Electronic Journal of Probability

Electron. J. Probab. **27** (2022), article no. 121, 1–99.
ISSN: 1083-6489 https://doi.org/10.1214/22-EJP845

Genealogies in bistable waves*

Alison Etheridge† Sarah Penington‡

Abstract

We study a model of selection acting on a diploid population (one in which each individual carries two copies of each gene) living in one spatial dimension. We suppose a particular gene appears in two forms (alleles) A and a, and that individuals carrying AA have a higher fitness than aa individuals, while Aa individuals have a lower fitness than both AA and aa individuals. The proportion of advantageous A alleles expands through the population approximately according to a travelling wave. We prove that on a suitable timescale, the genealogy of a sample of A alleles taken from near the wavefront converges to a Kingman coalescent as the population density goes to infinity. This contrasts with the case of directional selection in which the corresponding limit is thought to be the Bolthausen-Sznitman coalescent. The proof uses 'tracer dynamics'.

Keywords: coalescent process; travelling wave; selection.
MSC2020 subject classifications: 60J90; 92D10; 60J27.
Submitted to EJP on October 20, 2021, final version accepted on August 30, 2022.

1 Introduction and main results

Our interest in this work is in modelling the pattern of genetic variation left behind when a gene that is favoured by natural selection 'sweeps' through a spatially structured population in a travelling wave. The interaction between natural selection and spatial structure is a classical problem; the novelty of what we propose here is that we replace the simple directional selection considered in the majority of the mathematical work in this area by a model of selection acting on diploid individuals (carrying two copies of the gene in question) that provides a toy model for the dynamics of so-called hybrid zones. Hybrid zones are widespread in naturally occurring populations, [4], and there is a wealth of recent empirical work on their dynamics; see [1] for an example and a brief discussion. In our simple model, we shall suppose that the population is living in one spatial dimension, and that the gene has exactly two forms (alleles), A and a, and

*SP is supported by a Royal Society University Research Fellowship.
†Department of Statistics, University of Oxford, UK. E-mail: etheridg@stats.ox.ac.uk
‡Department of Mathematical Sciences, University of Bath, UK. E-mail: s.penington@bath.ac.uk

that type AA individuals are at a selective advantage over aa individuals, but that Aa individuals are at a selective disadvantage relative to both.

Our goal is to understand the genealogical trees that describe the relationships between individual genes sampled from the present day population. In the case of directional selection, there is a large body of work, of varying degrees of rigour, that suggests that if we take a sample of favoured individuals from close to the wavefront then, on suitable timescales, their genealogy is described by the so-called Bolthausen-Sznitman coalescent. In our models, where expansion of the favoured type is driven from the bulk of the wave, we shall see that the corresponding object is the classical Kingman coalescent.

Before giving a precise mathematical definition of our model in Section 1.1 and stating our main results in Section 1.2, we place our work in context.

Directional selection: the (stochastic) Fisher-KPP equation

The mathematical modelling of the way in which a genetic type favoured by natural selection spreads through a population that is distributed across space can be traced back at least to Fisher ([17]) and Kolmogorov, Petrovsky & Piscounov ([23]). They introduced the now classical Fisher-KPP equation,

$$\begin{aligned} \frac{\partial p}{\partial t}(t,x) = \frac{m}{2}\Delta p(t,x) + s_0 p(t,x)\big(1-p(t,x)\big) \qquad & \text{for } x\in\mathbb{R},\ t>0, \\ 0\le p(0,x)\le 1 \qquad & \forall x\in\mathbb{R}, \end{aligned} \tag{1.1}$$

as a model for the way in which the proportion $p(t,x)$ of genes that are of the favoured type changes with time. A shortcoming of this equation is that it does not take account of random genetic drift, that is, the randomness due to reproduction in a finite population. The classical way to introduce such randomness is through a Wright-Fisher noise term, so that the equation becomes

$$dp(t,x) = \frac{m}{2}\Delta p(t,x)dt + s_0 p(t,x)\big(1-p(t,x)\big)dt + \sqrt{\frac{1}{\rho_e}p(t,x)\big(1-p(t,x)\big)}W(dt,dx), \tag{1.2}$$

where W is a space-time white noise and ρ_e is an effective population density. This is a continuous space analogue of Kimura's stepping stone model [22], with the additional non-linear term capturing selection. This equation has the limitation that it only makes sense in one space dimension, but like (1.1) it exhibits travelling wave solutions ([27]) which can be thought of as modelling a selectively favoured type 'sweeping' through the population and, consequently, it has been the object of intensive study.

From a biological perspective, the power of mathematical models is that they can throw some light on the patterns of genetic variation that one might expect to see in the present day population if it has been subject to natural selection. Neither of the models above is adequate for this task. If it survives at all, one can expect a selectively favoured type to eventually be carried by all individuals in a population and from simply observing that type, we have no way of knowing whether it is fixed in the population as a result of natural selection, or purely by chance. However, in reality, it is not just a single letter in the DNA sequence that is modelled by the equation, but a whole stretch of genome that is passed down intact from parent to offspring, and on which we can expect some neutral mutations to arise. The pattern of *neutral* variation can be understood if we know how individuals sampled from the population are related to one another; that is, if we have a model for the genealogical trees relating individuals in a sample from the population. Equation (1.1) assumes an infinite population density everywhere so that a finite sample of individuals will be unrelated; in order to understand genealogies we have to consider (1.2). The first step is to understand the effect of the stochastic fluctuations on the forwards in time dynamics of the waves.

Any solution to (1.1) with a front-like initial condition $p(0,x)$ which decays sufficiently fast as $x \to \infty$ converges to the travelling wave solution with minimal wavespeed $\sqrt{2ms_0}$ ([34, 8]). Since the speed of this travelling wave is determined by the behaviour in the 'tip' of the wave, where the frequency of the favoured type is very low, it is very sensitive to stochastic fluctuations. A great deal of work has gone into understanding the effect of those fluctuations on the progress of the 'bulk' of the wave ([9, 10, 35, 11, 20, 26, 5]). The first striking fact is that the wave is significantly slowed by the noise ([11, 26]). The second ramification of the noise is that there really is a well-defined 'wavefront'; that is, assuming that the favoured type is spreading from left to right in our one-dimensional spatial domain, there will be a rightmost point of the support of the stochastic travelling wave ([27]). Moreover, the shape of the wavefront is well-approximated by a truncated Fisher wave ([9, 26]).

If we were to take a sample of favoured individuals from a population evolving according to the analogue of (1.2) without space, then, from [3], their genealogy would be given by a 'coalescent in a random background'; that is, it would follow a Kingman coalescent but with the instantaneous rate of coalescence of each pair of lineages at time t before the present given by $1/(N_0 \overleftarrow{p}(t))$, where $\overleftarrow{p}(t)$ is the proportion of the population that is of the favoured type at time t before the present, and N_0 is the total population size. This suggests that in the spatial context, as we trace back ancestral lineages, their instantaneous rate of coalescence on meeting at the point x should be proportional to $1/\overleftarrow{p}(t,x)$. In particular, this means that if several lineages are in the tip at the same time, then they can coalesce very quickly. In fact, principally because $p(t,x)$ is very rough, it is difficult to study the genealogy directly by tracking ancestral lineages and analysing when and where they meet. However, several plausible approximations (at least for the population close to the wavefront) have been proposed for which the frequencies of different types in the population are approximated by (1.2) and a consensus has emerged that for biologically reasonable models, over suitable timescales, the genealogy will be determined by a Bolthausen-Sznitman coalescent ([11, 5]). We emphasize that this arises as a further scaling of the Kingman coalescent in a random background. It reflects a separation of timescales. The 'multiple merger' events correspond to bursts of coalescence when several lineages are close to the tip of the wave. This then is the third ramification of adding genetic drift to (1.1); the genealogy of a sample of favoured alleles from the wavefront will be dominated by 'founder effects', resulting from the fluctuations in the wavefront. The idea is that from time to time a fortunate individual gets ahead of the wavefront, where its descendants can reproduce uninhibited by competition, at least until the rest of the population catches up, by which time they form a significant portion of the wavefront.

Other forms of selection: pushed and pulled waves of expansion

The Fisher-KPP equation, and its stochastic analogue (1.2), model a situation in which each individual in the population carries one copy of a gene that can occur in one of two types, usually denoted a and A and referred to as alleles. If the type A has a small selective advantage (in a sense to be made more precise when we describe our individual based model below), then in a suitable scaling limit, $p(t,x)$ represents the proportion of the population at location x at time t that carries the A allele. This can also be used as a model for the frequency of A alleles in a diploid population, provided that the advantage of carrying two copies of the A allele is twice that of carrying one. However, natural selection is rarely that simple; here our goal is to model a situation in which there is selection against heterozygotes, that is, individuals carrying one A allele and one a allele, and in which AA-homozygotes are fitter than aa. As we shall explain below, the analogue of the Fisher-KPP equation in this situation takes the form

$$\begin{aligned} \frac{\partial p}{\partial t}(t,x) &= \frac{m}{2}\Delta p(t,x) + s_0 f\big(p(t,x)\big) \qquad \text{for } x\in\mathbb{R},\, t>0, \\ 0 &\le p(0,x) \le 1 \qquad \forall x\in\mathbb{R}, \\ \text{where} \quad f(p) &= p(1-p)(2p-1+\alpha), \end{aligned} \tag{1.3}$$

with $\alpha > 0$ a parameter which depends on the relative fitnesses of AA, Aa and aa individuals.

In the case $\alpha \in (0,1)$, the non-linear term f is bistable (since $f(0) = 0 = f(1)$, $f'(0) < 0$, $f'(1) < 0$ and $f < 0$ on $(0,(1-\alpha)/2)$, $f > 0$ on $((1-\alpha)/2, 1)$) and the equation has a unique travelling wave solution given up to translation by the exact form

$$p(t,x) = g\big(x - \alpha\sqrt{\tfrac{ms_0}{2}}t\big), \quad \text{where } g(y) = \big(1 + e^{\sqrt{\frac{2s_0}{m}}y}\big)^{-1}. \tag{1.4}$$

For $\alpha \in [1,2)$, the travelling wave solution with minimal wavespeed is also given by (1.4). In both cases, solutions of (1.3) with suitable front-like initial conditions converge to the travelling wave (1.4) [16, 31]. The case $\alpha = 0$ corresponds to AA and aa being equally fit, in which case, for suitable initial conditions, there is a stationary 'hybrid zone' trapped between two regions composed almost entirely of AA and almost entirely of aa individuals respectively. As observed, for example, by Barton ([2]), when $\alpha > 2$ the symmetric wavefront of (1.4) is replaced by an asymmetric travelling wavefront moving at speed $\sqrt{2ms_0(\alpha-1)}$. This transition from symmetric to asymmetric wave corresponds to the transition from a 'pushed' wave to a 'pulled' wave, notions introduced by Stokes ([32]).

