability of voting 1 than 0 compared to the leaves of $\mathcal{T}(\boldsymbol{X}(t))$ rooted at $x > 0$ under majority voting (with the opposite being true for $x < 0$). This increases the tolerance of our interface, and can be viewed as the penalty one must pay for changing the distribution of the underlying tree from the heavy-tailed stable distribution to the relatively well behaved subordinated Brownian motion.

The intuition behind our new voting system, which we refer to as *marked majority voting*, is straightforward. However, formally proving a coupling of the majority and marked majority systems is more challenging. To aid us with this (at the slight risk of introducing complications) we introduce an intermediate voting system that can be readily compared to both voting systems. We call this the *exponentially marked* voting procedure. The definition of the exponentially marked system, together with a proof that it couples with the ordinary majority voting system in the appropriate sense (Theorem 3.10) make up the content of Section 3.2.1. After this, in Section 3.2.1, we define the marked majority voting system that will be carried through the one-dimensional proof, and prove (using the intermediate voting system) that it can be coupled to majority voting. Theorem 3.6 can then be restated in terms of the marked voting system on the subordinated Brownian tree (Theorem 3.14), and is proved using new techniques as well as those developed in [EFP17], [HD20].

Exponentially marked voting

In this section we aim to couple the majority voting system on $\mathcal{T}(\boldsymbol{X}(t))$ with a new voting system defined on a ternary branching subordinated Brownian motion. We will consider an $I(\varepsilon)^2$-truncated $\frac{\alpha}{2}$-stable subordinator denoted R_t^ε. Assumption 3.1 will also be adopted for all stable subordinators. To be precise, the Lévy measure of R_t^ε is given by

$$\sigma^{-2} I(\varepsilon)^{\alpha-2} \nu^{\frac{\alpha}{2}}(dx) \mathbb{1}_{\{0 \leq x \leq I(\varepsilon)^2\}},$$

for $\nu^{\frac{\alpha}{2}}$ the Lévy measure of the symmetric $\frac{\alpha}{2}$-stable process. Let $\boldsymbol{B}_{R^\varepsilon}(t)$ be the historical process of a ternary branching R_t^ε-subordinated Brownian motion (with branching rate ε^{-2}). Let us now make precise the form of the coupling that we desire. As before, let $\mathbb{V}(\boldsymbol{X}(t))$ denote the root vote of $\mathcal{T}(\boldsymbol{X}(t))$ under majority voting. We aim to define a voting system on $\mathcal{T}(\boldsymbol{B}_{R^\varepsilon}(t))$ with root vote $\widehat{\mathbb{V}}(\boldsymbol{B}_{R^\varepsilon}(t))$ satisfying

$$\mathbb{P}_x^\varepsilon[\mathbb{V}(\boldsymbol{X}(t)) = 1] \geq \mathbb{P}_x^\varepsilon[\widehat{\mathbb{V}}(\boldsymbol{B}_{R^\varepsilon}(t)) = 1] \quad \text{for all} \quad x \geq 0, \tag{3.7}$$

with the reverse inequality holding for $x \leq 0$. Having obtained this, it will suffice to prove the analogue of Theorem 3.6 with $\mathbb{V}(\boldsymbol{X})$ replaced by $\widehat{\mathbb{V}}(\boldsymbol{B}_{R^\varepsilon})$. In this way, we will have incorporated the problematic 'large jumps' of the α-stable process $\boldsymbol{X}(t)$ into the voting system $\widehat{\mathbb{V}}$.

Let us now define a voting system on $\mathcal{T}(\boldsymbol{B}_{R^\varepsilon}(t))$. Consider a collection of independent $\frac{\alpha}{2}$-stable subordinators, $\{R_i\}_{i\in M(t)}$, where $M(t)$ denotes the set of individuals that have ever been alive in $\mathcal{T}(\boldsymbol{X}(t))$. We will keep track of the random times $\{\widehat{S}_i\}_{i\in M(t)}$ that describe the first time R_i makes a jump of size larger than $I(\varepsilon)^2$ (these are the exponential times from which this voting system derives its name). Explicitly,

$$\widehat{S}_i := \inf\{t \geq 0 : |R_i(t) - R_i(t-)| > I(\varepsilon)^2\}.$$

The distribution of $\widehat{S}_i$ can be computed as follows. Let $\nu^{\frac{\alpha}{2}}(dx)$ be the Lévy measure of an $\frac{\alpha}{2}$-stable process. Then the first jump bigger than $I(\varepsilon)^2$ arrives at an exponential time with parameter

$$\sigma^{-2}I(\varepsilon)^{\alpha-2}\int_{\mathbb{R}}\nu^{\frac{\alpha}{2}}(dx) = \frac{C_\alpha\sigma^{-2}}{I(\varepsilon)^{2-\alpha}}\int_{I(\varepsilon)^2}^{\infty}\frac{1}{x^{\frac{\alpha}{2}+1}}dx = \frac{2C_\alpha\sigma^{-2}}{\alpha I(\varepsilon)^2}. \tag{3.8}$$

For simplicity we set $\widehat{I}(\varepsilon) := (2C_\alpha\sigma^{-2}\alpha^{-1})^{\frac{1}{2}}I(\varepsilon)$ and work on our computations with $\widehat{I}$ in what follows. As a consequence, the right hand side of (3.8) is equal to $\widehat{I}(\varepsilon)^2$ and so $\{\widehat{S}_i\}_{i\in M(t)}$ are i.i.d. $\mathtt{Exp}(\widehat{I}(\varepsilon)^{-2})$ distributed.

We further associate to each particle in $\mathcal{T}(\boldsymbol{B}_{R^\varepsilon}(t))$ its lifetime, $S_i \wedge t$, for $S_i \sim \mathtt{Exp}(\varepsilon^{-2})$. Now fix $\varepsilon > 0$ and $\alpha \in (1,2)$. For a fixed function $p : \mathbb{R}^{\mathrm{d}} \to [0,1]$, we define an *exponentially marked voting procedure* on $\mathcal{T}(\boldsymbol{B}_{R^\varepsilon}(t))$ as follows.

1. Each individual $B_i(R_i^\varepsilon)$ is said to be marked if $\widehat{S}_i < S_i$. Each marked individual votes 1 with probability $\frac{1}{2}$, and otherwise votes 0.

2. Each unmarked leaf i of $\mathcal{T}(\boldsymbol{B}_{R^\varepsilon}(t))$, independently, votes 1 with probability $p(B_i(R_i^\varepsilon(t)))$ and otherwise votes 0.

3. At each branch point in $\mathcal{T}(\boldsymbol{B}_{R^\varepsilon}(t))$, if the parent particle k is unmarked, it votes according to the majority vote of its three offspring $(k,1),(k,2)$ and $(k,3)$.

Definition 3.9 ($\widehat{\mathbb{V}}_p$). With the exponentially marked majority voting procedure described above, define $\widehat{\mathbb{V}}_p(\boldsymbol{B}_{R^\varepsilon}(t))$ to be the vote associated to the root $\emptyset$ of the ternary branching truncated stable tree.

The exponentially marked voting system can be understood intuitively as follows. First, since an $\frac{\alpha}{2}$-stable subordinated Brownian motion is equal in distribution to an α-stable process, we may, without loss of generality, consider the historical ternary branching $\frac{\alpha}{2}$-stable subordinated Brownian $\boldsymbol{B}_R(t)$ instead of $\boldsymbol{X}(t)$. Suppose the trees $\mathcal{T}(\boldsymbol{B}_R(t))$ and $\mathcal{T}(\boldsymbol{B}_{R^\varepsilon}(t))$ have been generated up to time t, and that these trees have the same branching structure, written as $\mathcal{T}(\boldsymbol{B}_{R^\varepsilon}(t)) = \mathcal{T}(\boldsymbol{B}_R(t))$. Then each individual $B_i(R_i)$ in $\mathcal{T}(\boldsymbol{B}_R(t))$ can be associated with the partner particle of i, $B_i(R_i^\varepsilon)$ in $\mathcal{T}(\boldsymbol{B}_{R^\varepsilon}(t))$. If the subordinator R_i has not made a large jump in the lifetime of particle i, we allow their partner particle $B_i(R_i^\varepsilon)$ to vote according to the majority vote of their three offspring (i.e. the partner particle is unmarked). If R_i does make a large jump in the lifetime of particle i, with equal probability $B_i(R_i)$ will jump to the left or right, thereby decreasing or increasing (respectively) the probability that the i-th particle votes one. To reflect this in the truncated system, we ask that the partner particle $B_i(R_i^\varepsilon)$ votes 1 or 0 with equal probability, regardless of its offspring's votes (i.e. the partner particle is marked). Provided the root, x, is non-negative, a jump of $B_i(R_i)$ is more likely to land in the right-half plane than the left. That is, the partner particle $B_i(R_i^\varepsilon)$ is being pessimistic when it votes one with probability $\frac{1}{2}$, and we expect (3.7) to hold.

As before, we will consider $p(x) = \mathbb{1}_{\{x\geq 0\}}$, and write $\widehat{\mathbb{V}} := \widehat{\mathbb{V}}_p$. Note that, for this choice of p, for any $x_1, x_2 \in \mathbb{R}$ with $x_1 \leq x_2$,

$$\mathbb{P}^\varepsilon_{x_1}[\widehat{\mathbb{V}}(\boldsymbol{B}_{R^\varepsilon}(t)) = 1] \leq \mathbb{P}^\varepsilon_{x_2}[\widehat{\mathbb{V}}(\boldsymbol{B}_{R^\varepsilon}(t)) = 1]. \tag{3.9}$$

Let $\mathcal{T}$ be a time-labelled ternary tree and define

$$\widehat{\mathbb{P}}^\varepsilon_x(\mathcal{T}) := \mathbb{P}^\varepsilon_x\left[\widehat{\mathbb{V}}(\boldsymbol{B}_{R^\varepsilon}(t)) = 1 | \mathcal{T} = \mathcal{T}(\boldsymbol{B}_{R^\varepsilon}(t))\right].$$

Then, since 0 and 1 are exchangeable in the exponential marked voting system,

$$\widehat{\mathbb{P}}^t_x(\mathcal{T}) = 1 - \widehat{\mathbb{P}}^t_{-x}(\mathcal{T}) \tag{3.10}$$

for all $x \in \mathbb{R}$ and $t \geq 0$. Setting $x = 0$ in equation (3.10) shows that $\widehat{\mathbb{P}}^t_0(\mathcal{T}) = \frac{1}{2}$ for all $t > 0$, and, together with monotonicity (3.9), this gives

$$\widehat{\mathbb{P}}^\varepsilon_x(\mathcal{T}) \geq \frac{1}{2} \quad \text{for} \quad x > 0, \quad \text{and} \quad \widehat{\mathbb{P}}^\varepsilon_x(\mathcal{T}) \leq \frac{1}{2} \quad \text{for} \quad x < 0.$$

To conclude this section, we prove the claimed coupling of voting systems.

Theorem 3.10. *Let $\varepsilon > 0$, $t \geq 0$ and $\alpha \in (1,2)$. Then*

1. *for all $x \geq 0$, $\mathbb{P}^{\varepsilon}_x[\mathbb{V}(\boldsymbol{X}(t)) = 1] \geq \mathbb{P}^{\varepsilon}_x[\widehat{\mathbb{V}}(\boldsymbol{B}_{R^\varepsilon}(t)) = 1]$,*

2. *for all $x \leq 0$, $\mathbb{P}^{\varepsilon}_x[\mathbb{V}(\boldsymbol{X}(t)) = 1] \leq \mathbb{P}^{\varepsilon}_x[\widehat{\mathbb{V}}(\boldsymbol{B}_{R^\varepsilon}(t)) = 1]$.*

Proof. We only prove the first inequality, since the second inequality will follow by the symmetry relations (3.5) and (3.10). First, by coupling the branching structure of $\mathcal{T}(\boldsymbol{B}_{R^\varepsilon}(t))$ and $\mathcal{T}(\boldsymbol{X}(t))$, it suffices to show that

$$\mathbb{P}^t_x(\mathcal{T}) \geq \widehat{\mathbb{P}}^t_x(\mathcal{T}) \text{ for all } x \geq 0. \tag{3.11}$$

We proceed by induction on the number of branching events in $\mathcal{T}(\boldsymbol{X}(t))$ and $\mathcal{T}(\boldsymbol{B}_{R^\varepsilon}(t))$. Let $\mathcal{T}_0$ denote the tree with a root and a single leaf. Each particle $B_i(R^\varepsilon_i)$ in $\mathcal{T}(\boldsymbol{B}_{R^\varepsilon}(t))$ has lifetime $S_i \wedge t$, for $S_i \sim \mathtt{Exp}(\varepsilon^{-2})$. We abuse notation and write the time of the first branching event as $S := S_0$. Let $\{\widehat{S}_i\}_{i\in M(t)}$ be a family of i.i.d. exponential random variables with rate $\widehat{I}(\varepsilon)^{-2}$. We introduce the shorthand notation

$$P^0_x(A) := \mathbb{P}^{\varepsilon}_x[A|\mathcal{T}(\boldsymbol{B}_{R^\varepsilon}(t)) = \mathcal{T}(\boldsymbol{X}(t)) = \mathcal{T}_0]$$

for all events A, and write

$$P^0_x(A|B) := \mathbb{P}^{\varepsilon}_x[A|B,\, \mathcal{T}(\boldsymbol{B}_{R^\varepsilon}(t)) = \mathcal{T}(\boldsymbol{X}(t)) = \mathcal{T}_0]$$

for all events A and B.

Note that, under the exponentially marked voting procedure, the vote of the root $\emptyset$ in $\mathcal{T}_0$ will be one if and only if either the root is unmarked (i.e. $\widehat{S}_0 \geq S_0$) and $B_0(R^\varepsilon_0(t)) > 0$, or if the root is marked (i.e. $\widehat{S}_0 < S_0$), in which case the ancestral particle votes one with probability $\frac{1}{2}$. Noting that the event $\{\mathcal{T}(\boldsymbol{B}_{R^\varepsilon}(t)) = \mathcal{T}_0\}$ can equivalently be written as $\{S > t\}$, we have, for all $x \geq 0$,

$$\begin{aligned}
\widehat{\mathbb{P}}^t_x(\mathcal{T}_0) &= P^0_x[B_0(R^\varepsilon_0(t)) > 0,\ \widehat{S}_0 \geq S] + \tfrac{1}{2}P^0_x[\widehat{S}_0 < S] \\
&= P^0_x[B_0(R^\varepsilon_0(t)) > 0]P^0_x[\widehat{S}_0 \geq S] + \tfrac{1}{2}P_x[\widehat{S}_0 < S] \\
&\leq P^0_x[B_0(R^\varepsilon_0(t)) > 0]\mathbb{P}^{\varepsilon}_x[\widehat{S}_0 \geq t] + \tfrac{1}{2}\mathbb{P}^{\varepsilon}_x[\widehat{S}_0 < t],
\end{aligned} \tag{3.12}$$

where in the second to last line we have used that, conditionally on the event $\{S > t\}$, the events $\{B_0(R_0^\varepsilon(t)) > 0\}$ and $\{\widehat{S}_0 \geq S\}$ are independent. In the last line we have used that $\mathbb{P}_x^\varepsilon[\widehat{S}_0 \geq S | S > t] \leq \mathbb{P}_x^\varepsilon[\widehat{S}_0 \geq t]$.

Continuing the proof of the base case, consider the leaf in $\boldsymbol{X}(t)$, which, by standard results (see [Boc55]), is equal in distribution to $B_0(R_0(t))$, for $(B_0(t))_{t\geq 0}$ a standard Brownian motion and $(R_0(t))_{t\geq 0}$ an $\frac{\alpha}{2}$-stable subordinator. Define the random variable

$$S_{R_0} = \inf\{t \geq 0 : |R_0(t) - R_0(t-)| > \widehat{I}(\varepsilon)^2\},$$

which describes the first time that $R_0(t)$ makes a jump of size greater than $\widehat{I}(\varepsilon)^2$ (this is defined with the understanding that if $S_{R_0} > S$ then we can keep running the process without branching). Of course, $S_{R_0} \overset{D}{=} \widehat{S}_0$, but S_{R_0} is defined in terms of the subordinator of the ancestral particle in $\boldsymbol{X}(t)$. Noting that S_{R_0} and S are independent, we have

$$\begin{aligned}\mathbb{P}_x^t(\mathcal{T}_0) &= P_x^0[B_0(R_0(t)) > 0 | S_{R_0} \geq t]\mathbb{P}_x^\varepsilon[S_{R_0} \geq t] \\ &\quad + P_x^0[B_0(R_0(t)) > 0 | S_{R_0} < t]\mathbb{P}_x^\varepsilon[S_{R_0} < t] \\ &\geq P_x^0[B_0(R_0^\varepsilon(t)) > 0]\,\mathbb{P}_x^\varepsilon[S_{R_0} \geq t] + \tfrac{1}{2}\mathbb{P}_x^\varepsilon[S_{R_0} < t] \qquad (3.13)\end{aligned}$$

using that, conditionally on $\{S_{R_0} > t\}$, it holds that $B_0(R_0(t)) \overset{D}{=} B_0(R_0^\varepsilon(t))$, and using also that, as a function of x, the quantity $\mathbb{P}_x^\varepsilon[p(B_0(R_0(t))) > 0 | S > t,\, S_{R_0} < t]$ satisfies a similar symmetry relation to (3.5) and (3.10), and hence is greater than or equal to $\frac{1}{2}$ for all $x \geq 0$. Combining (3.12) and (3.13), and noting that $S_{R_0} \overset{D}{=} \widehat{S}_0$, we obtain

$$\mathbb{P}_x^t(\mathcal{T}_0) \geq \widehat{\mathbb{P}}_x^t(\mathcal{T}_0)$$

for $x \geq 0$, proving the base case.

Now suppose that, for all trees with at most $n-1$ branching events, (3.11) holds. Let $\mathcal{T}^n$ be a tree with n branching events. Define the first three trees of descent, denoted $\mathcal{T}_1, \mathcal{T}_2$, and $\mathcal{T}_3$, to be the three subtrees of $\mathcal{T}^n$ generated at the time of the first branch. Assume $\mathcal{T}_1$, $\mathcal{T}_2$ and $\mathcal{T}_3$ have strictly fewer than n branching events. Let S be the time of the first branching event in $\mathcal{T}^n$. Write

$$g(\mathbb{P}_{X_S}^{t-S}(\mathcal{T}\star)) := g(\mathbb{P}_{X_S}^{t-S}(\mathcal{T}_1), \mathbb{P}_{X_S}^{t-S}(\mathcal{T}_2), \mathbb{P}_{X_S}^{t-S}(\mathcal{T}_3))$$

and similarly for $g(\widehat{\mathbb{P}}^{t-S}_{X_S}(\mathcal{T}\star))$. By (3.5)

$$g(\mathbb{P}^{t-S}_{X_S}(\mathcal{T}\star)) = 1 - g(\mathbb{P}^{t-S}_{-X_S}(\mathcal{T}\star)) \text{ and } g(\widehat{\mathbb{P}}^{t-S}_{X_S}(\mathcal{T}\star)) = 1 - g(\widehat{\mathbb{P}}^{t-S}_{-X_S}(\mathcal{T}\star)). \tag{3.14}$$

Define $P^n_x(A) := \mathbb{P}^\varepsilon_x(A|\mathcal{T}(\boldsymbol{B}_{R^\varepsilon}(t)) = \mathcal{T}(\boldsymbol{X}(t)) = \mathcal{T}^n)$, with corresponding expectation E^n_x. Then, by almost identical arguments to those used to obtain (3.13), but now conditioning on the event $\{S_{R_0} > S\}$, we have

$$\begin{aligned}
\mathbb{P}^t_x(\mathcal{T}) &= E^n_x\left[g(\mathbb{P}^{t-S}_{X_S}(\mathcal{T}\star))\right] \\
&= E_x\left[g(\mathbb{P}^{t-S}_{B(R^\varepsilon_S)}(\mathcal{T}\star))\mathbb{1}_{\{S_{R_0}>S\}}\right] + E^n_x\left[g(\mathbb{P}^{t-S}_{X_S}(\mathcal{T}\star))|S_{R_0} \le S\right] P^n_x[S_{R_0} \le S] \\
&\ge E^n_x\left[g(\mathbb{P}^{t-S}_{B(R^\varepsilon_S)}(\mathcal{T}\star))\mathbb{1}_{\{S_{R_0}>S\}}\right] + \tfrac{1}{2}P^n_x[S_{R_0} \le S],
\end{aligned} \tag{3.15}$$

using that, as a function of x, $E^n_x[g(\mathbb{P}^{t-S}_{X_S}(\mathcal{T}\star))|S_{R_0} \le S]$ satisfies a similar symmetry relation to (3.10), so is greater than or equal to $\frac{1}{2}$ for $x \ge 0$. By definition of the exponentially marked voting system, for $\widehat{S}_0$ as above, we also have

$$\widehat{\mathbb{P}}^t_x(\mathcal{T}) = E^n_x\left[g(\widehat{\mathbb{P}}^{t-S}_{B(R^\varepsilon_S)}(\mathcal{T}\star))\mathbb{1}_{\{\widehat{S}_0>S\}}\right] + \tfrac{1}{2}P^n_x[\widehat{S}_0 \le S]. \tag{3.16}$$

Therefore, since $S_{R_0} \stackrel{D}{=} \widehat{S}_0$, by (3.15) and (3.16), it suffices to show that

$$E^n_x\left[g(\mathbb{P}^{t-S}_{B(R^\varepsilon_S)}(\mathcal{T}\star))\mathbb{1}_{\{S_{R_0}>S\}}\right] \ge E^n_x\left[g(\widehat{\mathbb{P}}^{t-S}_{B(R^\varepsilon_S)}(\mathcal{T}\star))\mathbb{1}_{\{S_{R_0}>S\}}\right].$$

Now, for all $x > 0$,

$$\begin{aligned}
&E^n_x\left[g(\mathbb{P}^{t-S}_{B(R^\varepsilon_S)}(\mathcal{T}\star))\mathbb{1}_{\{S_{R_0}>S\}}\right] \\
&= E^n_x\left[g(\mathbb{P}^{t-S}_{B(R^\varepsilon_S)}(\mathcal{T}\star))\mathbb{1}_{\{S_{R_0}>S\}}|B(R^\varepsilon_S) > 0\right] P^n_x\left[B(R^\varepsilon_S) > 0\right] \\
&\quad + E^n_x\left[g(\mathbb{P}^{t-S}_{B(R^\varepsilon_S)}(\mathcal{T}\star))\mathbb{1}_{\{S_{R_0}>S\}}|B(R^\varepsilon_S) \le 0\right] P^n_x[B(R^\varepsilon_S) \le 0] \\
&= E^n_x\left[g(\mathbb{P}^{t-S}_{B(R^\varepsilon_S)}(\mathcal{T}\star))\mathbb{1}_{\{S_{R_0}>S\}}|B(R^\varepsilon_S) > 0\right] P^n_x[B(R^\varepsilon_S) > 0] + P^n_x[S_{R_0} > S]P^n_x[B(R^\varepsilon_S) \le 0] \\
&\quad - E^n_x\left[g(\mathbb{P}^{t-S}_{-B(R^\varepsilon_S)}(\mathcal{T}\star))\mathbb{1}_{\{S_{R_0}>S\}}|B(R^\varepsilon_S) \le 0\right] P^n_x[B(R^\varepsilon_S) \le 0] \\
&= E^n_x\left[g(\mathbb{P}^{t-S}_{B(R^\varepsilon_S)}(\mathcal{T}\star))\mathbb{1}_{\{S_{R_0}>S\}}|B(R^\varepsilon_S) > 0\right] (P^n_x[B(R^\varepsilon_S) > 0] - P^n_x[B(R^\varepsilon_S) \le 0]) \\
&\quad + P^n_x[S_{R_0} > S]P^n_x[B(R^\varepsilon_S) \le 0] \\
&\ge E^n_x\left[g(\widehat{\mathbb{P}}^{t-S}_{B(R^\varepsilon_S)}(\mathcal{T}\star))\mathbb{1}_{\{S_{R_0}>S\}}|B(R^\varepsilon_S) > 0\right] (P^n_x[B(R^\varepsilon_S) > 0] - P^n_x[B(R^\varepsilon_S) \le 0])
\end{aligned}$$

$$
\begin{aligned}
&+ P_x^n[S_{R_0} > S] P_x^n[B(R_S^\varepsilon) \le 0] \\
&= E_x^n\left[g(\widehat{\mathbb{P}}^{t-S}_{B(R_S^\varepsilon)}(\mathcal{T}\star))\mathbb{1}_{\{S_{R_0}>S\}}\right]
\end{aligned}
$$

where, in the second line, we have applied the symmetry (3.14). Furthermore, in the second to last line, we used monotonicity of g together with our inductive hypothesis, and the fact that, given $x \ge 0$, the difference $P_x^n[B(R_S^\varepsilon) > 0] - P_x^n[B(R_S^\varepsilon) \le 0]$ is non-negative. The final equality follows by reversing the arguments used above but for $\widehat{\mathbb{P}}^{t-S}_{B(R_S^\varepsilon)}(\mathcal{T}\star)$. $\square$

Marked majority voting

In this section, we describe what will be the final voting system in one-dimension, which we denote by $\mathbb{V}^\times$, defined on $\mathcal{T}(\boldsymbol{B}_{R^\varepsilon})$. This voting system will be carried throughout the one-dimensional proof. In spirit, $\mathbb{V}^\times$ is very similar to $\widehat{\mathbb{V}}$, but no longer relies on knowing the lifetime of particles in order to mark them. Instead, particles are marked (independently) when they are born. In Theorem 3.12, it will be shown that our new voting system $\mathbb{V}^\times$ can be easily coupled to the exponential marked voting system $\widehat{\mathbb{V}}$, so by Theorem 3.10, it can also be coupled to the original majority voting system $\mathbb{V}$. Our need for the intermediate system $\widehat{\mathbb{V}}$ to obtain a coupling of $\mathbb{V}^\times$ with $\mathbb{V}$ will be explained in Remark 3.13.

Under the marked majority voting system, particles will be marked (independently) with probability b_I, defined as follows. Recall that, for $i \in M(t)$, $\widehat{S}_i \sim \mathtt{Exp}(\widehat{I}(\varepsilon)^{-2})$ is equal in distribution to the first time the subordinator $(R_i(t))_{t\ge 0}$ makes a jump of size larger than $\widehat{I}(\varepsilon)^2$. Further recall that $S_i \sim \mathtt{Exp}(\varepsilon^{-2})$ is the lifetime of the particle $X_i \overset{D}{=} B_i(R_i)$ in $\boldsymbol{X}(t)$. Define

$$b_I := \mathbb{P}[\widehat{S}_i < S_i] \sim \frac{\widehat{I}(\varepsilon)^{-2}}{\widehat{I}(\varepsilon)^{-2} + \varepsilon^{-2}} \sim \frac{\varepsilon^2}{\widehat{I}(\varepsilon)^2}, \tag{3.17}$$

where $a \sim b$ means a is proportional to b with scaling constant independent of ε. This is the probability that particle i makes a large jump before the next branch time (provided particle i is not a leaf). By Assumption 1.37 (B), $b_I \to 0$ as $\varepsilon \to 0$.

Let $\varepsilon > 0$, $t \ge 0$ and $\alpha \in (1, 2)$. For a fixed function $p : \mathbb{R}^{\mathrm{d}} \to [0, 1]$, we define a *marked majority* voting procedure on $\mathcal{T}(\boldsymbol{B}_{R^\varepsilon}(t))$ as follows.

1. At each branch point in $\mathcal{T}(\boldsymbol{B}_{R^\varepsilon}(t))$, the parent particle j marks each of its three offspring $(j,1),(j,2)$ and $(j,3)$ independently with probability b_I. Each marked particle (independently) votes 1 with probability $\frac{1}{2}$ and otherwise votes 0.

2. Each unmarked leaf i of $\mathcal{T}(\boldsymbol{B}_{R^\varepsilon}(t))$, independently, votes 1 with probability $p(B_i(R_i^\varepsilon))$, and otherwise votes 0.

3. At each branch point in $\mathcal{T}(\boldsymbol{B}_{R^\varepsilon}(t))$, if the parent particle k is unmarked, it votes according to the majority vote of its three offspring $(k,1),(k,2)$ and $(k,3)$.

Observe that the initial ancestor of $\mathcal{T}(\boldsymbol{B}_{R^\varepsilon}(t))$ is never marked under this procedure.

Definition 3.11 ($\mathbb{V}_p^\times$). With the marked majority voting procedure described above, define $\mathbb{V}_p^\times$ to be the vote associated to the root $\emptyset$ of the ternary branching subordinated Brownian tree.

As usual, for our choice of initial condition $p(x) = \mathbb{1}_{\{x\geq 0\}}$, we write $\mathbb{V}^\times := \mathbb{V}_p^\times$. With this choice of initial condition, the marked majority voting system, $\mathbb{V}^\times$, retains many of the symmetry relations exploited in [EFP17], that have already used here for $\mathbb{V}$ and $\widehat{\mathbb{V}}$. Namely, for all $x_1, x_2 \in \mathbb{R}$, with $x_1 \leq x_2$,

$$\mathbb{P}_{x_1}^\varepsilon\left[\mathbb{V}^\times(\boldsymbol{B}_{R^\varepsilon}(t)) = 1\right] \leq \mathbb{P}_{x_2}^\varepsilon\left[\mathbb{V}^\times(\boldsymbol{B}_{R^\varepsilon}(t)) = 1\right], \tag{3.18}$$

and, for any time-labelled tree $\mathcal{T}$, if we set

$$\check{\mathbb{P}}_x^t(\mathcal{T}) = \mathbb{P}_x^\varepsilon\left[\mathbb{V}^\times(\boldsymbol{B}_{R^\varepsilon}(t)) = 1 | \mathcal{T}(\boldsymbol{B}_{R^\varepsilon}(t)) = \mathcal{T}\right],$$

then by symmetry of the historical stable process and exchangeability of 0 and 1 in the marked voting procedure,

$$\check{\mathbb{P}}_x^t(\mathcal{T}) = 1 - \check{\mathbb{P}}_{-x}^t(\mathcal{T}) \tag{3.19}$$

for all time-labelled trees $\mathcal{T}$, $x \in \mathbb{R}$, and $t \geq 0$. Setting $x = 0$ in (3.19) gives $\check{\mathbb{P}}_0^t(\mathcal{T}) = \frac{1}{2}$, so, by monotonicity (3.18), for any time-labelled ternary tree $\mathcal{T}$,

$$\check{\mathbb{P}}_x^t(\mathcal{T}) \geq \frac{1}{2} \;\text{ for } x > 0 \;\text{ and }\; \check{\mathbb{P}}_x^t(\mathcal{T}) \leq \frac{1}{2} \;\text{ for } x < 0. \tag{3.20}$$

We next introduce notation for our marked majority voting procedure. Recall that g is the majority voting function associated to $\mathbb{V}$. Define the *marked majority voting function* $g_\times : [0,1]^3 \to [0,1]$ by

$$g_\times(p_1, p_2, p_3) := g\left((1-b_I)p_1 + \tfrac{b_I}{2}, (1-b_I)p_2 + \tfrac{b_I}{2}, (1-b_I)p_3 + \tfrac{b_I}{2}\right).$$

This is the probability that an unmarked parent particle votes 1 under $\mathbb{V}^\times$, in the special case when the three offspring are independent and have probabilities p_1, p_2 and p_3 of voting 1 if they are unmarked. As before, we abuse notation and write $g_\times(p) := g_\times(p,p,p)$.

Let $\{u_-, \frac{1}{2}, u_+\}$ be the three solutions to $g_\times(p) = p$, satisfying $0 < u_- < \frac{1}{2} < u_+ < 1$. The exact derivation of the fixed points can be found in Proposition B.1 in Appendix B. It will be useful later to note that these fixed points can be approximated by Taylor expansion as

$$u_- = \frac{1}{2} - \frac{\sqrt{(1-b_I)^3(1-3b_I)}}{2(1-b_I)^3} \sim \frac{3}{4}b_I^2 + \mathcal{O}(b_I^3) \tag{3.21}$$

$$u_+ = \frac{1}{2} + \frac{\sqrt{(1-b_I)^3(1-3b_I)}}{2(1-b_I)^3} \sim 1 - \frac{3}{4}b_I^2 + \mathcal{O}(b_I^3). \tag{3.22}$$

We now return to the coupling of $\mathbb{V}_p^\times$ and the original voting system $\mathbb{V}_p$ via the intermediate system $\widehat{\mathbb{V}}_p$.

Theorem 3.12. *Let $\widehat{\mathbb{V}}$ the exponentially marked majority voting system and $\mathbb{V}^\times$ be as above. Then, for all $x \in \mathbb{R}$ and $t \geq 0$,*

$$\mathbb{P}_x^\varepsilon[\widehat{\mathbb{V}}(\boldsymbol{B}_{R^\varepsilon}(t)) = 1] = \mathbb{P}_x^\varepsilon[\mathbb{V}^\times(\boldsymbol{B}_{R^\varepsilon}(t)) = 1](1-b_I) + \frac{b_I}{2}. \tag{3.23}$$

Proof. Recall that, under the voting system $\widehat{\mathbb{V}}$, each particle i is marked if $\widehat{S}_i < S_i$. By definition of b_I, all particles in $\mathcal{T}(\boldsymbol{B}_{R^\varepsilon}(t))$ are marked with the same probability under both $\widehat{\mathbb{V}}$ and $\mathbb{V}^\times$, except for the ancestral particle, which remains unmarked under $\mathbb{V}_p^\times$ by definition. Conditioning on the marking of the ancestral particle in $\widehat{\mathbb{V}}$, we obtain

$$\begin{aligned}
&\mathbb{P}_x^\varepsilon[\widehat{\mathbb{V}}_p(\boldsymbol{B}_{R^\varepsilon}(t)) = 1] \\
&= \mathbb{P}_x^\varepsilon[\widehat{\mathbb{V}}_p(\boldsymbol{B}_{R^\varepsilon}(t)) = 1 | \widehat{S}_0 > S_0]\mathbb{P}_x^\varepsilon[\widehat{S}_0 > S_0] + \mathbb{P}_x^\varepsilon[\widehat{\mathbb{V}}_p(\boldsymbol{B}_{R^\varepsilon}(t)) = 1 | \widehat{S}_0 \leq S_0]\mathbb{P}_x^\varepsilon[\widehat{S}_0 \leq S_0] \\
&= \mathbb{P}_x^\varepsilon[\mathbb{V}_p^\times(\boldsymbol{B}_{R^\varepsilon}(t)) = 1](1-b_I) + \frac{b_I}{2},
\end{aligned}$$

where the last line follows from the definition of b_I. $\square$

Remark 3.13. The intermediate voting system $\widehat{\mathbb{V}}$ is crucial to obtain a coupling between the marked system $\mathbb{V}^{\times}$ and the original voting system $\mathbb{V}$. The marking of a particle in $\mathcal{T}(\boldsymbol{B}_{R^\varepsilon}(t))$ under $\widehat{\mathbb{V}}$ corresponds precisely to the event that its partner particle in $\mathcal{T}(\boldsymbol{X}(t))$ made a large jump, allowing for easy comparison of the two systems. However, under $\mathbb{V}^{\times}$, we do not have this direct comparison. Instead, we only know that the probability of marking in $\mathcal{T}(\boldsymbol{B}_{R^\varepsilon}(t))$ equals the probability of seeing large jumps in $\mathcal{T}(\boldsymbol{X}(t))$.

Finally to prove our main one-dimensional result, Theorem 3.6, it suffices, by Theorem 3.10 and Theorem 3.12, to show the following.

Theorem 3.14. *Let $\alpha \in (1,2)$, $k \in \mathbb{N}$ and $T^* \in (0,\infty)$. Suppose $I : \mathbb{R}_+ \to \mathbb{R}_+$ satisfies Assumptions 1.37 (A) and 1.37 (B). Let u_+, u_- be as in (3.21) and (3.22). Then there exist $c_1(\alpha, I, k), \varepsilon_1(\alpha, I, k) > 0$ such that, for all $t \in [0, T^*]$ and all $\varepsilon \in (0, \varepsilon_1(k))$,*

1. *for $x \geq c_1(k)\widehat{I}(\varepsilon)|\log\varepsilon|$, we have $\mathbb{P}^\varepsilon_x[\mathbb{V}^{\times}(\boldsymbol{B}_{R^\varepsilon}(t)) = 1] \geq u_+ - \varepsilon^k$;*
2. *for $x \leq -c_1(k)\widehat{I}(\varepsilon)|\log\varepsilon|$, we have $\mathbb{P}^\varepsilon_x[\mathbb{V}^{\times}(\boldsymbol{B}_{R^\varepsilon}(t)) = 1] \leq u_- + \varepsilon^k$.*

3.2.2 Proof of Theorem 3.14

We now prove Theorem 3.14, thereby proving the main one-dimensional result, Theorem 3.6. Throughout this section, let $p(x) = \mathbb{1}_{\{x \geq 0\}}$. Our next result verifies that the marked majority voting procedure cannot reduce the positive voting bias on the leaves when the root, x, is non-negative. By the symmetry in (3.19), the same result holds for the negative voting bias on the leaves when $x < 0$. Given this symmetry, we often state results only for positive x, when it is convenient to do so.

Lemma 3.15. *For any time-labelled ternary tree $\mathcal{T}$, any time $t > 0$, and any $x \geq 0$,*

$$\check{\mathbb{P}}^t_x(\mathcal{T}) \geq u_+ \mathbb{P}_x[B(R^\varepsilon_t) \geq 0] + u_- \mathbb{P}_x[B(R^\varepsilon_t) \leq 0].$$

The proof of Lemma 3.15 follows exactly as for [HD20, Lemma 3.1] with $g_\times$ instead of g therein, and is consequently omitted.

Remark 3.16. In [EFP17], the majority voting function g was considered throughout; this function enjoys the fixed points $0, \frac{1}{2}$, and 1. The analogous statement to Lemma 3.15 for g and a Brownian spatial motion was then stated with $u_+ = 1$ and $u_- = 0$. In our setting, the non-zero attractive fixed point u_- must be taken into account when leaves have a negative bias. To do this, we adapt ideas from [HD20], where more general voting functions than g were considered.

Remark 3.17. At this point, we can briefly remark upon why the symmetric marked system $\mathbb{V}_p^\times$ was chosen. One might think, as a first attempt, that asking all marked individuals to vote 0 with probability 1 (instead of 1/2) would simplify the coupling (3.7). While this is true, it is not obvious that the resulting system does not reduce the positive voting bias on leaves when the root x is non-negative, and an analogue of Lemma 3.15 would be challenging to obtain due to the lack of symmetry of the voting system.

We next aim to show that the iterative voting procedure amplifies a small positive bias at the leaves to a large positive voting bias at the root. To do this, we define $g_\times^{(n)}(p)$ inductively by

$$g_\times^{(1)}(p) = g_\times(p), \ g_\times^{(n+1)}(p) = g_\times^{(n)}(g_\times(p)).$$

Noting that the ancestral particle is never marked under $\mathbb{V}^\times$, we see that $g_\times^{(n)}(p)$ is the probability of voting 1 at the root of an n-level regular ternary tree under $\mathbb{V}^\times$ if the votes of the *unmarked* leaves are i.i.d. Bernoulli(p). We next consider the rate of convergence of $g_\times$ to its fixed points. Let $\widehat{I}(\varepsilon)$ be our scaling function, which we take to satisfy Assumptions 1.37 throughout.

Lemma 3.18. *For all $k \in \mathbb{N}$ there exist $A(k) < \infty$ and $\varepsilon_1(k) > 0$ such that, for all $\varepsilon \in (0, \varepsilon_1)$ and $n \geq A(k)|\log(\varepsilon)|$,*

$$g_\times^{(n)}\left(\frac{1}{2} + \varepsilon\right) \geq u_+ - \varepsilon^k \quad \text{and} \quad g_\times^{(n)}\left(\frac{1}{2} - \varepsilon\right) \leq u_- + \varepsilon^k.$$

Proof. We follow closely [HD20, Lemma 3.2], and note some slight differences arising due to the dependence of $g_\times$ on b_I (and, consequently, on ε and $\widehat{I}(\varepsilon)$), recalling that

$$g_\times(p) = g((1 - b_I)p + \tfrac{b_I}{2}),$$

where $b_I = \mathcal{O}(\frac{\varepsilon^2}{\widetilde{I}(\varepsilon)^2})$. We prove only the first inequality since the second follows by completely symmetric arguments. First, we show that there exists $C_k > 0$ and some fixed $q_0 > 0$ such that, after $n \geq C_k|\log\varepsilon|$ iterations,

$$g_\times^{(n)}(u_+ - x - q) \geq u_+ - \varepsilon^k \tag{3.24}$$

for all $q \geq q_0$. We then proceed to show that there exists $D > 0$ such that, after $n \geq D|\log\varepsilon|$ iterations,

$$g_\times^{(n)}\left(\frac{1}{2} + \varepsilon\right) \geq u_+ - k - q_0. \tag{3.25}$$

Combining (3.24) and (3.25) then gives the result.

To prove (3.24), choose ε_1 sufficiently small so that $b_I \leq \frac{1}{6}$ for all $\varepsilon \in (0, \varepsilon_1)$. Then

$$g_\times'\left(\frac{1}{2}\right) = (1 - b_I)g'\left(\frac{1}{2}\right) = \frac{3}{2}(1 - b_I) \geq \frac{5}{4} > 1. \tag{3.26}$$

Next, using that $g'(0) = 0$ and that g' is continuous, together with the estimate (3.21), we have, for ε_1 sufficiently small, that

$$g'\left(u_-(1 - b_I) + \frac{b_I}{2}\right) < \frac{1}{4}$$

for all $\varepsilon \in (0, \varepsilon_1)$. It follows, for this choice of ε_1, that $g_\times'(u_-) = (1-b_I)g'(u_-(1-b_I)+b_I/2) < \frac{1}{4}$. Since $g_\times'\left(\frac{1}{2}\right) > 1$ and $g_\times'$ is continuous,

$$q_0 := \inf\left\{q \geq 0 : g_\times'(u_- + q) \geq \tfrac{1}{2}\right\} > 0.$$

By the mean value theorem and the definition of q_0, for all $q < k_0$,

$$u_+ - g_\times(u_+ - q) = g_\times(u_- + q) - u_- \leq q\, g_\times'(u_- + q_0) = \frac{q}{2}.$$

Iterating this yields

$$u_+ - g_\times^{(n)}(u_+ - q) \leq \frac{1}{2^n}(u_+ - q) \leq \frac{1}{2^n}.$$

It follows that there exists $C_k > 0$ such that, if $n \geq C|\log(\varepsilon)|$ and $q \leq q_0$, then

$$g_\times^{(n)}(u_+ - q) \geq u_+ - \varepsilon^k,$$

thereby proving (3.24).

We now proceed to prove (3.25). As before (see equation 3.26), let ε_1 be sufficiently small that, for all $\varepsilon \in (0, \varepsilon_1)$, we have $g'_\times(\frac{1}{2}) > 1$. Since $g_\times$ is increasing and u_-, $\frac{1}{2}$ and u_+ are the only fixed points of $g_\times$, we have

$$g_\times(q) > q \text{ for all } q \in \left(\tfrac{1}{2}, u_+\right). \tag{3.27}$$

By the definition of q_0, and since $g'_\times$ is increasing on $(0, \frac{1}{2})$,

$$u_+ - q_0 - \tfrac{1}{2} = \tfrac{1}{2} - q_0 - u_- > 0.$$

Therefore

$$q_1 := \inf_{0 \le q \le u_+ - q_0 - \frac{1}{2}} \frac{g_\times(\frac{1}{2} + q) - (\frac{1}{2} + q)}{q} > 0,$$

and so, for $q \in [\varepsilon, u_+ - q_0 - \frac{1}{2}]$, by the definition of q_0 we have

$$g_\times\left(\tfrac{1}{2} + q\right) - \tfrac{1}{2} > (1 + k_1)q. \tag{3.28}$$

Reduce ε_1 if necessary so that $g_\times(\frac{1}{2} + \varepsilon) < u_+ - q_0$ for all $\varepsilon \in (0, \varepsilon_1)$. Then, applying (3.28) twice, we obtain

$$\begin{aligned} g^{(2)}_\times(\tfrac{1}{2} + \varepsilon) &= g_\times(\tfrac{1}{2} + (g_\times(\tfrac{1}{2} + \varepsilon) - \tfrac{1}{2})) \\ &\ge (1 + q_1)[g_\times(\tfrac{1}{2} + \varepsilon) - \tfrac{1}{2})] + \tfrac{1}{2} \\ &\ge (1 + q_1)^2 \varepsilon + \tfrac{1}{2}. \end{aligned}$$

Repeating this argument $n - 1$ times, we obtain

$$g^{(n)}_\times\left(\tfrac{1}{2} + \varepsilon\right) \ge (1 + q_1)^n \varepsilon + \tfrac{1}{2}.$$

It follows, for $D := \frac{1}{\log(1+q_1)}$, that after $n > D|\log(\varepsilon)|$ iterations,

$$g^{(n)}_\times\left(\tfrac{1}{2} + \varepsilon\right) \ge u_+ - q_0.$$

Setting $A = C_k + D$ proves the result. $\square$

We now briefly note a useful inequality for $g_\times$, that will be used in the proof of Theorem 3.14.

Lemma 3.19. *If $p_1, p_2, p_3 \geq \frac{1}{2}$, then*

$$g_\times(p_1, p_2, p_3) \geq \min(p_1, p_2, p_3, u_+).$$

If $p_1, p_2, p_3 \leq \frac{1}{2}$, then

$$g_\times(p_1, p_2, p_3) \leq \max(p_1, p_2, p_3, u_-).$$

Proof. We prove only the first inequality, since the second one follows by a symmetrical argument. Let us denote

$$p_{\min} = \min\{p_1, p_2, p_3, u_+\}.$$

If $p_1, p_2, p_3 \geq \frac{1}{2}$, it follows that $\frac{1}{2} \leq p_{\min} \leq u_+$. Since $g_\times$ is increasing in each variable,

$$g_\times(p_1, p_2, p_3) \geq g_\times(p_{\min}).$$

Therefore it suffices to show that $g_\times(p_{\min}) \geq p_{\min}$. For this, recall that u_-, $\frac{1}{2}$ and u_+ are the only fixed points of $g_\times$. Factorising $g_\times(p_{\min}) - p_{\min}$ yields

$$g_\times(p_{\min}) - p_{\min} = (p_{\min} - u_-)(2p_{\min} - 1)(u_+ - p_{\min}).$$

Since $\frac{1}{2} \leq p_{\min} \leq u_+$ we have $g_\times(p_{\min}) - p_{\min} \geq 0$, as required. $\square$

We next state a result which says there is a large enough regular ternary subtree of $\mathcal{T}(\boldsymbol{B}_{R^\varepsilon}(t))$ generated by time $t \geq a_1\varepsilon^2|\log\varepsilon|$ for Lemma 3.18 to apply. 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 $l \geq 0$, let $\mathcal{T}_l^{reg} = \mathcal{T}_{\lceil l \rceil}^{reg}$. For $\mathcal{T}$ a time-labelled ternary tree, we use $\mathcal{T}_l^{reg} \subseteq \mathcal{T}$ to mean that, as subtrees of $\mathcal{U}$, the tree $\mathcal{T}_l^{reg}$ is contained inside $\mathcal{T}$ (ignoring time labels).

Lemma 3.20. *Let $k \in \mathbb{N}$ and $A(k)$ be as in Lemma 3.18. There exist $a_1(\alpha, k) > 0$ and $\varepsilon_1(\alpha, k) > 0$ such that, for all $\varepsilon \in (0, \varepsilon_1)$ and $t \geq a_1(k)\varepsilon^2|\log\varepsilon|$,*

$$\mathbb{P}^\varepsilon\left[\mathcal{T}\left(\boldsymbol{B}_{R^\varepsilon}(t)\right) \supseteq \mathcal{T}_{A(k)|\log\varepsilon|}^{reg}\right] \geq 1 - \varepsilon^k. \tag{3.29}$$

The proof of (3.29) proceeds exactly as in [EFP17, Lemma 2.10], and is therefore omitted. We next aim to control the displacement of leaves in $\boldsymbol{B}_{R^\varepsilon}(t)$.

Lemma 3.21. *Fix $k \in \mathbb{N}$ and let $a_1(k)$ be as in Lemma 3.20. Then there exist $\varepsilon_1(\alpha, k) > 0$ and $l_1(\alpha, k) > 0$ such that, for all $x \in \mathbb{R}^{\mathrm{d}}$, $\varepsilon \in (0, \varepsilon_1)$ and $s \leq a_1(k)\varepsilon^2|\log\varepsilon|$,*

$$\mathbb{P}_x\left(\exists i \in N(s) : |B_i(R_i^\varepsilon(s)) - x| \geq l_1(k)\widehat{I}(\varepsilon)|\log\varepsilon|\right) \leq 2\varepsilon^k.$$

Remark 3.22. Lemma 3.21 highlights the importance of working with subordinated Brownian motions instead of the original stable processes. As we shall see in the proof of Lemma 3.21, to control leaves of the ternary branching stable tree, we will apply a many-to-one lemma. Using the same approach for stable processes (without truncation), the region in which we can control the leaves with probability ε^k grows without bound as ε goes to zero, which is clearly not strong enough for our goals.

Proof of Lemma 3.21. First note that for any $l > 0$

$$\begin{aligned}
\mathbb{P}_x\left[\exists i \in N(s) : |B(R_s^\varepsilon) - x| \geq l\widehat{I}(\varepsilon)|\log\varepsilon|\right] &\leq \mathbb{E}[N(s)]\mathbb{P}_0\left[|B(R_s^\varepsilon)| \geq l\widehat{I}(\varepsilon)|\log\varepsilon|\right] \\
&= e^{\frac{2s}{\varepsilon^2}}\mathbb{P}_0\left[|B(R_s^\varepsilon)| \geq l\widehat{I}(\varepsilon)|\log\varepsilon|\right] \\
&\leq \varepsilon^{-2a_1}\mathbb{P}_0\left[|B(R_s^\varepsilon)| \geq l\widehat{I}(\varepsilon)|\log\varepsilon|\right].
\end{aligned} \tag{3.30}$$

Denote the density of R_s^ε by f_ε. Then for any $h \geq 0$, partitioning over $\{t \leq h\widehat{I}(\varepsilon)^2|\log\varepsilon|\}$ allows us to obtain

$$\begin{aligned}
\mathbb{P}_0\left[|B(R_s^\varepsilon)| \geq l\widehat{I}(\varepsilon)|\log\varepsilon|\right] &\leq \int_0^{h\widehat{I}(\varepsilon)^2|\log\varepsilon|}\mathbb{P}_0\left[|B(t)| \geq l\widehat{I}(\varepsilon)|\log\varepsilon| \mid R_s^\varepsilon = t\right] f_\varepsilon(t)dt \\
&\quad + \mathbb{P}_0[B(R_s^\varepsilon) \geq h\widehat{I}(\varepsilon)^2|\log\varepsilon|] \\
&\leq \sup_{0 \leq t \leq h\widehat{I}(\varepsilon)^2|\log\varepsilon|}\mathbb{P}_0[|B(t)| \geq l\widehat{I}(\varepsilon)|\log\varepsilon|] \\
&\quad + \mathbb{P}_0[R_s^\varepsilon \geq h\widehat{I}(\varepsilon)^2|\log\varepsilon|].
\end{aligned}$$

We bound each term separately. First, by a Chernoff bound, for all $t \leq h\widehat{I}(\varepsilon)^2|\log\varepsilon|$,

$$\begin{aligned}\mathbb{P}_0\left[|B(t)| \geq l\widehat{I}(\varepsilon)|\log\varepsilon|\right] &= \mathbb{P}_0\left[\sqrt{2t}|Z| \geq l\widehat{I}(\varepsilon)|\log\varepsilon|\right]\\ &\leq \mathbb{P}_0\left[\sqrt{2h}|Z| \geq l|\log\varepsilon|^{1/2}\right]\\ &\leq \exp\left(-\tfrac{1}{4}\tfrac{l^2}{h}|\log\varepsilon|\right)\\ &= \varepsilon^{\frac{l^2}{4h}}.\end{aligned}$$

Now fix $h := k + 2a_1(k) + 1$. By Lemma B.3 in the appendix, there exists $\varepsilon_1(k) > 0$ such that, for all $\varepsilon \in (0, \varepsilon_1)$,

$$\mathbb{P}\left[R_s^\varepsilon \geq h\widehat{I}(\varepsilon)^2|\log\varepsilon|\right] \leq \varepsilon^{k+2a_1(k)}.$$

Putting this together, we obtain

$$\mathbb{P}\left[\exists i \in N(s) : |B_i(R_i^\varepsilon(s)) - x| \geq l\widehat{I}(\varepsilon)|\log\varepsilon|\right] \leq \varepsilon^k + \varepsilon^{\frac{l^2}{4h} - 2a_1(k)}$$

so the result holds by choosing $l = l_1(k)$ sufficiently large. □

With this we can prove Theorem 3.14, following a similar format to the proof of [EFP17, Theorem 2.6]. We suppose $z \geq 2l_1(k)\widehat{I}(\varepsilon)|\log\varepsilon|$ and $s^* := a_1(k)\varepsilon^2|\log\varepsilon|$, and consider the cases $t \leq s^*$ and $t \geq s^*$ separately. Firstly, if $t \leq s^*$, then, by Lemma 3.21, with high probability none of the particles in $\mathcal{T}(\boldsymbol{X}(t))$ have moved a distance further than $l_1(k)\widehat{I}(\varepsilon)|\log\varepsilon|$, and the result follows easily. Secondly, when $t \geq s^*$, we use Lemma 3.15 to show that the leaves of $\boldsymbol{B}_{R^\varepsilon}(t)$ have a positive voting bias, which, by Lemma 3.20, is magnified by $\mathcal{O}(|\log\varepsilon|)$ rounds of voting. Thus Lemma 3.18 applies and gives the result.

Proof of Theorem 3.14. The outline of this proof follows closely that of [EFP17, Theorem 2.6], with some computational differences due to our different choice of spatial motion and worse interface tolerance. Let $\varepsilon < \frac{1}{2}$ and define z_ε implicitly by the relation

$$\mathbb{P}_{z_\varepsilon}[B(R_{T^*}^\varepsilon) \geq 0] = \tfrac{1}{2} + (u_+ - u_-)^{-1}\varepsilon, \tag{3.31}$$

and note that $z_\varepsilon \sim \varepsilon\sqrt{4\pi T^*}$ as $\varepsilon \to 0$. Moreover, with arbitrarily high probability, $R_{T^*}^\varepsilon$ is bounded above by a constant (depending on T^*). Comparing the above estimate with the definition of z_ε, there exists a constant $C(T^*)$ such that $z_\varepsilon \leq C(T^*)\varepsilon$ as $\varepsilon \to 0$.

Suppose $\varepsilon_1(k) < \frac{1}{2}$ is sufficiently small that Lemmas 3.20 and 3.21 hold for all $\varepsilon \in (0, \varepsilon_1)$. Let $c_1(k) = 2l_1(k)$, for $l_1(k)$ as in Lemma 3.21. Then, since $\varepsilon \widehat{I}(\varepsilon)^{-1} \to 0$ as $\varepsilon \to 0$ by Assumption 1.37 (B), we can choose ε_1 sufficiently small that

$$l_1(k)\widehat{I}(\varepsilon)|\log\varepsilon| + z_\varepsilon \le c_1(k)\widehat{I}(\varepsilon)|\log\varepsilon|$$

for all $\varepsilon \in (0, \varepsilon_1)$. Let a_1 be as in Lemma 3.20 and define

$$s^*(\varepsilon, \alpha) = a_1(k)\varepsilon^2|\log\varepsilon|.$$

Let $t \in (0, s^*)$, let $z \ge c_1(k)\widehat{I}(\varepsilon)|\log\varepsilon|$, and, by Lemma 3.20, let $n \ge 1$ be sufficiently large that $\mathcal{T}(\boldsymbol{B}_{R^\varepsilon}(t)) \subseteq \mathcal{T}_n^{reg}$ with probability at least $1 - \varepsilon^k$. Then

$$\begin{aligned}
&\mathbb{P}_z(\mathbb{V}_p^\times(\boldsymbol{B}_{R^\varepsilon}(t)) = 0)\\
&\quad = \mathbb{P}_z\left(\{\mathbb{V}_p^\times(\boldsymbol{B}_{R^\varepsilon}(t)) = 0\} \cap \{\exists i \in N(t) : |B_i(R_i^\varepsilon(t)) - z| \ge c_1\widehat{I}(\varepsilon)|\log\varepsilon|\}\right)\\
&\qquad + \mathbb{P}_z\left(\{\mathbb{V}_p^\times(\boldsymbol{B}_{R^\varepsilon}(t)) = 0\} \cap \{\nexists i \in N(t) : |B_i(R_i^\varepsilon(t)) - z| \ge c_1\widehat{I}(\varepsilon)|\log\varepsilon|\}\right)\\
&\quad \le \varepsilon^k + 1 - g_\times^{(n)}(1)\\
&\quad = \varepsilon^k + g_\times^{(n)}(0)\\
&\quad \le \varepsilon^k + g_\times^{(n)}(u_-)\\
&\quad = \varepsilon^k + u_-,
\end{aligned}$$

where we have used that $g_\times^{(n)}(0) \le g_\times^{(n)}(u_-) = u_-$, since u_- is a fixed point of $g_\times$. Now suppose $t \in [s^*, T^*]$ and $z \ge c_1(k)\widehat{I}(\varepsilon)|\log\varepsilon|$. Let $\mathcal{T}_{s^*} = \mathcal{T}(\boldsymbol{B}_{R^\varepsilon}(s^*))$ be the time-labelled tree of the branching stable process at time s^*. Define

$$q_{t-s^*}(z) = \mathbb{P}_z[\mathbb{V}^\times(\boldsymbol{B}_{R^\varepsilon}(t - s^*)) = 1]$$

for all $z \in \mathbb{R}$. Write $\{\boldsymbol{B}_{R^\varepsilon}(s^*) > z_\varepsilon\}$ for the event $B_i(R_i^\varepsilon(s^*)) > z_\varepsilon$ for all $i \in N(s^*)$. Then

$$\begin{aligned}
\mathbb{P}_z[\mathbb{V}_p^\times(\boldsymbol{B}_{R^\varepsilon}(t)) = 1] &= \mathbb{P}_z\left[\mathbb{V}_{q_{t-s^*}(\cdot)}^\times(\boldsymbol{B}_{R^\varepsilon}(s^*)) = 1\right]\\
&\ge \mathbb{P}_z\left[\left\{\mathbb{V}_{q_{t-s^*}(z_\varepsilon)}^\times(\boldsymbol{B}_{R^\varepsilon}(s^*)) = 1\right\} \cap \{\boldsymbol{B}_{R^\varepsilon}(s^*) > z_\varepsilon\}\right]\\
&\ge \mathbb{P}_z\left[\mathbb{V}_{q_{t-s^*}(z_\varepsilon)}^\times(\boldsymbol{B}_{R^\varepsilon}(s^*)) = 1\right] - \varepsilon^k, \qquad (3.32)
\end{aligned}$$

where the first line follows by the Markov property of $\boldsymbol{B}_{R^\varepsilon}$ at time s^*, the second line follows by monotonicity (3.18), and the third line follows by Lemma 3.21 and our assumption $z \geq c_1\widehat{I}(\varepsilon)|\log\varepsilon|$.

Now, by Lemma 3.15 and (3.31) the definition of z_ε (noting that $t - s^* \leq T^*$), we have

$$\begin{aligned} q_{t-s^*}(z_\varepsilon) &\geq \mathbb{P}_{z_\varepsilon}[B(R^\varepsilon_{t-s^*}) \geq 0]u_+ + \mathbb{P}_{z_\varepsilon}[B(R^\varepsilon_{t-s^*}) \leq 0]u_- \\ &\geq u_+\left(\tfrac{1}{2} + (u_+ - u_0)^{-1}\varepsilon\right) + u_-\left(\tfrac{1}{2} - (u_+ - u_0)^{-1}\varepsilon\right) \\ &= \frac{1}{2} + \varepsilon. \end{aligned} \tag{3.33}$$

Substituting (3.33) into (3.32), we obtain

$$\mathbb{P}_z[\mathbb{V}_p^\times(\boldsymbol{B}_{R^\varepsilon}(t)) = 1] \geq \mathbb{P}_z\left[\mathbb{V}^\times_{\frac{1}{2}+\varepsilon}(\boldsymbol{B}_{R^\varepsilon}(s^*)) = 1\right] - \varepsilon^k. \tag{3.34}$$

Note that, if $p_i \geq \frac{1}{2} + \varepsilon$ for $i = 1,2,3$, then $g_\times(p_1,p_2,p_3) \geq \min(p_1,p_2,p_3,u_+)$ from Lemma 3.19. Therefore, if each leaf of $\mathcal{T}_{s^*}$ votes one independently with probability greater than $\frac{1}{2}+\varepsilon$, and $\mathcal{T}_{s^*} \supseteq \mathcal{T}^{reg}_{A|\log\varepsilon|}$, then each leaf in $\mathcal{T}^{reg}_{A|\log\varepsilon|}$ also votes one with probability greater than $\frac{1}{2}+\varepsilon$. By Lemma 3.20, $\mathcal{T}_{s^*} \supseteq \mathcal{T}^{reg}_{A|\log\varepsilon|}$ with probability at least $1-\varepsilon^k$, so by Lemma 3.18

$$\begin{aligned} \mathbb{P}_z\left[\mathbb{V}^\times_{\frac{1}{2}+\varepsilon}(\boldsymbol{B}_{R^\varepsilon}(s^*)) = 1\right] &\geq (1-\varepsilon^k)g_\times^{(\lceil A|\log\varepsilon|\rceil)}\left(\frac{1}{2}+\varepsilon\right) \\ &\geq (1-\varepsilon^k)(u_+ - \varepsilon^k) \\ &\geq u_+ - 2\varepsilon^k. \end{aligned}$$

Substituting this into (3.34) yields

$$\mathbb{P}_z[\mathbb{V}_p^\times(\boldsymbol{B}_{R^\varepsilon}(t)) = 1] \geq u_+ - 3\varepsilon^k,$$

thereby proving part (1) of Theorem 3.14. Part (2) of Theorem 3.14 follows by completely symmetric arguments. □

3.2.3 Slope of the interface

Just as in [EFP17], to prove Theorem 3.6 we make use of a lower bound on the 'slope' of the interface in dimension $\mathbb{d} = 1$. For a time-labelled ternary tree $\mathcal{T}$, recall that

$$\check{\mathbb{P}}^t_x(\mathcal{T}) := \mathbb{P}^\varepsilon_x[\mathbb{V}^\times(\boldsymbol{B}_{R^\varepsilon}(t)) = 1 | \mathcal{T}(\boldsymbol{B}_{R^\varepsilon}(t)) = \mathcal{T}].$$

Proposition 3.23. *Suppose $x \geq 0$ and $\eta > 0$. Then, for any time-labelled ternary tree $\mathcal{T}$ and any time t,*

$$\check{\mathbb{P}}^t_x(\mathcal{T}) - \check{\mathbb{P}}^t_{x-\eta}(\mathcal{T}) \geq \check{\mathbb{P}}^t_{x+\eta}(\mathcal{T}) - \check{\mathbb{P}}^t_x(\mathcal{T}). \tag{3.35}$$

Proof. This proof follows [EFP17] with some minor differences, because of our different choice of voting function $g_\times$. We proceed by induction on the number of branching events. Let $\mathcal{T}_0$ denote a time-labelled tree with a root and a single leaf. Recall that, under $\mathbb{V}^\times$, the initial ancestor is never marked. Therefore, since $x \geq 0$,

$$\check{\mathbb{P}}^t_x(\mathcal{T}_0) - \check{\mathbb{P}}^t_{x-\eta}(\mathcal{T}_0) = \int_{x-\eta}^{x} p_{0,t}(z)dz \geq \int_{x}^{x+\eta} p_{0,t}(z)dz = \check{\mathbb{P}}^t_{x+\eta}(\mathcal{T}_0) - \check{\mathbb{P}}^t_x(\mathcal{T}_0),$$

where $p_{0,t}$ denotes the transition density of $B(R^\varepsilon_t)$ started at 0. Now assume the inequality holds for all time-labelled ternary trees with at most n branch points. Let $\mathcal{T}$ be a time-labelled ternary tree with $n+1$ internal vertices, and denote the time of the first branching event in $\mathcal{T}$ by S. Let $\mathcal{T}_1, \mathcal{T}_2$ and $\mathcal{T}_3$ denote the three trees of descent from the first branching event in $\mathcal{T}$. Then

$$\begin{aligned}
\check{\mathbb{P}}^t_x(\mathcal{T}_0) - \check{\mathbb{P}}^t_{x-\eta}(\mathcal{T}_0) &= E_x\left[g_\times\left(\check{\mathbb{P}}^{t-S}_{B(R^\varepsilon_S)}(\mathcal{T}\star)\right)\right] - E_{x-\eta}\left[g_\times\left(\check{\mathbb{P}}^{t-S}_{B(R^\varepsilon_S)}(\mathcal{T}\star)\right)\right] \\
&= \int_{-\infty}^{\infty}\left[g_\times\left(\check{\mathbb{P}}^{t-S}_y(\mathcal{T}\star)\right) - g_\times\left(\check{\mathbb{P}}^{t-S}_{y-\eta}(\mathcal{T}\star)\right)\right] p_{x,S}(y)dy \\
&= \int_0^{\infty}\left[g_\times(\check{\mathbb{P}}^{t-S}_y(\mathcal{T}\star)) - g_\times\left(\check{\mathbb{P}}^{t-S}_{y-\eta}(\mathcal{T}\star)\right)\right](p_{x,S}(y) - p_{x,S}(-y))dy,
\end{aligned}$$

where the final line follows from the symmetry relation (3.19).