Considering the equation (1.3) for general monostable f (i.e. f satisfying $f(0) = 0 = f(1)$, $f'(0) > 0$, $f'(1) < 0$ and $f > 0$ on $(0,1)$), the travelling wave solution with minimal wavespeed c is called a pushed wave if $c > \sqrt{2ms_0 f'(0)}$, and is a pulled wave if $c = \sqrt{2ms_0 f'(0)}$. (Here, $\sqrt{2ms_0 f'(0)}$ is the spreading speed of solutions of the linearised equation.) The travelling wave solutions in the bistable case can also be seen as pushed waves (see [19]).

The natural stochastic version of (1.3), which was also discussed briefly by Barton ([2]), simply adds a Wright-Fisher noise as in (1.2). For $\alpha > 1$, this is a reparametrisation of an equation considered by Birzu et al. ([6]). Their model is framed in the language of ecology. Let $n(t,x)$ denote the population density at point x at time t. They consider

$$dn(t,x) = \frac{m}{2}\Delta n(t,x)dt + n(t,x)r\big(n(t,x)\big)dt + \sqrt{\gamma\big(n(t,x)\big)n(t,x)}W(dt,dx), \tag{1.5}$$

where W is a space-time white noise, $\gamma(n)$ quantifies the strength of the fluctuations, and $r(n)$ is the (density dependent) per capita growth rate. For example, for logistic growth, one would take $r(n) = r_0(1 - n/N)$ for some 'carrying capacity' N. A pushed wave arises when species grow best at intermediate population densities, known as an Allee effect in ecology. This effect is typically incorporated by adding a cooperative term to the logistic equation, for example by taking

$$r(n) = r_0\left(1 - \frac{n}{N}\right)\left(1 + \frac{Bn}{N}\right)$$

for some $B > 0$. If we write $p = n/N$, then, writing

$$s_0\left(1 - \frac{n}{N}\right)\left(\frac{2n}{N} - 1 + \alpha\right) = s_0(\alpha-1)\left(1 - \frac{n}{N}\right)\left(\frac{2}{\alpha-1}\frac{n}{N} + 1\right),$$

we see that for $\alpha > 1$ we can recover (1.5) from a stochastic version of (1.3) by setting $B = 2/(\alpha - 1)$ and $r_0 = s_0(\alpha - 1)$. Birzu et al. ([6]) define the travelling wave solution with minimal wavespeed to the deterministic equation with this form of r to be pulled if $B \leq 2$, 'semi-pushed' if $2 < B < 4$ and 'fully pushed' if $B \geq 4$ (see equation (7) in [6] for a more general definition). In our parametrisation this says that the wave is pulled for $\alpha \geq 2$ (as observed by [2]), semi-pushed for $3/2 < \alpha < 2$ and fully pushed for $\alpha \leq 3/2$. For $B \leq 2$ the wavespeed is determined by the growth rate in the tip (in particular it is independent of B), and just as for the Fisher wave, one can expect the behaviour to be very sensitive to stochastic fluctuations. For $B > 2$, the velocity of the wave increases with B, and also the region of highest growth rate shifts from the tip into the bulk of the wave. These waves should be much less sensitive to fluctuations in the tip. Moreover if we follow the ancestry of an allele of the favoured type A, that is we follow an ancestral lineage, then in the pulled case, we expect the lineage to spend most of its time in the tip of the wave, and in contrast, in the pushed case, it will spend more time in the bulk. Indeed, if the shape of the advancing wave is close to that of g in (1.4) and the speed is close to $\nu := \alpha\sqrt{ms_0/2}$, then we should expect the motion of the ancestral lineage *relative to the wavefront* to be approximately governed by the stochastic differential equation

$$dZ_t = \nu dt + \frac{m\nabla g(Z_t)}{g(Z_t)}dt + \sqrt{m}dB_t, \tag{1.6}$$

where $(B_t)_{t\geq 0}$ is a standard Brownian motion. (We shall explain this in more detail in the context of our model in Section 1.3 below.) The stationary measure of this diffusion (if it exists) will be the renormalised speed measure,

$$\pi(x) = \frac{C}{m}g(x)^2 \exp\left(2\nu x/m\right) = \frac{C}{m}e^{\frac{2\nu}{m}x}(1 + e^{\sqrt{\frac{2s_0}{m}}x})^{-2}. \tag{1.7}$$

Substituting for the wavespeed, $\nu = \alpha\sqrt{ms_0/2}$, we find that π is integrable for $0 < \alpha < 2$. In other words, the diffusion defined by (1.6) has a non-trivial stationary distribution when the wave is pushed, but not when it is pulled. The expression (1.7) appears in equation (S28) in [6], and earlier in [30] (where the authors study the deterministic equation (1.3)) and in Theorem 2 of [19] (in relation to pushed wave solutions of general reaction-diffusion equations). In [6], through a mixture of simulations and calculations, the authors also conjecture that the behaviour of the genealogical trees of a sample of A alleles from near the wavefront will change at $B = 4$ (corresponding to $\alpha = 3/2$) from being, on appropriate timescales, a Kingman coalescent for $\alpha \in (0, 3/2)$ to being a multiple merger coalescent for $\alpha > 3/2$.

Our calculation of the stationary distribution only tells us about a single ancestral lineage; to understand why there should be a further transition at $\alpha = 3/2$, we need to understand the behaviour of multiple lineages. We seek a 'separation of timescales' in which ancestral lineages reach stationarity on a faster timescale than coalescence; c.f. [29]. Recalling that we are sampling type A alleles from near the wavefront, then just as for the Fisher-KPP case, the instantaneous rate of coalescence of two lineages that meet at the position $x \in \mathbb{R}$ relative to the wavefront should be proportional to the inverse of the density of A alleles at x, which we approximate as $1/(2N_0 g(x))$ for a large constant N_0 (corresponding to the population density). If N_0 is sufficiently large, then the lineages will not coalesce before their spatial positions reach equilibrium, and so the probability that the two lineages are both at position x relative to the wavefront should be proportional to $\pi(x)^2$. This suggests that in this scenario the time to coalescence should be approximately exponential, with parameter proportional to $\int_{-\infty}^{\infty} \pi(x)^2/g(x)dx$ (this calculation appears in [6] in their equation (S119)). This quantity is finite precisely when $\alpha \in (0, 3/2)$. If we sample k lineages, one can conjecture that, because of the separation of timescales, once a first pair of lineages coalesces, the additional time until the next merger is the same as if the remaining $k-1$ lineages were started from points sampled independently according to the stationary distribution π. This then strongly suggests that in the regime $\alpha \in (0, 3/2)$, after suitable scaling, the genealogy of a sample will converge to a Kingman coalescent.

Although we believe that the suitably timescaled genealogy of lineages sampled from near the wavefront of the advance of the favoured type really will converge to Kingman's coalescent for all $\alpha \in (0, 3/2)$, our main results in this article will be restricted to the case $\alpha \in (0, 1)$. The difficulty is that for $\alpha > 1$, as $x \to \infty$, the stationary measure $\pi(x)$ does not decay as quickly as the wave profile $g(x)$. Consequently, a diffusion driven by (1.6) will spend a non-negligible proportion of its time in the region where g is very small, which is precisely where the fluctuations of p about g (or rather fluctuations of $1/p$ about $1/g$) become significant and our approximations break down. For this reason, in what follows, we shall restrict ourselves to the case $\alpha < 1$. Unlike the parameter range corresponding to (1.5), in this setting, the growth rate in the tip of the wave is actually negative, and the non-linear term f in (1.3) is bistable. In ecology this would correspond to a strong Allee effect; for us, it means that we can control the time that the ancestral lineage of an A allele spends in the tip of the wave (from which it is repelled). In Section 1.3 below, we will briefly discuss the case $\alpha \in [1, 3/2)$ in the context of our model.

Before discussing the definition of our model, we mention recent rigorous results of Tourniaire [33] on a related model. She studies a model that mimics a population expanding according to a travelling wave, and her model also exhibits fully pushed, semi-pushed and pulled regimes. The model is a branching Brownian motion with space-dependent branching rate and negative drift in which particles are killed if they hit the origin; she shows that in the semi-pushed regime, the number of particles evolves approximately according to an α-stable continuous-state branching process, suggesting that the genealogy is governed by a beta coalescent (a multiple merger coalescent).

Some biological considerations

Our goal is to write down a mathematically tractable, but biologically plausible, individual based model for a spatially structured population subject to selection acting on diploids, and to show that when suitably scaled the genealogy of a sample from near the wavefront of expansion of A alleles converges to a Kingman coalescent. As we will see below, for this model the proportion of A alleles will be governed by a discrete space stochastic analogue of (1.3) with $0 < \alpha < 1$.

The model that we define and analyse below will be a modification of a classical Moran model for a spatially structured population with selection in which we treat each allele as an individual. In order to justify this choice, we first follow a more classical approach by considering a variant of a model that is usually attributed to Fisher and Wright, for a large (diploid) population, evolving in discrete generations.

First we explain the form of the nonlinearity in (1.3). For simplicity, let us temporarily consider a population without spatial structure. We are following the fate of a gene with two alleles, a and A. Individuals in the population each carry two copies of the gene. During reproduction, each individual produces a very large number of germ cells (containing a copy of all the genetic material of the parent) which then split into gametes (each carrying just one copy of the gene). All the gametes produced in this way are pooled and, if the population is of size N_0, then $2N_0$ gametes are sampled (without replacement) from the pool. The sampled gametes fuse at random to form the next generation of diploid individuals. To model selection, we suppose that the numbers of germ cells produced by individuals are in the proportion $1 + 2\alpha s : 1 + (\alpha - 1)s : 1$ for genetic types AA, Aa, aa respectively. Here $\alpha \in (0,1)$ is a positive constant and $s > 0$ is small, with $(\alpha+1)s < 1$. Notice in particular that type AA homozygotes are 'fitter' than type aa homozygotes, in that they contribute more gametes to the pool (fecundity selection). Both are fitter than the heterozygotes (Aa individuals).

Suppose that the proportion of type A alleles in the population is w. If the population is in Hardy-Weinberg proportions, then the proportions of AA, Aa and aa individuals are w^2, $2w(1-w)$ and $(1-w)^2$ respectively. Hence the proportion of type A in the (effectively infinite) pool of gametes produced during reproduction is

$$\begin{aligned}
&\frac{(1+2\alpha s)w^2 + \frac{1}{2}(1+(\alpha-1)s)2w(1-w)}{1+2\alpha s w^2 + (\alpha-1)s \cdot 2w(1-w)} \\
&= (1+\alpha s - s)w + (3-\alpha)sw^2 - 2sw^3 + \mathcal{O}(s^2) \\
&= (1-(\alpha+1)s)w + \alpha s(2w-w^2) + s(3w^2-2w^3) + \mathcal{O}(s^2) && (1.8) \\
&= w + \alpha s w(1-w) + sw(1-w)(2w-1) + \mathcal{O}(s^2). && (1.9)
\end{aligned}$$

We will assume that s is sufficiently small that terms of $\mathcal{O}(s^2)$ are negligible. If the population were infinite, then the frequency of A alleles would evolve deterministically, and if $s = s_0/K$ for some large K, then measuring time in units of K generations, we see that w will evolve approximately according to the differential equation

$$\frac{dw}{dt} = \alpha s_0 w(1-w) + s_0 w(1-w)(2w-1) = s_0 w(1-w)(2w-1+\alpha), \tag{1.10}$$

and we recognise the nonlinearity in (1.3).