By identical arguments applied to the right hand side of (3.35), we obtain

$$\begin{aligned}
&\left(\check{\mathbb{P}}^t_x(\mathcal{T}_0) - \check{\mathbb{P}}^t_{x-\eta}(\mathcal{T}_0)\right) - \left(\check{\mathbb{P}}^t_{x+\eta}(\mathcal{T}_0) - \check{\mathbb{P}}^t_{x-\eta}(\mathcal{T}_0)\right) \\
&= \int_0^{\infty}\left(g_\times\left(\check{\mathbb{P}}^{t-S}_y(\mathcal{T}\star)\right) - g_\times\left(\check{\mathbb{P}}^{t-S}_{y-\eta}(\mathcal{T}\star)\right)\right)(p_{x,S}(y) - p_{x,S}(-y))dy \\
&\quad - \int_0^{\infty}\left(g_\times\left(\check{\mathbb{P}}^{t-S}_{y+\eta}(\mathcal{T}\star)\right) - g_\times\left(\check{\mathbb{P}}^{t-S}_y(\mathcal{T}\star)\right)\right)(p_{x,S}(y) - p_{x,S}(-y))dy.
\end{aligned}$$

Since $x \geq 0$, we have $p_{x,S}(y) - p_{x,S}(-y) \geq 0$ for $y \geq 0$, and it suffices to show that, for $y \geq 0$,

$$\left(g_\times\left(\check{\mathbb{P}}^{t-S}_y(\mathcal{T}\star)\right) - g_\times\left(\check{\mathbb{P}}^{t-S}_{y-\eta}(\mathcal{T}\star)\right)\right) - \left(g_\times\left(\check{\mathbb{P}}^{t-S}_{y+\eta}(\mathcal{T}\star)\right) - g_\times\left(\check{\mathbb{P}}^{t-S}_y(\mathcal{T}\star)\right)\right) \geq 0. \tag{3.36}$$

By the inductive hypothesis,

$$\left(\check{\mathbb{P}}_y^{t-S}(\mathcal{T}_i) - \check{\mathbb{P}}_{y-\eta}^{t-S}(\mathcal{T}_i)\right) - \left(\check{\mathbb{P}}_{y+\eta}^{t-S}(\mathcal{T}_i) - \check{\mathbb{P}}_y^{t-S}(\mathcal{T}_i)\right) \geq 0 \tag{3.37}$$

for $y \geq 0$. By the monotonicity of $g_\times$ and (3.37),

$$g_\times\left(\check{\mathbb{P}}_{y-\eta}^{t-S}(\mathcal{T}\star)\right) \leq g_\times\left(2\check{\mathbb{P}}_y^{t-S}(\mathcal{T}\star) + \check{\mathbb{P}}_{y+\eta}^{t-S}(\mathcal{T}\star)\right). \tag{3.38}$$

Substituting (3.38) into (3.36), it suffices to show

$$g_\times\left(\check{\mathbb{P}}_{y+\eta}^{t-S}(\mathcal{T}\star)\right) - 2g_\times\left(\check{\mathbb{P}}_y^{t-S}(\mathcal{T}\star)\right) + g_\times\left(2\check{\mathbb{P}}_y^{t-S}(\mathcal{T}\star) - \check{\mathbb{P}}_{y+\eta}^{t-S}(\mathcal{T}\star)\right) \leq 0. \tag{3.39}$$

By definition of $g_\times$, (3.39) is equivalent to

$$\begin{aligned} g\left((1-b_I)\check{\mathbb{P}}_{y+\eta}^{t-S}(\mathcal{T}\star) + \frac{b_I}{2}\right) - 2g\left((1-b_I)\check{\mathbb{P}}_y^{t-S}(\mathcal{T}\star) + \frac{b_I}{2}\right) \\ +g\left((1-b_I)(2\check{\mathbb{P}}_y^{t-S}(\mathcal{T}\star) - \check{\mathbb{P}}_{y+\eta}^{t-S}(\mathcal{T}\star)) + \frac{b_I}{2}\right) \leq 0. \end{aligned} \tag{3.40}$$

To see that (3.40) holds, note that

$$\begin{aligned} &g(p_1+\eta_1, p_2+\eta_2, p_3+\eta_3) - 2g(p_1,p_2,p_3) + g(p_1-\eta_1, p_2-\eta_2, p_3-\eta_3) \\ &= 2\eta_1\eta_2(1-2p_3) + 2\eta_2\eta_3(1-2p_1) + 2\eta_3\eta_1(1-2p_2). \end{aligned}$$

Setting $p_i = (1-b_I)\check{\mathbb{P}}_y^{t-S}(\mathcal{T}_i) + \frac{b_I}{2}$ and $\eta_i = (1-b_I)(\check{\mathbb{P}}_{y+\eta}^{t-S}(\mathcal{T}_i) - \check{\mathbb{P}}_y^{t-S}(\mathcal{T}_i))$, we see that (3.40) will hold if $p_i \geq 0$, or equivalently, if $\check{\mathbb{P}}_y^{t-S}(\mathcal{T}_i) \geq \frac{1}{2}$, which holds by (3.20) since $y \geq 0$. □

Corollary 3.24. *Let $\varepsilon_1(I,\alpha)$ and $c_1(I,\alpha)$ be as in Theorem 3.14. Let $\varepsilon < \min(\varepsilon_1(\alpha), \frac{1}{24})$. Suppose that, for some $t \in [0, T^*]$ and $z \in \mathbb{R}$,*

$$\left|\mathbb{P}_z^\varepsilon[\mathbb{V}^\times(\boldsymbol{B}_{R^\varepsilon}(t)) = 1] - \tfrac{1}{2}\right| \leq \tfrac{5}{12},$$

and let $w \in \mathbb{R}$ with $|z-w| \leq c_1(\alpha)\widehat{I}(\varepsilon)|\log\varepsilon|$. Then

$$\left|\mathbb{P}_z^\varepsilon[\mathbb{V}^\times(\boldsymbol{B}_{R^\varepsilon}(t)) = 1] - \mathbb{P}_w^\varepsilon[\mathbb{V}^\times(\boldsymbol{B}_{R^\varepsilon}(t)) = 1]\right| \geq \frac{|z-w|}{48c_1(\alpha)\widehat{I}(\varepsilon)|\log\varepsilon|}.$$

This follows exactly by the proof of [EFP17, Corollary 2.12], replacing the interface width $\varepsilon|\log\varepsilon|$ with $\widehat{I}(\varepsilon)|\log\varepsilon|$, and we refer the reader there.

3.2.4 A coupling argument

In this section, we aim to construct a coupling of the one-dimensional and multi-dimensional voting systems, so that the results of Section 3.2 can be applied to the multidimensional setting. To accomplish this, we consider a preliminary coupling of $d(W(R_s^\varepsilon), t-s)$, the distance between a $\mathbb{d}$-dimensional subordinated Brownian motion and Γ_{t-s}, to a one-dimensional subordinated Brownian motion (this is the content of Theorem 3.25). The proof of this result follows closely [EFP17, Proposition 2.13]. We will use this result to motivate (and prove) the true coupling that will be used in our proof of the multidimensional result, Theorem 3.27. The necessity for not one but *two* coupling results, though laborious, allows us to overcome a key computational difficulty that arises in this work but not in the Brownian setting of [EFP17].

To begin, we consider a coupling of $d(W(R_s^\varepsilon), t-s)$ to $B(R_s^\varepsilon)$ (Theorem 3.25). We use directly the Brownian coupling from [EFP17], together with regularity of $(\Gamma_t)_{0\le t<\mathscr{T}}$. We then define a new process, proving it satisfies the desired coupling (Theorem 3.27) and has a root vote close to that of $\boldsymbol{W}_{R^\varepsilon}$ (Proposition 3.28). Let us first state some regularity properties also used in [EFP17], which follow from the Assumptions 1.16.

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
$$|\nabla d(x,t)| = 1. \tag{3.41}$$
Moreover, d is a $C^{a,\frac{a}{2}}$ function in $\{(x,t) : |d(x,t)| \le c_0, t \le T^*\}$ for some $a > 2$.

2. If we view $\mathbf{n} := \nabla d$ as the positive normal direction, then for $x \in \Gamma_t$, the normal velocity of Γ_t at x is $-d(x,t)$, and the curvature of Γ_t at x is $-\Delta d(x,t)$. Thus, (1.10) becomes
$$\dot{d}(x,t) = \Delta d(x,t) \tag{3.42}$$
for all x such that $d(x,t) = 0$.

3. There exists $C_0 > 0$ such that, for all $t \in [0, T^*]$ and all x such that $|d(x,t)| \le c_0$,
$$\left|\nabla\left(\dot{d}(x,t) - \Delta d(x,t)\right)\right| \le C_0. \tag{3.43}$$

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

$$|d(x,t) - d(x,s)| \leq V_0(s-t). \tag{3.44}$$

Theorem 3.25. *Let $k \in \mathbb{N}$. For $\alpha \in (1,2)$, let $(W_t)_{t\geq 0}$ be a d-dimensional standard Brownian motion started at $x \in \mathbb{R}^{\mathrm{d}}$, and let $(R_t^\varepsilon)_{t\geq 0}$ be an $\widehat{I}(\varepsilon)^2$-truncated $\frac{\alpha}{2}$-stable subordinator satisfying Assumption 3.1. Fix $t \leq T^*$ and $\beta < c_0$, for c_0 as in (3.43). Define the stopping time*

$$T_\beta = \inf(\{s \in [0, (k+1)\varepsilon^2 |\log \varepsilon|) : |d(W_s, t-s)| > \beta\} \cup \{t\}).$$

If $R_s^\varepsilon < T_\beta \wedge t$, then there exists a one-dimensional standard Brownian motion $(B_t)_{t\geq 0}$ started at $d(x,t)$, and there exist constants $C_0, D_0 > 0$ and $\varepsilon_1(k) > 0$ such that, for all $\varepsilon \in (0, \varepsilon_1(k))$, with probability at least $1 - \varepsilon^{k+1}$,

$$|d(W(R_s^\varepsilon), t-s) - B(R_s^\varepsilon)| \leq C_0\beta s + D_0(k+2)\widehat{I}(\varepsilon)^2 |\log \varepsilon|. \tag{3.45}$$

Proof. We first rewrite $d(W(R_s^\varepsilon), t-s)$ as

$$d(W(R_s^\varepsilon), t - R_s^\varepsilon) + \left[d(W(R_s^\varepsilon), t-s) - d(W(R_s^\varepsilon), t - R_s^\varepsilon)\right]. \tag{3.46}$$

By the Brownian coupling result [EFP17, Proposition 2.13], since $R_s^\varepsilon \leq T_\beta \wedge t$ by assumption, there exists a one-dimensional Brownian motion $(B_t)_{t\geq 0}$ started from $d(x,t)$ and $C_0 > 0$ such that

$$B(R_s^\varepsilon) - C_0\beta R_s^\varepsilon \leq d(W(R_s^\varepsilon), t - R_s^\varepsilon) \leq B(R_s^\varepsilon) + C_0\beta R_s^\varepsilon. \tag{3.47}$$

By (3.44), the second term in (3.46) is bounded above by

$$|d(W(R_s^\varepsilon), t-s) - d(W(R_s^\varepsilon), t - R_s^\varepsilon)| \leq V_0|R_s^\varepsilon - s|. \tag{3.48}$$

Combining (3.46), (3.47) and (3.48),

$$\begin{aligned} |d(W(R_s^\varepsilon), t-s) - B(R_s^\varepsilon)| &\leq C_0\beta R_s^\varepsilon + V_0|R_s^\varepsilon - s| \\ &\leq C_0\beta s + D_0|R_s^\varepsilon - s| \end{aligned}$$

where $D_0 := V_0 + C_0 c_0$. By Lemma B.3, there exists $\varepsilon_1(k) > 0$ such that, for all $\varepsilon \in (0, \varepsilon_1(k))$,

$$\mathbb{P}(|R_s^\varepsilon - s| > (k+2)\widehat{I}(\varepsilon)^2 |\log \varepsilon|) \leq \varepsilon^{k+1}. \qquad \square$$

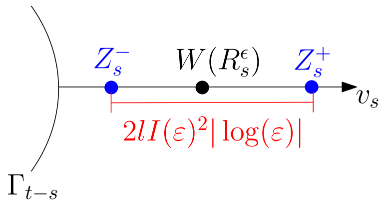

Figure 3.1: Possible configuration of Z_s^+, Z_s^- and $W(R_s^\varepsilon)$

The key difference between the coupling in Theorem 3.25 and that of the Brownian setting [EFP17, Proposition 2.14] is the additional $\mathcal{O}(\widehat{I}(\varepsilon)^2|\log\varepsilon|)$ error obtained in (3.45). This seemingly innocent error causes strife when one attempts to adapt the multidimensional Brownian proof to our setting, as it propagates with each iteration of voting. To overcome this technical difficulty, we introduce the auxiliary $\mathbb{d}$-dimensional processes $(Z_t^+)_{t\geq 0}$ and $(Z_t^-)_{t\geq 0}$ that can be coupled to our one-dimensional motion $(B(R_t^\varepsilon))_{t\geq 0}$ as in (3.45) without this additional error term. With this, it is straightforward to adapt the techniques of [EFP17] to prove the multidimensional result for the historical ternary branching motions $\boldsymbol{Z}^+$ and $\boldsymbol{Z}^-$. Finally, by controlling the error between the root votes of $\boldsymbol{Z}^-$ and $\boldsymbol{Z}^+$ to that of the $\mathbb{d}$-dimensional subordinated Brownian motion, we obtain the desired result. We define Z^+ and Z^- as follows.

Definition 3.26 (Z_s^+, Z_s^-)**.** Let $(W(R_t^\varepsilon))_{t\geq 0}$ be a $\mathbb{d}$-dimensional subordinated Brownian motion. Fix $0 < t, T < \mathscr{T}$, $l > 0$ and $\beta < c_0$, for c_0 as in (3.41). Let $x_s \in \Gamma_{t-s}$ be the unique point on Γ_{t-s} that is the shortest distance from $W(R_s^\varepsilon)$, and v_s be the outward facing unit vector perpendicular to the tangent hypersurface of Γ_{t-s} at x_s. Then we define the processes $(Z_s^+)_{0\leq s\leq T}$ and $(Z_s^-)_{0\leq s\leq T}$ by

$$Z_s^+ = \begin{cases} W(R_s^\varepsilon) + l\widehat{I}(\varepsilon)^2|\log\varepsilon|v_s & \text{if } |d(W(R_s^\varepsilon), t-s)| \leq \beta \\ W(R_s^\varepsilon) & \text{otherwise.} \end{cases} \tag{3.49}$$

$$Z_s^- = \begin{cases} W(R_s^\varepsilon) - l\widehat{I}(\varepsilon)^2|\log\varepsilon|v_s & \text{if } |d(W(R_s^\varepsilon), t-s)| \leq \beta \\ W(R_s^\varepsilon) & \text{otherwise.} \end{cases} \tag{3.50}$$

Observe that we may choose ε sufficiently small that any point x on the line segment between $W(R_s^\varepsilon)$ and Z_s^+ (or Z_s^-) satisfies $d(x, t-s) \leq c_0$. It then follows by (3.41) that

$|\nabla d(x, t-s)| = 1$, so

$$d(Z_s^-, t-s) = d(W(R_s^\varepsilon), t-s) - l\widehat{I}(\varepsilon)^2 |\log(\varepsilon)|, \tag{3.51}$$

$$d(Z_s^+, t-s) = d(W(R_s^\varepsilon), t-s) + l\widehat{I}(\varepsilon)^2 |\log(\varepsilon)|. \tag{3.52}$$

Consequently, we obtain the following important restatement of Theorem 3.25.

Theorem 3.27. *Let $k \in \mathbb{N}$. For $\alpha \in (1,2)$ and D_0 as in (3.45), let Z_t^+ and Z_t^- be as in Definition 3.26 for $l := D_0(k+2)$, started at $x \in \mathbb{R}^{\mathrm{d}}$. Let $(R_t^\varepsilon)_{t\geq 0}$ be an $\widehat{I}(\varepsilon)^2$-truncated $\frac{\alpha}{2}$-stable subordinator satisfying Assumption 3.1. Fix $t \leq T^*$, and fix $\beta < c_0$ for c_0 as in (3.43). Define the stopping time*

$$T_\beta = \inf(\{s \in [0, (k+1)\varepsilon^2 |\log \varepsilon|) : |d(W_s, t-s)| > \beta\} \cup \{t\}).$$

If $R_s^\varepsilon < T_\beta \wedge t$, then there exists a one-dimensional standard Brownian motion $(B_t)_{t\geq 0}$ started at $d(x,t)$, and there exists $C_0 > 0$ such that, with probability at least $1 - \varepsilon^{k+1}$,

$$d(Z_s^+, t-s) \leq B(R_s^\varepsilon) + C_0 \beta s \tag{3.53}$$

$$d(Z_s^-, t-s) \geq B(R_s^\varepsilon) + C_0 \beta s. \tag{3.54}$$

This follows immediately from Theorem 3.25 and equations (3.52), (3.51) and so we omit a formal proof.

Notation. *As we have seen in Theorem 3.27, Z^+ and Z^- satisfy (3.53) and (3.54) when $l := D_0(k+2)$ in (3.49) and (3.50). For the remainder of this work, we shall take $l := D_0(k+2)$ in the definition of Z^+ and Z^-, where the choice of k will be clear in the given context.*

To conclude this section, we compare the root votes of $\mathbb{V}_p^\times(\boldsymbol{Z}^-)$ and $\mathbb{V}_p^\times(\boldsymbol{Z}^+)$ to that of $\mathbb{V}_p^\times(\boldsymbol{W}_{R^\varepsilon})$. We denote by $\boldsymbol{Z}^+$ the ternary branching process (with branching rate ε^{-2}) in which individuals independently travel according to $(Z_s^+)_{0\leq s\leq \mathscr{T}}$. Define $\boldsymbol{Z}^-$ similarly.

Proposition 3.28. *Let $\alpha \in (1,2)$, $0 < t < \mathscr{T}$, $0 \leq \beta < c_0$, and $k \in \mathbb{N}$ and let F be as in (1.35). Let $\boldsymbol{Z}^+(t)$ and $\boldsymbol{Z}^-(t)$ be the historical path of branching processes defined above. Then, for any $x \in \mathbb{R}^{\mathrm{d}}$, there exist $M_1, M_2 > 0$ such that*

$$\left|\mathbb{P}_x[\mathbb{V}_p^\times(\boldsymbol{Z}^-(t)) = 1)] - \mathbb{P}_x[\mathbb{V}_p^\times(\boldsymbol{W}_{R^\varepsilon}(t)) = 1)]\right| \leq M_1 e^{-\frac{t}{\varepsilon^2}} + F(\varepsilon, M_2) \tag{3.55}$$

$$\left|\mathbb{P}_x[\mathbb{V}_p^\times(\boldsymbol{Z}^+(t)) = 1)] - \mathbb{P}_x[\mathbb{V}_p^\times(\boldsymbol{W}_{R^\varepsilon}(t)) = 1)]\right| \leq M_1 e^{-\frac{t}{\varepsilon^2}} + F(\varepsilon, M_2). \tag{3.56}$$

Proposition 3.28 is integral to the proof of the main multidimensional result. While intuitively the root votes in (3.55) and (3.56) should be close, obtaining the precise bound above requires lengthy calculations which are not illuminating. To not disrupt our flow as we begin the proof of the multidimensional result, we defer the proof of Proposition 3.28 to Section B.2 of Appendix B.

Remark 3.29. Observe that Proposition 3.28 marks the first appearance of the term $F(\varepsilon)$ that will later contribute to the tolerance of the interface in Theorem 1.39. Recall that our requirement that $F(\varepsilon) \to 0$ as $\varepsilon \to 0$ corresponds to the assumption that $\alpha \in (1, 2)$.

3.3 Marked majority voting, for $\mathbb{d} \geq 2$

Let us briefly recall the notation introduced in Section 3.2. We write X for the one-dimensional α-stable process ($\boldsymbol{X}$ for the associated historical ternary branching process), and Y for the $\mathbb{d}$-dimensional α-stable process ($\boldsymbol{Y}$ for the associated historical ternary branching process). The one-dimensional R_t^ε-subordinated standard Brownian motion is denoted $B(R_t^\varepsilon)$ (with associated historical ternary branching process $\boldsymbol{B}_{R^\varepsilon}$), and the R_t^ε-subordinated $\mathbb{d}$-dimensional Brownian motion is denoted $W(R_t^\varepsilon)$ (with associated historical ternary branching process $\boldsymbol{W}_{R^\varepsilon}$). As ever, all stable processes and subordinators are assumed to satisfy Assumption 3.1.

To begin our proof of the $\mathbb{d}$-dimensional result, we show that it suffices to prove our main theorem for the marked voting system on $\boldsymbol{Z}^+$ and $\boldsymbol{Z}^-$. We accomplish this through a series of coupling results, namely

$$\begin{aligned}
\mathbb{P}_x^\varepsilon \left[\mathbb{V}_p(\boldsymbol{Y}(t)) = 1 \right] &\overset{Prop\ 3.31}{\approx} \mathbb{P}_x^\varepsilon \left[\mathbb{V}_p^+(\boldsymbol{W}_{R^\varepsilon}(t)) = 1 \right] \\
&\overset{Prop\ 3.33}{\approx} \mathbb{P}_x^\varepsilon \left[\mathbb{V}_p^\times(\boldsymbol{W}_{R^\varepsilon}(t)) = 1 \right] \\
&\overset{Prop\ 3.28}{\approx} \mathbb{P}_x^\varepsilon \left[\mathbb{V}_p^\times(\boldsymbol{Z}^+(t)) = 1 \right].
\end{aligned}$$

We then proceed as in [EFP17] to show that, by an $\mathcal{O}(\varepsilon^2 |\log \varepsilon|)$ time, an interface is generated (Section 3.3.2), and that the region about the interface propagates, by comparison to our one-dimensional result (Section 3.3.3).

3.3.1 A coupling of voting systems, for $\mathbb{d} \geq 2$

In this section, we aim to obtain a coupling of the majority voting system on $\mathcal{T}(\boldsymbol{Y}(t))$ to the marked majority voting system on $\mathcal{T}(\boldsymbol{W}_{R^\varepsilon}(t))$. However, unlike the analogous coupling in one dimension, we can no longer exploit the symmetry (3.5) of the systems afforded to us by the initial condition $p(x) = \mathbb{1}_{\{x \geq 0\}}$. Instead, in $\mathbb{d}$ dimensions, the presence of curvature means that any individual that makes a large jump near the interface will have a positive (resp. negative) bias to cross over to the other side of the interface, in regions where the interface is convex (resp. concave). Consequently, there is no longer a straightforward way to relate particles in $\boldsymbol{Y}$ to marked particles in $\boldsymbol{W}_{R^\varepsilon}$ that have an equal probability of voting one or zero. To overcome this, we consider in Proposition 3.31 an intermediate coupling between the majority voting system on $\mathcal{T}(\boldsymbol{Y}(t))$ and the *asymmetric* marked majority voting systems on $\mathcal{T}(\boldsymbol{W}_{R^\varepsilon}(t))$. The latter systems will then be coupled to the symmetric marked majority voting system of interest in Proposition 3.33, and, combining these results, we obtain the desired coupling in Corollary 3.34.

Let $\varepsilon > 0$, $t \geq 0$ and $\alpha \in (1, 2)$. Let b_I be as in (3.17). For a fixed function $p : \mathbb{R}^{\mathbb{d}} \to [0, 1]$, we define the *positively biased* (resp. *negatively biased*) asymmetric marked voting procedures on $\mathcal{T}(\boldsymbol{W}_{R^\varepsilon}(t))$ as follows.

1. At each branch point in $\mathcal{T}(\boldsymbol{W}_{R^\varepsilon}(t))$, the parent particle j marks each of its three offspring $(j, 1), (j, 2)$ and $(j, 3)$ independently with probability b_I. All marked particles vote 1 with probability 1 (resp. 0).

2. Independently, each unmarked leaf i of $\mathcal{T}(\boldsymbol{W}_{R^\varepsilon}(t))$ votes 1 with probability $p(W_i(R_i^\varepsilon(t))$, and otherwise votes 0.

3. At each branch point in $\mathcal{T}(\boldsymbol{W}_{R^\varepsilon}(t))$, if the parent particle k is unmarked, it votes according to the majority vote of its three offspring $(k, 1), (k, 2)$ and $(k, 3)$.

Definition 3.30 **($\mathbb{V}_p^+, \mathbb{V}_p^-$).** Define $\mathbb{V}_p^+$ (resp. $\mathbb{V}_p^-$) to be the vote associated to the root $\emptyset$ of the ternary branching tree with truncated stable motion under the positively biased (resp. negatively biased) asymmetric marked voting procedure described above.

With this definition, we can obtain the following bound on the root vote of $\mathcal{T}(\boldsymbol{Y}(t))$.

Proposition 3.31. *For all $\varepsilon > 0$, $x \in \mathbb{R}^{\mathrm{d}}$, $t \geq 0$ and $p : \mathbb{R}^{\mathrm{d}} \to [0,1]$ we have*

$$(1-b_I)\mathbb{P}_x^\varepsilon[\mathbb{V}_p^-(\boldsymbol{W}_{R^\varepsilon}(t))=1] \leq \mathbb{P}_x^\varepsilon[\mathbb{V}_p(\boldsymbol{Y}(t))=1] \leq (1-b_I)\mathbb{P}_x^\varepsilon[\mathbb{V}_p^+(\boldsymbol{W}_{R^\varepsilon}(t))=1]+b_I.$$

This proof of this result proceeds almost identically to the proof of the one-dimensional coupling of voting systems, Theorems 3.10 and 3.12 using an intermediate asymmetric exponentially marked voting system, we refer the reader there for details.

Remark 3.32. It is important to note that Theorem 3.10 cannot be adapted to the symmetric marked voting system $\mathbb{V}_p^\times$ in d dimensions. This is because, for general p, we cannot conclude $\mathbb{E}[p(B(R_t^\varepsilon))] \geq \frac{1}{2}$: a crucial step in the proof of Theorem 3.10, seen in, for example, equation (3.13). However, the proof of Theorem 3.10 can be adapted for the asymmetric marked voting systems, since (3.13) becomes, simply, $\mathbb{E}[p(B(R_t^\varepsilon))] \geq 0$. Our ability to adapt Theorem 3.10 to the asymmetric marked voting systems is our main motivation in using them as intermediate systems. We further compare $\mathbb{V}_p^+$ and $\mathbb{V}_p^-$ to $\mathbb{V}_p^\times$ so that we can exploit the symmetry of $\mathbb{V}_p^\times$ in our later proofs.

The positively and negatively marked systems can then be compared to our (symmetric) marked system as follows.

Proposition 3.33. *There exists $C > 0$ such that, for all $\varepsilon > 0$, $x \in \mathbb{R}^{\mathrm{d}}$, $t \geq 0$ and $p : \mathbb{R}^{\mathrm{d}} \to [0,1]$,*

$$\sup_{x\in\mathbb{R}^{\mathrm{d}}}\left(\mathbb{P}_x^\varepsilon[\mathbb{V}_p^+(\boldsymbol{W}_{R^\varepsilon}(t))=1]-\mathbb{P}_x^\varepsilon[\mathbb{V}_p^\times(\boldsymbol{W}_{R^\varepsilon}(t))=1]\right) \leq Cb_I \tag{3.57}$$

$$\sup_{x\in\mathbb{R}^{\mathrm{d}}}\left(\mathbb{P}_x^\varepsilon[\mathbb{V}_p^\times(\boldsymbol{W}_{R^\varepsilon}(t))=1]-\mathbb{P}_x^\varepsilon[\mathbb{V}_p^-(\boldsymbol{W}_{R^\varepsilon}(t))=1]\right) \leq Cb_I. \tag{3.58}$$

Proof. We prove only (3.57), noting that (3.58) follows by symmetric arguments. Define $g_+ : [0,1] \to [0,1]$ by $g_+(q) = g((1-b_I)q + b_I)$, where g is the ordinary majority voting function. This is the probability that an unmarked parent particle votes 1 under $\mathbb{V}_p^+$ in the special case when the three offspring are independent, and each has probability q of voting 1 if it is unmarked. Write S for the time of the first branch in $\boldsymbol{W}_{R^\varepsilon}(\cdot)$. To ease notation, set

$$u_\times^\varepsilon(x,t) = \mathbb{P}_x^\varepsilon\left[\mathbb{V}_p^\times(\boldsymbol{W}_{R^\varepsilon}(t))=1\right] \quad \text{and} \quad u_+^\varepsilon(x,t) = \mathbb{P}_x^\varepsilon\left[\mathbb{V}_p^+(\boldsymbol{W}_{R^\varepsilon}(t))=1\right].$$

Then, by the Markov property at time $t \wedge S$ and the definitions of $\mathbb{V}_p^\times$ and $\mathbb{V}_p^+$ (noting that the initial ancestor is never marked in both schemes), we have

$$u_\times^\varepsilon(x,t) = \mathbb{E}_x^\varepsilon\left[g_\times(u_\times^\varepsilon(W(R_S^\varepsilon), t-S))\mathbb{1}_{\{S\leq t\}}\right] + \mathbb{E}_x^\varepsilon\left[p(W(R_t^\varepsilon))\mathbb{1}_{\{S>t\}}\right]$$
$$u_+^\varepsilon(x,t) = \mathbb{E}_x^\varepsilon\left[g_+\left(u_+^\varepsilon(W(R_S^\varepsilon), t-S)\right)\mathbb{1}_{\{S\leq t\}}\right] + \mathbb{E}_x^\varepsilon[p(W(R_t^\varepsilon))\mathbb{1}_{\{S>t\}}\right].$$

It follows that

$$|u_\times^\varepsilon(x,t) - u_+^\varepsilon(x,t)| \leq \mathbb{E}_x^\varepsilon\left[\left|g_\times(u_\times^\varepsilon(W(R_S^\varepsilon), t-s)) - g_+(u_+^\varepsilon(W(R_S^\varepsilon), t-s))\right|\mathbb{1}_{\{S\leq t\}}\right].$$

By the definitions of $g_\times$ and g_+, and since g is Lipschitz with constant $\frac{3}{2}$, we have

$$\begin{aligned}
&|u_\times^\varepsilon(x,t) - u_+^\varepsilon(x,t)| \\
&\quad\leq \tfrac{3}{2}\mathbb{E}_x^\varepsilon\left[\left|(1-b_I)(u_\times^\varepsilon(W(R_S^\varepsilon), t-S) - u_+^\varepsilon(W(R_S^\varepsilon), t-S)) - \tfrac{b_I}{2}\right|\mathbb{1}_{\{S\leq t\}}\right] \\
&\quad\leq \tfrac{3}{4}b_I + \tfrac{3}{2}(1-b_I)\mathbb{E}_x^\varepsilon\left[\left|u_\times^\varepsilon(W(R_S^\varepsilon), t-S) - u_+^\varepsilon(W(R_S^\varepsilon), t-S)\right|\mathbb{1}_{\{S\leq t\}}\right] \\
&\quad= \tfrac{3}{4}b_I + \tfrac{3}{2}(1-b_I)\int_0^t \frac{e^{-\rho\varepsilon^{-2}}}{\varepsilon^2}\mathbb{E}_x^\varepsilon\left[|u_\times^\varepsilon(W(R_u^\varepsilon), t-\rho) - u_+^\varepsilon(W(R_u^\varepsilon), t-\rho)|\right]d\rho \\
&\quad\leq \tfrac{3}{4}b_I + \tfrac{3}{2}(1-b_I)\int_0^t \frac{e^{-\rho\varepsilon^{-2}}}{\varepsilon^2}\|u_\times^\varepsilon(\cdot, t-\rho) - u_+^\varepsilon(\cdot, t-\rho)\|_\infty d\rho,
\end{aligned}$$

where $\|\cdot\|_\infty$ denotes the supremum norm, and we have used that $S \sim \texttt{Exp}(\varepsilon^{-2})$ is independent of the spatial motion. Noting that the above inequality holds for all $x \in \mathbb{R}^{\mathbb{d}}$, and applying the change of variables $\rho \mapsto t-\rho$, we obtain

$$\|u_\times^\varepsilon(\cdot,t) - u_+^\varepsilon(\cdot,t)\|_\infty \leq \tfrac{3}{4}b_I + \tfrac{3}{2}e^{-t\varepsilon^{-2}}\int_0^t e^{\rho\varepsilon^{-2}}\varepsilon^{-2}\|u_\times^\varepsilon(\cdot,\rho) - u_+^\varepsilon(\cdot,\rho)\|_\infty d\rho.$$

By an adaptation of Grönwall's inequality as proved in [Dra02, Theorem 15],

$$\begin{aligned}
\|u_\times^\varepsilon(\cdot,t) - u_+^\varepsilon(\cdot,t)\|_\infty &\leq \tfrac{3}{4}b_I\exp\left(\tfrac{3}{2}\left(\int_0^t \exp\left(-s\varepsilon^{-2}\right)\varepsilon^{-2}\right)ds\right) \\
&= \tfrac{3}{4}b_I\exp\left(\frac{3}{2}\mathbb{P}[S\leq t]\right) \\
&\leq \tfrac{3}{4}b_I\exp\left(\tfrac{3}{2}\right).
\end{aligned}$$

Setting $C := \frac{3}{4}\exp\left(\frac{3}{2}\right)$ gives the result. □

The following result then follows by combining Propositions 3.31 and 3.33.

Corollary 3.34. *Let $\varepsilon > 0$, $x \in \mathbb{R}^{\mathrm{d}}$, $t \geq 0$, and $p : \mathbb{R}^{\mathrm{d}} \to [0,1]$. Let C be as in Proposition 3.33 and let $\bar{C} = C + 1$. Write $u^\varepsilon(x,t) = \mathbb{P}_x^\varepsilon[\mathbb{V}_p(\boldsymbol{Y}(t)) = 1]$. Then*

$$(1-b_I)\mathbb{P}_x^\varepsilon[\mathbb{V}_p^\times(\boldsymbol{W}_{R^\varepsilon}(t)) = 1] - Cb_I \leq u^\varepsilon(x,t) \leq (1-b_I)\mathbb{P}_x^\varepsilon[\mathbb{V}_p^\times(\boldsymbol{W}_{R^\varepsilon}(t)) = 1] + \bar{C}b_I.$$

With this coupling of voting systems in mind, we can now state the theorem for the d-dimensional marked system that in turn proves Theorem 1.39.

Theorem 3.35. *Let $\alpha \in (1,2)$, and fix a function $I : \mathbb{R}_+ \to \mathbb{R}_+$ satisfying Assumptions 1.37. Suppose the initial condition p satisfies Assumptions 1.16. Let $\mathscr{T}$ and $d(x,t)$ be as in Section 1.5, let F be as in (1.35), and fix $T^* \in (0, \mathscr{T})$. Let u_+, u_- be as in (3.21) and (3.22). Then there exist $\varepsilon_{\mathrm{d}}(\alpha, I)$, $a_{\mathrm{d}}(\alpha, I)$, $c_{\mathrm{d}}(\alpha, I)$, $M > 0$ such that, for $\varepsilon \in (0, \varepsilon_{\mathrm{d}})$ and $a_{\mathrm{d}}\varepsilon^2|\log\varepsilon| \leq t \leq T^*$,*

1. *for x with $d(x,t) \geq c_{\mathrm{d}}\widehat{I}(\varepsilon)|\log\varepsilon|$ we have $\mathbb{P}_x\left[\mathbb{V}_p^\times(\boldsymbol{W}_{R^\varepsilon}(t)) = 1\right] \geq u_+ - F(\varepsilon, M)$,*
2. *for x with $d(x,t) \leq -c_{\mathrm{d}}\widehat{I}(\varepsilon)|\log\varepsilon|$, we have $\mathbb{P}_x\left[\mathbb{V}_p^\times(\boldsymbol{W}_{R^\varepsilon}(t)) = 1\right] \leq u_- + F(\varepsilon, M)$.*

3.3.2 Generation of interface

We now show that, in time $\mathcal{O}(\varepsilon^2|\log\varepsilon|)$, an interface of width $\mathcal{O}(\widehat{I}(\varepsilon)|\log\varepsilon|)$ is generated, and, outside this interface, the root vote of the marked voting system is at a distance at most ε^k from one of the fixed points u_+ and u_-.

Proposition 3.36. *For fixed $k \in \mathbb{N}$, there exists $\rho_{\mathrm{d}}(k) > 0$ such that for $t \geq \rho_{\mathrm{d}}(k)\varepsilon^2|\log\varepsilon|$ and $x \in \mathbb{R}$,*

$$u_- - \varepsilon^k \leq \mathbb{P}_x^\varepsilon[\mathbb{V}_p^\times(\boldsymbol{W}_{R^\varepsilon}(t)) = 1] \leq u_+ + \varepsilon^k \tag{3.59}$$

for any initial condition $p : \mathbb{R}^{\mathrm{d}} \to [0,1]$.

Proof. This is a special case of [HD20, Lemma 3.5], the proof of which we follow closely. We prove the right hand inequality in (3.59) and note that the left hand inequality follows by very similar arguments. It is easy to verify that $g_\times'(u_+) < 1$ and, since u_+ is a fixed point of $g_\times$, it is easy to verify that, for ε sufficiently small, $g_\times(q) < q$ for $q \in (u_+, 1]$. Therefore

$$q_2 := \inf_{q \in (0, 1-u_+]} \frac{(u_+ + q) - g_\times(u_+ + q)}{q} \in (0,1)$$

and, for $\delta \in [0, 1-u_+]$,

$$g_\times(u_+ + \delta) - u_+ \leq (1-q_2)\delta.$$

Iterating this as in the proof of Lemma 3.18, we obtain

$$q_\times^{(n)}(u_+ + \delta) - u_+ \leq (1-q_2)^n\delta.$$

In particular,

$$g_\times^{(n)}(u_+ + (1-u_+)) - u_+ \leq (1-q_2)^n(1-u_+) \leq \varepsilon^k$$

after $n \geq C(k)|\log\varepsilon|$ iterations, for some $C(k) > 0$. That is, $g_\times^{(n)}(1) \leq u_+ + \varepsilon^k$ after $n \geq C(k)|\log\varepsilon|$ iterations. We note that, since $g_\times$ is increasing, the largest value of the iterates of $g_\times$ will be when $p = 1$. Finally, by identical arguments to Lemma 3.20, there exists $\rho_{\mathbb{d}}(k) > 0$ such that, for $t \geq \rho_{\mathbb{d}}(k)\varepsilon^2|\log\varepsilon|$,

$$\mathbb{P}^\varepsilon\left[\mathcal{T}(\boldsymbol{W}_{R^\varepsilon}(t)) \supset \mathcal{T}^{reg}_{C(k)|\log\varepsilon|}\right] \geq 1-\varepsilon^k.$$

Therefore, when $t \geq \rho_{\mathbb{d}}\varepsilon^2|\log\varepsilon|$, it holds that $\mathbb{P}^\varepsilon_x[\mathbb{V}_p^\times(\boldsymbol{W}_{R^\varepsilon}(t)) = 1] \leq u_+ + 2\varepsilon^k$, $\square$

To begin, we first note that Proposition 3.36 holds for $\boldsymbol{Z}^+$ and $\boldsymbol{Z}^-$. That is, for $k \in \mathbb{N}$, there exists $\rho_{\mathbb{d}}^+(k)$ and $\rho_{\mathbb{d}}^-(k)$ such that, for $t \geq \rho_{\mathbb{d}}^+(k)\varepsilon^2|\log\varepsilon|$,

$$u_- - \varepsilon^k \leq \mathbb{P}^\varepsilon_x[\mathbb{V}_p^\times(\boldsymbol{Z}^+(t)) = 1] \leq u_+ + \varepsilon^k, \tag{3.60}$$

and, for $t \geq \rho_{\mathbb{d}}^-(k)\varepsilon^2|\log\varepsilon|$,

$$u_- - \varepsilon^k \leq \mathbb{P}^\varepsilon_x[\mathbb{V}_p^\times(\boldsymbol{Z}^-(t)) = 1] \leq u_+ + \varepsilon^k, \tag{3.61}$$

for any initial condition $p : \mathbb{R}^{\mathbb{d}} \to [0,1]$. We let $\rho_{\mathbb{d}}(k) := \rho_{\mathbb{d}}^-(k) \vee \rho_{\mathbb{d}}^+(k)$.

Proposition 3.37. *Let $\alpha \in (1,2)$, $k \in \mathbb{N}$ and $\rho_{\mathbb{d}}(k)$ be as above. Fix $I : \mathbb{R}_+ \to \mathbb{R}_+$ satisfying Assumptions 1.37 (A) and 1.37 (B). Let u_+ and u_- be as in (3.21), (3.22). Then there exist $\varepsilon_{\mathbb{d}}(\alpha, I), a_{\mathbb{d}}(\alpha, I), b_{\mathbb{d}}(\alpha, I) > 0$ such that, for all $\varepsilon \in (0, \varepsilon_{\mathbb{d}})$, if*

$$t_{\mathbb{d}}(k,\varepsilon) := (a_{\mathbb{d}}(k) \vee \rho_{\mathbb{d}}(k))\varepsilon^2|\log\varepsilon|,$$
$$t'_{\mathbb{d}}(k,\varepsilon) := (2(a_{\mathbb{d}}(k) \vee \rho_{\mathbb{d}}(k)) + k + 1)\varepsilon^2|\log\varepsilon|,$$

then for $t \in [t_{\mathbb{d}}, t'_{\mathbb{d}}]$,

1. *for* $d(x,t) \geq b_{\mathrm{d}}(k)\widehat{I}(\varepsilon)|\log\varepsilon|$*, we have* $\mathbb{P}^{\varepsilon}_x[\mathbb{V}^{\times}_p(\boldsymbol{Z}^+(t)) = 1] \geq u_+ - \varepsilon^k$*,*

2. *for* $d(x,t) \leq -b_{\mathrm{d}}(k)\widehat{I}(\varepsilon)|\log\varepsilon|$*, we have* $\mathbb{P}^{\varepsilon}_x[\mathbb{V}^{\times}_p(\boldsymbol{Z}^+(t)) = 1] \leq u_- + \varepsilon^k$*.*

Remark 3.38. By almost identical arguments Proposition 3.37 holds when $\boldsymbol{Z}^+$ is replaced with $\boldsymbol{Z}^-$.

Proof. We follow the proof of [EFP17, Proposition 2.16] closely, and consider the multidimensional analogues of Lemmas 3.20 and 3.21. First, by an identical argument to Lemma 3.20, there exist $a_{\mathrm{d}}(k) > 0$ and $\varepsilon_{\mathrm{d}}(\alpha) > 0$ such that, for all $\varepsilon \in (0, \varepsilon_{\mathrm{d}})$ and $t \geq a_{\mathrm{d}}(k)\varepsilon^2|\log\varepsilon|$,

$$\mathbb{P}^{\varepsilon}\left[\mathcal{T}\left(\boldsymbol{Z}^+(t)\right) \supseteq \mathcal{T}^{\mathrm{reg}}_{A|\log\varepsilon|}\right] \geq 1 - \varepsilon^k, \tag{3.62}$$

for $A(k)$ as in Lemma 3.18. Moreover, by standard estimates for the multidimensional standard normal variable, the proof of Lemma 3.21 can be adapted to show that there exist $h_{\mathrm{d}}(k) > 0$ and $\varepsilon_{\mathrm{d}}(I, \alpha) > 0$ such that, for all $\varepsilon \in (0, \varepsilon_{\mathrm{d}})$ and $t \in [t_{\mathrm{d}}, t'_{\mathrm{d}}]$,

$$\mathbb{P}_x\left[\exists i \in N(s) : \|W_i(R^{\varepsilon}_i(s)) - x\| \geq h_{\mathrm{d}}(k)\widehat{I}(\varepsilon)|\log\varepsilon|\right] \leq \varepsilon^k. \tag{3.63}$$

By the definition of Z^+_s, $\|Z^+_s\| \leq \|W(R^{\varepsilon}_s)\| + D_0(k+2)\widehat{I}(\varepsilon)^2|\log\varepsilon|$, and so (3.63) implies that, for ε sufficiently small, there exists $l_{\mathrm{d}}(k) > h_{\mathrm{d}}(k)$ for which

$$\mathbb{P}_x\left[\exists i \in N(s) : \|Z^+_s - x\| \geq l_{\mathrm{d}}(k)\widehat{I}(\varepsilon)|\log\varepsilon|\right] \leq \varepsilon^k. \tag{3.64}$$

Set $b_{\mathrm{d}} = 2l_{\mathrm{d}}$. Recall that $d(x,t)$ is the signed distance between $x \in \mathbb{R}^{\mathrm{d}}$ and Γ_t. By the regularity assumption on Γ_t (3.44), there exist $v_0, V_0 > 0$ such that, for $t \leq v_0$ and $x \in \mathbb{R}^{\mathrm{d}}$, $|d(x,0) - d(x,t)| \leq V_0 t$. Reduce ε_{d} if necessary so that $t'_{\mathrm{d}} \leq v_0$ for all $\varepsilon \in (0, \varepsilon_{\mathrm{d}})$. Let $\varepsilon \in (0, \varepsilon_{\mathrm{d}})$, let $t \in [t_{\mathrm{d}}, t'_{\mathrm{d}}]$ and let x be such that $d(x,t) \geq b_{\mathrm{d}}\widehat{I}(\varepsilon)|\log\varepsilon|$ and $\|Z^+_i(t) - x\| \leq l_{\mathrm{d}}\widehat{I}(\varepsilon)|\log\varepsilon|$. It follows by the triangle inequality and the Lipschitz continuity of $d(\cdot, t)$ that

$$\begin{aligned} d(Z^+_i(t), 0) &\geq d(x,t) - \left|d(x,t) - d(Z^+_i(t), t)\right| - \left|d(Z^+_i(t), t) - d(Z^+_i(t), 0)\right| \\ &\geq b_{\mathrm{d}}\widehat{I}(\varepsilon)|\log\varepsilon| - l_{\mathrm{d}}\widehat{I}(\varepsilon)|\log\varepsilon| - V_0 t'_{\mathrm{d}} \\ &= \tfrac{1}{2}b_{\mathrm{d}}\widehat{I}(\varepsilon)|\log\varepsilon| - V_0(2(a_{\mathrm{d}} \vee \rho_{\mathrm{d}}) + k + 1)\varepsilon^2|\log\varepsilon|. \end{aligned}$$

By Assumption 1.37 (B) we may reduce $\varepsilon_{\mathbb{d}}$ if necessary so that

$$d(Z_i^+(t), 0) \geq \tfrac{1}{4} b_{\mathbb{d}} \widehat{I}(\varepsilon) |\log \varepsilon|$$

for all $\varepsilon \in (0, \varepsilon_{\mathbb{d}})$. Finally we have that, reducing $\varepsilon_{\mathbb{d}}$ if necessary, by Assumptions 1.16 (B) and1.16 (C),

$$\begin{aligned} p(Z_i^+(t)) &\geq \tfrac{1}{2} + \mu\left(\tfrac{1}{4} b_{\mathbb{d}} \widehat{I}(\varepsilon)|\log \varepsilon| \wedge \eta\right) \\ &\geq \tfrac{1}{2} + \varepsilon \end{aligned} \tag{3.65}$$

for all $\varepsilon \in (0, \varepsilon_{\mathbb{d}})$. We then combine (3.62), (3.63) and (3.65) exactly as in the proof of Theorem 3.14, to obtain that that, for $\varepsilon \in (0, \varepsilon_{\mathbb{d}})$, $t \in [t_{\mathbb{d}}, t'_{\mathbb{d}}]$, and x such that $d(x,t) \geq b_{\mathbb{d}} \widehat{I}(\varepsilon)|\log \varepsilon|$,

$$\mathbb{P}_x^\varepsilon[\mathbb{V}_p^\times(\boldsymbol{Z}^+(t)) = 1] \geq u_+ - 3\varepsilon^k.$$

The upper bound is obtained using the same approach. □

3.3.3 Propagation of the interface

In this section, we consider the propagation of the interface for the marked majority voting systems. Throughout this section, define

$$\gamma(t) := K_1 e^{K_2 t} \widehat{I}(\varepsilon) |\log \varepsilon|,$$

where the choice of K_1 and K_2 will be clear in the given context.

Proposition 3.39. *Let $\alpha \in (1,2)$, let $l \in \mathbb{N}$ with $l \geq 4$ and fix $I : \mathbb{R}_+ \to \mathbb{R}_+$ satisfying Assumptions 1.37. Let $a_{\mathbb{d}}$ and $t_{\mathbb{d}}$ be as in Proposition 3.37. Then there exist $K_1, K_2 > 0$ such that, for all $\varepsilon \in (0, \varepsilon_{\mathbb{d}})$ and $t \in [t_{\mathbb{d}}, T^*]$, we have*

$$\sup_{x \in \mathbb{R}^{\mathbb{d}}} \left(\mathbb{P}_x^\varepsilon[\mathbb{V}_p^\times(\boldsymbol{Z}^+(t)) = 1] - \mathbb{P}_{d(x,t)+\gamma(t)}^\varepsilon[\mathbb{V}_p^\times(\boldsymbol{B}_{R^\varepsilon}(t)) = 1]\right) \leq \varepsilon^l \tag{3.66}$$

and

$$\sup_{x \in \mathbb{R}^{\mathbb{d}}} \left(\mathbb{P}_x^\varepsilon[\mathbb{V}_p^\times(\boldsymbol{Z}^-(t)) = 0] - \mathbb{P}_{d(x,t)-\gamma(t)}^\varepsilon[\mathbb{V}_p^\times(\boldsymbol{B}_{R^\varepsilon}(t)) = 0]\right) \leq \varepsilon^l. \tag{3.67}$$

By following the pattern already seen in the proof of Proposition 2.15, Proposition 3.39 follows from Lemma 3.40 below. The proof of Theorem 3.35 will then follow easily, and is stated at the end of this section. We defer the lengthy proof of Lemma 3.40 to Section 3.3.4.

Throughout this section, we will extend the domain of $g_\times : [0,1] \to [0,1]$ to all of $\mathbb{R}$. Namely, we set

$$g_\times(p) = \begin{cases} g_\times(0) & \text{if } p < 0 \\ g_\times(p) & \text{if } p \in [0,1] \\ g_\times(1) & \text{if } p > 1. \end{cases}$$

Lemma 3.40. *Let $K_1 > 0$, $\alpha \in (1,2)$, $l \in \mathbb{N}$ with $l \geq 4$, and $I : \mathbb{R}_+ \to \mathbb{R}_+$, and assume Assumptions 1.37 hold. Let t'_{d} be as in Proposition 3.37. Then there exist $K_2 = K_2(K_1, l) > 0$ and $\varepsilon_{\mathrm{d}}(K_1, K_2, l) > 0$ such that, for all $\varepsilon \in (0, \varepsilon_{\mathrm{d}})$, $x \in \mathbb{R}^{\mathrm{d}}$, $s \in [0, (l+1)\varepsilon^2 |\log \varepsilon|]$ and $t \in [t'_{\mathrm{d}}, T^*]$,*

$$\begin{aligned} &E_x\left[g_\times\left(\mathbb{P}^\varepsilon_{d(Z^+_s, t-s)+\gamma(t-s)}[\mathbb{V}^\times(\boldsymbol{B}_{R^\varepsilon}(t-s)) = 1] + \varepsilon^l\right)\right] \\ &\qquad \leq \frac{3}{4}\varepsilon^l + E_{d(x,t)}\left[g_\times\left(\mathbb{P}^{\varepsilon, t-s}_{B(R^\varepsilon_s)+\gamma(t)}[\mathbb{V}^\times(\boldsymbol{B}_{R^\varepsilon}(t-s)) = 1]\right)\right] + \mathbb{1}_{\{s \leq \varepsilon^3\}} \end{aligned} \tag{3.68}$$

and

$$\begin{aligned} &E_x\left[g_\times\left(\mathbb{P}^\varepsilon_{d(Z^-_s, t-s)-\gamma(t-s)}[\mathbb{V}^\times(\boldsymbol{B}_{R^\varepsilon}(t-s)) = 0] + \varepsilon^l\right)\right] \\ &\qquad \leq \frac{3}{4}\varepsilon^l + E_{d(x,t)}\left[g_\times\left(\mathbb{P}^{\varepsilon, t-s}_{B(R^\varepsilon_s)-\gamma(t)}[\mathbb{V}^\times(\boldsymbol{B}_{R^\varepsilon}(t-s)) = 0]\right)\right] + \mathbb{1}_{\{s \leq \varepsilon^3\}}. \end{aligned} \tag{3.69}$$

Proposition 3.39, together with our coupling of $\boldsymbol{W}_{R^\varepsilon}$ to $\boldsymbol{Z}^+$ and $\boldsymbol{Z}^-$, allows us to compare the root vote of the multidimensional subordinated Brownian motion to that of the one-dimensional subordinated Brownian motion.

Corollary 3.41. *Let $\alpha \in (1,2)$, let $l \in \mathbb{N}$ with $l \geq 4$, and fix $I : \mathbb{R}_+ \to \mathbb{R}_+$ satisfying Assumption 1.37. Define a_{d} and t_{d} as above and let $F(\varepsilon)$ be as in (1.35). There exist $K_1, K_2, M > 0$ and $\varepsilon_{\mathrm{d}}(l) > 0$ such that, for all $\varepsilon \in (0, \varepsilon_{\mathrm{d}})$ and $t \in [t_{\mathrm{d}}, T^*]$, we have*

$$\sup_{x \in \mathbb{R}^{\mathrm{d}}} \left(\mathbb{P}^\varepsilon_x[\mathbb{V}^\times_p(\boldsymbol{W}_{R^\varepsilon}(t) = 1] - \mathbb{P}^\varepsilon_{d(x,t)+\gamma(t)}[\mathbb{V}^\times_p(\boldsymbol{B}_{R^\varepsilon}(t)) = 1]\right) \leq F(\varepsilon, M) + \varepsilon^l \tag{3.70}$$

and

$$\sup_{x \in \mathbb{R}^{\mathbb{d}}} \Big(\mathbb{P}_x^\varepsilon[\mathbb{V}_p^\times(\boldsymbol{W}_{R^\varepsilon}(t)) = 0] - \mathbb{P}_{d(x,t)-\gamma(t)}^\varepsilon[\mathbb{V}_p^\times(\boldsymbol{B}_{R^\varepsilon}(t)) = 0] \Big) \leq F(\varepsilon, M) + \varepsilon^l. \tag{3.71}$$

Proof. This follows immediately from Proposition 3.28 and Proposition 3.39 using that, for $t \geq \mathcal{O}(\varepsilon^2|\log\varepsilon|)$, we have $e^{-\frac{t}{\varepsilon^2}} \leq \varepsilon$, so, for M_2 as in Proposition 3.28, there exists $M > 0$ such that, for ε sufficiently small,

$$M_1 e^{-\frac{t}{\varepsilon^2}} + F(\varepsilon, M_2) \leq F(\varepsilon, M). \qquad \square$$

With this, we can now prove Theorem 3.35.

Proof of Theorem 3.35. Set $c_{\mathbb{d}}(l) := c_1(l) + K_1 e^{K_2 T^*}$. Then, for any $x \in \mathbb{R}^{\mathbb{d}}$, and for any $t \in [t_{\mathbb{d}}, T^*]$ such that $d(x,t) \leq -c_{\mathbb{d}}(l)\widehat{I}(\varepsilon)|\log\varepsilon|$, we have

$$d(x,t) + K_1 e^{K_2 t}\widehat{I}(\varepsilon)|\log\varepsilon| \leq -c_1(l)\widehat{I}(\varepsilon)|\log\varepsilon|.$$

Then, by Theorem 3.14, reducing $\varepsilon_{\mathbb{d}}$ if necessary so that $\varepsilon_{\mathbb{d}} < \varepsilon_1(l)$, we have, by Proposition 3.41, that

$$\mathbb{P}_x^\varepsilon[\mathbb{V}_p^\times(\boldsymbol{W}_{R^\varepsilon}(t)) = 1)] \leq u_- + \varepsilon^l + F(\varepsilon, M)$$

for all such x and t.

Reduce ε as necessary so that $\varepsilon^l + F(\varepsilon, M) \leq u_- + F(\varepsilon, M+1)$. Similarly, for x and t such that $d(x,t) \geq c_{\mathbb{d}}(l)\widehat{I}(\varepsilon)|\log\varepsilon|$, we have, by Theorem 3.14 and (3.67), that $\mathbb{P}_x^\varepsilon[\mathbb{V}_p^\times(\boldsymbol{W}_{R^\varepsilon}(t)) = 1)] \geq u_+ - F(\varepsilon, M+1)$. $\square$

Proof of Theorem 1.39. This follows immediately by combining Theorem 3.35 and Corollary 3.34. $\square$

3.3.4 Proof of Lemma 3.40

Following the original proof of [EFP17, Lemma 2.18], we consider the two cases $|d(x,t)| \leq D\widehat{I}(\varepsilon)|\log\varepsilon|$ and $|d(x,t)| \geq D\widehat{I}(\varepsilon)|\log\varepsilon|$ separately, for some large $D > 0$. Since neither the one-dimensional process $B(R_s^\varepsilon)$ nor the multidimensional process Z^+ (or Z^-) travel further than a distance $\mathcal{O}(\widehat{I}(\varepsilon)|\log\varepsilon|)$ in time $s = \mathcal{O}(\varepsilon^2|\log(\varepsilon)|)$ (with high probability), if D is sufficiently large and $d(x,t) \leq -D\widehat{I}(\varepsilon)|\log\varepsilon|$, then Theorem 3.14 implies the left hand side of (3.68) is bounded above by ε^k, and if $d(x,t) \geq D\widehat{I}(\varepsilon)|\log\varepsilon|$, the right hand side of (3.68) will at least 1. When $|d(x,t)| \leq D\widehat{I}(\varepsilon)|\log\varepsilon|$, we apply Theorem 3.27 so that, with probability at least $1-\varepsilon^{l+1}$,

$$d(Z_s^+, t-s) \leq B(R_s^\varepsilon) + \mathcal{O}(\widehat{I}(\varepsilon)|\log\varepsilon|)s.$$

By the monotonicity of $g_\times$, we can then bound the left hand side of (3.68) by

$$E_{d(x,t)}\left[g_\times\left(\mathbb{P}^\varepsilon_{B(R_s^\varepsilon)+\gamma(t-s)+\mathcal{O}(s)\widehat{I}(\varepsilon)|\log\varepsilon|}[\mathbb{V}_p^\times(\boldsymbol{B}_{R^\varepsilon}(t))=1]+\varepsilon^l\right)\right]+\mathcal{O}(\varepsilon^{l+1}). \tag{3.72}$$

We then control (3.72) by considering two cases: when the argument of $g_\times$ in (3.72) is bounded away from $\frac{1}{2}$, and when it is close to $\frac{1}{2}$. In the former case, we use that $|g_\times'(p)| < \frac{2}{3}$ when p is bounded far enough away from $\frac{1}{2}$, together with the monotonicty of $g_\times$, to obtain (3.68). In the second case, we apply the slope of the interface result, Corollary 3.24, to bound $\mathbb{P}_z^\varepsilon[\mathbb{V}_p^\times(\boldsymbol{B}_{R^\varepsilon}(t))=1] - \mathbb{P}_w^\varepsilon[\mathbb{V}_p^\times(\boldsymbol{B}_{R^\varepsilon}(t))=1]$ when $z = B(R_s^\varepsilon)+\gamma(t-s)+\mathcal{O}(s)\widehat{I}(\varepsilon)|\log\varepsilon|$ and $w = z+\gamma(t)$.

Proof of Lemma 3.40. For all $u \geq 0$ and $z \in \mathbb{R}$, let

$$\mathbb{Q}_z^{\varepsilon,u} = \mathbb{P}_z^\varepsilon[\mathbb{V}^\times(\boldsymbol{B}_{R^\varepsilon}(u))=1].$$

Let C_0 be as in (3.43), and and c_1 be defined as in Theorem 3.14. Let

$$R := 2c_1(l) + 4(l+1)\mathbb{d} + 1, \tag{3.73}$$

and fix K_2 such that

$$K_1(K_2 - C_0) - C_0R - C_1 = c_1(1). \tag{3.74}$$

At first, we let $\varepsilon_{\mathbb{d}}(l) = \varepsilon_1(l)$ where $\varepsilon_1(l)$ is defined in Theorem 3.14.

Following the proof of [EFP17, Lemma 2.18], we begin by estimating the probability that a $\mathbb{d}$-dimensional subordinated Brownian motion moves further than at distance of order $\widehat{I}(\varepsilon)|\log\varepsilon|$ in time $s \leq (l+1)\varepsilon^2|\log\varepsilon|$. Define the event

$$A_x = \left\{\sup_{u\in[0,R_s^\varepsilon]} ||W_u - x|| \leq 2(l+1)\widehat{I}(\varepsilon)|\log\varepsilon|\right\}.$$

Then, bounding $\|W(u)-x\|$ by the sum of the moduli of $\mathbb{d}$ one-dimensional Brownian motions, and applying Lemma B.3,

$$
\begin{aligned}
P_x[A_x^c] &\leq 2\mathbb{d}P_0\left[\sup_{u\in[0,R_s^\varepsilon]} B_u > 2(l+1)\widehat{I}(\varepsilon)|\log\varepsilon|\right] \\
&\leq 2\mathbb{d}P_0\left[\sup_{u\in[0,(l+2)\widehat{I}(\varepsilon)^2|\log\varepsilon|]} B_u > 2(l+1)\widehat{I}(\varepsilon)|\log\varepsilon|\right] + 2\mathbb{d}\varepsilon^{l+1} \\
&\leq 4\mathbb{d}P_0\left[B_1 > 2((l+1)|\log\varepsilon|)^{1/2}\right] + 2\mathbb{d}\varepsilon^{l+1} \\
&\leq 6\mathbb{d}\varepsilon^{l+1}, \qquad (3.75)
\end{aligned}
$$

where the second inequality follows by the reflection principle, and the final inequality follows by identical arguments to those in the proof of Lemma 3.21. Now consider the cases

(i) $d(x,t) \leq -(2c_1(l) + 2(l+1)\mathbb{d} + K_1 e^{K_2(t-s)})\widehat{I}(\varepsilon)|\log\varepsilon|$

(ii) $d(x,t) \geq (2c_1(l) + 2(l+1)\mathbb{d} + K_1 e^{K_2(t-s)})\widehat{I}(\varepsilon)|\log\varepsilon|$

(iii) $|d(x,t)| \leq (2c_1(l) + 2(l+1)\mathbb{d} + K_1 e^{K_2(t-s)})\widehat{I}(\varepsilon)|\log\varepsilon|$.

Case (i): By (3.43), there exist $v_0, V_0 > 0$ such that, if $s \leq v_0$ and $x \in \mathbb{R}^{\mathbb{d}}$, then

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

Reduce $\varepsilon_{\mathbb{d}}$ if necessary to ensure that, for all $\varepsilon \in (0,\varepsilon_{\mathbb{d}})$, $(l+1)\varepsilon^2|\log\varepsilon| \leq v_0$. Then if A_x occurs,

$$
\begin{aligned}
&d(W(R_s^\varepsilon), t-s) + K_1 e^{K_2(t-s)}\widehat{I}(\varepsilon)|\log\varepsilon| \qquad (3.76)\\
&\quad\leq -(2c_1(l) + 2(l+1)\mathbb{d})\widehat{I}(\varepsilon)|\log\varepsilon| + |d(W(R_s^\varepsilon), t-s) - d(x,t)| \\
&\quad\leq -(2c_1(l) + 2(l+1)\mathbb{d})\widehat{I}(\varepsilon)|\log\varepsilon| + |d(x,t) - d(x,t-s)| + ||W(R_s^\varepsilon) - x|| \\
&\quad\leq -2c_1(l)\widehat{I}(\varepsilon)|\log\varepsilon| + V_0(l+1)\varepsilon^2|\log\varepsilon|. \qquad (3.77)
\end{aligned}
$$

By Assumption 1.37 (B), we may reduce $\varepsilon_{\mathbb{d}}$ if necessary so that

$$d(W(R_s^\varepsilon), t-s) + K_1 e^{K_2(t-s)}\widehat{I}(\varepsilon)|\log\varepsilon| \leq -\tfrac{3}{2}c_1(l)\widehat{I}(\varepsilon)|\log\varepsilon|,$$

for all $\varepsilon \in (0, \varepsilon_{\mathbb{d}})$. Then, since $d(Z_s^+, t-s) \le d(W(R_s^\varepsilon), t-s) + D_0(l+2)\widehat{I}(\varepsilon)^2|\log\varepsilon|$, reducing ε further if necessary, we have

$$d(Z_s^+, t-s) + K_1 e^{K_2(t-s)}\widehat{I}(\varepsilon)|\log\varepsilon| \le -c_1(l)\widehat{I}(\varepsilon)|\log\varepsilon|.$$

So, by Theorem 3.14 and the definition of $g_\times$,

$$\begin{aligned} E_x\left[g_\times\left(\mathbb{Q}^{\varepsilon,t-s}_{d(Z_s^+,t-s)+\gamma(t-s)} + \varepsilon^l\right)\right] &\le E_x[g_\times(u_- + 2\varepsilon^l)\mathbb{1}_{A_x}] + P_x(A_x^c) \\ &\le u_- + 6\mathbb{d}\varepsilon^{l+1} + 12\varepsilon^l b_I, \end{aligned}$$

where the last line follows by calculating $g_\times(u_- + 2\varepsilon^l)$ explicitly and reducing $\varepsilon_{\mathbb{d}}$ if necessary.

Next, recall that $g_\times(p) = g((1-b_I)p + \frac{b_I}{2})$ for $p \in [0,1]$. So

$$g_\times'(p) = 6(1-b_I)\left((1-b_I)p + \tfrac{b_I}{2}\right)\left(1 - \left((1-b_I)p + \tfrac{b_I}{2}\right)\right).$$

Hence, if

$$(1-b_I)(p+\delta) + \tfrac{b_I}{2} \le \tfrac{1}{9} \quad \text{or} \quad (1-b_I)p + \tfrac{b_I}{2} \ge \tfrac{8}{9} \tag{3.78}$$

then

$$g_\times(p+\delta) \le g_\times(p) + \tfrac{2}{3}\delta. \tag{3.79}$$

From Proposition 3.36 (since $t - s \ge \rho_{\mathbb{d}}^+(l)\varepsilon^2|\log\varepsilon|$) and (3.79), if we reduce ε as necessary, we obtain that

$$E_{d(x,t)}\left[g_\times\left(\mathbb{Q}^{\varepsilon,t-s}_{Z_s^+ + \gamma(t)}\right)\right] \ge u_- - \tfrac{2}{3}\varepsilon^l.$$

By choosing ε small enough that $6\mathbb{d}\varepsilon^{l+1} + 12\varepsilon^l b_I + \frac{2}{3}\varepsilon^l \le \frac{3}{4}\varepsilon^l$, equation (3.68) holds in this case.

Case (ii): Suppose now that $d(x,t) \ge (2c_1(\alpha) + 2(l+1)\mathbb{d} + K_1 e^{K_2(t-s)})\widehat{I}(\varepsilon)|\log\varepsilon|$. Using this, together with a similar argument to that used to obtain (3.75), we have

$$P_{d(x,t)}[|B(R_s^\varepsilon)| \le (c_1(l)-1)\widehat{I}(\varepsilon)|\log\varepsilon|] \le \varepsilon^{l+1}. \tag{3.80}$$

It follows that

$$\begin{aligned}
E_{d(x,t)}\left[g_\times\left(\mathbb{Q}^{\varepsilon,t-s}_{B(R^\varepsilon_s)+\gamma(t)}\right)\right] &\geq E_{d(x,t)}\left[g_\times\left(\mathbb{Q}^{\varepsilon,t-s}_{B(R^\varepsilon_s)+\gamma(t)}\right)\mathbb{1}_{\{B(R^\varepsilon_s)\geq c_1(l)\widehat{I}(\varepsilon)|\log\varepsilon|\}}\right] \\
&\geq g_\times(u_+-\varepsilon^l)-\varepsilon^{l+1} \\
&\geq u_+-\varepsilon^{l+1}-12\varepsilon^l b_I,
\end{aligned}$$

where the second inequality follows by Theorem 3.14, and in the third inequality we expand $g_\times(u_+-\varepsilon^l)$ and reduce $\varepsilon_{\mathbb{d}}$ if necessary. From (3.60) and Proposition 3.36, we can get that, for ε small enough,

$$E_x\left[g_\times\left(\mathbb{Q}^{\varepsilon,t-s}_{d(Z^+_s,t-s)+\gamma(t-s)}+\varepsilon^l\right)\right]\leq u_+ + \frac{2}{3}\varepsilon^l$$

Hence (reducing ε if necessary) it is clear (3.68) holds in this case.