The easiest way to incorporate spatial structure into the Wright-Fisher model described above is to suppose that the population is subdivided into demes (islands of population) which we can, for example, take to be the vertices of a lattice, and in each generation a proportion of the gametes produced in a deme is distributed to its neighbours (plausible, for example, for a population of plants). If we assume that this dispersal is symmetric, the population size in each deme is the same, and the proportion of gametes that migrate scales as $1/K$, then this will result in the addition of a term involving the discrete Laplacian to the equation (1.10).

Since we are interested in understanding the interplay of selection, spatial structure, and random genetic drift, we must consider a population with finite population size in each deme. We shall nonetheless assume that the population in each deme is large, so that our assumption that the population is in Hardy-Weinberg equilibrium remains valid. When this assumption is satisfied, to specify the evolution of the proportions of the types AA, Aa, aa, it suffices to track the proportion of A gametes in each deme. Moreover, because we assume that the chosen gametes fuse at random to form the next generation, the genealogical trees relating a sample of alleles from the population can also be recovered from tracing just single types. The only role that pairing of genes in individuals plays is in determining what proportion of the gamete pool will be contributed by a given allele in the parental population.

Returning to our non-spatial model, suppose that the proportion of A alleles in some generation t is w and recall that the population consists of $2N_0$ alleles. The probability that two type A alleles sampled from generation $t+1$ are both descendants of the same parental allele is approximately $1/(2N_0 w)$ since s is small, while the probability that three or more are all descended from the same parent is $\mathcal{O}(1/N_0^2)$. Recalling that $s = s_0/K$ for some large K, if now we measure time in units of K generations, the forwards in time model for allele frequencies will be approximated by a stochastic differential equation,

$$dw = s_0 w(1-w)(2w-1+\alpha)dt + \sqrt{\frac{K}{2N_0} w(1-w)}dB_t,$$

where $(B_t)_{t\geq 0}$ is a Brownian motion, and the genealogy of a sample of type A alleles from our population will be well-approximated by a time-changed Kingman coalescent in which the instantaneous rate of coalescence, when the proportion of type A alleles in the population is w, is $K/(2N_0 w)$.

The Wright-Fisher model is inconvenient mathematically, but we now see that for the purpose of understanding the genealogy, we can replace it by any other model in which, over large timescales, the allele frequencies evolve in (approximately) the same way and in which, as we trace backwards in time, the genealogy of a sample of favoured alleles is (approximately) the same (time-changed) Kingman coalescent. This will allow us to replace the discrete generation (diploid) 'Wright-Fisher' model by a much more mathematically convenient 'Moran model', in which changes in allele frequencies in each deme will be driven by Poisson processes of reproduction events in which exactly one allele is born and exactly one dies.

Because our Moran model deals directly with alleles, from now on we shall refer to alleles as *individuals*. To understand the form that our Moran model should take, let us first consider the non-spatial setting. Once again we trace $2N_0$ individuals (alleles), but now we label them $1, 2, \ldots, 2N_0$. Reproduction events will take place at the times of a rate $2N_0 K$ Poisson process. Inspired by (1.9), we divide events into three types: neutral events, which will take place at rate $2N_0 K(1-(\alpha+1)s)$, events capturing directional selection at rate $2N_0 K\alpha s$, and events capturing selection against heterozygosity, at rate $2N_0 Ks$. In a neutral event, an ordered pair of individuals is chosen uniformly at random from the population; the first dies and is replaced by an offspring of the second (and this offspring inherits the label of the first individual). At an event corresponding to directional selection, an ordered pair of individuals is chosen uniformly at random from the population; if the type of the second is A, then it produces an offspring which replaces the first. At an event corresponding to selection against heterozygosity, an ordered triplet of individuals is picked from the population; if the second and third are of the same type, then the second produces an offspring that replaces the first. (Note that in such an event, the first individual is either replaced by or remains a type A if and only if at least two of the triplet of individuals picked were type A.)

Note that if X_1, X_2 and X_3 are i.i.d. Bernoulli(w) random variables then

$$\mathbb{P}\left(X_1+X_2\geq 1\right)=2w-w^2 \quad \text{and} \quad \mathbb{P}\left(X_1+X_2+X_3\geq 2\right)=3w^2-2w^3,$$

and recall that $s=s_0/K$. Then using (1.8), we see that for large K, the proportion of A alleles under this model will be close to that under our time-changed Wright-Fisher model. Moreover, since there is at most one birth event at a time, coalescence of ancestral lineages is necessarily pairwise. If in a reproduction event the parent is type A, then the probability that a pair of type A ancestral lineages corresponds to the parent and its offspring (and therefore merges in the event) is $2/(2N_0 w(2N_0 w-1))$, where w is the proportion of A alleles in the population. Since s is very small, the instantaneous rate at which events with a type A parent fall is approximately $2N_0 Kw$. Thus, the probability that a particular pair of two type A individuals sampled from the population at time $t+\delta t$ are descended from the same type A individual at time t is (up to a lower order error) $K\delta t/(N_0 w)$. Therefore (after rescaling time by a factor $1/2$, and replacing s_0 by $2s_0$) the genealogy and changes in allele frequencies under this model will be (up to a small error) the same as under the Wright-Fisher model.

In what follows, to avoid too many factors of two, we are going to write $N=2N_0$ for the number of individuals in our Moran model.

1.1 Definition of the model

We now give a precise definition of our model. Take $\alpha \in (0,1)$, $s_0 > 0$ and $m > 0$. Let $n, N \in \mathbb{N}$. We are going to define our (structured) Moran model on $\frac{1}{n}\mathbb{Z}$ in such a way that there are N individuals in each site (or deme) and they are indexed by $[N] := \{1, \ldots, N\}$. We shall denote the type of the ith individual at site x at time t by $\xi_t^n(x,i) \in \{0,1\}$, with $\xi_t^n(x,i) = 1$ meaning that the individual is type A, and $\xi_t^n(x,i) = 0$ meaning that the individual is type a. For $x \in \frac{1}{n}\mathbb{Z}$ and $t \geq 0$, let

$$p_t^n(x) = \frac{1}{N}\sum_{i=1}^{N} \xi_t^n(x,i)$$

be the proportion of type A at x at time t. We shall reserve the symbol x for space and i, j, k for the label of an individual.

Let

$$s_n = \frac{2s_0}{n^2} \quad \text{and} \quad r_n = \frac{n^2}{2N}. \tag{1.11}$$

(Here, s_n is a selection parameter which determines the space scaling needed to see a non-trivial limit, and r_n is a time scaling parameter.)

To specify the dynamics of the process, we define four independent families of i.i.d. Poisson processes. These will govern neutral reproduction, directional selection, selection against heterozygotes and migration respectively. Let $((\mathcal{P}_t^{x,i,j})_{t\geq 0})_{x\in\frac{1}{n}\mathbb{Z}, i\neq j\in[N]}$ be i.i.d. Poisson processes with rate $r_n(1-(\alpha+1)s_n)$. Let $((\mathcal{S}_t^{x,i,j})_{t\geq 0})_{x\in\frac{1}{n}\mathbb{Z}, i\neq j\in[N]}$ be i.i.d. Poisson processes with rate $r_n \alpha s_n$. Let $((\mathcal{Q}_t^{x,i,j,k})_{t\geq 0})_{x\in\frac{1}{n}\mathbb{Z}, i,j,k\in[N] \text{ distinct}}$ be i.i.d. Poisson processes with rate $\frac{1}{N} r_n s_n$. Let $((\mathcal{R}_t^{x,i,y,j})_{t\geq 0})_{x,y\in\frac{1}{n}\mathbb{Z}, |x-y|=n^{-1}, i,j\in[N]}$ be i.i.d. Poisson processes with rate $m r_n$.

For a given initial condition $p_0^n : \frac{1}{n}\mathbb{Z} \to \frac{1}{N}\mathbb{Z} \cap [0,1]$, we assign labels to the type A individuals in each site uniformly at random. That is, we define $(\xi_0^n(x,i))_{x\in\frac{1}{n}\mathbb{Z}, i\in[N]}$ as follows. For each $x \in \frac{1}{n}\mathbb{Z}$ independently, take $I_x \subseteq [N]$, where I_x is chosen uniformly at random from $\{A \subseteq [N] : |A| = Np_0^n(x)\}$. For $i \in [N]$, let $\xi_0^n(x,i) = \mathbb{1}_{\{i\in I_x\}}$.

The process $(\xi_t^n(x,i))_{x\in\frac{1}{n}\mathbb{Z}, i\in[N], t\geq 0}$ evolves as follows.

1. If t is a point in $\mathcal{P}^{x,i,j}$, then at time t, the individual at (x,i) is replaced by offspring of the individual at (x,j), i.e. we let $\xi_t^n(x,i) = \xi_{t-}^n(x,j)$.
2. If t is a point in $\mathcal{S}^{x,i,j}$, then at time t, if the individual at (x,j) is type A then the individual at (x,i) is replaced by offspring of the individual at (x,j), i.e. we let
$$\xi_t^n(x,i) = \begin{cases} \xi_{t-}^n(x,j) & \text{if } \xi_{t-}^n(x,j) = 1, \\ \xi_{t-}^n(x,i) & \text{otherwise.} \end{cases}$$
3. If t is a point in $\mathcal{Q}^{x,i,j,k}$, then at time t, if the individuals at (x,j) and (x,k) have the same type then the individual at (x,i) is replaced by offspring of the individual at (x,j), i.e. we let
$$\xi_t^n(x,i) = \begin{cases} \xi_{t-}^n(x,j) & \text{if } \xi_{t-}^n(x,j) = \xi_{t-}^n(x,k), \\ \xi_{t-}^n(x,i) & \text{otherwise.} \end{cases}$$
4. If t is a point in $\mathcal{R}^{x,i,y,j}$, then at time t, the individual at (x,i) is replaced by offspring of the individual at (y,j), i.e. we let $\xi_t^n(x,i) = \xi_{t-}^n(y,j)$.

Ancestral lineages will be represented in the form of a pair with the first coordinate recording the spatial position and the second the label of the ancestor. More precisely, for $T \geq 0$, $t \in [0,T]$, $x_0 \in \frac{1}{n}\mathbb{Z}$ and $i_0 \in [N]$, if the individual at site y with label j is the ancestor at time $T-t$ of the individual at site x_0 with label i_0 at time T, then we let $(\zeta_t^{n,T}(x_0,i_0), \theta_t^{n,T}(x_0,i_0)) = (y,j)$. The pair $(\zeta_t^{n,T}(x_0,i_0), \theta_t^{n,T}(x_0,i_0))_{t\in[0,T]}$ is a jump process with

$$(\zeta_0^{n,T}(x_0,i_0), \theta_0^{n,T}(x_0,i_0)) = (x_0,i_0),$$

which evolves as follows. For some $t \in (0,T]$, suppose that $(\zeta_{t-}^{n,T}(x_0,i_0), \theta_{t-}^{n,T}(x_0,i_0)) = (x,i)$. Then if $T-t$ is a point in $\mathcal{P}^{x,i,j}$ for some $j \neq i$, we let $(\zeta_t^{n,T}(x_0,i_0), \theta_t^{n,T}(x_0,i_0)) = (x,j)$. If instead $T-t$ is a point in $\mathcal{S}^{x,i,j}$ for some $j \neq i$, we let

$$(\zeta_t^{n,T}(x_0,i_0), \theta_t^{n,T}(x_0,i_0)) = \begin{cases} (x,j) & \text{if } \xi^n_{(T-t)-}(x,j) = 1, \\ (x,i) & \text{otherwise.} \end{cases}$$

If instead $T-t$ is a point in $\mathcal{Q}^{x,i,j,k}$ for some $j \neq k \in [N] \setminus \{i\}$, we let

$$(\zeta_t^{n,T}(x_0,i_0), \theta_t^{n,T}(x_0,i_0)) = \begin{cases} (x,j) & \text{if } \xi^n_{(T-t)-}(x,j) = \xi^n_{(T-t)-}(x,k), \\ (x,i) & \text{otherwise.} \end{cases}$$

Finally, if $T-t$ is a point in $\mathcal{R}^{x,i,y,j}$ for some $y \in \{x-n^{-1}, x+n^{-1}\}$, $j \in [N]$, we let $(\zeta_t^{n,T}(x_0,i_0), \theta_t^{n,T}(x_0,i_0)) = (y,j)$. These are the only times at which the ancestral lineage process $(\zeta_s^{n,T}(x_0,i_0), \theta_s^{n,T}(x_0,i_0))_{s\in[0,T]}$ jumps.

1.2 Main results

Recall from (1.4) that $g : \mathbb{R} \to \mathbb{R}$ is given by

$$g(x) = (1 + e^{\sqrt{\frac{2s_0}{m}}x})^{-1}. \tag{1.12}$$

In our main results, we will make the following assumptions on the initial condition p_0^n, for $b_1, b_2 > 0$ to be specified later:

$$\begin{aligned} &p_0^n(x) = 0 \ \forall x \geq N, \quad p_0^n(x) = 1 \ \forall x \leq -N, \\ &\sup_{x\in\frac{1}{n}\mathbb{Z}} |p_0^n(x) - g(x)| \leq b_1 \qquad \text{and} \qquad \sup_{z_1,z_2\in\frac{1}{n}\mathbb{Z}, |z_1-z_2|\leq n^{-1/3}} |p_0^n(z_1) - p_0^n(z_2)| \leq n^{-b_2}. \end{aligned} \tag{A}$$

These assumptions ensure that p_0^n is a front-like initial condition which is fairly close to the travelling wave profile g and is not too rough. We will assume throughout that there exists $a_0 > 0$ such that $(\log N)^{a_0} \leq \log n$ for n sufficiently large. The idea is that we need $N \gg n \gg 1$, in order that p_t^n is close to the deterministic limit, but we do not want N to tend to infinity so quickly that we don't see the effect of the stochastic perturbation at all.

For $t \geq 0$, define the position of the random travelling front at time t by letting

$$\mu_t^n = \sup\{x \in \tfrac{1}{n}\mathbb{Z} : p_t^n(x) \geq 1/2\}. \tag{1.13}$$

For $t \geq 0$ and $R > 0$, let

$$G_{R,t} = \{(x,i) \in \tfrac{1}{n}\mathbb{Z} \times [N] : |x - \mu_t^n| \leq R,\ \xi_t^n(x,i) = 1\}, \tag{1.14}$$

the set of type A individuals which are near the front at time t.

Our first main result says that if at a large time T_n we sample a type A individual from near the front, then the position of its ancestor relative to the front at a much earlier time $T_n - T_n'$ has distribution approximately given by π (as defined in (1.15) below).

Theorem 1.1. *Suppose* $\alpha \in (0,1)$ *and, for some* $a_1 > 1$, $N \geq n^{a_1}$ *for* n *sufficiently large. There exists* $b_1 > 0$ *such that for* $b_2 > 0$ *and* $K_0 < \infty$ *the following holds. Suppose condition* (A) *holds,* $T_n \leq N^2$ *and* $T'_n \to \infty$ *as* $n \to \infty$ *with* $T_n - T'_n \geq (\log N)^2$*. Let* $(X_0, J_0) \in \frac{1}{n}\mathbb{Z} \times [N]$ *be measurable with respect to* $\sigma((\xi^n_{T_n}(x,i))_{x \in \frac{1}{n}\mathbb{Z}, i \in [N]})$ *with* $(X_0, J_0) \in G_{K_0, T_n}$*. Then*

$$\zeta^{n,T_n}_{T'_n}(X_0, J_0) - \mu^n_{T_n - T'_n} \xrightarrow{d} Z \quad \text{as } n \to \infty,$$

where Z *is a random variable with density*

$$\pi(x) = \frac{g(x)^2 e^{\alpha\sqrt{\frac{2s_0}{m}}x}}{\int_{-\infty}^{\infty} g(y)^2 e^{\alpha\sqrt{\frac{2s_0}{m}}y} dy}. \tag{1.15}$$

Our second main result says that the genealogy of a sample of type A individuals from near the front at a large time T_n is approximately given by a Kingman coalescent (under a suitable time rescaling).

Theorem 1.2. *Suppose* $\alpha \in (0,1)$ *and, for some* $a_2 > 3$, $N \geq n^{a_2}$ *for* n *sufficiently large. There exists* $b_1 > 0$ *such that for* $b_2 > 0$, $k_0 \in \mathbb{N}$ *and* $K_0 < \infty$*, the following holds. Suppose condition* (A) *holds, and take* $T_n \in [N, N^2]$*. Let* $(X_1, J_1), \ldots, (X_{k_0}, J_{k_0})$ *be measurable with respect to* $\sigma((\xi^n_{T_n}(x,i))_{x \in \frac{1}{n}\mathbb{Z}, i \in [N]})$ *and distinct, with* $(X_i, J_i) \in G_{K_0, T_n}$ $\forall i \in [k_0]$*.*

For $i, j \in [k_0]$*, let* $\tau^n_{i,j}$ *denote the time at which the* i^{th} *and* j^{th} *ancestral lineages coalesce, i.e. let*

$$\tau^n_{i,j} = \inf\{t \geq 0 : (\zeta^{n,T_n}_t(X_i, J_i), \theta^{n,T_n}_t(X_i, J_i)) = (\zeta^{n,T_n}_t(X_j, J_j), \theta^{n,T_n}_t(X_j, J_j))\}.$$

Then

$$\left(\frac{(2m+1)n}{N} \frac{\int_{-\infty}^{\infty} g(x)^3 e^{2\alpha\sqrt{\frac{2s_0}{m}}x} dx}{\left(\int_{-\infty}^{\infty} g(x)^2 e^{\alpha\sqrt{\frac{2s_0}{m}}x} dx\right)^2} \tau^n_{i,j} \right)_{i,j \in [k_0]} \xrightarrow{d} (\tau_{i,j})_{i,j \in [k_0]} \quad \text{as } n \to \infty,$$

where $\tau_{i,j}$ *is the time at which the* i^{th} *and* j^{th} *ancestral lineages coalesce in the Kingman* k_0*-coalescent.*

We now state two further results that follow easily from the proofs of Theorems 1.1 and 1.2. The first result says that at large times, the proportion of type A in the population expands approximately according to the travelling wave solution (1.4) of the partial differential equation (1.3).

Theorem 1.3. *Suppose* $\alpha \in (0,1)$ *and, for some* $a_1 > 1$, $N \geq n^{a_1}$ *for* n *sufficiently large. For* $\ell \in \mathbb{N}$*, there exist* $b_1, c > 0$ *such that for* $b_2 > 0$ *the following holds. Suppose condition* (A) *holds; then for* n *sufficiently large,*

$$\mathbb{P}\left(\sup_{x \in \frac{1}{n}\mathbb{Z},\, t \in [\log N, N^2]} |p^n_t(x) - g(x - \mu^n_t)| > e^{-(\log N)^c} \right) \leq \left(\frac{n}{N}\right)^\ell \qquad \text{and}$$

$$\mathbb{P}\left(\exists t \in [\log N, N^2], s \in [0, 1 \wedge (N^2 - t)] : |\mu^n_{t+s} - \mu^n_t - \alpha\sqrt{\frac{ms_0}{2}}s| > e^{-(\log N)^c} \right) \leq \left(\frac{n}{N}\right)^\ell.$$

The second additional result is closely related to Theorem 1.1. It says that for any fixed $t_0 > 0$, if at a large time T_n we sample a type A individual from some location near the front, then the position of its ancestor relative to the front at time $T_n - t_0$ has distribution approximately given by Z_{t_0}, where $(Z_t)_{t \geq 0}$ is the diffusion given in (1.6) with Z_0 given by the position relative to the front of the sampled individual at time T_n.

Theorem 1.4. *Suppose $\alpha \in (0,1)$ and, for some $a_1 > 1$, $N \geq n^{a_1}$ for n sufficiently large. There exists $b_1 > 0$ such that for $b_2 > 0$, $t_0 > 0$, $\delta > 0$ and $K_0 < \infty$ the following holds for n sufficiently large. Suppose condition* (A) *holds and take $(\log N)^2 + t_0 \leq T_n \leq N^2$ and $X_0 \in \frac{1}{n}\mathbb{Z}$ with $|X_0 - \mu^n_{T_n}| \leq K_0$. Let $J_0 \in [N]$ be measurable with respect to $\sigma((\xi^n_{T_n}(x,i))_{x\in\frac{1}{n}\mathbb{Z}, i\in[N]})$ with $\xi^n_{T_n}(X_0, J_0) = 1$. Then for $y_0 \in \mathbb{R}$,*

$$\left|\mathbb{P}\left(\zeta^{n,T_n}_{t_0}(X_0,J_0) - \mu^n_{T_n - t_0} \leq y_0\right) - \mathbb{P}_{X_0 - \mu^n_{T_n}}(Z_{t_0} \leq y_0)\right| < \delta,$$

where under $\mathbb{P}_{z_0}$, $(Z_t)_{t\geq 0}$ solves the SDE

$$dZ_t = \alpha\sqrt{\tfrac{ms_0}{2}}dt + m\frac{\nabla g(Z_t)}{g(Z_t)}dt + \sqrt{m}dB_t, \quad Z_0 = z_0.$$

A stronger result would be to show convergence of the process $(\zeta^{n,T_n}_t(X_0,J_0) - \mu^n_{T_n-t})_{t\geq 0}$ to the diffusion $(Z_t)_{t\geq 0}$, but our results do not give us sufficient control of the increments of $\zeta^{n,T_n}_t(X_0,J_0)$ over short time intervals.