Case (iii): Finally, suppose $|d(x,t)| \leq (2c_1(\alpha)+2(l+1)\mathbb{d}+K_1e^{K_2(t-s)})\widehat{I}(\varepsilon)|\log(\varepsilon)|$. If A_x occurs and $u\in[0,(l+2)\widehat{I}(\varepsilon)^2|\log\varepsilon|]$,

$$\begin{aligned}
|d(W_u,t-u)| &\leq ||W_u-x||+|d(x,t)|+|d(x,t)-d(x,t-u)| \\
&\leq (2c_1(1)+4(l+1)\mathbb{d}+K_1e^{K_2(t-s)})\widehat{I}(\varepsilon)|\log\varepsilon|+V_0(l+2)\widehat{I}(\varepsilon)^2|\log(\varepsilon)|.
\end{aligned}$$

Therefore (reducing ε if necessary), since $\mathbb{P}(R^\varepsilon_s > (l+2)\widehat{I}(\varepsilon)^2|\log\varepsilon|)\leq\varepsilon^{l+1}$, we have that, with probability at least $1-\varepsilon^{l+1}$,

$$|d(W(R^\varepsilon_s),t-s)|\leq(R+K_1e^{K_2(t-s)})\widehat{I}(\varepsilon)|\log\varepsilon|$$

for R as in (3.73). Set

$$\beta=(R+K_1e^{K_2(t-s)})\widehat{I}(\varepsilon)|\log\varepsilon|. \tag{3.81}$$

Reduce $\varepsilon_{\mathbb{d}}$ if necessary so that, for all $\varepsilon\in(0,\varepsilon_{\mathbb{d}})$, we have $\beta\leq c_0/2$, for c_0 as in (3.41). Recall that

$$T_\beta=\inf(\{s\in[0,(l+1)\varepsilon^2|\log\varepsilon|):|d(W_s,t-s)|>\beta\}\cup\{t\}).$$

Note that $\mathbb{P}(R^\varepsilon_s>T_\beta)\leq 2\varepsilon^{l+1}$. Indeed, if A_x occurs, then $T_\beta>R^\varepsilon_s$ with probability at least $1-\varepsilon^{l+1}$. Therefore, by Theorem 3.27 (and by reducing ε if necessary so that $T_\beta<t$),

$$d(Z^+_s,t-s)\leq B(R^\varepsilon_s)+C_0\beta s \tag{3.82}$$

with probability at least $1-2\varepsilon^{l+1}$. Then, by monotonicity of $g_\times$ and by (3.82), if we partition over $\{R_s^\varepsilon > T_\beta\}$ and A_x, we obtain

$$E_x\left[g_\times\left(\mathbb{Q}^{\varepsilon,t-s}_{d(Z_s^+,t-s)+\gamma(t-s)}+\varepsilon^l\right)\right] \le E_{d(x,t)}\left[g_\times\left(\mathbb{Q}^{\varepsilon,t-s}_{B(R_s^\varepsilon)+C_0\beta s+\gamma(t-s)}+\varepsilon^l\right)\right] + (2+6\mathbb{d})\varepsilon^{l+1}. \tag{3.83}$$

Let

$$D := \left\{\left|\mathbb{Q}^{\varepsilon,t-s}_{B(R_s^\varepsilon)+C_0\beta s+\gamma(t-s)} - \frac{1}{2}\right| \le \frac{5}{12}\right\}.$$

We consider D and D^c separately to bound the right hand side of (3.83). First suppose the event D occurs. Then, by the definition of β (3.81),

$$\begin{aligned}
K_1 e^{K_2 t}\widehat{I}(\varepsilon)|\log\varepsilon| &- \left(C_0\beta s + K_1 e^{K_2(t-s)}\widehat{I}(\varepsilon)|\log\varepsilon|\right)\\
&= \left(K_1 e^{K_2(t-s)}\left(e^{K_2 s}-1-C_0 s\right) - C_0 R s\right)\widehat{I}(\varepsilon)|\log\varepsilon|\\
&\ge \left(K_1(K_2-C_0)-C_0 R\right) s\widehat{I}(\varepsilon)|\log\varepsilon|\\
&= c_1(1) s\widehat{I}(\varepsilon)|\log\varepsilon|
\end{aligned} \tag{3.84}$$

where the final equality follows by (3.74). Reducing $\varepsilon_{\mathbb{d}}$ if necessary so that $\varepsilon_{\mathbb{d}} < \min(\varepsilon_1(1), \frac{1}{24})$, for $\varepsilon \in (0,\varepsilon_{\mathbb{d}})$ we may apply Corollary 3.24 with $z = B(R_s^\varepsilon) + C_0\beta s + K_1 e^{K_2(t-s)}\widehat{I}(\varepsilon)|\log\varepsilon|$ and $w = z + c_1(1)s\widehat{I}(\varepsilon)|\log\varepsilon| \le B(R_s^\varepsilon)+\gamma(t)$ to give

$$\mathbb{Q}^{\varepsilon,t-s}_{B(R_s^\varepsilon)+C_0\beta s+\gamma(t-s)}\mathbb{1}_D \le \left(\mathbb{Q}^{\varepsilon,t-s}_{B(R_s^\varepsilon)+\gamma(t)} - \tfrac{1}{48}s\right)\mathbb{1}_D. \tag{3.85}$$

Now suppose the event D^c occurs. Reducing $\varepsilon_{\mathbb{d}}$ if necessary so that $\frac{1}{12} < \frac{1}{9} - \varepsilon^l(1-b_I) - \frac{b_I}{2}$, which implies (3.78) for $\delta = \varepsilon^l$. Thus, for $\varepsilon \in (0,\varepsilon_{\mathbb{d}})$, we have

$$\begin{aligned}
g_\times\left(\mathbb{Q}^{\varepsilon,t-s}_{B(R_s^\varepsilon)+C_0\beta s+\gamma(t-s)}+\varepsilon^l\right)\mathbb{1}_{D^c} &\le \left(g_\times\left(\mathbb{Q}^{\varepsilon,t-s}_{B(R_s^\varepsilon)+C_0\beta s+\gamma(t-s)}\right)+\tfrac{2}{3}\varepsilon^l\right)\mathbb{1}_{D^c}\\
&\le \left(g_\times\left(\mathbb{Q}^{\varepsilon,t-s}_{B(R_s^\varepsilon)+\gamma(t)}\right)+\tfrac{2}{3}\varepsilon^l\right)\mathbb{1}_{D^c},
\end{aligned} \tag{3.86}$$

where the first inequality follows by (3.79) and the second inequality by (3.84) and the monotonicity of $g_\times$. Putting (3.85) and (3.86) into (3.83), and since $2+6\mathbb{d} \le 8\mathbb{d}$, we obtain

$$\begin{aligned}
E_x\left[g_\times(\mathbb{Q}^{\varepsilon,t-s}_{d(Z_s^+,t-s)+\gamma(t-s)}+\varepsilon^l)\right] &\le E_{d(x,t)}\left[g_\times\left(\mathbb{Q}^{\varepsilon,t-s}_{B(R_s^\varepsilon)+\gamma(t)} - \tfrac{1}{48}s+\varepsilon^l\right)\mathbb{1}_D\right]\\
&\quad + E_{d(x,t)}\left[\left(g_\times\left(\mathbb{Q}^{\varepsilon,t-s}_{B(R_s^\varepsilon)+\gamma(t)}\right)+\tfrac{2}{3}\varepsilon^l\right)\mathbb{1}_{D^c}\right] + 8\mathbb{d}\varepsilon^{l+1}\\
&\le E_{d(x,t)}\left[g_\times\left(\mathbb{Q}^{\varepsilon,t-s}_{B(R_s^\varepsilon)+\gamma(t)}\right)\right] + \tfrac{2}{3}\varepsilon^l + \varepsilon^l\mathbb{1}_{\{\frac{1}{48}s\le\varepsilon^l\}} + 8\mathbb{d}\varepsilon^{l+1},
\end{aligned}$$

where, in the final inequality, we use that $g'_\times(p) \leq \frac{3}{2}$ for all $p \in [0,1]$. Further reducing $\varepsilon_{\mathbb{d}}$ if necessary so that $8\mathbb{d}\varepsilon^{l+1} \leq \frac{1}{12}\varepsilon^l$ and $48\varepsilon^l \leq \varepsilon^3$ for all $\varepsilon \in (0, \varepsilon_{\mathbb{d}})$ gives the result. $\square$

4

Conclusion

We now present some points on which the results of this thesis are inconclusive, and which could be the subject of future research.

In Chapter 2, and in particular Theorem 1.34, to study the effect of genetic drift we have focused on two scaling regimes, the weak noise/selection ratio regime and the strong noise/selection ratio regime. However, this clearly does not cover all the possible scalings one can consider. A first generalisation is to write explicitly how to relax the conditions on the strong noise/selection regime, as we mentioned in Remark 1.35. That is, we scale as in the strong noise/selection regime but change the scaling limits of $\boldsymbol{s}_n$ to be

$$\begin{cases} \boldsymbol{s}_n n^{2\beta} \to K & \lim_{n\to\infty} u_n \log n < \infty \text{ or } \mathbb{d} \geq 3, \\ \frac{\boldsymbol{s}_n n^{2\beta}}{u_n \log n} \to K & u_n \log n \to \infty \text{ and } \mathbb{d} = 2, \end{cases} \tag{4.1}$$

for some $K > 0$. We call this the *mild noise/selection ratio regime.*

We expect that, in this case, we can recover a version of Theorem 1.34 but with $\rho_* < (\mathbb{d} - 1)/\nu$. The fact that the limits in (4.1) are greater than 0 means we expect to see more than one dual particle undergoing a meaningful excursion, but the fact that $u_n \log(n)$ has a non-trivial limit means that a large proportion of particles should coalesce. Thus, we expect selection to have a weaker effect on the population, meaning that we see an *effective* asymmetric selection towards aa that is strictly smaller than ν, supporting our statement. In order to obtain a formal proof of this, we believe the starting point would be to adapt the techniques in [Eth+17], which deals with scaling regimes in which noise and selection both have an effect present in the limit. The main difficulty in such an adaptation is that the selection is not the same, and it needs great care to check which parts of the proofs need rewriting.

Another point for further research would be to prove an analogue of Theorem 1.39 with genetic drift, in the same vein as Theorem 1.20. For that, we can consider the SLFVS with long-range dispersal: this means the radius of events is chosen according to

$$\mu(dr) = \frac{1_{r\geq 1}}{r^{d+\alpha+1}}dr.$$

Once again we consider the dual process, $\Xi^n(t)$, given by the obvious analogue of the one in Definition 1.13. One then should scale, space, time, the impact parameters u_n and the selection parameter $\boldsymbol{s}_n$ so that,

1. ancestral lineages $\xi_t^{n,j}$ move, approximately, as α-stable processes running at speed $I(\varepsilon_n)^{\alpha-2}$,
2. coalescence events in $\Xi^n(t)$ occur with *arbitrarily low* probability,
3. selective events in $\Xi^n(t)$ occur at rate ε_n^{-2}.

Comparing with Theorem 1.39, we can convince ourselves that Theorem 1.20 should also hold for the SLFV with long-range dispersal in the said scaling. The difficulty then is just choosing the parameters in the right way. Work on this is underway and we expect it to be in the DPhil thesis of Kimberly Becker, with whom we worked on the results of Chapter 3.

Furthermore, it would be interesting to see if there are other possible scaling regimes for the SLFVS with long-range dispersal. A first way of exploring this would be to try and get a strong noise/selection ratio regime, as we did in Section 1.6.1. However, our methods are not easy to adapt to the long-range case. In our proof of Proposition 2.38 we exploit that if two particles of the dual process are at a distance bigger than $2\mathcal{R}_n$, then they evolve independently. This is a consequence of the SLFVS having events with radii bounded by $\mathcal{R}_n$. For the SFLVS with long-range dispersal, this no longer holds, as events can have an unbounded radius. We believe that a way of trying to understand the relationship of dual particles, in this case, is

Lemma 5.3 in [BEV13], which roughly characterises the coalescence time of dual particles under a particular scaling. If one could generalise this result, one could understand which scalings produce particles that coalesce *instantly*. With this, a strong noise/selection regime for the SFLVS with long-range dispersal would follow.

Another interesting objective would be to obtain a generalisation of Theorem 1.39 when $\alpha < 1$. However, we do not expect our techniques to be adaptable to this case, since, for $\alpha < 1$, big jumps (which our techniques erase) are now the main force of translation of the α-stable process. More interestingly, we no longer expect to see motion by mean curvature flow, but rather by fractional curvature flow [CS10], [Imb09], suggesting even more that new techniques are needed.

Bibliography

[AC79] S. M. Allen and J. W. Cahn. “A microscopic theory for antiphase boundary motion and its application to antiphase domain coarsening”. In: *Acta metallurgica* 27.6 (1979), pp. 1085–1095.

[AHM08] M. Alfaro, H. Hilhorst, and H. Matano. “The singular limit of the Allen–Cahn equation and the FitzHugh–Nagumo system”. In: *J. Differ. Equ.* 245.2 (2008), pp. 505–565.

[Alf+18] M. Alfaro, D. Antonopoulou, G. Karali, and H. Matano. “Generation of fine transition layers and their dynamics for the stochastic Allen-Cahn equation”. 2018. URL: `arXiv:1812.03804v3`.

[AR01] S. Asmussen and J. Rosiński. “Approximations of small jumps of Lévy processes with a view towards simulation”. In: *J. Appl. Probab.* 38.2 (2001), pp. 482–493. DOI: `10.1017/s0021900200019987`.

[Bar96] S. C. H. Barrett. “The reproductive biology and genetics of island plants”. In: *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences* 351.1341 (1996), pp. 725–733.

[BBC16] H. Berestycki, J. Bouhours, and G. Chapuisat. “Front blocking and propagation in cylinders with varying cross section”. In: *Cacl. Var. Partial Dif.* 55.3 (2016), p. 44. DOI: `10.1007/s00526-016-0962-2`.

[BEV10] N. Barton, A. M. Etheridge, and A. Véber. “A new model for evolution in spatial continuum”. In: *Electron. J. Probab.* 15 (2010), pp. 162–216.

[BEV13] N. Berestycki, A. M. Etheridge, and A. Véber. “Large scale behaviour of the spatial Λ-Fleming–Viot process”. In: *Ann. Inst. H. Poincaré Probab. Statist.* 49.2 (2013), pp. 374–401. DOI: `doi.org/10.1214/11-AIHP471`.

[BH89] N. Barton and G. M. Hewitt. “Adaptation, speciation and hybrid zones”. In: *Nature* 341 (1989), pp. 497–503.

[BH91] R. F. Bass and P. Hsu. “Some potential theory for reflecting Brownian motion in Holder and Lipschitz domains.” In: *Ann. Probab.* 19.2 (1991), pp. 486–508.

[Bil95] P. Billingsley. *Probability and Measure.* Wiley, 1995.

[BK91] L. Bronsard and R. V. Kohn. “Motion by mean curvature as the singular limit of Ginzburg-Landau dynamics”. In: *J. Differ. Equ.* 90.2 (1991), pp. 211–237.

[Bla+20] J. Blath, A. González, N. Kurt, and M. Wilke-Berenger. “The seed bank coalescent with simultaneous switching”. In: *Electron. J. Probab.* 25 (2020), pp. 1–21. DOI: `10.1214/19-EJP401`.

[Boc55] S. Bochner. *Harmonic analysis and the theory of probability.* University of California Press, Berkeley-Los Angeles, Calif., 1955, pp. viii+176.

[Bro90] L. Bronsard. “On the slowness of phase boundary motion in one space dimension”. In: *STIN* 91 (1990), p. 21480.

[Buc05] J.A. Buckland. “Mean curvature flow with free boundary on smooth hypersurfaces”. In: *J. fur Reine Angew. Math.* 586 (2005), pp. 71–90.

[Bug07] R. Buggs. “Empirical study of hybrid zone movement”. In: *Heredity* 99 (2007), pp. 301–312.

[Bus07] J. Buschbom. “Migration between continents: Geographical structure and long-distance gene flow in *Porpidia flavicunda* (lichen-forming *Ascomycota*).” In: *Mol. Ecol.* 16.9 (2007), pp. 1835–1846.

[CDP99] J. T. Cox, R. Durrett, and E. A. Perkins. “Rescaled particle systems converging to super–Brownian motion”. In: *Perplexing problems in probability.* Vol. 44. Progress in Probability. Birkhäuser, Boston, MA, 1999, pp. 269–284.

[CG05] G. Chapuisat and E. Grenier. “Existence and nonexistence of traveling wave solutions for a bistable reaction-diffusion equation in an infinite cylinder whose diameter is suddenly increased”. In: *Commun. Part. Diff. Eq.* 30.12 (2005), pp. 1805–1816. DOI: 10.1080/03605300500300006.

[Che92] X. Chen. “Generation and propagation of interfaces for reaction-diffusion equations”. In: *J. Diff. Equ.* 96 (1992), pp. 116–141.

[CK03] Z. Chen and T. Kumagai. “Heat kernel estimates for stable-like processes on d-sets”. In: *Stoch. Proc. Appl.* 108.1 (2003), pp. 27–62. ISSN: 0304-4149. DOI: https://doi.org/10.1016/S0304-4149(03)00105-4.

[CMS00] M. L. Cain, B. G. Milliga, and A. E. Strand. “Long-distance seed dispersal in plant populations”. In: *Am. J. Bot.* 87.9 (2000), pp. 1217–1227.

[Coh+99] I. Cohen, I Golding, Y. Kozlovsky, E. Ben-Jacob, and I. Ron. “Continuous and discrete models of cooperation in complex bacterial colonies”. In: *Fractals* 7 (1999), pp. 235–247.

[Cor+20] F. Cordero, A. González Casanova, J. Schweinsberg, and M. Wilke-Berenguer. “Λ-coalescents arising in populations with dormancy”. 2020. URL: arXiv:2009.09418v1.

[CP05] J. T. Cox and E. A. Perkins. “Rescaled Lotka–Volterra models converge to super–Brownian motion”. In: *Ann. Probab.* 33 (2005), pp. 904–947.

[CR07] S. Cohen and J. Rosinski. “Gaussian approximation of multivariate Lévy processes with applications to simulation of tempered stable processes”. In: *Bernoulli* 13.1 (2007), pp. 195–210.

[CS10] L. A. Caffarelli and P. E. Souganidis. “Convergence of Nonlocal Threshold Dynamics Approximations to Front Propagation”. In: *Arch. Rational Mech. Anal.* 195 (2010), pp. 1–23. DOI: 10.1007/s00205-008-0181-x.

[CV19] F. Cordero and G. Véchambre. "Moran model and Wright-Fisher diffusion with selection and mutation in a one-sided random environment". 2019. URL: arXiv:1911.12089v2.

[DFL86] A. De Masi, P. A. Ferrari, and J. L. Lebowitz. "Reaction-diffusion equations for interacting particle systems". In: *J. Stat. Phys.* 44 (1986), pp. 589–644. DOI: 10.1007/BF01011311.

[Dia13] E. H. A. Dia. "Error bounds for small jumps of Lévy processes". In: *Adv. Appl. Probab.* 45.1 (2013), pp. 86–105.

[DK99] P. Donnelly and T.G. Kurtz. "Genealogical processes for Fleming-Viot models with selection and recombination". In: *Ann. Appl. Probab.* 9.4 (1999), pp. 1091–1148.

[Dra02] S. S. Dragomir. "Some Gronwall type inequalities and applications". 2002. URL: https://rgmia.org/papers/monographs/standard.pdf.

[EFP17] A. M. Etheridge, N. Freeman, and S. Penington. "Branching Brownian motion, mean curvature flow and the motion of hybrid zones". In: *Electron. J. Probab.* 22 (2017).

[EK19] A. M. Etheridge and T. G. Kurtz. "Genealogical constructions of population models". In: *Ann. Probab.* 47.4 (2019), pp. 1827–1910.

[EK86] S. N. Ethier and T. G. Kurtz. *Markov processes: characterization and convergence.* Wiley, 1986. DOI: 10.1002/9780470316658.

[EPB16] M. Esser, S. Probst, and E. Baake. "Partitioning, duality, and linkage disequilibria in the Moran model with recombination". In: *J. Math. Biol.* 73 (2016), pp. 161–197. DOI: 10.1007/s00285-015-0936-6.

[Eth+17] A.M. Etheridge, N. Freeman, S. Penington, and D. Straulino. "Branching Brownian motion and selection in the spatial Λ-Fleming–Viot Process". In: *Ann. Appl. Probab.* 27.5 (2017), pp. 2605–2645. DOI: 10.1007/s00208-008-0258-8.

[Eth08] A. M. Etheridge. "Drift, draft and structure: some mathematical models of evolution". In: *Banach Center Publ.* 80 (2008), pp. 121–144.

[Eth11] A. M. Etheridge. *Some Mathematical Models from Population Genetics: École d'Été de Probabilités de Saint-Flour XXXIX-2009.* Vol. 39. Springer Science & Business Media, 2011. DOI: 10.1007/978-3-642-16632-7.

[Eth19] A. M. Etheridge. "Spatial population models". In: *Ensaios Mat.* Vol. 34. Sociedade Brasileira de Matemática, 2019.

[EV12] A. M. Etheridge and A. Véber. "The spatial Λ-Fleming–Viot process on a large torus: Genealogies in the presence of recombination." In: *Ann. Appl. Probab.* 22.6 (2012), pp. 2165–2209. DOI: 10.1214/12-AAP842.

[EVY20] A. M. Etheridge, A. Véber, and F. Yu. "Rescaling limits of the spatial Lambda-Fleming-Viot process with selection". In: *Electron. J. Probab.* 25 (2020), pp. 1–89. DOI: 10.1214/20-EJP523.

[Fel75] J. Felsenstein. "A pain in the torus: some difficulties with models of isolation by distance". In: *Am. Nat.* (1975), pp. 359–368.

[FLS20] F. Foutel-Rodier, A. Lambert, and E. Schertzer. "Kingman's coalescent with erosion". In: *Electron. J. Probab.* 25 (2020), pp. 1–33. DOI: 10.1214/20-EJP450.

[Fun99] T. Funaki. "Singular limit for stochastic reaction-diffusion equation and generation of random interfaces". In: *Acta Mech. Sin.* 15 (1999), pp. 407–438. DOI: https://doi.org/10.1007/BF02650735.

[GH86] M. Gage and R. Hamilton. "The heat equation shrinking convex plane curves". In: *J. Diff. Geom.* 23 (1986), pp. 417–491.

[Gon11] P. Gonçalves. "Microscopic Dynamics for the Porous Medium Equation". In: *Dynamics, Games and Science II.* Springer Berlin Heidelberg, 2011, pp. 387–392. ISBN: 978-3-642-14788-3.

[Goo18] M. Gooding. "Long term behaviour of spatial population models with heterozygous or asymmetric homozygous selection". PhD thesis. Oxford University, 2018. URL: https://ora.ox.ac.uk/objects/uuid:ef35b918-35ad-4f90-98ea-325fc692f1eb.

[Gra87] M. A. Grayson. "The heat equation shrinks embedded plane curves to round points". In: *J. Diff. Geom.* 26 (1987), pp. 285–311.

[GS02] A. Grigor'yan and L. Saloff-Coste. "Hitting probabilities for Brownian motion on Riemannian manifolds". In: *J. Math. Pures. Appl.* 81.2 (2002), pp. 115–142. DOI: https://doi.org/10.1016/S0021-7824(01)01244-2.

[GZ15] C. Gui and M. Zhao. "Traveling wave solutions of Allen–Cahn equation with a fractional Laplacian". In: *Ann. I. H. Poincare-AN.* Vol. 32. 4. Elsevier. 2015, pp. 785–812.

[Har08] G. H. Hardy. "Mendelian proportions in a mixed population". In: *Science* 28.706 (1908), pp. 49–50.

[HD20] X. Huang and R. Durrett. "Motion by mean curvature in interacting particle systems". 2020. URL: arXiv:2006.03090.

[HRW12] M. Hairer, M. D. Ryser, and H. Weber. "On the triviality of the 2D Allen-Cahn equation". In: *Electron. J. Probab.* 17.39 (2012), p. 14.

[HS73] W. G. Hunt and R. K. Selander. "Biochemical genetics of hybridisation in European house mice". In: *Heredity* 31.1 (1973), pp. 11–33.

[Hsu86] P. Hsu. "On Excursions of Reflecting Brownian Motion". In: *Trans. Am. Math. Soc.* 296 (1986), pp. 239–264. DOI: 10.2307/2000572.

[Hui89] G. Huisken. "Nonparametric mean curvature evolution with boundary conditions". In: *J. Differ. Equ.* 77.2 (1989), pp. 369–378.

[Imb09] C. Imbert. "Level set approach for fractional mean curvature flows". In: *Interfaces Free Bound* 11.1 (2009), pp. 153–176.

[IS09] C. Imbert and P. E. Souganidis. “Phasefield theory for fractional diffusion-reaction equations and applications”. 2009. arXiv: 0907.5524 [math.AP].

[KCO02] M. Koslowski, A. M. Cuitino, and M. Ortiz. “A phase-field theory of dislocation dynamics, strain hardening and hysteresis in ductile single crystals”. In: *J. Mech. Phys. Solids.* 50.12 (2002), pp. 2597–2635.

[Kim53] M. Kimura. “Stepping stone model of population”. In: *Ann. Rep. Nat. Inst. Genetics Japan* 3 (1953), pp. 62–63.

[Kin82] J. F. C. Kingman. “The coalescent”. In: *Stoch. Proc. Appl.* 13 (1982), pp. 235–248.

[KT10] A. Kohatsu-Higa and P. Tankov. “Jump-adapted discretization schemes for Lévy-driven SDEs”. In: *Stoch. Proc. Appl.* 120.11 (2010), pp. 2258–2285.

[Lee18] K. Lee. “Generation of interfaces for multi-dimensional stochastic Allen–Cahn equation with a noise smooth in space”. In: *Stochastics* 90.6 (2018), pp. 836–860. DOI: 10.1080/17442508.2018.1426586.

[Lis+20] A. Lischke, G. Pang, M. Gulian, F. Song, C. Glusa, X. Zheng, Z. Mao, W. Cai, M. M. Meerschaert, M. Ainsworth, and G. Karniadakis. “What is the fractional Laplacian? A comparative review with new results”. In: *J. Comput. Phys.* 404 (2020). DOI: 10.1016/j.jcp.2019.109009.

[Mal48] G. Malécot. “Les mathématiques de l’hérédité”. In: *Masson et Cie* (1948).

[Man11] C. Mantegazza. “Lecture notes on mean curvature flow”. In: *Prog. Math.* Vol. 20. Birkäuser, 2011. ISBN: 978-3-0348-0145-4.

[McK67] H. P. McKean. “Propagation of chaos for a class of non-linear parabolic equations”. In: *Lecture Series in Differential Equations* 2 (1967), pp. 177–194.

[McK75] H. P. McKean. “Application of Brownian motion to the equation of Kolmogorov-Petrovski-Piskunov”. In: *Common. Pure Appl. Math.* 28 (1975), pp. 323–331.

[MF72] H. Movljankulov and A. Filatov. “Ob odnom priblizennom metode postroenija resenii integral’nyh uravnenii, Tr”. In: *ta Kibern. AN UzSSR, Taskent* 12 (1972), pp. 11–18.

[MNL06] H. Matano, K. Nakamura, and B. Lou. “Periodic traveling waves in a two-dimensional cylinder with saw-toothed boundary and their homogenization limit”. In: *Netw. Heterog. Media* 1.4 (2006), pp. 537–568.

[Mor58] P. A. P. Moran. “Random processes in genetics”. In: *Proc. Camb. Philos. Soc.* Vol. 54. Cambridge Philosophical Society, 1958. DOI: `10.1017/S0305004100033193`.

[MS89] P. D. Mottoni and M. Schatzman. “Evolution geometrique d’interface”. In: *CR Acad. Sci. Ser. I. Math* 309 (1989), pp. 453–458.

[MS90] P. D. Mottoni and M. Schatzman. “Development of interfaces in $\mathbb{R}^N$”. In: *P. Roy. Soc. Edinb. A.* 116.3-4 (1990), pp. 207–220.

[Per02] E. A. Perkins. “Dawson-Watanabe Superprocesses and Measure-valued diffusions”. In: *École d’été de probabilités de Saint Flour XXIX-1999*. Vol. 1781. Springer-Verlag, 2002.

[Pit99] J. Pitman. “Coalescents With Multiple Collisions”. In: *Ann. Probab.* 27.4 (1999), pp. 1870–1902. DOI: `10.1214/aop/1022874819`.

[Rea21] B. S. Reatini. “The influence of hybridization on range dynamics”. PhD thesis. 2021. DOI: `10.17615/n20t-dk06`.

[Sag99] S. Sagitov. “The General Coalescent with Asynchronous Mergers of Ancestral Lines”. In: *J. Appl. Probab.* 36.4 (1999), pp. 1116–1125.

[Sat99] K. Sato. *Lévy Processes and Infinitely Divisible Distributions*. Cambridge University Press, 1999. ISBN: 9780521553025.

[Sch02] K. Schürger. “Laplace transforms and suprema of stochastic processes”. In: *Adv. Finance. Stoch.* Springer, 2002, pp. 285–294.

[SK14] J. Sabine and N. Kurt. “On the notion(s) of duality for Markov processes”. In: *Probab. Surv.* 11 (2014), pp. 59–120. DOI: `10.1214/12-PS206.`.

[Sko64] A. V. Skorohod. “Branching diffusion processes.” In: *Th. Prob. Appl.* 9 (1964), pp. 492–497.

[Smi21] A. Smith. “Probability models of population structure and gene flow”. PhD thesis. University of Oxford, 2021.

[Smy88] W.B. Smythe. *Static and dynamic electricity*. United States: Hemisphere Publishing., 1988. ISBN: 0-89116-916-4.

[SS80] T. Shiga and A. Shimizu. “Infinite dimensional stochastic differential equations and their applications”. In: *J. Math. Kyoto Univ.* 20.3 (1980), pp. 395–416. DOI: 10.1215/kjm/1250522207.

[Stu03] A. Sturm. “On Convergence of Population Processes in Random Environments to the Stochastic Heat Equation with Colored Noise”. In: *Electron. J. Probab.* 8 (2003), pp. 1–39. DOI: 10.1214/EJP.v8-129.

[Tay09] J. E. Taylor. “The Genealogical Consequences of Fecundity Variance Polymorphism”. In: *Genetics* 182.3 (2009), pp. 813–837. DOI: 10.1534/genetics.109.102368.

[Toe+18] D. PL. Toews, I. J. Lovette, D. E. Irwin, and A. Brelsford. “Similar hybrid composition among different age and sex classes in the Myrtle–Audubon’s warbler hybrid zone”. In: *The Auk: Ornithological Advances* 135.4 (2018), pp. 1133–1145.

[Wri43] S. Wright. “Isolation by distance”. In: *Genetics* 28.2 (1943), p. 114.

Appendix A

A.1 A geometric computation

Proposition A.1. *Let* $\Omega = \{(x_1, x'), x_1 \in \mathbb{R}, x' \in \phi(x_1) \subseteq \mathbb{R}^{\mathrm{d}-1}\}$ *with* $\phi(x_1) = H - h(x_1)$. *Let* $z > 0$, *and let* $a = -z + \sqrt{\mathrm{r}^2 - (H + h(z))^2}$. *Let* $x \in \partial\Omega$, *let* $v_1(x)$ *be the vector pointing from* x *to* $(a, 0, .., 0)$ *and let* $v_2(x)$ *be the (inward) normal vector of* $\partial\Omega$ *at* x. *Then*

$$\langle v_1(x), v_2(x)\rangle = H + h(-x_1) + h'(-x_1)(-z - x_1 - \sqrt{\mathrm{r}^2 - (H + h(z))^2}). \tag{A.2}$$

Proof. This is mostly a matter of writing the vectors and then the scalar product explicitly. Without loss of generality, by rotational symmetry, we can assume $x = (x_1, x_2, 0, \ldots, 0)$. In fact, since x is in the boundary of $\partial\Omega$, we can assume that $x = (x_1, H + h(-x_1), 0, \ldots, 0)$. Thus, we can write $v_1(x)$ as

$$v_1(x) = (a - x_1, -(H + h(-x_1)), 0, \ldots, 0).$$

On the other hand, since the boundary of the domain has a contour given by $h(-x_1)$, one can see $v_2(x)$ is given by

$$v_2(x) = (h'(-x_1), -1, 0, \ldots, 0).$$

Hence we have that

$$\begin{aligned}\langle v_1(x), v_2(x)\rangle &= h'(-x_1)(a - x_1) + (H - h(-x_1)) \\ &= H + h(-x_1) + h'(-x_1)(-z - x_1 - \sqrt{\mathrm{r}^2 - (H + h(z))^2}).\end{aligned}$$

□

A.2 A SLFVS with reflecting boundary condition

Since our main focus in the deterministic setting was on the interaction between selection and the shape of the domain we would like, for completeness, to obtain analogous results in the stochastic setting. As is evident from the results in the deterministic setting, such results should depend on local effects around a spherical shell, and therefore follow from our work in Section 2.2. However, the SLFVS has only previously been studied on the whole of Euclidean space, or on a torus. In this section we suggest a way in which the SLFVS can be extended to a process with reflecting boundary conditions, at least for a class of domains that includes Ω as in Figure 1.1. The idea is simple: when an event overlaps the boundary of the domain, we sample parents in such a way that the transition densities of the motion of ancestral lineages will be what we obtain by *reflecting* the motion each time it intersects the domain. To be more precise, to obtain a reflected motion we will use the method of images, a classical way of obtaining solutions of equations with reflecting boundary conditions (see e.g. [Smy88]). In the interests of simplifying the notation, we present the details in $\mathbb{d} = 2$, but the ideas are easily adapted to $\mathbb{d} \geq 3$.

We reserve the symbol Ω for the domain of Figure 1.1, and use $\mathcal{D}$ to refer to the class of domains for which we now define the SLFVS with reflecting boundary conditions.

Geometric definitions.

We will need two notions: *reflected points* at the boundary, and a classification of balls depending on the way in which they intersect the boundary of $\mathcal{D}$. Both will rest on a decomposition of the boundary of $\mathcal{D}$ into lines, and we only consider domains for which this is possible (which of course includes Ω). In higher dimensions, we would replace 'line' with 'plane'. Below we shall use the term 'affine hyperplane' for a doubly infinite straight line, and 'line' to mean a segment of such a line.

We define a line (in dimension 2) as a closed convex subset of an affine hyperplane. We suppose that $\partial\mathcal{D}$ can be decomposed as a finite union of lines. That is, there exist lines $(L_i)_{i=1}^{N}$

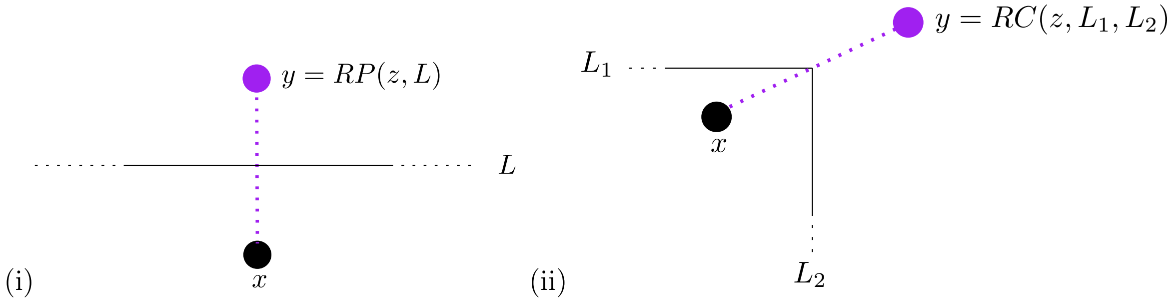

Figure A.1: (i) Example of how we reflect a point with respect to a line. (ii) Example of how we reflect a point with respect to a corner.

such that

$$\partial\mathcal{D} = \bigcup_{i=1}^{N} L_i. \tag{A.3}$$

We suppose that the lines are maximal, in the sense that

$$|L_j \cap L_i| \leq 1, \text{ and if } L_j \cap L_i \neq \emptyset \text{ then } H_{L_i} \neq H_{L_j} \quad \forall j \neq i, \tag{A.4}$$

where H_{L_i} is the affine hyperplane that contains L_i. Note that $\partial\Omega$ admits a unique decomposition into lines. Indeed, in that case, it is elementary to write down explicit expressions; for example, one line and its corresponding hyperplane are given by

$$L = \{(x_1, R_0) \in \Omega | x_1 \leq 0\} \subseteq H_L = \{(x_1, R_0) | x_1 \in \mathbb{R}\}.$$

From now on we write $\mathcal{L}_\mathcal{D}$ for the set of lines such that (A.3) and (A.4) hold for a domain $\mathcal{D}$.

Let $L \in \mathcal{L}_\mathcal{D}$, then let H_L be the affine hyperplane such that $L \subseteq H_L$, and write $\widehat{n}_{H_L}$ for the corresponding normal vector. The orthogonal projection of z into H_L is given by $z - \langle z, \widehat{n}_{H_L}\rangle \widehat{n}_{H_L}$. If this element lies in L, then we define the reflection of z with respect to L as its reflection with respect to H_L; that is the unique (other) point with the same orthogonal projection and distance to H_L as z. We also need to define reflection with respect to a corner of our domain. A formal definition is below, and we provide an illustration in Figure A.1.

Definition A.2 (Reflected points). For $L \in \mathcal{L}_\mathcal{D}$ and $z \in \mathcal{D}$ we define $\mathfrak{R}_P(z, L)$, the reflected point of z with respect to L, as

$$\mathfrak{R}_P(z, L) := \{w \in \mathcal{D}^c | z - \langle z, \widehat{n}_{H_L}\rangle \widehat{n}_{H_L} = w - \langle w, \widehat{n}_{H_L}\rangle \widehat{n}_{H_L} \in L \text{ and } d(z, H_L) = d(w, H_L)\}.$$

For $L_1, L_2 \in \mathcal{L}_\Omega$ with $L_1 \cap L_2 \neq \emptyset$ we define the set of reflected points with respect to the corner, $\mathfrak{R}_C(z, L_1, L_2)$, as

$$\mathfrak{R}_C(z, L_1, L_2) = \{w \in \mathcal{D}^c | (w \notin \mathfrak{R}_P(z, L_1) \cup \mathfrak{R}_P(z, L_2)) \wedge (\exists \widehat{z} \in \mathfrak{R}_P(z, L_2); \\ \widehat{z} - \langle \widehat{z}, \widehat{n}_{H_{L_1}} \rangle \widehat{n}_{H_{L_1}} = w - \langle w, \widehat{n}_{H_{L_1}} \rangle \in \Omega^c) \text{ and } (d(\widehat{z}, H_{L_1}) = d(w, H_{L_1}))\}.$$

We also define the 'completed' versions:

$$\overline{\mathfrak{R}_P(z, L)} = \mathfrak{R}_P(z, L) \cup \{z\},$$

$$\overline{\mathfrak{R}_C(z, L_1, L_2)} = \mathfrak{R}_C(z, L_1, L_2) \cup \mathfrak{R}_C(z, L_2, L_1) \cup \mathfrak{R}_P(z, L_1) \cup \mathfrak{R}_P(z, L_2) \cup \{z\}.$$

We note that $|\mathfrak{R}_P(z, L)| \leq 1$, as, for any hyperplane, there are two points with the same orthogonal projection and distance to it. Also we remark that, for $L_1, L_2 \in \mathcal{L}_\mathcal{D}$ with $L_1 \cap L_2 \neq \emptyset$, and for $w \in \mathbb{R}$, if we set

$$\overline{\mathfrak{R}_P}^{-1}(w, L_1) = \{z \in \Omega | z \in \overline{\mathfrak{R}_P}(w, L_1)\},$$

$$\overline{\mathfrak{R}_C}^{-1}(w, L_1, L_2) = \{z \in \Omega | z \in \overline{\mathfrak{R}_C}(w, L_1, L_2)\},$$

then for any $z \in \Omega$ we have that

$$\overline{\mathfrak{R}_P}^{-1}(z, L_1) = \overline{\mathfrak{R}_C}^{-1}(z, L_1, L_2) = \{z\}. \tag{A.5}$$

We now define a classification of balls depending on their intersection with $\mathcal{D}$. For any $A \subset \mathbb{R}^d$ we set

$$A \cap \mathcal{L}_\mathcal{D} := \{L \in \mathcal{L}_\mathcal{D} | A \cap L \neq \emptyset \text{ or } A \cap H_L \cap \mathcal{D}^c \neq \emptyset\}.$$

Definition A.3. [Classification of balls in $\mathbb{R}^2$] Let $x \in \mathbb{R}^2$ and $r > 0$.

1. We say $B(x, r)$ is inside $\mathcal{D}$ if $B(x, r) \subset \mathcal{D}$.

2. We say $B(x, r)$ is outside $\mathcal{D}$ if $B(x, r) \subset \mathcal{D}^c$.

3. We say $B(x, r)$ intersects a side of $\mathcal{D}$ if $|B(x, r) \cap \mathcal{L}_\mathcal{D}| = 1$ and $B(x, r)$ is not outside of $\mathcal{D}$.

4. We say $B(x,r)$ intersects a corner if $|B(x,r) \cap \mathcal{L}_\mathcal{D}| = |\{L_1, L_2\}| = 2$ with $L_1 \cap L_2 \neq \emptyset$ and $B(x,r)$ is not outside of $\mathcal{D}$.

5. We say $B(x,r)$ intersects a hallway if $|B(x,r) \cap \mathcal{L}_\mathcal{D}| = |\{L_1, L_2\}| = 2$ with $L_1 \cap L_2 = \emptyset$ and $B(x,r)$ is not outside of $\mathcal{D}$.

Note that Definition A.3 provides a complete classifications of balls in $\mathbb{R}^2$, of radius at most r, given that $|B(x,r) \cap \mathcal{L}_\mathcal{D}| \leq 2$ for all $x \in \mathbb{R}^2$. With this we can define the concept of reflected sampling in $B(x,r)$.

Definition A.4 (Reflected sampling). Let $x \in \mathbb{R}^2$, and let $r > 0$ such that $B(x,r)$ is not outside $\mathcal{D}$, does not intersect a hallway, and satisfies $|B(x,r) \cap \mathcal{L}_\mathcal{D}| \leq 2$. We define reflected sampling in $B(x,r)$ as follows:

1. If $B(x,r)$ is inside $\mathcal{D}$ then it is simply the uniform sampling in $B(x,r)$.

2. If $B(x,r)$ intersects a side L of $\mathcal{D}$, that is $B(x,r) \cap \mathcal{L}_\mathcal{D} = \{L\}$, then we pick a point z uniformly at random in $B(x,r)$, and take the only point in $\overline{\mathfrak{R}_P}^{-1}(z,P)$.

3. If $B(x,r)$ intersects a corner $\{L_1, L_2\}$ of $\mathcal{D}$, that is $B(x,r) \cap \mathcal{L}_\mathcal{D} = \{L_1, L_2\}$, then we pick a point z uniformly at random in $B(x,r)$, and subsequently pick a point uniformly at random in $\overline{\mathfrak{R}_C}^{-1}(z, L_1, L_2)$.

Reflected sampling can be thought of as sampling from the complete ball and then, if the ball intersects the boundary, *folding it along the boundary* in such a way that the content is all inside $\mathcal{D}$. This is illustrated for the domain Ω in Figure A.2.

The SLFVS on Ω.

The SLFVS on Ω can now be defined in the obvious way. It is driven by the same Poisson point process Π of events as in Definition 1.32, but now, for a point $(t,x,r) \in \Pi$, if $B(x,r) \cap \Omega = \emptyset$ then nothing happens. However, if $B(x,r)$ intersects the boundary $\partial\Omega$, parental locations are chosen according to *reflected sampling*, with the allele frequencies, of course, only being updated in Ω. This construction will work for any domain $\mathcal{D}$ that fulfils the following conditions:

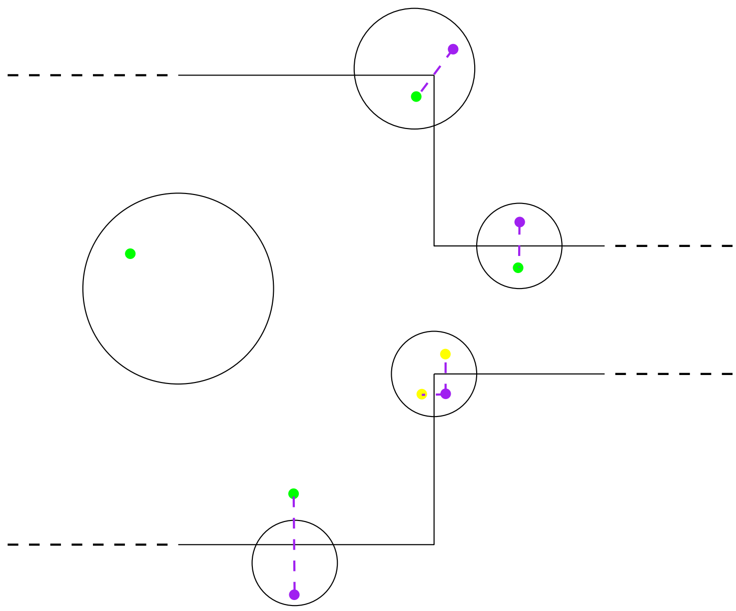

Figure A.2: *Reflected sampling in* Ω. Each of the balls intersects Ω. For the ball contained completely within Ω, the green point is uniformly sampled in the usual way. For balls intersecting the boundary, points in purple are sampled uniformly from the intersection of the ball with Ω^c, corresponding to a uniformly sampled point falling outside the domain. The points in green correspond to the cases in which reflected sampling leads to a unique point within the domain. The points in yellow correspond to a case in which reflection gives two possible outcomes for the reflected sampling and we need to choose one of them uniformly at random. The purple lines join sampled points to the corresponding possible outcomes of reflected sampling.

Assumptions A.5. We impose the following conditions on the domain $\mathcal{D}$.

1. There exists a set of lines $\mathcal{L}_\mathcal{D}$ such that $\mathcal{D}$ fulfils (A.3) and (A.4).

2. There is $\mathcal{R}_\mathcal{D} > 0$ such that, for all $x \in \mathbb{R}^2$, we have $|B(x, \mathcal{R}_\mathcal{D}) \cap \mathcal{L}_\mathcal{D}| \leq 2$ and $B(x, \mathcal{R}_\mathcal{D})$ does not intersect a hallway.

3. For any $r < \mathcal{R}_\mathcal{D}$ and $x \in \mathbb{R}^2$, if $B(x, r)$ intersects a corner given by $\{L_1, L_2\}$ then $\overline{\mathfrak{R}_C}^{-1}(w, L_1, L_2) \neq \emptyset$ for all $w \in B(x, r)$. If $B(x, r)$ intersects a side given by the line L, then $\overline{\mathfrak{R}_P}^{-1}(w, P) \neq \emptyset$ for all $w \in B(x, r)$.

We can then construct the SLFVS on $\mathcal{D}$, provided that events are assumed to have radius bounded by $\mathcal{R}_{\mathcal{D}}$. These conditions are sufficient to ensure that the reflected sampling that we use to choose a parental location will be well defined. The third condition prevents there not being a pre-image through the reflection of the parental location; the second removes the possibility of an event intersecting two walls, for which we have not defined the corresponding reflected points. We could, of course, extend our definitions further, but we are primarily interested in the domain Ω, and one can check that Ω fulfils these conditions with $\mathcal{R}_\Omega \leq r_0 \wedge R_0 - r_0$.

Blocking in the stochastic setting

Using the approach adopted in Theorem 1.34, we can prove a stochastic analogue of Theorem 1.26. While we omit the details, the key tool is a branching and coalescing dual for the SLFVS on Ω. This mirrors the dual for the process on the whole Euclidean space, introduced in Definition 2.22, except that uniform sampling of offspring locations is replaced by reflected sampling.

The key ideas in the proof of Theorem 1.34 can then be adapted in an obvious way to this setting. Since we deliberately constructed the reflected sampling in such a way that transition probabilities of the jump process that is followed by a particular lineage could be obtained from those on Euclidean space via the method of images, it should be no surprise that Lemma 2.29 can be translated to this setting after replacing Brownian motion with reflected Brownian motion.

A.2.1 The dual process.

We proceed to state a duality result for the SFLVS on Ω, which will be the core of the proof of Theorem 1.36. The duality relation will be once again with respect to a historical process and expressed through a voting procedure.

We begin by describing the dual process of branching and coalescing lineages. It is driven by the same Poisson point process of 'events' that drives the SLFVS on Ω. Recall that Π^n is a Poisson point process on $\mathbb{R}^+ \times \mathbb{R}^d \times (0, \mathcal{R}^n_\Omega)$, with $\mathcal{R}^n_\Omega = n^{-\beta}\mathcal{R}_\Omega$, whose intensity is given by (1.25) in the weak noise/selection ratio regime and by (1.27) in the strong noise/selection ratio regime. The impact, asymmetry and selection parameters, $(u_n)_{n\mathbb{N}}$, $(\gamma_n)_{n\mathbb{N}}$, and $(\boldsymbol{s}_n)_{n\mathbb{N}}$ respectively are given in (1.26) in the weak noise/selection ratio case or fulfil the condition (1.28) in the strong noise/selection ratio regime.

Definition A.6 (SFLVS on Ω, dual process). The dual of the SFLVS on Ω is the process obtained when in Definition 2.22 we sample with reflected sampling in $B(x,r)$ (instead of sampling uniformly).We will denote the particles of this process by $\mathcal{P}^n(t) = (\xi_1^n(t), ..., \xi_{N(t)}^n(t))$, and the historical process of branching and coalescing lineages by $\Xi^n(t)$.

Again, it is straightforward to obtain an analogue of Theorem 2.25 for the SLFVS on Ω and its dual. Hence, to prove Theorem 1.36 it is enough to prove the following result.

Theorem A.7. *Denote $\rho_* = (\mathbb{d}-1)/\nu$ and suppose $r_0 \leq \rho_*$.*

1. *Under the weak noise/selection ratio regime, the following holds for the dual process. For any $k \in \mathbb{N}$ there exist $n_*(k) < \infty$, and $a_*(k), d_*(k) \in (0,\infty)$ such that for all $n \geq n_*$ and all t satisfying $a_*\varepsilon_n^2|\log\varepsilon_n| \leq t < \infty$,*

$$\text{for } x \text{ such that } \widehat{d}(x) \leq -d_*\varepsilon_n|\log\varepsilon_n|, \text{ we have } \mathbb{P}_x\left[\mathbb{V}_{1_{x_1\geq 0}}(\Xi^n(t)) = 1\right] \leq \varepsilon_n^k.$$

2. *Under the strong noise/selection ratio regime, as n tends to infinity, the dual process behaves approximately as a reflected Brownian motion. That is, for almost every $x \in \Omega$, every $\varepsilon > 0$ and every $t \geq 0$, there is a reflected Brownian motion, $(W_t)_{t\geq 0}$ and there is $n_* \in \mathbb{N}$, such that for all $n \geq n_*$,*

$$\left|\mathbb{P}_x\left[\mathbb{V}_{1_{x_1\geq 0}}(\Xi^n(t)) = 1\right] - \mathbb{P}_x[W_1(t\sigma^2) \geq 0]\right| \leq \varepsilon \tag{A.6}$$

A.2.2 Motion of a single ancestral lineage

To prove Theorem A.7, one of the main tools is following an ancestral lineage and proving it is, approximately, a reflected Brownian motion in Ω. We need to write the transition rates of the motion of an ancestral lineage. For this, with an abuse of notation, write:

$$\mathfrak{R}(z, B(x,r)) = \begin{cases} \{z\} & \text{if } B(x,r) \text{ is inside } \Omega, \\ \emptyset & \text{if } B(x,r) \text{ is outside } \Omega, \\ \overline{\mathfrak{R}_P}(z, L) & \text{if } B(x,r) \cap \mathcal{L}_\Omega = \{L\}, \\ \overline{\mathfrak{R}_C}(z, L_1, L_2) & \text{if } B(x,r) \cap \mathcal{L}_\Omega = \{L_1, L_2\}. \end{cases} \tag{A.7}$$

Let χ_n be 1 in the weak noise/selection ratio regime and $1/\widehat{u}_n$ in the strong noise/selection ratio regime. Then the motion of an ancestral lineage has transition rate from y to z given by

$$\widehat{m}_n(y, dz) = \\ \chi_n u_n n n^{\mathrm{d}\beta} \int_0^{\mathcal{R}_\Omega^n} \int_{\mathbb{R}^{\mathrm{d}}} 1_{|x-y|<r} \left(\sum_{\widehat{z} \in R(z, B(x,r))} \frac{1_{|\widehat{z}-x|<r}}{|B(x,r)||\mathfrak{R}^{-1}(\widehat{z}, B(x,r))|} \right) dz dx d\mu^n(r), \tag{A.8}$$

where $\mathfrak{R}^{-1}(\widehat{z}, B(x,r))$ is understood to be the inverse of the corresponding term in (A.7): $\overline{\mathfrak{R}_P}(z, L)$ or $\overline{\mathfrak{R}_C}(z, L_1, L_2)$. The sum over $\emptyset$ is understood always to be 0. To see why expression (A.8) holds, note that if the ancestral lineage is at y and it jumps to z, it must be affected by an event that covers y and some reflected point of z: $\widehat{z}$, say. If the event has radius r, then the set of possible centres x of such events is the set of those x for which $|x-y| < r$ and $|\widehat{z} - x| < r$, for $\widehat{z}$ a reflected point of z. Each point in that set is selected with intensity $\chi_n n \, n^{\mathrm{d}\beta}$, and the radius is selected with intensity $\mu^n(dr)$. The parental location is then chosen uniformly in the preimage of the corresponding reflection of $\widehat{z}$, where $\widehat{z}$ is chosen uniformly on the ball $B(x,r)$; thus, choosing z has probability $\frac{dz}{|B(x,r)||\mathfrak{R}^{-1}(\widehat{z}, B(x,r))|}$. Finally the probability that our lineage is actually affected by the event is u_n. Putting these together yields (A.8).

Even though $\widehat{m}_n(y, dz)$ is not homogeneous in space we can still compute the total jump rate and note that it is independent of the actual position. Indeed:

$$\begin{aligned}
&\int_\Omega \widehat{m}_n(y,dz)\\
&= \chi_n u_n n n^{\mathrm{d}\beta} \int_0^{\mathcal{R}_\Omega^n} \int_{\mathbb{R}^{\mathrm{d}}} 1_{|x-y|<r} \int_\Omega \left(\sum_{\widehat{z}\in R(z,B(x,r))} \frac{1_{|\widehat{z}-x|<r}}{|B(x,r)||\mathfrak{R}^{-1}(\widehat{z},B(x,r))|} \right) dz dx d\mu^n(r)\\
&= \chi_n u_n n n^{\mathrm{d}\beta} \int_0^{\mathcal{R}_\Omega^n} \int_{\mathbb{R}^{\mathrm{d}}} 1_{|x-y|<r} \int_{\mathbb{R}^{\mathrm{d}}} \frac{1_{|\widehat{z}-x|<r}}{|B(x,r)|} d\widehat{z} dx d\mu^n(r)\\
&= \int_0^{\mathcal{R}_\Omega^n} n\chi_n u_n n^{\mathrm{d}\beta} V_r d\mu^n(dr)\\
&= u_n n \chi_n |B(0,1)| \int_0^{\mathcal{R}_\Omega} r^{\mathrm{d}} d\mu(r), \qquad \text{(A.9)}
\end{aligned}$$

where the second equality follows from the fact that, by the third condition of Assumptions A.5, every point in a ball that intersects the boundary of the domain has a pre-image through $\mathfrak{R}$. Recall that we denote by $m_n(y,dz)$ the transition rate of an ancestral lineage of the homogeneous process: that is when, instead of having Ω, we have domain $\mathbb{R}^{\mathrm{d}}$. Note that then

$$\int_\Omega \widehat{m}_n(y,dz) = \int_\Omega m_n(y,dz). \qquad \text{(A.10)}$$

Suppose now that y is distance at least $2\mathcal{R}_\Omega^n$ from every line. Hence it is only intersected by balls inside Ω. In this case, the reflected sampling is simply a uniform sampling, which implies that

$$\widehat{m}_n(y,dz) = m_n(y,dz). \qquad \text{(A.11)}$$

Now suppose that y is at a distance less than $2\mathcal{R}_\Omega^n$ from a line L, but at a distance greater than $2\mathcal{R}_\Omega^n$ from any intersection between two lines. Then every ball that contains y only intersects at most one line of Ω. Note also that for every point in the ball, there is a unique point with the same projection and distance to L, and so $|\overline{\mathfrak{R}_P}^{-1}(w,L)| = 1$ for every w in the ball. Using these facts, some cumbersome computations give

$$\widehat{m}_n(y,dz) = m_n(y,dz) + m_n(y,\mathfrak{R}_P(dz,L)). \qquad \text{(A.12)}$$

Furthermore, in the particular case of Ω, if we are at distance less than $2\mathcal{R}_\Omega^n$ from the corners $(0,R_0)$ or $(0,-R_0)$ then, denoting by L_1 and L_2 the lines which intersects this corner, we obtain by a similar computation that

$$\begin{aligned}\widehat{m}_n(y,dz) &= m_n(y,dz) + m_n(y,\mathfrak{R}_P(dz,L_1)) \\ &\quad + m_n(y,\mathfrak{R}_P(dz,L_2)) + m_n(y,\mathfrak{R}_C(dz,L_1,L_2)).\end{aligned} \tag{A.13}$$

Relationships (A.10), (A.11), (A.12), (A.13) are exactly what we would expect from a version of m_n reflected on the boundary of Ω. From Lemma 2.29, the finite dimensional distributions of the jump process with transition rates m_n converge to those of a Brownian motion with diffusion coefficient

$$\sigma^2 = \frac{u}{2\mathbb{d}} \int_0^{\mathcal{R}} \int_{\mathbb{R}^{\mathbb{d}}} |z|^2 \frac{|B(0,r)\cap B(z,r)|}{|B(0,r)|} dz d\mu(dr). \tag{A.14}$$

Thus we expect the finite dimensional distributions of a jump process with transition rates $\widehat{m}_n$ to converge to those of a reflected Brownian motion with the same diffusion coefficient σ^2. This is presented in the next section.

A.2.3 Weak noise/selection ratio regime, coupling of motion of an ancestral lineage to a reflected Brownian motion

In this section, we sketch the proof of the weak noise/selection ratio regime of Theorem A.7. Most of the work for Theorem 2.27 applies to the dual process for the SLFVS on Ω. The main difference is proving an analogue of Lemma 2.29 for the process in Ω. This is one of the main points we deal with in this section, and the tool we use to conclude Theorem A.7.

From the identities presented in the last section, we expect that a process with transition rates given by $\widehat{m}_n$ converges in some sense to a reflected Brownian motion. To formalise this, let $\zeta^n(t)$ be a jump process with transition rates given by $\widehat{m}_n$. We set $\eta = u|B(0,1)|\int_0^{\mathcal{R}} r^{\mathbb{d}}\mu(dr)$, and note that $\eta(1+\gamma_n)\varepsilon_n^{-2}$ is the rate at which an ancestral lineage is affected by a selective event. For $u > 0$, we define

$$\Omega_-^u := \{x \in \Omega : \|x\| \geq u \text{ and } x_1 < 0\}$$

In this section we prove the following coupling result.

Theorem A.8. *Fix $\mathbb{r}$ such that $r_0 < \mathbb{r} < R_0$. Let $x \in \Omega_-^{(\mathbb{r}-r_0)/2}$, $a \geq 0$ and $k \in \mathbb{N}$. Then for n sufficiently large, and for every $t \in [0, (a + \eta^{-1}(k+1))\varepsilon_n^2 |\log(\varepsilon_n)|]$, there is a coupling of ξ^n and a reflected Brownian motion W under which:*

$$\mathbb{P}_x\left[|\xi^n(t) - W(\sigma^2 t)| \geq n^{-\beta/6}\right] = \mathcal{O}(n^{-\beta}(t \vee 1)) + \varepsilon_n^{k+1},$$

where ζ^n and W both have the same starting point x.

Of course, since we are experiencing reflection there is much that could go wrong in Theorem A.8. For instance, we have corners that do not give the relationship (A.13), which could distort the behaviour we state in Theorem A.8. The next lemma will be used to show that the distortion caused by those *bad* corners is negligible on the time scale of interest.

Lemma A.9. *Let $\rho < \frac{r_0}{2} \wedge \frac{\mathbb{r}-r_0}{4}$, and define $K := B((0, r_0), \rho) \cup B((0, -r_0), \rho)$. Denote by τ_K the hitting time of K for the reflected Brownian motion W. Let $x \in \Omega_-^{(\mathbb{r}-r_0)/2}$, $a > 0$ and $k \in \mathbb{N}$. Then, for n sufficiently large and for $t \in [0, \sigma^2(a + \eta^{-1}(k+2))\varepsilon_n^2 |\log(\varepsilon_n))|]$, we have*

$$\mathbb{P}_x[\tau_K \leq t] \leq \varepsilon_n^{k+1}.$$

Proof. First note that, for any $d(k)$, computing as in the derivation of (2.8),

$$\mathbb{P}_x^{\varepsilon}\left[\sup_{0 \leq s \leq t} \|W(s) - x\| \geq d(k)\varepsilon_n |\log(\varepsilon_n)|\right] \leq 2\mathbb{d}\varepsilon_n^{\frac{1}{4}\left(\frac{d(k)^2}{\mathbb{d}\sigma^2(a+\eta^{-1}(k+2))} - \frac{\mathbb{d}}{d(k)^2\sigma^2(a+\eta^{-1}(k+2))}\right)}. \tag{A.15}$$

Choose $d(k)$ such that

$$\frac{1}{4}\left(\frac{d(k)^2}{\mathbb{d}\sigma^2(a+\eta^{-1}(k+2))} - \frac{\mathbb{d}}{d(k)^2\sigma^2(a+\eta^{-1}(k+2))}\right) \geq k+2.$$

Then, for n big enough, we have

$$\mathbb{P}_x^{\varepsilon}[\sup_{0 \leq s \leq t} \|W(s) - x\| \geq d(k)\varepsilon_n |\log(\varepsilon_n)|] \leq \varepsilon_n^{k+1}.$$

Furthermore, choose n to be large enough that $\mathbb{r} - \rho > d(k)\varepsilon_n |\log(\varepsilon_n)|$. Then, since $x \in \Omega_-^{(\mathbb{r}-r_0)/2}$ implies that we start at distance at least $\mathbb{r} - \rho$ from K

$$\mathbb{P}_x[\tau_K \leq t] \leq \mathbb{P}_x^{\varepsilon}\left[\sup_{0 \leq s \leq t} \|W(s) - x\| \geq d(k)\varepsilon_n |\log(\varepsilon_n)|\right] \leq \varepsilon_n^{k+1}.$$

□

We start the construction of the coupling for Theorem A.8. The starting point will be the homogeneous version of the processes. That is, let $(\zeta^n(t))_{t\geq 0}$ be a jump process with transition kernel given by m_n. We first use this process to construct a Brownian motion $(B_t)_{t\geq 0}$ in $\mathbb{R}^{\mathrm{d}}$.

For $i \geq 1$, let $X_i = \zeta^n(\frac{i}{n^{2\beta}}) - \zeta^n(\frac{i-1}{n^{2\beta}})$. Then $X_1, X_2, ...$ are i.i.d rotationally symmetric random variables. Furthermore, as seen in [EFP17], we have that $\mathbb{E}_x[|X_1|^2] = 2\mathrm{d}\sigma^2 n^{-2\beta}$. By (A.9) and (A.10) the number of jumps made by ζ^n in the time interval $[0, n^{-2\beta}]$ is a Poisson random variable with mean $\Theta(1)$, so since each jump has magnitude at most $2\mathcal{R}^n_\Omega$, we obtain $\mathbb{E}[|X_1|^4] = \mathcal{O}(n^{-4\beta})$. Then, by the Skorohod Embedding Theorem (see [Bil95], Theorem 37.7), there is a Brownian motion B started at x and a sequence of stopping times $v_1, v_2, ..$ such that, setting $v_0 = 0$, we have that $(v_i - v_{i-1})_{i\geq 1}$ are i.i.d and

$$B(v_i) = \zeta^n\left(\frac{i}{n^{2\beta}}\right), \;\; \mathbb{E}_x[v_i - v_{i-1}] = \frac{1}{2\mathrm{d}}\mathbb{E}_x[|X_1|^2] = \sigma^2 n^{-2\beta}, \;\; \mathbb{E}_x[(v_i - v_{i-1})^2] = \mathcal{O}(n^{-4\beta}). \tag{A.16}$$

Note that, in particular, $\mathbb{E}_x[v_{\lfloor tn^{2\beta}\rfloor}] = \sigma^2 n^{-2\beta}\lfloor tn^{2\beta}\rfloor$, and $\mathrm{Var}[v_{\lfloor tn^{2\beta}\rfloor}] = \mathcal{O}(tn^{-2\beta})$. Also, by Chebyshev's inequality,

$$\mathbb{P}[|v_{\lfloor tn^{2\beta}\rfloor} - \sigma^2 t| \geq n^{-\beta/2}] = \mathcal{O}(tn^{-\beta}). \tag{A.17}$$

The key for Theorem A.8 is showing that the stopping times $(v_1, v_2, ...)$ can be used to construct the coupling of ξ^n and W. For this, we firstly need the next result, which follows easily by our choice of time intervals and Chebyshev's inequality.

Lemma A.10. *Let $t > 0$ be fixed. Then*

$$\mathbb{P}_x\left[\exists i \leq \lfloor n^{2\beta}t\rfloor \;\; s.t. \;\; v_i - v_{i-1} > \frac{\sigma^2}{n^{2\beta}} + \frac{1}{n^{\beta/2}}\right] = \mathcal{O}(tn^{-\beta}).$$

Note that, from Chebyshev's inequality and (A.16), it follows that

$$\mathbb{P}[|v_{\lfloor tn^{2\beta}\rfloor} - \sigma^2 t| \geq n^{-\beta/2}] = O(tn^{-\beta}). \tag{A.18}$$

Hence Lemma A.10 tells us that, with high probability, up to time approximately $t\sigma^2$, all the stopping times v_i are concentrated around their means. With this, the following can be obtained in the same manner as Lemma 2.13.

Lemma A.11. *Let $t \in [0,(a+\eta^{-1}(k+1))\varepsilon_n^2|\log(\varepsilon_n)|]$. Consider $(W_t)_{t\geq 0}$ a reflected Brownian motion in Ω. Then, for all $x \in \Omega_-^{(\mathbb{r}-r_0)/2}$ and for all ε sufficiently small, we have that*

$$\mathbb{P}_x\left[\exists i < \lfloor n^{2\beta}t\rfloor \ s.t. \sup_{u,s\in[v_{i-1},v_i]} |W(s)-W(u)| > n^{-\beta/6}\right] = \mathcal{O}(n^{-\beta}t).$$

We will proceed now to prove Theorem A.8. The idea is to reflect the paths given by ζ and B in (A.16), obtaining ξ and W. For this, we basically cut the paths of ζ and B into smaller pieces, each of size bounded by $n^{-\beta/6}$. We will then prove that when those pieces are reflected, which is what gives us ξ and W, they coincide at the times $v_1, v_2, ..,$ as long as we are away from the corners $(0,r_0)$ and $(0,-r_0)$. By Lemma A.9, we expect that with high probability the ancestral lineage and the Brownian motion never come within a distance less than $n^{-\beta/6}$ of those corners. That is, this pathological case does not occur and thus the coupling holds, with high probability. To properly formalise this we will need some notation and an additional lemma.

Recall that we use $(\zeta_t^n)_{t\geq 0}$ for an ancestral lineage of the homogeneous SLFVS, that is, a jump process with transition rates given by m_n. We will assume this process always starts at some $x \in \Omega_-^{(\mathbb{r}-r_0)/2}$. Consider then a pair of processes, $(\widehat{\xi}_t, s_t)_{t\geq 0} \in \Omega \times \{-1,1\}^2$, defined in terms of ζ as follows.

Definition A.12 (Ancestral lineage reflected by following a line segment). First set $\widehat{\xi}_0^n = \zeta_0^n$ and $s_t = (1,1)$. Now define

$$\Delta_t^n = ((\Delta_1^n)_t, (\Delta_2^n)_t) := \zeta_t - \zeta_{t^-}. \tag{A.19}$$

Consider also the line segment between $\widehat{\xi}_{t^-}$ and $y_1 := \widehat{\xi}_{t^-} + ((\Delta_1^n)_t(s_1)_{t^-}, \Delta_2^n)_t(s_2)_{t^-})$, that is

$$L_t := \left\{ y_\lambda \in \mathbb{R}^2 | y_\lambda = (1-\lambda)\widehat{\xi}_{t^-} + \lambda(\widehat{\xi}_{t^-} + ((\Delta_1^n)_t(s_1)_{t^-}, \Delta_2^n)_t(s_2)_{t^-})), \lambda \in (0,1)\right\}.$$

The processes $(\widehat{\xi}_t, s_t)_{t\geq 0}$ jump at the times where $(\zeta_t)_{t\geq 0}$ jumps. Let t be a jump time of ζ_t. Then at t we update $(\widehat{\xi}_t, s_t)$ as follows.

1. If $L_t \cap \mathcal{L}_\Omega = \emptyset$, then $\widehat{\xi}_t^n = \widehat{\xi}_{t^-}^n + ((\Delta_1^n)_t(s_1)_{t^-}, (\Delta_2^n)_t(s_2)_{t^-})$.
2. If $L_t \cap \mathcal{L}_\Omega = \{L_1\}$ and $\widehat{y}_1 := \overline{\mathfrak{R}_P}^{-1}(\widehat{\xi}_{t^-}^n + ((\Delta_1^n)_t(s_1)_{t^-}, (\Delta_2^n)_t(s_2)_{t^-}), L_1) \in \Omega$, then set $\widehat{\xi}_t^n = \widehat{y}_1$ and $(s_i)_t = (-1)(s_i)_{t^-}$, where i is such that $|(\widehat{n}_{L_1})_j| = 1_{i=j}$.

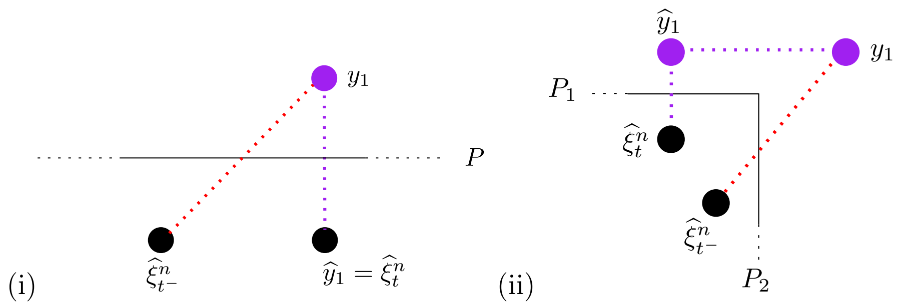

Figure A.3: (i) Example of how we obtain $\widehat{\xi}_t$ in the second case of the Definition A.12. (ii) Example of how we obtain $\widehat{\xi}_t$ in the third case of the Definition A.12.