1.3 Strategy of the proof

We will show that if $N \gg n$, then if n is large and T_0 is not too large, $(p^n_t)_{t\in[0,T_0]}$ is approximately given by a solution of the PDE

$$\frac{\partial u}{\partial t} = \tfrac{1}{2}m\Delta u + s_0 u(1-u)(2u-1+\alpha). \tag{1.16}$$

(Recall from our discussion of a non-spatial Moran model before Section 1.1 that the non-linear term in (1.16) comes from the events corresponding to the Poisson processes $(\mathcal{S}^{x,i,j})_{x,i,j}$ and $(\mathcal{Q}^{x,i,j,k})_{x,i,j,k}$. The Laplacian term comes from the Poisson processes $(\mathcal{R}^{x,i,y,j})_{x,i,y,j}$ which cause migration between neighbouring sites and whose rate was chosen to coincide with the diffusive rescaling.)

As noted in (1.4), $u(t,x) := g(x - \alpha\sqrt{\frac{ms_0}{2}}t)$ is a travelling wave solution of (1.16). In the case $\alpha \in (0,1)$, work of Fife and McLeod [16] shows that for a front-like initial condition u_0 satisfying $\limsup_{x\to\infty} u_0(x) < \frac{1}{2}(1-\alpha)$ and $\liminf_{x\to-\infty} u_0(x) > \frac{1}{2}(1-\alpha)$, the solution of (1.16) converges to a moving front with shape g and wavespeed $\alpha\sqrt{\frac{ms_0}{2}}$. We can use this to show that if $N \gg n$, then for large n, with high probability,

$$p^n_t(x) \approx g(x - \mu^n_t)\ \forall x \in \tfrac{1}{n}\mathbb{Z}, t \in [\log N, N^2] \quad \text{and} \quad \frac{\mu^n_t - \mu^n_s}{t-s} \approx \alpha\sqrt{\tfrac{ms_0}{2}}\ \forall s < t \in [\log N, N^2], \tag{1.17}$$

where μ^n_t is the front location defined in (1.13) (recall Theorem 1.3; this result will be proved in Proposition 3.1).

Suppose the event in (1.17) occurs, and sample a type A individual at a large time T_n by taking (X_0, J_0) with $\xi^n_{T_n}(X_0,J_0) = 1$. We will show that the recentred ancestral lineage process $(\zeta^{n,T_n}_t(X_0,J_0) - \mu^n_{T_n-t})_{t\in[0,T_n]}$ moves approximately according to the diffusion

$$dZ_t = \alpha\sqrt{\tfrac{ms_0}{2}}dt + \frac{m\nabla g(Z_t)}{g(Z_t)}dt + \sqrt{m}dB_t,$$

where $(B_t)_{t\geq 0}$ is a Brownian motion (recall Theorem 1.4; the connection to the diffusion $(Z_t)_{t\geq 0}$ will be established in Lemma 4.3). This can be explained heuristically as follows. Observe first that $(\mu^n_{T_n-t} - \mu^n_{T_n-t-s})/s \approx \alpha\sqrt{\frac{ms_0}{2}}$ for $s > 0$. Then if $\zeta^{n,T_n}(X_0,J_0)$ jumps at some time t, and $\zeta^{n,T_n}_{t-}(X_0,J_0) = x_0$, the conditional probability that $\zeta^{n,T_n}_t(X_0,J_0) = x_0 + n^{-1}$ is

$$\frac{p^n_{T_n-t}(x_0+n^{-1})}{p^n_{T_n-t}(x_0-n^{-1}) + p^n_{T_n-t}(x_0+n^{-1})} \approx \frac{1}{2} + \frac{1}{2}\frac{\nabla g(x_0 - \mu^n_{T_n-t})}{g(x_0-\mu^n_{T_n-t})}n^{-1}.$$

Finally, the total rate at which $\zeta^{n,T_n}(X_0, J_0)$ jumps is given by $2mr_nN = mn^2$, and the jumps have increments $\pm n^{-1}$.

As we observed before in (1.7), $(Z_t)_{t\geq 0}$ has a unique stationary distribution given by π, as defined in (1.15). In Theorem 1.1, we show rigorously that for large t, $\zeta_t^{n,T_n}(X_0, J_0) - \mu^n_{T_n - t}$ has distribution approximately given by π. Theorem 1.1 is not strong enough to give the precise estimates that we need for Theorem 1.2, and so in fact we prove Theorem 1.2 first and then Theorem 1.1 will follow from results that we have obtained along the way.

A pair of ancestral lineages can only coalesce if they are distance at most n^{-1} apart. Take a pair of type A individuals at time T_n by sampling $(X_1, J_1) \neq (X_2, J_2)$ with $\xi^n_{T_n}(X_1, J_1) = 1 = \xi^n_{T_n}(X_2, J_2)$. Suppose that at some time $T_n - t$, their ancestral lineages are at the same site but have not coalesced, i.e. $\zeta_t^{n,T_n}(X_1, J_1) = x = \zeta_t^{n,T_n}(X_2, J_2)$ for some $x \in \frac{1}{n}\mathbb{Z}$. For $\delta_n > 0$ sufficiently small, on the time interval $[T_n - t - \delta_n, T_n - t]$, each type A individual at x produces offspring at x at rate approximately r_nN, and not many individuals produce more than one offspring. Hence the number of pairs of type A individuals at x at time $T_n - t$ which have common ancestors at time $T_n - t - \delta_n$ is approximately $r_nN^2\delta_n p^n_{T_n - t - \delta_n}(x)$ (see Lemma 5.2). Therefore, the probability that our pair of lineages coalesce within time δ_n (backwards in time), which is the same as the probability that it is one such pair, is approximately

$$\frac{r_nN^2\delta_n p^n_{T_n-t-\delta_n}(x)}{\binom{Np^n_{T_n-t}(x)}{2}} \approx \frac{n^2\delta_n}{Np^n_{T_n-t}(x)}. \tag{1.18}$$

Similarly, if $\zeta_t^{n,T_n}(X_1, J_1) = x$ and $\zeta_t^{n,T_n}(X_2, J_2) = x + n^{-1}$ then, since an individual at x produces offspring at $x + n^{-1}$ at rate mr_nN and vice-versa, the probability that the pair of lineages coalesce within time δ_n is approximately

$$\frac{mr_nN^2\delta_n(p^n_{T_n-t-\delta_n}(x) + p^n_{T_n-t-\delta_n}(x+n^{-1}))}{Np^n_{T_n-t}(x) \cdot Np^n_{T_n-t}(x+n^{-1})} \approx \frac{mn^2\delta_n}{Np^n_{T_n-t}(x)}. \tag{1.19}$$

These heuristics suggest that for $x_0 \in \frac{1}{n}\mathbb{Z}$, since $\pi(x_0)\pi(x_0 + n^{-1})^{-1} \approx 1$ and $\pi(x_0)\pi(x_0 - n^{-1})^{-1} \approx 1$, the rate at which the pair of ancestral lineages of (X_1, J_1) and (X_2, J_2) coalesce and the ancestral lineage of (X_1, J_1) is at location x_0 relative to the front should be approximately

$$n^{-2}\pi(x_0)^2 \cdot \frac{n^2}{Ng(x_0)} + 2n^{-2}\pi(x_0)^2 \cdot \frac{mn^2}{Ng(x_0)} = (2m+1)\frac{\pi(x_0)^2}{Ng(x_0)}.$$

Note that for some constants $C_1, C_2 > 0$,

$$\frac{\pi(x_0)^2}{g(x_0)} \sim C_1 e^{(2\alpha-3)\sqrt{\frac{2s_0}{m}}x_0} \to 0 \text{ as } x_0 \to \infty \text{ and } \frac{\pi(x_0)^2}{g(x_0)} \sim C_2 e^{2\alpha\sqrt{\frac{2s_0}{m}}x_0} \to 0 \text{ as } x_0 \to -\infty. \tag{1.20}$$

This suggests that coalescence only occurs (fairly) close to the front. If a pair of lineages coalesce close to the front, then the rate at which they subsequently coalesce with another given lineage is $\mathcal{O}(n^2N^{-1})$, which suggests that if $N \gg n^2$, their location relative to the front will have distribution approximately given by π before any more coalescence occurs. Hence the genealogy of a sample of type A individuals from near the front should be approximately given by a Kingman coalescent with rate

$$\sum_{x_0 \in \frac{1}{n}\mathbb{Z}} (2m+1)\frac{\pi(x_0)^2}{Ng(x_0)} \approx (2m+1)\frac{n}{N}\int_{-\infty}^{\infty} \frac{\pi(y)^2}{g(y)}dy.$$

This result is proved in Theorem 1.2 (with the additional technical assumption that $N \gg n^3$).

For $\alpha \in [1,2)$, work of Rothe [31] shows that for the PDE (1.16), if the initial condition $u_0(x)$ decays sufficiently quickly as $x \to \infty$ then the solution converges to a moving front with shape g and wavespeed $\alpha\sqrt{\frac{ms_0}{2}}$. Moreover, (1.20) holds for any $\alpha \in (0,3/2)$, which suggests that Theorem 1.2 should hold for any $\alpha \in (0,3/2)$. The main difficulty in proving the theorem for this range of α is that $p_t^n(x)^{-1}$ is hard to control when $x - \mu_t^n$ is very large, i.e. far ahead of the front. This in turn makes it hard to control the motion of ancestral lineages if they are far ahead of the front. For $\alpha \in (0,1)$, the non-linear term $f(u) = u(1-u)(2u-1+\alpha)$ in the PDE (1.16) satisfies $f(u) < 0$ for $u \in (0, \frac{1}{2}(1-\alpha))$, which means that far ahead of the front, the proportion of type A decays. This allows us to show that with high probability, no lineages of type A individuals stay far ahead of the front for a long time (see Proposition 6.1), which then gives us upper bounds on the probabilities of lineages being far ahead of the front at a fixed time (see Proposition 2.5). A proof of Theorem 1.2 for $\alpha \in [1,3/2)$ would require a different method to bound these tail probabilities, along with more delicate estimates on $p_t^n(x)$ for large x in order to apply [31] and ensure that $p_t^n(\cdot) \approx g(\cdot - \mu_t^n)$ with high probability at large times t.

One of the main tools in the proofs of Theorems 1.1 and 1.2 is the notion of tracers. In population genetics, this corresponds to labelling a subset of individuals by a neutral genetic marker, which is passed down from parent to offspring, and which has no effect on the fitness of an individual by whom it is carried. Such markers allow us to deduce which individuals in the population are descended from a particular subset of ancestors (c.f. [12]). The idea of using these markers, or 'tracers', in the context of expanding biological populations goes back at least to Hallatschek and Nelson [20], and has subsequently been used, for example, by Durrett and Fan [14], Birzu et al. [6] and Biswas et al. [7]. The idea is that at some time t_0, a subset of the type A individuals are labelled as 'tracers'. At a later time t, we can look at the subset of type A individuals which are descended from the original set of tracers. In particular, for $0 \le t_0 \le t$ and $x_1, x_2 \in \frac{1}{n}\mathbb{Z}$, we can record the proportion of individuals at x_2 at time t which are descended from type A individuals at x_1 at time t_0. This tells us the conditional probability that the time-t_0 ancestor of a randomly chosen type A individual at x_2 at time t was at x_1. For $x_1, x_2 \in \frac{1}{n}\mathbb{Z}$ and $t \ge 0$, and taking $\delta_n > 0$ very small, we can also record the number of pairs of type A individuals at x_1 and x_2 at time $t + \delta_n$ which have the same ancestor at time t. This tells us the conditional probability that a randomly chosen pair of type A lineages at x_1 and x_2 at time $t + \delta_n$ coalesce in the time interval $[t, t+\delta_n]$.