3. If $L_t \cap \mathcal{L}_\Omega = \{L_1\}$ and $\overline{\mathfrak{R}_P}^{-1}(\widehat{\xi}^n_{t-} + ((\Delta^n_1)_t(s_1)_{t-}, (\Delta^n_2)_t(s_2)_{t-}), L_1)$ is empty, we proceed as follows. We take $\widehat{y}_1$ to be the reflection of y_1 with respect to H_{L_1}, and note $\widehat{y}_1$ is not in Ω. Let L_2 be the line that intersects the line segment between $\widehat{\xi}^n_{t-}$ and $\widehat{y}_1$. Set $\widehat{\xi}^n_t = \overline{\mathfrak{R}_C}^{-1}(y_1, L_1, L_2)$, $(s_1)_t = (-1)(s_1)_{t-}$ and $(s_2)_t = (-1)(s_2)_{t-}$.

We present examples of how we proceed in the second case and the third case of the definition of $(\widehat{\xi}_t)$ in Figure A.3.

By definition, $\widehat{\xi}_t \in \Omega$ for all $t \geq 0$. Furthermore, from the the definitions of $\mathfrak{R}_C, \mathfrak{R}_P$ and relationships (A.12), (A.13) it should hold that $\widehat{\xi}^n$ has transition rates given by $\widehat{m}_n$, until it is at distance less than $2\mathcal{R}^n_\Omega$ of $(0, r_0)$ or $(0, -r_0)$. Figure A.3 should help this intuition. The next lemma helps formalise this fact.

We let $\kappa = \{(0, r_0)\} \cup \{(0, -r_0)\}$ and, for $x \in \Omega$, denote by $d(x, \kappa)$ the distance between κ and x.

Lemma A.13. *Let $i \in \mathbb{N}$. Suppose that $d(\widehat{\xi}(\frac{i}{n^{2\beta}}), \partial\Omega) \leq n^{-\beta/6}$, that $|\widehat{\xi}_2(i/n^{2\beta}) - R_0| \leq n^{-\beta/6}$ and that $d(\widehat{\xi}(\frac{j}{n^{2\beta}}), \kappa) > n^{-\beta/6}$ for all $1 \leq j \leq i$. Then*

$$\left| \left(\widehat{\xi}\left(\frac{i+1}{n^{2\beta}} \right) \right)_2 - R_0 \right| = \left| \left(\widehat{\xi}\left(\frac{i}{n^{2\beta}} \right) \right)_2 + (\Delta^n_2)_i (s_2)_{in^{-2\beta}} - R_0 \right|. \tag{A.20}$$

Recall that $(\Delta^n_t)_{t \geq 0}$ are the discontinuities of the jump process $(\zeta^n_t)t \geq 0$. Hence, Lemma A.13 states that, as long as $\widehat{\xi}_t$ is in $\{x \in \Omega : |x - R_0| \leq n^{-\beta/6}\}$ and away from κ it fulfils relationship (A.12). It is not hard to see analogous results apply when $\widehat{\xi}_t$ is at distance less than $n^{-\beta/6}$ to any other line that make the boundary of the domain Ω (as long as the process is at distance greater than $n^{-\beta/6}$ of κ). Thus, as long as the process $\widehat{\xi}_t$ stays away from κ, it is a jump process with transition rates given by $\widehat{m}_n$.

Proof of Lemma A.13. First, let $\delta_0 = 0$ and $\delta_1, .., \delta_m$ be the (random number of) jumps of $(\zeta(t))_2$ in $[i/n^{2\beta}, (i+1)/n^{2\beta})$. Hence, by the definition of Δ_i^n, we have that

$$(\Delta_i^n)_2 = \sum_{l=1}^{m} \delta_l. \tag{A.21}$$

Let $t_0 = i/n^{2\beta}$, and let t_l be the time at which the jump δ_l occurs. First note that, for all $1 \le l \le m$, we have that

$$\left|\widehat{\xi}(t_l)_2 - R_0\right| = \left|\widehat{\xi}(t_{l-1}) + (s_2)_{t_{l-1}}\delta_l - R_0\right|. \tag{A.22}$$

Indeed, first suppose that δ_l causes a reflection of $\widehat{\xi}$ at time t_l, meaning that $L_t \cap \mathcal{L}_\Omega \neq \emptyset$. For $a \in \mathbb{R}^{\mathrm{d}}$, use $\mathfrak{R}_P(a_2, L_1)$ to denote the second component of the point $\mathfrak{R}_P(a, L_1)$. Hence, by the definition of $\widehat{\xi}$, we have

$$\left|\widehat{\xi}(t_l)_2 - R_0\right| = \left|\mathfrak{R}_P(\widehat{\xi}(t_{l-1})_2 + \delta_l(s_2)_{t_{l-1}}, L_1) - R_0\right| = \left|\widehat{\xi}(t_{l-1})_2 + \delta_l(s_2)_{t_{l-1}} - R_0\right|,$$

where the last equality holds because, by definition, $\mathfrak{R}_P$ gives the point inside of Ω that has the same distance to the line L_1 as the original point, and, by the definition of L_1, this implies that the distance of the second coordinate to R_0 must be the same. Now suppose δ_l does not cause a reflection of $\widehat{\xi}$. Then, by the definition of $\widehat{\xi}$, (A.22) follows immediately. Hence we conclude that (A.22) always holds.

With (A.22) we can show (A.20). For this we prove that, for all $1 \le l \le m$,

$$\left|\left(\widehat{\xi}(t_{l-1})\right)_2 + (s_2)_{t_{l-1}}\delta_l - R_0\right| = \left|\left(\widehat{\xi}\left(\frac{i}{n^{2\beta}}\right)\right)_2 + \left(\sum_{j=1}^{l}\delta_j\right)(s_2)_{in^{-2\beta}} - R_0\right|. \tag{A.23}$$

Armed with this, if we evaluate at $l = m$, use (A.21) and (A.22), and use that t_m is the last jump before t we obtain (A.20).

We prove (A.23) by induction. The case $l = 1$ is trivial. Now suppose that (A.23) holds for $l - 1$. Under this assumption we prove the equation holds for l.

Suppose that $(s_2)_{t_{l-1}} = (s_2)_{in^{-2\beta}}$. By definition, $(s_2)_t$ was multiplied by (-1) each time $\widehat{\xi}$ was reflected between $in^{-2\beta}$ and t_{l-1}. Since they both are the same, it follows that $\widehat{\xi}$ was reflected an even number of times. Since (A.23) holds for every jump before t_l, it means that

$$\operatorname{sgn}\left(\left(\widehat{\xi}(t_{l-1})\right)_2 + (s_2)_{t_{l-1}}\delta_l - R_0\right) = \operatorname{sgn}\left(\left(\widehat{\xi}\left(\frac{i}{n^{2\beta}}\right)\right)_2 + \left(\sum_{j=1}^{l-1}\delta_j\right)(s_2)_{in^{-2\beta}} - R_0\right).$$

This, together with (A.23), implies that

$$\begin{aligned}\left|\left(\widehat{\xi}(t_{l-1})\right)_2 + \delta_l(s_2)_{t_{l-1}} - R_0\right| &= \left|\left(\widehat{\xi}\left(\frac{i}{n^{2\beta}}\right)\right)_2 + \left(\sum_{j=1}^{l-1}\delta_j\right)(s_2)_{in^{-2\beta}} + \delta_l(s_2)_{t_l} - R_0\right| \\ &= \left|\left(\widehat{\xi}\left(\frac{i}{n^{2\beta}}\right)\right)_2 + \left(\sum_{j=1}^{l}\delta_j\right)(s_2)_{in^{-2\beta}} - R_0\right|,\end{aligned}$$

where the last equality holds since we assume $(s_2)_{t_{l-1}} = (s_2)_{in^{-2\beta}}$.

Now suppose $(s_2)_{t_{l-1}} \neq (s_2)_{in^{-2\beta}}$. By definition, $(s_2)_t$ was multiplied by (-1) each time $\widehat{\xi}$ was reflected between $in^{-2\beta}$ and t_{l-1}. Since they both are different, it follows that $\widehat{\xi}$ was reflected an odd number of times. Since (A.23) holds for every jump before t_l, it means that

$$\operatorname{sgn}\left(\left(\widehat{\xi}(t_{l-1})\right)_2 + (s_2)_{t_{l-1}}\delta_l - R_0\right) \neq \operatorname{sgn}\left(\left(\widehat{\xi}\left(\frac{i}{n^{2\beta}}\right)\right)_2 + \left(\sum_{j=1}^{l}\delta_j\right)(s_2)_{in^{-2\beta}} - R_0\right).$$

This, together with (A.23), implies that

$$\begin{aligned}\left|\left(\widehat{\xi}(t_{l-1})\right)_2 + \delta_l(s_2)_{t_{l-1}} - R_0\right| &= \left|\left(\widehat{\xi}\left(\frac{i}{n^{2\beta}}\right)\right)_2 + \left(\sum_{j=1}^{l-1}\delta_j\right)(s_2)_{in^{-2\beta}} - \delta_l(s_2)_{t_l} - R_0\right| \\ &= \left|\left(\widehat{\xi}\left(\frac{i}{n^{2\beta}}\right)\right)_2 + \left(\sum_{j=1}^{l}\delta_j\right)(s_2)_{in^{-2\beta}} - R_0\right|,\end{aligned}$$

where the last equality holds since we assume $(s_2)_{t_{l-1}} \neq (s_2)_{in^{-2\beta}}$ and hence $(s_2)_{in^{-2\beta}} = (-1)(s_2)_{t_{l-1}}$. Thus we have shown that (A.23) holds. □

We finally have all the ingredients to prove Theorem A.8.

Proof of Theorem A.8. Consider the event:

$$J_i = \left\{ \sup_{u,s\in\left[\frac{i}{n^{2\beta}},\frac{i+1}{n^{2\beta}}\right]} |\widehat{\xi}(u) - \widehat{\xi}(s)| \leq n^{-\beta/6} \right\}. \tag{A.24}$$

We note that this event has high probability. Indeed, looking at the complement we see

$$\begin{aligned}
\mathbb{P}_x\left[\sup_{u,z\in\left[\frac{i-1}{n^{2\beta}},\frac{i}{n^{2\beta}}\right]}|\widehat{\xi}(u)-\widehat{\xi}(z)|>n^{-\beta/6}\right] &\leq \mathbb{P}_x\left[2\mathcal{R}^n_\Omega Z>n^{-\beta/6}\right]\\
&= \mathbb{P}_x\left[2^4(\mathcal{R}^n_\Omega)^4 Z^4>n^{-2\beta/3}\right]\\
&\leq \frac{2^4(\mathcal{R}^n_\Omega)^4\mathcal{O}(\mathbb{E}(Z^4))}{n^{-2\beta/3}}\\
&= \mathcal{O}(n^{-10\beta/3}), \qquad \text{(A.25)}
\end{aligned}$$

with Z a Poisson random variable of mean $\Theta(1)$. In the last computation the first inequality is due to the number of jumps of $\widehat{\xi}$ in an interval of size $n^{-2\beta}$ being a Poisson random variable of parameter $\Theta(1)$, and the size of the jumps being bounded by $2\mathcal{R}^n_\Omega$. The second inequality is just Markov's inequality.

Now consider the Brownian motion $(B_t)_{t\geq 0}$ in (A.16) and reflect all the excursions in the boundary of Ω such that they remain inside Ω. By doing this we obtain a process $(W_t)_{t\geq 0}$. By Theorem 7.3 in [Hsu86] this process is a reflected Brownian motion.

Fix n large enough so that Lemma A.9 holds, and such that

$$\rho_n := n^{-\beta/6}+n^{-2\beta}<\frac{r_0}{4}\wedge\frac{\mathbb{r}-r_0}{2}. \qquad \text{(A.26)}$$

Let $K_n := B((0,r_0),\rho_n)\cup B((0,-r_0),\rho_n)$. Denote by τ_{K_n} the hitting time of K_n for the reflected Brownian motion $(W_t)_{t\geq 0}$ and consider the events

$$A^n_i = \left\{\sup_{u,s\in[v_{i-1},v_i]}|W(s)-W(u)|\leq n^{-\beta/6}\right\}, \qquad \text{(A.27)}$$

$$C^n = \{\tau_{K_n}>t\sigma^2+n^{-\beta/2}\}. \qquad \text{(A.28)}$$

We prove, by induction and under the events $(A_i)_{i=1}^{k+1}, (J_i)_{i=0}^{k}, C^n$, $\{v_k\leq\sigma^2 t+n^{-\beta/2}\}$, that the following family of events hold for every $0\leq i\leq k$. Let $\mathcal{E}_0 := \{W(v_0)=\widehat{\xi}(0)\}$, and define, for $1\leq i\leq k$,

$$\mathcal{E}_i := \left\{W(v_i)=\widehat{\xi}\left(\frac{i}{n^{2\beta}}\right) \text{ and } \inf_{s\in[(i-1)/n^{2\beta},i/n^{2\beta}],a\in\{r_0,-r_0\}} d(\widehat{\xi}(s),(0,a))>n^{-2\beta}\right\}. \qquad \text{(A.29)}$$

We proceed with the induction. We note that the base case, $\mathcal{E}_0$, is trivial since $v_0 = 0$ and we assume $W_0 = \zeta(0) = \widehat{\xi}(0) = x$. Now let us show that $(\mathcal{E}_j)_{j=0}^i$ implies $\mathcal{E}_{i+1}$. Note that, under C^n and $\mathcal{E}_i$, we have that, for $a \in \{r_0, -r_0\}$,

$$d(\widehat{\xi}(i/n^{2\beta}), (0,a)) = d(W(v_i), (0,a)) \geq \rho_n.$$

Under J_i^n, we have $\widehat{\xi}$ makes jumps of magnitude at most $n^{-\beta/6}$. Hence, from inequality (A.26), we obtain

$$\inf_{s \in [i/n^{2\beta}, (i+1)/n^{2\beta}], a \in \{r_0, -r_0\}} d(\widehat{\xi}(s), (0,a)) > \rho_n - n^{-\beta/6} > n^{-2\beta}.$$

Thus, to conclude that $\mathcal{E}_{i+1}$ holds, we simply need to show that

$$W(v_{i+1}) = \widehat{\xi}^n \left(\frac{i+1}{n^{2\beta}} \right). \tag{A.30}$$

For this recall the notation $\Delta_i^n = \zeta(\frac{i+1}{n^{2\beta}}) - \zeta(\frac{i}{n^{2\beta}})$. Then, by equation (A.16), we have that:

$$((\Delta_1^n)_i(s_1)_{in^{-2\beta}}, (\Delta_2^n)_i(s_2)_{in^{-2\beta}}) = \widehat{B}^i(v_{i+1}) - \widehat{\xi}\left(\frac{i}{n^{2\beta}} \right) = \widehat{B}^i(v_{i+1}) - W(v_i), \tag{A.31}$$

for some Brownian motion $(\widehat{B}_t^i)_{t \in [v_i, v_{i+1}]}$ (which is, in fact, a rotation and translation of B from (A.16)). The second equality in (A.31) is from our assumption that $\mathcal{E}_i$ holds. Furthermore, if there were no reflection in the boundary, then we would have that:

$$\widehat{\xi}\left(\frac{i+1}{n^{2\beta}} \right) = \widehat{\xi}\left(\frac{i}{n^{2\beta}} \right) + ((\Delta_1^n)_i(s_1)_{in^{-2\beta}}, (\Delta_2^n)_i(s_2)_{in^{-2\beta}}) = \widehat{B}^i(v_{i+1}), \tag{A.32}$$

$$W(v_{i+1}) = W(v_i) + ((\Delta_1^n)_i(s_1)_{in^{-2\beta}}, (\Delta_2^n)_i(s_2)_{in^{-2\beta}}) = \widehat{B}^i(v_{i+1}), \tag{A.33}$$

and the result would hold. Of course, these relationships are not true when there is reflection. However, they are the starting point of our next argument.

Consider the events

$$\widehat{A}_i^n = \left\{ \sup_{u,s \in [v_{i-1}, v_i]} |\widehat{B}^i(s) - \widehat{B}^i(u)| \leq n^{-\beta/6} \right\}, \tag{A.34}$$

and suppose that $(\widehat{A}_i^n)_{i=1}$ holds for $i = 1, .., k$. Note that then the three processes involved in each step $\widehat{B}^i$, W and $\widehat{\xi}$, under the $\widehat{A}_i^n$, A_i^n and J_i^n at most move $n^{-\beta/6}$ between the times of the inductive step. To make use of this, we partition the domain Ω as

$$\Omega = \Omega_1^n \cup \Omega_2^n \cup \Omega_3^n,$$

where

$$\Omega_1^n = \{x \in \Omega | d(x, \partial\Omega) > n^{-\beta/6}\}, \tag{A.35}$$

$$\Omega_2^n = \{x \in \Omega | d(x, \partial\Omega) \leq n^{-\beta/6} \text{ and } d(x, \kappa) > n^{-\beta/6}\}, \tag{A.36}$$

$$\Omega_3^n = \{x \in \Omega | d(x, \kappa) \leq n^{-\beta/6}\}. \tag{A.37}$$

We note that $\Omega_i^n \cap \Omega_j^n = \emptyset$ for $i, j \in \{1, 2, 3\}$, $i \neq j$.

Suppose first that $W(v_i) = \widehat{\xi}(i/n^{2\beta}) \in \Omega_1^n$. In that case, under $J_i^n, A_i^n, \widehat{A}_i^n$, there would be no reflection (because we are at distance at least $n^{-\beta/6}$ from the boundary), and hence relationships (A.32), (A.33) actually hold, giving equality (A.30).

Now suppose that $W(v_i) = \widehat{\xi}(i/n^{2\beta}) \in \Omega_2^n$, and that $d(W(v_i), (0, R_0)), d(W(v_i), (0, -R_0)) > n^{-\beta/6}$. That is, we are close to a line of the boundary, but not to a corner. To fix some notation, we will suppose this line is given by

$$L = \{x = (x_1, R_0) \in \partial\Omega | x_1 < 0\},$$

but the following argument will work with any line. Since L only gives a reflection in the y-axis, and we are far from any other line, we already have, by equations (A.31), (A.32), (A.33), that

$$(W(v_{i+1}))_1 = (\widehat{B}^i(v_{i+1}))_1 = \left(\widehat{\xi}\left(\frac{i+1}{n^{2\beta}}\right)\right)_1,$$

and hence, we simply need to prove the equality on the other axis. By definition W is just the process B with all the excursions flipped so that it stays inside the domain, and $\widehat{B}$ is a translation and rotation of B, starting at time v_i from the same point as W, but without reflection. Hence, $|W(v_{i+1})_2 - R_0| = |\widehat{B}(v_{i+1})_2 - R_0|$, and since there is only one point inside the domain that fulfils that identity we just need to prove that $|\left(\widehat{\xi}\left(\frac{i+1}{n^{2\beta}}\right)\right)_2 - R_0| = |\widehat{B}^i(v_{i+1})_2 - R_0|$. However, this is immediate from Lemma A.13 and (A.31), so (A.30) holds.

Suppose now $W(v_i) = \widehat{\xi}(i/n^{2\beta}) \in \Omega_2^n$ and that $d(W(v_i), (0, R_0)) \wedge d(W(v_i), (0, -R_0)) \leq n^{-\beta/6}$. In that case the argument made in Lemma A.13 is valid in each component of the process separately, since the reflection in this case occurs independently for each component. Indeed, the x-axis is always reflected at 0, no matter what happens with the y-axis. On the other hand, the y-axis is always reflected at R_0 (or $-R_0$), independently of the x-axis. The notation becomes cumbersome so we omit the complete proof, but writing the argument for each component gives relationship (A.30).

In the last case, $W(v_i) = \widehat{\xi}(i/n^{2\beta}) \in \Omega_3^n$, the arguments we have used so far no longer work. This is due to the reflection in one component being dependent on the value of the other component. However, under C^n and $\{v_k \leq t\sigma^2 + n^{-\beta/2}\}$, we have $W(v_i) \notin K_n$ and $\Omega_3^n \subseteq K_n$, so under C^n and $\{v_k \leq t\sigma^2 + n^{-\beta/2}\}$ we know $W(v_i) = \widehat{\xi}(i/n^{2\beta}) \notin \Omega_3^n$.

In synthesis, we have concluded (A.30) holds under $(A_i)_{i=1}^{k+1}, (\widehat{A}_i)_{i=1}^{k+1}, (J_i)_{i=0}^{k}, C^n$, $\{v_k \leq \sigma^2 t + n^{-\beta/2}\}$, and hence the inductive step is complete and we have that the family of events $(\mathcal{E})_{i=0}^{k}$ holds.

Finally, we consider $(\xi_t)_{t\geq 0}$, a process with transition rates $\widehat{m}_n$, by setting it equal to $(\widehat{\xi}_t)_{t\geq 0}$ until it is at distance less than $2\mathcal{R}_\Omega^n$ from $(0, r_0)$ or $(0, -r_0)$, and afterwards letting it evolve independently. Note that be definition the processes $(\xi_t)_{t\geq 0}$ and $(\widehat{\xi}_t)_{t\geq 0}$ are equal under $(\mathcal{E}_i)_{i=1}^{k}$ up to time v_k. We now write

$$U_i := \left(\bigcap_{j=1}^{i} A_j \cap \widehat{A}_j \cap J_j \right) \cap C^n \cap \{v_{\lfloor tn^{2\beta} \rfloor} \leq \sigma^2 t + n^{-\beta/2}\}.$$

Note that we have proven that $\mathcal{E}_i$ holds under U_i. Now we can perform the following computation:

$$\begin{aligned}
&\mathbb{P}_x \left[|\widehat{\xi}^n(t) - W(\sigma^2 t)| \geq n^{-\beta/6} \right] \\
&\quad \leq \mathbb{P}_x \left[|\widehat{\xi}^n(t) - W(\sigma^2 t)| \geq n^{-\beta/6}, U_{\lfloor tn^{2\beta} \rfloor} \right] + \mathbb{P}_x \left[((\widehat{A}_i^n)_{i=1}^{\lfloor tn^{2\beta} \rfloor})^c \right] + \mathbb{P}_x \left[((A_i^n)_{i=1}^{\lfloor tn^{2\beta} \rfloor})^c \right] \\
&\qquad + \sum_{i=1}^{\lfloor tn^{2\beta} \rfloor} \mathbb{P}_x[(J_i^n)^c] + \mathbb{P}_x[(C_n)^c] + \mathbb{P}_x[v_{\lfloor tn^{2\beta} \rfloor} > \sigma^2 t + n^{-\beta/2}] + \mathcal{O}(n^{-\beta} t),
\end{aligned} \tag{A.38}$$

where the last identity is a union bound. Now note that

$$\mathbb{P}_x \left[((A_i^n)_{i=1}^{\lfloor tn^{2\beta} \rfloor})^c \right] = \mathbb{P}_x \left[\exists i < n^{2\beta} t \text{ s.t. } \sup_{u,s \in [v_{i-1}, v_i]} |W(s) - W(u)| > n^{-\beta/6} \right] = \mathcal{O}(n^{-\beta} t).$$

given Lemma A.11. Furthermore, by the same proof as Lemma A.11, we can show that

$$\mathbb{P}_x \left[((\widehat{A}_i^n)_{i=1}^{\lfloor tn^{2\beta} \rfloor})^c \right] \leq \mathbb{P}_x \left[\exists i < n^{2\beta} t \text{ s.t. } \sup_{u,s \in [v_{i-1}, v_i]} |\widehat{B}^i(s) - \widehat{B}^i(u)| > n^{-\beta/6} \right] = \mathcal{O}(n^{-\beta} t).$$

On the other hand, by (A.25), we have that

$$\sum_{i=1}^{\lfloor tn^{2\beta}\rfloor} \mathbb{P}_x[(J_i^n)^c] = \mathcal{O}(tn^{2\beta}n^{-10\beta/3}) = \mathcal{O}(n^{-4\beta/3}t).$$

By (A.18) we also have that

$$\mathbb{P}_x[v_{\lfloor tn^{2\beta}\rfloor} > t\sigma^2 + n^{-\beta/2}] \le \mathbb{P}_x[|v_{\lfloor tn^{2\beta}\rfloor} - t\sigma^2| > n^{-\beta/2}] = \mathcal{O}(tn^{-\beta}).$$

Now take n sufficiently large so that $n^{-\beta/2} \le \sigma^2\eta^{-1}\varepsilon_n^2|\log(\varepsilon)|$. Note this is possible since $\varepsilon_n^{-2} = o(\log(n))$. Then $\sigma^2 t + n^{-\beta/2} \le \sigma^2(a+\eta^{-1}(k+2))\varepsilon_n^2|\log(\varepsilon_n)|$. By Lemma A.9, we have that

$$\mathbb{P}_x[(C_n)^c] = \mathbb{P}_x[\tau_{K_n} \le t\sigma^2 + n^{-\beta/2}] = \varepsilon_n^{k+1}.$$

From this, and the fact that $U_{\lfloor tn^{2\beta}\rfloor}$ implies $(\mathcal{E}_i)_{i=1}^{\lfloor tn^{2\beta}\rfloor}$, we have that the right hand side of (A.38) can be bounded by

$$\begin{aligned}
&\mathbb{P}_x\left[|\widehat{\xi}^n(t) - W(\sigma^2 t)| \ge n^{-\beta/6}\right] \\
&\le \mathbb{P}_x\left[|\widehat{\xi}^n(t) - W(\sigma^2 t)| \ge n^{-\beta/6}, U_{\lfloor tn^{2\beta}\rfloor}, (\mathcal{E}_i)_{i=1}^{\lfloor tn^{2\beta}\rfloor}\right] + \varepsilon_n^{k+1} + \mathcal{O}(n^{-\beta}t) \\
&\le \mathbb{P}_x\left[\left|\widehat{\xi}^n(t) - \widehat{\xi}\left(\frac{\lfloor tn^{2\beta}\rfloor}{n^{2\beta}}\right)\right| + |W(v_{\lfloor tn^{2\beta}\rfloor}) - W(\sigma^2 t)| \ge n^{-\beta/6}, U_{\lfloor tn^{2\beta}\rfloor}, (\mathcal{E}_i)_{i=1}^{\lfloor tn^{2\beta}\rfloor}\right] \\
&\quad + \varepsilon_n^{k+1} + \mathcal{O}(n^{-\beta}t) \\
&\le \mathbb{P}_x\left[|W(v_{\lfloor tn^{2\beta}\rfloor}) - W(\sigma^2 t)| \ge n^{-\beta/6}/2\right] + \mathbb{P}_x\left[\left|\widehat{\xi} - \widehat{\xi}\left(\frac{\lfloor tn^{2\beta}\rfloor}{n^{2\beta}}\right)\right| \ge n^{-\beta/6}/2\right] \\
&\quad + \varepsilon_n^{k+1} + \mathcal{O}(n^{-\beta}t).
\end{aligned} \tag{A.39}$$

In the above, the the second inequality holds by the fact that $\xi(\frac{\lfloor tn^{2\beta}\rfloor}{n^{2\beta}}) = \widehat{\xi}(\frac{\lfloor tn^{2\beta}\rfloor}{n^{2\beta}}) = W(v_{\lfloor tn^{2\beta}\rfloor})$ under $(\mathcal{E}_i)_{i=1}^{\lfloor tn^{2\beta}\rfloor}$. The first term on the right hand side of (A.39) can be bounded as follows. By computing once again as we did in the proof of Proposition 2.13, we deduce that

$$\mathbb{P}_x[|W(v_{\lfloor sn^{2\beta}\rfloor}) - W(\sigma^2 t)| \ge n^{-\beta/6}/2] \le \mathcal{O}\left(\exp\left(-\frac{1}{8\mathrm{d}^2}n^{\beta/6}\right)\right) + \mathcal{O}(tn^{-\beta}). \tag{A.40}$$

Now for the second term on the right hand side of (A.39), we can use Chebyshev's inequality to obtain

$$\mathbb{P}_x\left[\left|\widehat{\xi}^n(t) + \widehat{\xi}\left(\frac{\lfloor tn^{2\beta}\rfloor}{n^{2\beta}}\right)\right| \ge n^{-\beta/6}/2\right] \le \mathcal{O}(\mathbb{E}_x[|X_1|^2]n^{\beta/3}) = \mathcal{O}(n^{-5\beta/3}), \tag{A.41}$$

here the equality holds from the fact that the number of jumps of ξ in a time of length $n^{2\beta}$ is $\Theta(1)$. Using (A.40) and (A.41) on the right hand side of (A.39) finishes the proof. $\square$

Theorem A.8 is the only important ingredient that was missing to obtain the proof for the weak noise/selection ratio regime of Theorem A.7. The result then follows by using Theorem A.8 instead of Lemma 2.29 in the proof of Theorem 2.27.

A.2.4 Strong noise/selection ratio regime

In the same way that we proved Theorem 2.27 using Lemma 2.37, the strong noise/selection ratio regime of Theorem A.7 follows from the following result.

Lemma A.14. *Let $(\mathcal{P}^n_s)_{0\leq s} := (\zeta^n_1(s), ..., \zeta^n_{N(s)}(s))_{0\leq s}$ be the dual process of the SLFVS on Ω, under the strong noise regime. Let $x \in \Omega$ and $T > 0$. Then there is a Brownian motion reflected in the boundary of Ω, $(W_t)_{t\geq 0}$, and a coupling of $\mathcal{P}^n$ and W such that, for any $\varepsilon > 0$, there is $n_* \in \mathbb{N}$ such that, for all $n \geq n_*$, we have*

$$\mathbb{P}_x\left[(N(T) = 1) \text{ and } (|\zeta^n(T) - W(T\sigma^2)| < n^{-\beta/6})\right] \geq 1 - \varepsilon. \tag{A.42}$$

The proof of Lemma A.14 has two ingredients: proving that with high probability we have only one ancestral lineage at time T, and then proving that it moves, approximately, as a reflected Brownian motion. However, the first part follows from the same proof as Lemma 2.37. Indeed, let $\widehat{\zeta}^{n,1}, \widehat{\zeta}^{n,2}$ denote two ancestral lineages of the SLFVS, coupled to $\zeta^{n,1}, \zeta^{n,2}$ in the same way as in the proof of Theorem A.8. Define $\widehat{\eta}^n := \widehat{\zeta}^{n,1} - \widehat{\zeta}^{n,2}$ and $\eta^n := \zeta^{n,1} - \zeta^{n,2}$. Then, for any $x \in \Omega$ with distance more than $2\log(n)^{-3}$ from $(0, r_0)$ and $(0, -r_0)$, it follows from the geometry of the domain that

$$\mathbb{P}_x[\widehat{\eta}^n \text{ reaches value } \log(n)^{-4}] \leq \mathbb{P}_x[\eta^n \text{ reaches values } 4\log(n)^{-4}]. \tag{A.43}$$

The right hand side of equation (A.43) tends to 0 as we proved in Proposition 2.38. Hence, we just need to prove that the (only) ancestral lineage does indeed evolve like a reflecting Brownian motion with high probability. Note that Theorem A.8 is not enough, since this Theorem only works for a small window of time, where in this case, we need a coupling valid for an arbitrary time. This is obtained by using the following result instead of Lemma A.9 in the proof of Theorem A.8.

As before, we use $\kappa = \{(0, r_0)\} \cup \{(0, -r_0)\}$ and, for $x \in \Omega$, denote by $d(x, \kappa)$ the distance between κ and x.

Lemma A.15. *Let $x \in \Omega$ with $d(x, \kappa) > 0$ and let $\rho > 0$ be such that $\rho < d(x, \kappa)$ and $B((0, r_0), \rho) \cap B((0, -r_0), \rho) = \emptyset$. Let $K = B((0, r_0), \rho) \cup B((0, -r_0), \rho)$ and denote by τ_K the hitting time of K for the reflected Brownian motion W. Then*

$$\mathbb{P}_x(\tau_K \leq t) = \mathcal{O}(|\log(\rho)|^{-1} t) \ \text{ as } \rho \to 0.$$

Proof. Let $\delta > 0$ be sufficiently small that

$$\rho + \delta < d(x, \kappa) \text{ and } B((0, r_0), \rho + \delta) \cap B((0, -r_0), \rho + \delta) = \emptyset.$$

Set $\widehat{K} = B((0, r_0), \rho+\delta) \cup B((0, -r_0), \rho+\delta)$. Since $x \notin \widehat{K}$, we have that, by [GS02, Theorem 3.7],

$$\mathbb{P}_x(\tau_K \leq t) \leq 2\mathrm{cap}(K, \widehat{K}) t, \tag{A.44}$$

where

$$\mathrm{cap}(K, \widehat{K}) = \int_\Omega |\nabla \phi(x)|^2 dx, \tag{A.45}$$

and where ϕ is given by,

$$\phi(s) = \frac{\log(s)}{\log(\rho) - \log(\rho + \delta)} + C$$

for some $C \in \mathbb{R}$. Since $B((0, r_0), \rho + \delta)$ and $B((0, -r_0), \rho + \delta)$ are disjoint, we can bound ϕ on each of those balls to bound $\mathrm{cap}(K, \widehat{K})$. From this, we compute that

$$\begin{aligned} \mathrm{cap}(K, \widehat{K}) &= \frac{3}{2} \frac{\pi}{(\log(\rho) - \log(\rho + \delta))^2} \int_\rho^{\rho+\delta} \frac{1}{s} ds, \\ &= \frac{3}{2} \frac{\pi}{|\log(\rho + \delta) - \log(\rho)|} = \mathcal{O}(|\log(\rho)|^{-1}), \end{aligned}$$

from which we see that $t\mathrm{cap}(K, \widehat{K}) = \mathcal{O}(t|\log(\rho)|^{-1})$ and from this the result follows. □

By changing the proof of Theorem A.8 to use Lemma A.15 instead of Lemma A.9, we obtain the following result. Once we have this, the proof of Lemma A.14 follows the same pattern as the proof of Lemma 2.37.

Theorem A.16. *Let $x \in \Omega$ with $d(x, \kappa) > 0$, and fix $T > 0$. Then there exists $n_* \in \mathbb{N}$ such that, for all $n \geq n_*$, there is a coupling between ζ^n and a reflected Brownian motion W under which:*

$$\mathbb{P}_x \left[|\zeta^n(T) - W(\sigma^2 T)| \geq n^{-\beta/6} \right] \leq \mathcal{O} \left(n^{-\beta}(T \vee 1) + T|\log(n)|^{-1} \right),$$

where ζ^n and W both have the same starting point x.

Remark A.17. One could be tempted to think that Theorem A.16 is stronger than Theorem A.8. We want to emphasise that this is not true. Since $|\log(n)|^{-1}$ is of a bigger order than ε_n^k for $k \geq 3$, Theorem A.8 gives a better bound at the expense of having a small time window where the coupling is valid. Meanwhile, Theorem A.16 is valid for an arbitrary time, paying the price of having a worse bound.

Appendix B

B.1 Fixed points of $g_\times$

Proposition B.1. *The function $g_\times$ has fixed points u_-, $\frac{1}{2}$, and u_+, where*

$$u_- = \frac{1}{2} - \frac{\sqrt{(1-b_I)^3(1-3b_I)}}{2(1-b_I)^3} \tag{B.46}$$

$$u_+ = \frac{1}{2} + \frac{\sqrt{(1-b_I)^3(1-3b_I)}}{2(1-b_I)^3} \tag{B.47}$$

Proof. We aim to find a such that $g_\times\left(\frac{1}{2}+a\right) = \frac{1}{2}+a$. Note that,

$$\begin{aligned} g_\times\left(\tfrac{1}{2}+a\right) &= 3\left((1-b_I)(\tfrac{1}{2}+a)+\tfrac{b_I}{2}\right)^2 - 2\left((1-b_I)(\tfrac{1}{2}+a)+\tfrac{b_I}{2}\right)^3 \\ &= 2(b_I-1)^3a^3 + \tfrac{3}{2}(1-b_I)a + \tfrac{1}{2}. \end{aligned}$$

Setting this equal to $\frac{1}{2}+a$ we obtain the quadratic equation

$$2(1-b_I)^3a^2 + \tfrac{3}{2}(1-b_I) = 0,$$

for which $u_- - \frac{1}{2}$ and $u_+ - \frac{1}{2}$ are clearly solutions. To see that $g_\times(\frac{1}{2}) = \frac{1}{2}$, note that

$$g_\times\left(\tfrac{1}{2}\right) = g\left((1-b_I)\tfrac{1}{2}+\tfrac{b_I}{2}\right) = g\left(\tfrac{1}{2}\right) = \tfrac{1}{2}.$$

□

B.2 Proof of Proposition 3.28

Proof. We only prove the stated upper bound for (3.56), noting that the bound for (3.55) follows by almost identical arguments. To begin, we will construct a coupling of $\boldsymbol{Z}^+(t)$ to a historical ternary branching subordinated Brownian motion, $\boldsymbol{W}_{R^\varepsilon}(t)$. Define

$$u(x,t) := \mathbb{P}_x[\mathbb{V}^\times(\boldsymbol{Z}^+(t)) = 1] \;\text{ and }\; v(x,t) := \mathbb{P}_x[\mathbb{V}^\times(\boldsymbol{W}\boldsymbol{R}^\varepsilon(t)) = 1].$$

Abuse notation and let S denote the time of the first branching event in both $\boldsymbol{Z}^+(t)$ and $\boldsymbol{W}_{R^\varepsilon}(t)$. By the Markov property at time S, u and v can be written as

$$u(x,t) = \mathbb{E}_x[g_\times(u(Z_S^+, t-S))\mathbb{1}_{S\le t}] + \mathbb{E}_x[p(Z_t^+)\mathbb{1}_{S>t}] \tag{B.48}$$

$$v(x,t) = \mathbb{E}_x[g_\times(v(W(R_S^\varepsilon), t-S))\mathbb{1}_{S\le t}] + \mathbb{E}_x[p(W(R_t^\varepsilon))\mathbb{1}_{S>t}]. \tag{B.49}$$

To bound the difference between u and v, we control the difference of the first terms and second terms in (B.48) and (B.49) separately. First, since $S \sim \exp(\varepsilon^{-2})$,

$$\mathbb{E}_x[p(Z_t^+)\mathbb{1}_{S>t}] - \mathbb{E}_x[p(W(R_t^\varepsilon))\mathbb{1}_{S>t}] \le \mathbb{P}[t \le S] \le e^{-\frac{t}{\varepsilon^2}}. \tag{B.50}$$

To bound the difference of the first terms in (B.48) and (B.49), set $t^* := k\varepsilon^2|\log(\varepsilon)|$ and $\delta_\varepsilon := D_0(k+2)I^2(\varepsilon)|\log(\varepsilon)|$. Denote the transition density of $W(R_t^\varepsilon)$ by $p_t(x,y)$. Then, since $g_\times$ is bounded above by one and, by the definition of Z_t^+,

$$\begin{aligned}
&\mathbb{E}_x\left[g_\times(u(Z_S^+, t-S))\mathbb{1}_{S\le t}\right] \\
&\quad\le \mathbb{E}_x\left[g_\times(u(Z_S^+, t-S))\mathbb{1}_{S\le t\wedge t^*}\right] + \mathbb{P}[S > t^*] \\
&\quad\le \sup_{\|w\|\le\delta_\varepsilon} \mathbb{E}_x\left[g_\times(u(W(R_S^\varepsilon)+w, t-S))\mathbb{1}_{S\le t\wedge t^*}\right] + \varepsilon^k \\
&\quad= \sup_{\|w\|\le\delta_\varepsilon} \mathbb{E}\left[\int_{\mathbb{R}^{\mathrm{d}}} g_\times(u(z+w, t-S))p_S(x,z)dz\mathbb{1}_{S\le t\wedge t^*}\right] + \varepsilon^k \\
&\quad= \sup_{\|w\|\le\delta_\varepsilon} \mathbb{E}\left[\int_{\mathbb{R}^{\mathrm{d}}} g_\times(u(z, t-S))p_S(x,z-w)dz\mathbb{1}_{S\le t\wedge t^*}\right] + \varepsilon^k \\
&\quad\le \mathbb{E}\left[\int_{\mathbb{R}^{\mathrm{d}}} g_\times(u(z, t-S))p_S(x,z)dz\mathbb{1}_{S\le t\wedge t^*}\right] + \varepsilon^k \\
&\qquad + \sup_{\|w\|\le\delta_\varepsilon} \mathbb{E}\left[\int_{\mathbb{R}^{\mathrm{d}}} g_\times(u(z, t-S))[p_S(x,z-w) - p_S(x,z)]\mathbb{1}_{S\le t\wedge t^*}dz\right].
\end{aligned} \tag{B.51}$$

Consider the d-ball $B_x(S^{1/\alpha}+\delta_\varepsilon) := \{z \in \mathbb{R}^{\mathrm{d}} : \|z-x\| \le S^{1/\alpha}+\delta_\varepsilon\}$. To ease notation, let $B_x := B_x(S^{1/\alpha}+\delta_\varepsilon)$. Then, to bound the last term in (B.51), we use that $g(u(z,t-S))$ is bounded by 1 to get

$$\begin{aligned}
&\sup_{\|w\|\le\delta_\varepsilon} \mathbb{E}\left[\int_{\mathbb{R}^{\mathrm{d}}} g(u(z, t-S))[p_S(x,z-w) - p_S(x,z)]\mathbb{1}_{S\le t\wedge t^*}dz\right] \\
&\quad\le \sup_{\|w\|\le\delta_\varepsilon} \mathbb{E}\left[\int_{\mathbb{R}^{\mathrm{d}}} |p_S(x,z-w) - p_S(x,z)|\mathbb{1}_{S\le t\wedge t^*}dz\right]
\end{aligned}$$

$$\leq \sup_{\|w\|\leq\delta_\varepsilon} \mathbb{E}\left[\int_{B_x} |p_S(x,z-w)-p_S(x,z)|dz \mathbb{1}_{S\leq t\wedge t^*}\right]$$
$$+ \sup_{\|w\|\leq\delta_\varepsilon} \mathbb{E}\left[\int_{B_x^c} |p_S(x,z-w)-p_S(x,z)|dz \mathbb{1}_{S\leq t\wedge t^*}\right] \tag{B.52}$$

To bound the first term on the right hand side of (B.52), first note that

$$\mathbb{P}[S<\delta_\varepsilon^\alpha] \leq 1-\exp\left(-\frac{\delta_\varepsilon^\alpha}{\varepsilon^2}\right) \leq \frac{\widehat{I}(\varepsilon)^{2\alpha}}{\varepsilon^2}|\log(\varepsilon)|^\alpha.$$

By Proposition B.5, using that the first integral is bounded above by two, and allowing the constant C to change from line to line,

$$\begin{aligned}
&\sup_{\|w\|\leq\delta_\varepsilon} \mathbb{E}\left[\int_{B_x} |p_S(x,z-w)-p_S(x,z)|dz \mathbb{1}_{S\leq t\wedge t^*}\right]\\
&\leq \sup_{\|w\|\leq\delta_\varepsilon} C\mathbb{E}\left[\int_{B_x} \|w\|(R_S^\varepsilon)^{-(\mathbb{d}+1)/2} dz \mathbb{1}_{S\leq t\wedge t^*}\mathbb{1}_{S\geq\delta_\varepsilon^\alpha}\right] + 2\frac{\widehat{I}(\varepsilon)^{2\alpha}}{\varepsilon^2}|\log(\varepsilon)|^\alpha\\
&\leq C\delta_\varepsilon\mathbb{E}\left[V(B_x)(R_S^\varepsilon)^{-(\mathbb{d}+1)/2}\mathbb{1}_{S\geq\delta_\varepsilon^\alpha}\right] + 2\frac{\widehat{I}(\varepsilon)^{2\alpha}}{\varepsilon^2}|\log(\varepsilon)|^\alpha\\
&\leq C\delta_\varepsilon\mathbb{E}\left[(2S)^{\mathbb{d}/\alpha}(R_S^\varepsilon)^{-(\mathbb{d}+1)/2}\mathbb{1}_{S\geq\delta_\varepsilon^\alpha}\right] + 2\frac{\widehat{I}(\varepsilon)^{2\alpha}}{\varepsilon^2}|\log(\varepsilon)|^\alpha\\
&\leq C\delta_\varepsilon\mathbb{E}\left[S^{\mathbb{d}/\alpha-\mathbb{d}/2-1/2}\mathbb{1}_{S\geq\delta_\varepsilon^\alpha}\right] + 2\frac{\widehat{I}(\varepsilon)^{2\alpha}}{\varepsilon^2}|\log(\varepsilon)|^\alpha\\
&\leq C\delta_\varepsilon\mathbb{E}\left[S^{\mathbb{d}/\alpha-\mathbb{d}/2-1/2}\right] + 2\frac{\widehat{I}(\varepsilon)^{2\alpha}}{\varepsilon^2}|\log(\varepsilon)|^\alpha\\
&= CD_0(k+2)\widehat{I}(\varepsilon)^2|\log(\varepsilon)|\varepsilon^{\frac{2\mathbb{d}}{\alpha}-\mathbb{d}-1} + 2\frac{\widehat{I}(\varepsilon)^{2\alpha}}{\varepsilon^2}|\log(\varepsilon)|^\alpha
\end{aligned} \tag{B.53}$$

where $V(B_x)$ denotes the volume of B_x. $V(B_x)$, in turn, is proportional to $(S^{1/\alpha}+\delta_\varepsilon)^{\mathbb{d}}$, which, conditionally on $S\geq\delta_\varepsilon^\alpha$, is bounded above by $(2S)^{\mathbb{d}/\alpha}$. In the final equality we apply Lemma B.4. Note that $\frac{\mathbb{d}}{\alpha}-\frac{\mathbb{d}}{2}-\frac{1}{2}>-1$ for all $\mathbb{d}\geq 2$, and $\alpha\leq 2$, so $\mathbb{E}\left[S^{\frac{\mathbb{d}}{\alpha}-\frac{(\mathbb{d}+1)}{2}}\right]$ is finite, with

$$\mathbb{E}\left[S^{\frac{\mathbb{d}}{\alpha}-\frac{(\mathbb{d}+1)}{2}}\right] = \Gamma\left(\tfrac{\mathbb{d}}{\alpha}-\tfrac{\mathbb{d}}{2}+\tfrac{1}{2}\right)\varepsilon^{\frac{2\mathbb{d}}{\alpha}-\mathbb{d}-1}.$$

Then, by Assumption 1.37 (C), (B.53) tends to 0 with ε.

To bound the second term in (B.52), we use [CK03, Theorem 1.1], which provides upper bounds on the transition density of subordinated Brownian motion, as long this has a truncated stable subordinator with truncation level independent of ε Thus, to use said result we need to scale in such a way that our subordinator does not depend on ε. That is, we rewrite $W(R^\varepsilon_s)$ in terms of a 1-truncated subordinated Brownian motion as follows. Let $(U^a_s)_{s\geq 0}$ denote an $\alpha/2$-stable subordinator with truncation level a (and no speed change). Let $\overset{D}{=}$ denote equality in distribution. Then

$$R^\varepsilon_s \overset{D}{=} U^{\widehat{I}(\varepsilon)^2}_{s\widehat{I}(\varepsilon)^{\alpha-2}} \overset{D}{=} \widehat{I}(\varepsilon)^2 U^1_{s\widehat{I}(\varepsilon)^{-2}}$$

where the first equality follows by definition and the second equality is straightforward to show using the jump measure of the truncated subordinator. As such, by the scaling property of the Brownian motion,

$$W(R^\varepsilon_s) \overset{D}{=} \widehat{I}(\varepsilon) W(U^1_{t\widehat{I}(\varepsilon)^{-2}}). \tag{B.54}$$

Denote the transition density of $(W(U^1_t))_{t\geq 0}$ by $\hat{p}_t(x,y)$. By (B.54), $\hat{p}_t(x,y)$ is related to the transition density of $W(R^\varepsilon_s)$ by

$$p_t(x,y) = \hat{p}_{\widehat{I}(\varepsilon)^{-2}t}(\widehat{I}(\varepsilon)^{-1}x, \widehat{I}(\varepsilon)^{-1}y).$$

Therefore, the second term in (B.52) can be rewritten as

$$\sup_{\|w\|\leq\delta_\varepsilon} \mathbb{E}\left[\int_{B^c_x} |\hat{p}_{\widehat{I}(\varepsilon)^{-2}S}(\widehat{I}(\varepsilon)^{-1}x, \widehat{I}(\varepsilon)^{-1}(z-w)) - \hat{p}_{\widehat{I}(\varepsilon)^{-2}S}(\widehat{I}(\varepsilon)^{-1}x, \widehat{I}(\varepsilon)^{-1}(z))| dz \mathbb{1}_{S\leq t\wedge t^*}\right],$$

which, by [CK03, Theorem 1.1], is bounded above by

$$\begin{aligned}
&\sup_{\|w\|\leq\delta_\varepsilon} \mathbb{E}\left[\int_{B^c_x} \frac{DS\widehat{I}(\varepsilon)^{-2}}{((z-w)\widehat{I}(\varepsilon)^{-1})^{\mathrm{d}+\alpha}} dz\right] + \mathbb{E}\left[\int_{B^c_x} \frac{DS\widehat{I}(\varepsilon)^{-2}}{(z\widehat{I}(\varepsilon)^{-1})^{\mathrm{d}+\alpha}} dz\right] \\
&\quad = \sup_{\|w\|\leq\delta_\varepsilon} \mathbb{E}\left[\int_{B^c_x} \frac{DS\widehat{I}(\varepsilon)^{\alpha-1}}{(z-w)^{\mathrm{d}+\alpha}} dz\right] + \mathbb{E}\left[\int_{B^c_x} \frac{DS\widehat{I}(\varepsilon)^{\alpha-1}}{z^{\mathrm{d}+\alpha}} dz\right] \\
&\quad \leq D\widehat{I}(\varepsilon)^{\alpha-1}\mathbb{E}\left[S\int_{S^{1/\alpha}}^\infty \frac{1}{z^{\alpha+1}} dz\right] \\
&\quad \leq D\widehat{I}(\varepsilon)^{\alpha-1}\mathbb{E}[S(S^{-1})] \\
&\quad = D\widehat{I}(\varepsilon)^{\alpha-1},
\end{aligned} \tag{B.55}$$

for some $D > 0$ that we allow to change from line to line. Here we have applied the change of variables $z - w \mapsto z$ in the fourth line, and Lemma B.4 in the final inequality. Note that, since $\alpha > 1$, the last quantity tends to 0 as $\varepsilon \to 0$. Recall from (1.35) that

$$F(\varepsilon, M) = M\left(\widehat{I}(\varepsilon)^2 |\log(\varepsilon)| \varepsilon^{\frac{2\mathbb{d}}{\alpha} - \mathbb{d} - 1} + \frac{\widehat{I}(\varepsilon)^{2\alpha}}{\varepsilon^2} |\log(\varepsilon)|^\alpha + \widehat{I}(\varepsilon)^{\alpha-1}\right).$$

Then, choosing $M > 0$ and ε sufficiently small that

$$F(\varepsilon, M) \geq \varepsilon^k + D\widehat{I}(\varepsilon)^{\alpha-1} + CD_0(k+2)\widehat{I}(\varepsilon)^2 |\log(\varepsilon)| \varepsilon^{\frac{2\mathbb{d}}{\alpha} - \mathbb{d} - 1} + 2\frac{\widehat{I}(\varepsilon)^{2\alpha}}{\varepsilon^2} |\log(\varepsilon)|^\alpha,$$

by (B.55) and (B.53) we can bound (B.52) to obtain

$$\mathbb{E}_x\left[g_\times(u(Z_S^+, t-S))\mathbb{1}_{S\leq t}\right] \leq \mathbb{E}\left[\int_{\mathbb{R}^{\mathbb{d}}} g_\times(u(z, t-S))p_S(x,z)dz \mathbb{1}_{S\leq t\wedge t^*}\right] + F(\varepsilon, M). \quad \text{(B.56)}$$

Finally, we note that

$$\begin{aligned}
\mathbb{E}_x\left[g_\times(u(W(R_S^\varepsilon), t-S)\mathbb{1}_{S\leq t}\right] &\geq \mathbb{E}_x\left[g_\times(u(W(R_S^\varepsilon), t-S)\mathbb{1}_{S\leq t\wedge t^*}\right] \\
&= \mathbb{E}\left[\int_{\mathbb{R}^{\mathbb{d}}} g_\times(u(z, t-S))p_S(x,z)dz \mathbb{1}_{S\leq t\wedge t^*}\right],
\end{aligned}$$

which, together with (B.56) and (B.50) gives

$$\begin{aligned}
u(x,t) - v(x,t) \leq \mathbb{E}&\left[\int g_\times(u(z,t-S))p_S(x,z)dz \mathbb{1}_{S\leq t\wedge t^*}\right] \\
&- \mathbb{E}\left[\int g_\times(v(z,t-S))p_S(x,z)dz \mathbb{1}_{S\leq t\wedge t^*}\right] + e^{-\frac{t}{\varepsilon^2}} + F(\varepsilon, M).
\end{aligned}$$

Using the same approach, we can obtain the lower bound

$$\begin{aligned}
u(x,t) - v(x,t) \geq \mathbb{E}&\left[\int g_\times(u(z,t-S))p_S(x,z)dz \mathbb{1}_{S\leq t\wedge t^*}\right] \\
&- \mathbb{E}\left[\int g_\times(v(z,t-S))p_S(x,z)dz \mathbb{1}_{S\leq t\wedge t^*}\right] \\
&- \left(e^{-\frac{t}{\varepsilon^2}} + F(\varepsilon, M)\right).
\end{aligned}$$

Thus, using that $g_\times$ is Lipschitz with constant $\frac{3}{2}$, we obtain

$$\begin{aligned}
|u(x,t) - v(x,t)| \leq \mathbb{E}&\left[\left|\int (g(u(z,t-S)) - g(v(z,t-S)))p_S(x,z)dz\right| \mathbb{1}_{S\leq t}\right] \\
&+ e^{-\frac{t}{\varepsilon^2}} + F(\varepsilon, M) \\
\leq\ & \frac{3}{2}\mathbb{E}\left[|u(W(R_S^\varepsilon), t-S) - v(W(R_S^\varepsilon), t-S)|\mathbb{1}_{S\leq t}\right]
\end{aligned}$$

$$
\begin{aligned}
&+ e^{-\frac{t}{\varepsilon^2}} + F(\varepsilon, M) \\
\leq\ & \frac{3}{2}\int_0^t \|u(\cdot,\rho) - v(\cdot, t-\rho)\|_\infty e^{-(t-\rho)\varepsilon^{-2}} \varepsilon^{-2} d\rho \\
&+ e^{-\frac{t}{\varepsilon^2}} + F(\varepsilon, M).
\end{aligned}
$$

Finally, using Gronwall's inequality [MF72] (see also [Dra02, Theorem 15]) we deduce that

$$
\begin{aligned}
\|u(\cdot,t) - v(\cdot,t)\|_\infty &\leq \left(e^{-\frac{t}{\varepsilon^2}} + F(\varepsilon, M)\right) \exp\left(\frac{3}{2}\int_0^t \exp(-u\varepsilon^{-2})\varepsilon^{-2} du\right), \\
&\leq \left(e^{-\frac{t}{\varepsilon^2}} + F(\varepsilon, M)\right) \exp\left(\frac{3}{2}\right),
\end{aligned}
$$

which gives the desired bound by choosing appropriate values $M_1, M_2 > 0$. □

B.3 Truncated subordinator calculations

Proposition B.2. *Let $(R^\varepsilon_t)_{t\geq 0}$ be the $\widehat{I}(\varepsilon)^2$-truncated stable subordinator run at speed $\sigma^{-2}\widehat{I}(\varepsilon)^{\alpha-2}$. Then*

$$
\mathbb{E}(R^\varepsilon_s) = s.
$$

Proof. Let R^ε_s be the truncated subordinator. Then

$$
\mathbb{E}(e^{-\lambda R^\varepsilon_s}) = \exp\left(C_\alpha \sigma^{-2} s \widehat{I}(\varepsilon)^{\alpha-2} \int_0^{\widehat{I}(\varepsilon)^2} \frac{e^{-\lambda y} - 1}{y^{\frac{\alpha}{2}+1}} dy\right).
$$

Using integration by parts, the above exponent can be written as

$$
\int_0^{\widehat{I}(\varepsilon)^2} \frac{e^{-\lambda y} - 1}{y^{\frac{\alpha}{2}+1}} dy = -\frac{2}{\alpha}\widehat{I}(\varepsilon)^{-\alpha}(e^{-\lambda \widehat{I}(\varepsilon)^2} - 1) - \frac{2}{\alpha}\lambda^{\alpha/2}\gamma(1-\tfrac{\alpha}{2}, \lambda\widehat{I}(\varepsilon)^2)
$$

where γ denotes the lower incomplete gamma function. Writing the Laplace transform of R^ε_s as $\mathbb{E}(e^{-\lambda R^\varepsilon_s}) = \phi(\lambda)$, we have

$$
\mathbb{E}(e^{-\lambda R^\varepsilon_s}) = \exp\left(-\frac{2}{\alpha}C_\alpha s \sigma^{-2}\widehat{I}(\varepsilon)^{-2}(e^{-\lambda\widehat{I}(\varepsilon)^2} - 1) - \frac{2}{\alpha}C_\alpha s\sigma^{-2}\widehat{I}(\varepsilon)^{\alpha-2}\lambda^{\alpha/2}\gamma(1-\tfrac{\alpha}{2}, \lambda\widehat{I}(\varepsilon)^2)\right).
$$

Recall that $\mathbb{E}(R_s) = -\frac{d}{d\lambda}\phi(\lambda)|_{\lambda=0}$. Here,

$$
-\frac{d}{d\lambda}\phi(\lambda)|_{\lambda=0} = \frac{2}{\alpha}C_\alpha\sigma^{-2}s\,\frac{d}{d\lambda}\left(\widehat{I}(\varepsilon)^{-2}(e^{-\lambda\widehat{I}(\varepsilon)^2} - 1) + \widehat{I}(\varepsilon)^{\alpha-2}\lambda^{\alpha/2}\gamma(1-\tfrac{\alpha}{2}, \lambda\widehat{I}(\varepsilon)^2)\right)\Bigg|_{\lambda=0}. \tag{B.57}
$$

This derivative can be calculated by considering the following identities for γ. First, by the definition of γ and a change of variables

$$\gamma(1-\tfrac{\alpha}{2},\lambda\widehat{I}(\varepsilon)^2)=\widehat{I}(\varepsilon)^{2-\alpha}\int_0^{\lambda} z^{-\alpha/2}e^{-z\widehat{I}(\varepsilon)^2}dz.$$

Therefore

$$\frac{d}{d\lambda}\gamma(1-\tfrac{\alpha}{2},\lambda\widehat{I}(\varepsilon)^2)=\widehat{I}(\varepsilon)^{2-\alpha}\lambda^{-\alpha/2}e^{-\lambda\widehat{I}(\varepsilon)^2}.$$

Under a different change of variables, γ also satisfies

$$\lambda^{\alpha/2-1}\gamma(1-\tfrac{\alpha}{2},\lambda\widehat{I}(\varepsilon)^2)=\int_0^{\widehat{I}(\varepsilon)^2} z^{-\alpha/2}e^{-z\lambda}dz,$$

which, by Taylor's theorem, is equal to $\frac{2\widehat{I}(\varepsilon)^{2-\alpha}}{2-\alpha}$ when $\lambda=0$. Substituting this into (B.57), we obtain

$$-\frac{d}{d\lambda}\phi(\lambda)|_{\lambda=0}=\frac{2}{2-\alpha}C_\alpha\sigma^{-2}s=s.$$

$\square$

Lemma B.3. *Let R_s^ε be an $\widehat{I}(\varepsilon)^2$-truncated $\frac{\alpha}{2}$-stable subordinator, run at speed $\sigma^{-2}\widehat{I}(\varepsilon)^{\alpha-2}$. Then there exists $\varepsilon_1>0$ such that, for all $\varepsilon\in(0,\varepsilon_1)$ and $s\le\varepsilon^2|\log\varepsilon|$,*

$$\mathbb{P}_x[|R_s^\varepsilon-s|>(k+1)\widehat{I}(\varepsilon)^2|\log\varepsilon|]\le\varepsilon^k.$$

Proof. Let $y\in\mathbb{R}$ and consider

$$\mathbb{P}_x[|R_s^\varepsilon-s|>y]=\mathbb{P}[R_s^\varepsilon>y+s]+\mathbb{P}[R_s^\varepsilon<s-y]. \tag{B.58}$$

Recall that the Lévy measure of R_s^ε is

$$\nu(dv)=C_\alpha\sigma^{-2}\widehat{I}(\varepsilon)^{\alpha-2}\frac{1}{\|v\|^{1+\frac{\alpha}{2}}}\mathbb{1}_{\{0\le v\le\widehat{I}(\varepsilon)^2\}}dv.$$

By considering the Laplace transform of R_s^ε, for $\lambda>0$

$$\begin{aligned}\mathbb{E}_x[e^{\lambda R_s^\varepsilon-\lambda\mathbb{E}(R_s^\varepsilon)}]&\le\exp\left(C_\alpha\sigma^{-2}s\widehat{I}(\varepsilon)^{\alpha-2}\int_0^{\widehat{I}(\varepsilon)^2}\frac{e^{\lambda v}-1-\lambda v}{v^{1+\frac{\alpha}{2}}}dv\right)\\&\le\exp\left(2\sigma^{-2}C_\alpha s\widehat{I}(\varepsilon)^{\alpha-2}\lambda^2\int_0^{\widehat{I}(\varepsilon)^2}v^{1-\frac{\alpha}{2}}dv\right)\\&=\exp\left(m_\alpha s\widehat{I}(\varepsilon)^2\lambda^2\right),\end{aligned}$$

where $m_\alpha := \sigma^{-2}\frac{4C_\alpha}{4-\alpha}$ and the second inequality holds for $\widehat{I}(\varepsilon)^2\lambda \leq 1$. By Proposition B.2 and Markov's inequality,

$$\begin{aligned}\mathbb{P}_x[R_s^\varepsilon > y+s] &= \mathbb{P}_x[R_s^\varepsilon - \mathbb{E}(R_s^\varepsilon) > y] \\ &= \mathbb{P}_x[e^{\lambda R_s^\varepsilon - \lambda\mathbb{E}(R_s^\varepsilon)} > e^{\lambda y}] \\ &\leq \mathbb{E}[e^{\lambda R_s^\varepsilon - \lambda\mathbb{E}(R_s^\varepsilon)}]e^{-\lambda y} \\ &\leq \exp\left(m_\alpha s\widehat{I}(\varepsilon)^2\lambda^2 - \lambda y\right).\end{aligned}$$

The latter can be bounded above by ε^k if $\lambda - m_\alpha s\widehat{I}(\varepsilon)^2\lambda^2 > k|\log\varepsilon|$. Setting $\lambda := \widehat{I}(\varepsilon)^{-2}$ and $y = (k+1)\widehat{I}(\varepsilon)^2|\log\varepsilon|$, this becomes

$$(k+1)|\log\varepsilon| - m_\alpha\frac{s}{\widehat{I}(\varepsilon)^2} > k|\log\varepsilon|,$$

which holds for ε sufficiently small by Assumption 1.37 (B). This gives us control over the first term of (B.58). For the second term, note that, for ε sufficiently small, $s - y < 0$. Thus, by non-negativity of the subordinator, this term is zero. $\square$

Lemma B.4. *Let R_t^ε be the $\widehat{I}(\varepsilon)^2$-truncated $\frac{\alpha}{2}$-stable subordinator run at speed $\sigma^{-2}\widehat{I}(\varepsilon)^{\alpha-2}$. Then, for any $p > 0$, there is some $C > 0$ such that*

$$Cs^{-p} \leq \mathbb{E}\left((R_s^\varepsilon)^{-p}\right) \leq Cs^{-p}.$$

Proof. Similar to the proof of Proposition B.2, the Laplace transform of R_s^ε is given by

$$\begin{aligned}\mathbb{E}(e^{-\lambda R_s^\varepsilon}) &= \exp\left(-\frac{2}{\alpha}C_\alpha\sigma^{-2}s\widehat{I}(\varepsilon)^{-2}(e^{-\lambda\widehat{I}(\varepsilon)^2}-1) - \frac{2}{\alpha}C_\alpha\sigma^{-2}s\widehat{I}(\varepsilon)^{\alpha-2}\lambda^{\alpha/2}\gamma(1-\tfrac{\alpha}{2}, \lambda\widehat{I}(\varepsilon)^2)\right) \\ &\leq \exp\left(-\frac{2}{\alpha}C_\alpha\sigma^{-2}s(\lambda + O(\lambda^2\widehat{I}(\varepsilon)^2))\right).\end{aligned}$$

By [Sch02, Theorem 1.1] and the above bound,

$$\begin{aligned}\mathbb{E}\left((R_s^\varepsilon)^{-1}\right) &\leq \int_0^\infty \exp\left(-\frac{2}{\alpha}C_\alpha\sigma^{-\alpha}s(\lambda + O\left(\lambda^2\widehat{I}(\varepsilon)^2\right))\right)d\lambda \\ &\leq \frac{\alpha\sigma^\alpha}{2C_\alpha s},\end{aligned}$$

using that $e^{-x-y} < e^{-x}$ for $x, y > 0$. More generally,

$$\begin{aligned}\mathbb{E}\left((R_s^\varepsilon)^{-p}\right) &\le \int_0^\infty \exp\left(-\frac{2}{\alpha}C_\alpha \sigma^{-\alpha} s(\lambda^{1/p} + O\left(\lambda^{2/p}\widehat{I}(\varepsilon)^2\right))\right) d\lambda \\ &\le \int_0^\infty \exp\left(-\frac{2}{\alpha}C_\alpha\sigma^{-\alpha}s\lambda^{1/p}\right) d\lambda \\ &= p(\tfrac{2}{\alpha}C_\alpha\sigma^{-\alpha}s)^{-p}\Gamma(p).\end{aligned}$$

The lower bound follows by the same approach. ◻

Proposition B.5. *Let $p(x, y, t)$ be the transition density of $W(R_t^\varepsilon)$, and let S be an exponential random variable of parameter ε^{-2}. Then there is a constant $C > 0$ such that, conditionally on R_S^ε, we have*

$$|p(x, y, S) - p(x, y + z, S)| \le C(R_S^\varepsilon)^{-\frac{\mathrm{d}+1}{2}} \|z\|,$$

for all $x, y, z \in \mathbb{R}^\mathrm{d}$.

Proof. Let $q(x, y, t)$ be the transition density of the Brownian motion $(W_t)_{t\ge 0}$. By definition, conditionally on R_S^ε, we have that $p(x, y, S) = q(x, y, R_S)$. Hence, by the law of total probability, it will be enough to prove that

$$|q(x, y, R_S^\varepsilon) - q(x, y + z, R_S^\varepsilon)| \le C(R_S^\varepsilon)^{-\frac{\mathrm{d}+1}{2}} \|z\|.$$

Recall that $q(x, y, R_S^\varepsilon)$ is given explicitly by

$$q(x, y, R_S^\varepsilon) = \frac{1}{(4\pi R_S^\varepsilon)^{\frac{\mathrm{d}}{2}}} \exp\left(-\frac{\|x - y\|^2}{4R_S^\varepsilon}\right).$$

Therefore, by the Mean Value Theorem,

$$|q(x, y, R_S^\varepsilon) - q(x, y + z, R_S^\varepsilon)| \le \frac{1}{(4\pi R_S^\varepsilon)^{\frac{\mathrm{d}}{2}}} \|z\| \left\|\nabla \exp\left(-\frac{\|\xi\|^2}{4R_S^\varepsilon}\right)\right\|$$

for some ξ in the line segment between $y - x$ and $y + z - x$. Now

$$\left\|\nabla \exp\left(-\frac{\xi^2}{4R_S^\varepsilon}\right)\right\| = \frac{\|\xi\|}{2R_S^\varepsilon} \exp\left(-\frac{\|\xi\|^2}{4R_S^\varepsilon}\right) \le \frac{1}{(R_S^\varepsilon)^{\frac{1}{2}}} C$$

for some $C > 0$, since $x\exp(-x^2)$ is uniformly bounded. Hence there is some $C > 0$ such that

$$|q(x, y, R_S^\varepsilon) - q(x, y + z, R_S^\varepsilon)| \le \frac{C\|z\|}{(R_S^\varepsilon)^{\frac{\mathrm{d}+1}{2}}},$$

as claimed. ◻