In Section 2, we will define a 'good' event E in terms of these 'tracer' random variables, and in Sections 3-6, we will show that the event E occurs with high probability. The proof of Theorem 1.3 will appear in Section 3. In Section 2, we will show that conditional on the tracer random variables, if the event E occurs, the locations of ancestral lineages relative to the front approximately have distribution π (see Lemma 2.7), pairs of nearby lineages coalesce at approximately the rates given in (1.18) and (1.19) (see Proposition 2.8), and we are unlikely to see two pairs of lineages coalesce in a short time (see Proposition 2.9). We can also prove bounds on the tail probabilities of lineages being far ahead of or far behind the front (see Propositions 2.5 and 2.6). These results combine to give a proof of Theorem 1.2. In Section 7, we use results from the earlier sections to complete the proofs of Theorems 1.1 and 1.4. Finally, in Section 8, we give a glossary of frequently used notation.

2 Proof of Theorem 1.2

Throughout Sections 2-7, we suppose $\alpha \in (0,1)$. We let

$$\kappa = \sqrt{\frac{2s_0}{m}} \quad \text{and} \quad \nu = \alpha\sqrt{\frac{ms_0}{2}}. \tag{2.1}$$

For $k \in \mathbb{N}$, let $[k] = \{1, \dots, k\}$. For $0 \le t_1 \le t_2$ and $x_1, x_2 \in \frac{1}{n}\mathbb{Z}$, let

$$q^n_{t_1,t_2}(x_1,x_2) = \frac{1}{N}|\{i \in [N] : \xi^n_{t_2}(x_2,i) = 1,\, \zeta^{n,t_2}_{t_2-t_1}(x_2,i) = x_1\}|, \tag{2.2}$$

the proportion of individuals at x_2 at time t_2 which are type A and are descended from an individual at x_1 at time t_1. Similarly, for $0 \le t_1 \le t_2$ and $x_1 \in \mathbb{R}$, $x_2 \in \frac{1}{n}\mathbb{Z}$, let

$$\begin{aligned} q^{n,+}_{t_1,t_2}(x_1,x_2) &= \frac{1}{N}|\{i \in [N] : \xi^n_{t_2}(x_2,i) = 1,\, \zeta^{n,t_2}_{t_2-t_1}(x_2,i) \ge x_1\}| \\ \text{and} \quad q^{n,-}_{t_1,t_2}(x_1,x_2) &= \frac{1}{N}|\{i \in [N] : \xi^n_{t_2}(x_2,i) = 1,\, \zeta^{n,t_2}_{t_2-t_1}(x_2,i) \le x_1\}|. \end{aligned} \tag{2.3}$$

Fix a large constant $C > 2^{13}\alpha^{-2}$, and let

$$\delta_n = \lfloor N^{1/2}n^2 \rfloor^{-1},\ \epsilon_n = \lfloor (\log N)^{-2}\delta_n^{-1} \rfloor \delta_n,\ \gamma_n = \lfloor (\log\log N)^4 \rfloor \text{ and } d_n = \kappa^{-1}C\log\log N. \tag{2.4}$$

For $t \ge 0$, $\ell \in \mathbb{N}$ and $x_1, \dots, x_\ell \in \frac{1}{n}\mathbb{Z}$, let

$$\begin{aligned} &\mathcal{C}^n_t(x_1,x_2,\dots,x_\ell) \\ &= \Big\{(i_1,\dots,i_\ell) \in [N]^\ell : (x_j,i_j) \ne (x_{j'},i_{j'})\, \forall j \ne j' \in [\ell],\ \xi^n_{t+\delta_n}(x_j,i_j) = 1\, \forall j \in [\ell], \\ &\qquad (\zeta^{n,t+\delta_n}_{\delta_n}(x_j,i_j), \theta^{n,t+\delta_n}_{\delta_n}(x_j,i_j)) = (\zeta^{n,t+\delta_n}_{\delta_n}(x_1,j_1), \theta^{n,t+\delta_n}_{\delta_n}(x_1,j_1))\, \forall j \in [\ell]\Big\}, \end{aligned} \tag{2.5}$$

the set of ℓ-tuples of distinct type A individuals at $x_1, \dots, x_\ell$ at time $t + \delta_n$ which all have a common ancestor at time t. Recall the definition of μ^n_t in (1.13). For $y, \ell > 0$, $0 \le s \le t$ and $x \in \frac{1}{n}\mathbb{Z}$, let

$$r^{n,y,\ell}_{s,t}(x) = \frac{1}{N}\big|\{i \in [N] : \xi^n_t(x,i) = 1,\ \zeta^{n,t}_{t'}(x,i) \ge \mu^n_{t-t'} + y\ \ \forall t' \in \ell\mathbb{N}_0 \cap [0,s]\}\big|, \tag{2.6}$$

the proportion of individuals at x at time t which are type A and whose ancestor at time $t - t'$ was to the right of $\mu^n_{t-t'} + y$ for each $t' \in \ell\mathbb{N}_0 \cap [0,s]$.

Fix $T_n \in [(\log N)^2, N^2]$ and define the σ-algebra

$$\begin{aligned} \mathcal{F} = \sigma\Big(&(p^n_t(x))_{x \in \frac{1}{n}\mathbb{Z}, t \le T_n}, (\xi^n_{T_n}(x,i))_{x \in \frac{1}{n}\mathbb{Z}, i \in [N]}, \\ &(q^n_{T_n-t_1,T_n-t_2}(x_1,x_2))_{x_1,x_2 \in \frac{1}{n}\mathbb{Z}, t_1,t_2 \in \delta_n\mathbb{N}_0, t_2 \le t_1 \le T_n}, \\ &\quad (\mathcal{C}^n_{T_n-t}(x_1,x_2))_{x_1,x_2 \in \frac{1}{n}\mathbb{Z}, t \in \delta_n\mathbb{N}, t \le T_n}, (\mathcal{C}^n_{T_n-t}(x_1,x_2,x_3))_{x_1,x_2,x_3 \in \frac{1}{n}\mathbb{Z}, t \in \delta_n\mathbb{N}, t \le T_n}\Big). \end{aligned} \tag{2.7}$$

We now define some 'good' events, which occur with high probability, as we will show later. Take $c_1, c_2 > 0$ small constants, and $t^*, K \in \mathbb{N}$ large constants, to be specified later. The first event will allow us to show that the probability a lineage at x_2 at time $t + \gamma_n$ has an ancestor at x_1 at time t is approximately $n^{-1}\pi(x_1 - \mu^n_t)$. For $x_1, x_2 \in \frac{1}{n}\mathbb{Z}$ and $0 \le t \le T_n$, define the event

$$A^{(1)}_t(x_1,x_2) = \left\{\left|\frac{q^n_{t,t+\gamma_n}(x_1,x_2)}{p^n_{t+\gamma_n}(x_2)} - n^{-1}\pi(x_1 - \mu^n_t)\right| \le n^{-1}(\log N)^{-3C}\right\}.$$

The next two events will allow us to control the probability that a lineage is far ahead of, or far behind, the front. For $x_1, x_2 \in \frac{1}{n}\mathbb{Z}$ and $0 \le t \le T_n$, define the events

$$\begin{aligned} A_t^{(2)}(x_1,x_2) &= \left\{ \frac{q^{n,+}_{t,t+t^*}(x_1,x_2)}{p^n_{t+t^*}(x_2)} \le c_1 e^{-(1+\frac{1}{2}(1-\alpha))\kappa(x_1-(x_2-\nu t^*)\vee(\mu^n_t+K)+2)} \right\} \\ \text{and} \qquad A_t^{(3)}(x_1,x_2) &= \left\{ \frac{q^{n,-}_{t,t+t^*}(x_1,x_2)}{p^n_{t+t^*}(x_2)} \le c_1 e^{-\frac{1}{2}\alpha\kappa((x_2-\nu t^*)-x_1+1)} \right\}. \end{aligned}$$

The next two events will give us a useful bound on the probability that a lineage is at the site x at time t, conditional on its location at time $t+\epsilon_n$, and will allow us to show that lineages do not move more than distance 1 in time ϵ_n. For $x \in \frac{1}{n}\mathbb{Z}$ and $0 \le t \le T_n$, define the events

$$\begin{aligned} A_t^{(4)}(x) &= \left\{ q^n_{t,t+\epsilon_n}(x,x') \le n^{-1}\epsilon_n^{-1} p^n_{t+\epsilon_n}(x') \, \forall x' \in \tfrac{1}{n}\mathbb{Z} \right\} \\ \text{and} \qquad A_t^{(5)}(x) &= \left\{ q^n_{t,t+\epsilon_n}(x',x) \le \mathbb{1}_{|x-x'|\le 1} \, \forall x' \in \tfrac{1}{n}\mathbb{Z} \right\}. \end{aligned}$$

The next event will allow us to show that lineages do not move more than distance $(\log N)^{2/3}$ in time t^*. For $x \in \frac{1}{n}\mathbb{Z}$ and $0 \le t \le T_n$, define the event

$$A_t^{(6)}(x) = \left\{ q^n_{t,t+k\delta_n}(x',x) \le \mathbb{1}_{|x-x'|\le (\log N)^{2/3}} \, \forall k \in [t^*\delta_n^{-1}], x' \in \tfrac{1}{n}\mathbb{Z} \right\}.$$

The next four events will give us estimates on the probability that a pair of lineages at the same site or neighbouring sites coalesce in time δ_n, and bounds on the probabilities that a pair of lineages further apart coalesce, or a set of three lineages coalesce. For $x \in \frac{1}{n}\mathbb{Z}$ and $0 \le t \le T_n$, define the events

$$\begin{aligned} B_t^{(1)}(x) &= \left\{ \frac{\big| |\mathcal{C}^n_t(x,x)| - n^2 N\delta_n p^n_t(x)\big|}{n^2 N \delta_n p^n_t(x)} \le 2n^{-1/5} \right\}, \\ B_t^{(2)}(x) &= \left\{ \frac{\big| |\mathcal{C}^n_t(x,x+n^{-1})| - \frac{1}{2}mn^2 N\delta_n (p^n_t(x)+p^n_t(x+n^{-1}))\big|}{\frac{1}{2}mn^2 N\delta_n(p^n_t(x)+p^n_t(x+n^{-1}))} \le 2n^{-1/5} \right\}, \\ B_t^{(3)}(x) &= \left\{ \frac{|\mathcal{C}^n_t(x,x')|}{n^2 N\delta_n p^n_t(x)} \le n^{-1/5}\mathbb{1}_{|x-x'|<Kn^{-1}} \, \forall x' \in \tfrac{1}{n}\mathbb{Z} \text{ with } |x'-x|>n^{-1} \right\}, \\ \text{and} \qquad B_t^{(4)}(x) &= \left\{ \frac{|\mathcal{C}^n_t(x,y,y')|}{n^2 N\delta_n p^n_t(x)} \le n^{-1/5}\mathbb{1}_{|y-x|\vee|y'-x|<Kn^{-1}} \, \forall y,y' \in \tfrac{1}{n}\mathbb{Z} \right\}. \end{aligned}$$

Fix $c_0 > 0$ sufficiently small that $(1+\frac{1}{4}(1-\alpha))(1-2c_0) > 1$. Let

$$D_n^+ = (1/2 - c_0)\kappa^{-1}\log(N/n) \quad \text{and} \quad D_n^- = -26\kappa^{-1}\alpha^{-1}\log N \tag{2.8}$$

and for $t \ge 0$ and $\epsilon \in (0,1)$, recalling (2.4), let

$$\begin{aligned} I^n_t = \tfrac{1}{n}\mathbb{Z} \cap [\mu^n_t - N^4, \mu^n_t + D^+_n], \; I^{n,\epsilon}_t &= \tfrac{1}{n}\mathbb{Z} \cap [\mu^n_t + D^-_n, \mu^n_t + (1-\epsilon)D^+_n] \\ \text{and} \quad i^n_t &= \tfrac{1}{n}\mathbb{Z} \cap [\mu^n_t - d_n, \mu^n_t + d_n]. \end{aligned} \tag{2.9}$$

We will show that with high probability, a pair of lineages are never both more than D_n^+ ahead of the front before they coalesce, and neither lineage is ever more than $|D_n^-|$ behind the front.

We now define an event which says that $(p^n_t)_{t\in[0,N^2]}$ is close to a moving front with shape g and wavespeed approximately ν. Let

$$\begin{aligned}
E_1 &= E_1(c_2)\\
&= \Big\{ \sup_{x\in\frac{1}{n}\mathbb{Z}, t\in[\log N, N^2]} |p_t^n(x) - g(x-\mu_t^n)| \le e^{-(\log N)^{c_2}} \Big\}\\
&\quad \cap \big\{ p_t^n(x) \in [\tfrac{1}{5}g(x-\mu_t^n), 5g(x-\mu_t^n)] \ \forall t\in[\tfrac{1}{2}(\log N)^2, N^2], x \le \mu_t^n + D_n^+ + 2\big\}\\
&\quad \cap \big\{ p_t^n(x) \le 5g(D_n^+) \ \forall t \in [\tfrac{1}{2}(\log N)^2, N^2], x \ge \mu_t^n + D_n^+\big\}\\
&\quad \cap \big\{ |\mu_{t+s}^n - \mu_t^n - \nu s| \le e^{-(\log N)^{c_2}} \ \forall t \in [\log N, N^2], s\in[0, 1\wedge(N^2-t)]\big\}\\
&\quad \cap \big\{ |\mu^n_{\log N}| \le 2\nu \log N\big\}.
\end{aligned}$$

Let $T_n^- = T_n - (\log N)^2$ and define the event

$$\begin{aligned}
E_2 &= E_2(c_1, t^*, K)\\
&= E_2' \cap \bigcap_{t\in\delta_n\mathbb{N}_0\cap[0,T_n^-]} \Bigg(\bigcap_{x_1\in i^n_{T_n-t-\gamma_n},\, x_2\in i^n_{T_n-t}} A^{(1)}_{T_n-t-\gamma_n}(x_1,x_2) \cap \bigcap_{x\in I^n_{T_n-t-\epsilon_n}} A^{(4)}_{T_n-t-\epsilon_n}(x)\Bigg),
\end{aligned} \tag{2.10}$$

where

$$\begin{aligned}
E_2' = E_2'(c_1,t^*,K) = &\bigcap_{t\in\delta_n\mathbb{N}_0\cap[0,T_n^-]} \ \bigcap_{x_1\in I^n_{T_n-t-t^*},\, x_2\in I^n_{T_n-t},\, x_1-\mu^n_{T_n-t-t^*}\ge K} A^{(2)}_{T_n-t-t^*}(x_1,x_2)\\
&\cap \bigcap_{t\in\delta_n\mathbb{N}_0\cap[0,T_n^-]} \ \bigcap_{x_1\in I^n_{T_n-t-t^*},\, x_2\in I^n_{T_n-t},\, x_1-\mu^n_{T_n-t-t^*}\le -K} A^{(3)}_{T_n-t-t^*}(x_1,x_2)\\
&\cap \bigcap_{t\in\delta_n\mathbb{N}_0\cap[0,T_n^-+t^*]} \ \bigcap_{x\in\frac{1}{n}\mathbb{Z}\cap[-N^5,N^5]} (A^{(5)}_{T_n-t-\epsilon_n}(x) \cap A^{(6)}_{T_n-t-\delta_n}(x)).
\end{aligned} \tag{2.11}$$

Define the event

$$E_3 = E_3(K) = \bigcap_{t\in\delta_n\mathbb{N}_0\cap[0,T_n^-]} \ \bigcap_{x\in I^n_{T_n-t}} \bigcap_{j=1}^{4} B^{(j)}_{T_n-t-\delta_n}(x). \tag{2.12}$$

Finally, we define an event which says that with high probability, no lineages stay distance K ahead of the front for time $K\log N$. Recalling (2.6), let

$$E_4 = E_4(t^*,K) = \bigcap_{t\in\delta_n\mathbb{N}_0\cap[0,T_n^-]} \left\{ \mathbb{P}\left(r^{n,K,t^*}_{K\log N, T_n-t}(x) = 0 \ \forall x\in \tfrac{1}{n}\mathbb{Z} \middle| \mathcal{F}\right) \ge 1 - \left(\frac{n}{N}\right)^2 \right\}, \tag{2.13}$$

and let $E = \cap_{j=1}^4 E_j$. Note that $E\in\mathcal{F}$ (and thus $E\in\mathcal{F}_t$ for all t) because the events $A_t^{(i)}$ and $B_t^{(j)}$ only involve p, q, and $\mathcal{C}$.

The following result will be proved in Sections 3-6.

Proposition 2.1. *Suppose for some $a_2 > 3$, $N \ge n^{a_2}$ for n sufficiently large. Take $c_1 > 0$. There exist $t^*, K \in \mathbb{N}$ (with $K > 104\kappa^{-1}\alpha^{-1}t^*$) and $b_1, c_2 > 0$ such that for $b_2 > 0$, if condition* (A) *holds, for n sufficiently large,*

$$\mathbb{P}(E^c) \le \frac{n}{N}.$$

From now on in this section, we will take $c_1 \in (0,1)$ sufficiently small that letting

$\lambda = \frac{1}{4}(1-\alpha)$,

$$\begin{aligned} c_1((e^{\lambda\kappa}-1)^{-1}e^{\lambda\kappa}+e^{-(1+\lambda)\kappa}(1-e^{-(1+\lambda)\kappa})^{-1})^2+e^{-2(1+\lambda)\kappa} &< 1, \\ c_1(e^{\lambda\kappa}-1)^{-1}e^{\lambda\kappa}+e^{-(1+\lambda)\kappa} &< 1, \\ c_1(1+e^{3\alpha\kappa/4}(e^{\alpha\kappa/4}-1)^{-1})+e^{-\alpha\kappa/4} &< 1, \\ \text{and} \qquad e^{-\alpha\kappa/4}+c_1(1-e^{-\alpha\kappa/4})^{-1} &< e^{-\alpha\kappa/5}, \end{aligned} \tag{2.14}$$

and then take t^*, K, b_1, b_2 and c_2 as in Proposition 2.1.

Take $K_0<\infty$, $k_0\in\mathbb{N}$ and (X_1,J_1), $(X_2,J_2),\dots,(X_{k_0},J_{k_0})\in\frac{1}{n}\mathbb{Z}\times[N]$ measurable with respect to $\sigma((\xi^n_{T_n}(x,i))_{x\in\frac{1}{n}\mathbb{Z},i\in[N]})$ and distinct, with $(X_i,J_i)\in G_{K_0,T_n}$ $\forall i\in[k_0]$. For $t\in[0,T_n]$ and $i\in[k_0]$, let

$$\zeta^{n,i}_t=\zeta^{n,T_n}_t(X_i,J_i) \quad \text{and} \quad \tilde{\zeta}^{n,i}_t=\zeta^{n,T_n}_t(X_i,J_i)-\mu^n_{T_n-t}, \tag{2.15}$$

the location of the i^{th} ancestral lineage at time T_n-t, and its location relative to the front. For $i,j\in[k_0]$, let

$$\tau^n_{i,j}=\inf\{t\ge 0:(\zeta^{n,T_n}_t(X_i,J_i),\theta^{n,T_n}_t(X_i,J_i))=(\zeta^{n,T_n}_t(X_j,J_j),\theta^{n,T_n}_t(X_j,J_j))\},$$

the time at which the i^{th} and j^{th} lineages coalesce. Recall (2.7), and for $t\in[0,T_n]$, define the σ-algebra

$$\mathcal{F}_t=\sigma\big(\mathcal{F},\sigma((\zeta^{n,j}_s)_{s\le t,j\in[k_0]},(\mathbb{1}_{\tau^n_{i,j}\le s})_{s\le t,i,j\in[k_0]})\big). \tag{2.16}$$

Then $((\zeta^{n,j}_{k\delta_n})_{j\in[k_0]},(\mathbb{1}_{\tau^n_{i,j}\le k\delta_n})_{i,j\in[k_0]})_{k\in\mathbb{N}_0,k\le T_n\delta_n^{-1}}$ is a strong Markov process with respect to the filtration $(\mathcal{F}_{k\delta_n})_{k\in\mathbb{N}_0,k\le T_n\delta_n^{-1}}$.

For $k\in\mathbb{N}_0$, let $t_k=k\lfloor(\log N)^C\rfloor$. For $i,j\in[k_0]$, let

$$\tilde{\tau}^n_{i,j}=\begin{cases}\tau^n_{i,j} & \text{if } \tau^n_{i,j}\notin(t_k,t_k+2K\log N]\,\forall k\in\mathbb{N}_0 \text{ and } |\tilde{\zeta}^{n,i}_{\lfloor\tau^n_{i,j}\delta_n^{-1}\rfloor\delta_n}|\wedge|\tilde{\zeta}^{n,j}_{\lfloor\tau^n_{i,j}\delta_n^{-1}\rfloor\delta_n}|\le\frac{1}{64}\alpha d_n,\\ T_n & \text{otherwise,}\end{cases} \tag{2.17}$$

i.e. $\tilde{\tau}^n_{i,j}$ only counts coalescence which happens fairly near the front and not too soon after t_k (backwards in time from time T_n) for any k. Let

$$\beta_n=(1+2m)\frac{n}{N}t_1\frac{\int_{-\infty}^{\infty}g(y)^3e^{2\alpha\kappa y}dy}{\left(\int_{-\infty}^{\infty}g(y)^2e^{\alpha\kappa y}dy\right)^2}=(1+2m)\frac{n}{N}t_1\int_{-\infty}^{\infty}\pi(y)^2g(y)^{-1}dy. \tag{2.18}$$

Along with Proposition 2.1, the following three propositions are the main intermediate results in the proof of Theorem 1.2, and will be proved in Section 2.1. The first proposition says that if a pair of lineages i and j have not coalesced by time t_k, and one of them is not too far from the front, then the probability that $\tilde{\tau}^n_{i,j}\le t_{k+1}$ is approximately β_n.

Proposition 2.2. *Suppose for some $a_2>3$, $N\ge n^{a_2}$ for n sufficiently large. For $\epsilon\in(0,1)$, on the event E, for $i,j\in[k_0]$ and $k\in\mathbb{N}_0$ with $t_{k+1}\le T_n^-$, if $\zeta^{n,i}_{t_k}\wedge\zeta^{n,j}_{t_k}\in I^{n,\epsilon}_{T_n-t_k}$ and $\tau^n_{i,j}>t_k$ then*

$$\mathbb{P}\left(\tilde{\tau}^n_{i,j}\in(t_k,t_{k+1}]\middle|\mathcal{F}_{t_k}\right)=\beta_n(1+\mathcal{O}((\log N)^{-2})).$$

The second proposition says that two pairs of lineages are unlikely to coalesce in the same time interval $(t_k,t_{k+1}]$.

Proposition 2.3. *Suppose for some $a_2>3$, $N\ge n^{a_2}$ for n sufficiently large. For $\epsilon\in(0,1)$, there exists $\epsilon'>0$ such that on the event E, for $k\in\mathbb{N}_0$ with $t_{k+1}\le T_n^-$ the following* *holds. For* $i, j_1, j_2 \in [k_0]$ *distinct, if* $\zeta^{n,\ell}_{t_k} \wedge \zeta^{n,\ell'}_{t_k} \in I^{n,\epsilon}_{T_n - t_k}$ *and* $\tau^n_{\ell,\ell'} > t_k$ $\forall \ell \neq \ell' \in \{i, j_1, j_2\}$ *then*

$$\mathbb{P}\left(\tilde{\tau}^n_{i,j_1}, \tilde{\tau}^n_{i,j_2} \in (t_k, t_{k+1}] \middle| \mathcal{F}_{t_k}\right) = \mathcal{O}(n^{1-\epsilon'} N^{-1}). \tag{2.19}$$

For $i_1, i_2, j_1, j_2 \in [k_0]$ *distinct, if* $\zeta^{n,\ell}_{t_k} \wedge \zeta^{n,\ell'}_{t_k} \in I^{n,\epsilon}_{T_n - t_k}$ *and* $\tau^n_{\ell,\ell'} > t_k$ $\forall \ell \neq \ell' \in \{i_1, i_2, j_1, j_2\}$ *then*

$$\mathbb{P}\left(\tilde{\tau}^n_{i_1,j_1}, \tilde{\tau}^n_{i_2,j_2} \in (t_k, t_{k+1}] \middle| \mathcal{F}_{t_k}\right) = \mathcal{O}(n^{1-\epsilon'} N^{-1}). \tag{2.20}$$

The last proposition says that for a pair of lineages i and j, with high probability $\tilde{\tau}^n_{i,j} = \tau^n_{i,j}$, and at least one of the lineages is fairly near the front until they have coalesced.

Proposition 2.4. *Suppose* $T_n \geq N$ *and, for some* $a_2 > 3$, $N \geq n^{a_2}$ *for* n *sufficiently large. For* $\epsilon \in (0,1)$ *sufficiently small, for* n *sufficiently large, on the event* E*, for* $i \neq j \in [k_0]$,

$$\mathbb{P}\left(\tau^n_{i,j} \neq \tilde{\tau}^n_{i,j} \middle| \mathcal{F}_0\right) \leq (\log N)^{-2}$$

and

$$\mathbb{P}\left(\exists t \in \delta_n \mathbb{N}_0 \cap [0, Nn^{-1} \log N] : \zeta^{n,i}_t \wedge \zeta^{n,j}_t \notin I^{n,\epsilon}_{T_n - t},\ \tau^n_{i,j} > t \middle| \mathcal{F}_0\right) \leq (\log N)^{-2}.$$

Before proving Propositions 2.2-2.4, we show how they can be combined with Proposition 2.1 to prove Theorem 1.2.

Proof of Theorem 1.2. Let $(B_{i,j,k})_{i<j\in[k_0],k\in\mathbb{N}_0}$ be i.i.d. Bernoulli random variables with

$$\mathbb{P}\left(B_{i,j,k} = 1\right) = \beta_n,$$

and let $B_{j,i,k} = B_{i,j,k}$ for $i < j \in [k_0]$. For $k \in \mathbb{N}_0$, let

$$P_k = \{i \in [k_0] \setminus \{1\} : \tau^n_{i,j} > t_k\ \forall j \in [i-1]\} \cup \{1\},$$

the set of lineages at time $T_n - t_k$ which have not coalesced with a lineage of lower index. Take $\epsilon > 0$ sufficiently small that Proposition 2.4 holds, and take $\epsilon' > 0$ as in Proposition 2.3. Define the event

$$A_k = \left\{\zeta^{n,i}_{t_k} \wedge \zeta^{n,j}_{t_k} \in I^{n,\epsilon}_{T_n - t_k}\ \forall i \neq j \in P_k\right\}.$$

Take $k \in \mathbb{N}_0$ with $t_{k+1} \leq T_n^-$, and suppose the event $E \cap A_k$ occurs. Then by Proposition 2.2, for each pair of lineages $i \neq j \in P_k$,

$$\mathbb{P}\left(\tilde{\tau}^n_{i,j} \in (t_k, t_{k+1}] \middle| \mathcal{F}_{t_k}\right) = \beta_n(1 + \mathcal{O}((\log N)^{-2})),$$

and by Proposition 2.3,

$$\mathbb{P}\left(|\{(i,j) : i < j \in P_k \text{ and } \tilde{\tau}^n_{i,j} \in (t_k, t_{k+1}]\}| \geq 2 \middle| \mathcal{F}_{t_k}\right) = \mathcal{O}(n^{1-\epsilon'} N^{-1}) = o(\beta_n (\log N)^{-2})$$

by the definition of β_n in (2.18). Therefore, conditional on $\mathcal{F}_{t_k}$, we can couple $(\tilde{\tau}^n_{i,j})_{i,j\in P_k}$ and $(B_{i,j,k})_{i<j\in[k_0]}$ in such a way that if $E \cap A_k$ occurs then

$$\mathbb{P}\left(\exists i \neq j \in P_k : B_{i,j,k} \neq \mathbb{1}_{\tilde{\tau}^n_{i,j} \in (t_k, t_{k+1}]} \middle| \mathcal{F}_{t_k}\right) = \mathcal{O}(\beta_n (\log N)^{-2}). \tag{2.21}$$

Note that for n sufficiently large, if the event E occurs, then by Proposition 2.4,

$$\mathbb{P}\left(\bigcup_{k=0}^{\lfloor Nn^{-1} t_1^{-1} \log N \rfloor} (A_k)^c \middle| \mathcal{F}_0\right) \leq \binom{k_0}{2} (\log N)^{-2}. \tag{2.22}$$

Now define $(\sigma^n_{i,j,k})_{i,j\in[k_0],k\in\mathbb{N}_0}$ inductively as follows. Let $\sigma^n_{i,i,0}=0\ \forall i\in[k_0]$, and $\sigma^n_{i,i',0}=t_1$ $\forall i\neq i'\in[k_0]$. For $k\in\mathbb{N}_0$, we define $(\sigma^n_{i,j,k+1})_{i,j\in[k_0]}$ using $(\sigma^n_{i,j,k})_{i,j\in[k_0]}$ as follows. For $i\in[k_0]$, let $\pi_k(i)=\min\{i'\in[k_0]:\sigma^n_{i',i,k}\leq t_k\}$. Then for each pair $i,j\in[k_0]$, set

$$\sigma^n_{i,j,k+1}=\begin{cases}\sigma^n_{i,j,k} & \text{if } \sigma^n_{i,j,k}\leq t_k,\\ t_{k+1} & \text{if } \sigma^n_{i,j,k}>t_k \text{ and } B_{\pi_k(i),\pi_k(j),k}=1,\\ t_{k+2} & \text{if } \sigma^n_{i,j,k}>t_k \text{ and } B_{\pi_k(i),\pi_k(j),k}=0.\end{cases}$$

Note that $\sigma^n_{i,j,k}$ is non-decreasing in k, and set $\sigma^n_{i,j}=\lim_{k\to\infty}\sigma^n_{i,j,k}$ for each pair $i,j\in[k_0]$, so $\sigma^n_{i,j}=\sigma^n_{i,j,k}$ for all k such that $t_k\geq\sigma^n_{i,j}$.

Suppose $\tilde\tau^n_{i,j}=\tau^n_{i,j}\ \forall i,j\in[k_0]$. For some $k\in\mathbb{N}_0$, suppose $\{(i,j):\tau^n_{i,j}>t_k\}=\{(i,j):\sigma^n_{i,j}>t_k\}$ and $B_{i,j,k}=\mathbb{1}_{\tilde\tau^n_{i,j}\in(t_k,t_{k+1}]}\ \forall i\neq j\in P_k$. Then for $i,j\in[k_0]$ with $\tau^n_{i,j}>t_k$ we have that $\tau^n_{\pi_k(i),i}\leq t_k$ and $\tau^n_{\pi_k(j),j}\leq t_k$, and so

$$\mathbb{1}_{\tau^n_{i,j}\in(t_k,t_{k+1}]}=\mathbb{1}_{\tilde\tau^n_{i,j}\in(t_k,t_{k+1}]}=\mathbb{1}_{\tilde\tau^n_{\pi_k(i),\pi_k(j)}\in(t_k,t_{k+1}]}=B_{\pi_k(i),\pi_k(j),k}=\mathbb{1}_{\sigma^n_{i,j}=t_{k+1}},$$

since $\pi_k(i),\pi_k(j)\in P_k$. In particular, $\{(i,j):\tau^n_{i,j}>t_{k+1}\}=\{(i,j):\sigma^n_{i,j}>t_{k+1}\}$. By induction, it follows that for $k^*\in\mathbb{N}$, if for each $k\in\{0\}\cup[k^*]$ we have $B_{i,j,k}=\mathbb{1}_{\tilde\tau^n_{i,j}\in(t_k,t_{k+1}]}$ $\forall i\neq j\in P_k$ then

$$\{(i,j):\tau^n_{i,j}\in(t_k,t_{k+1}]\}=\{(i,j):\sigma^n_{i,j}=t_{k+1}\}\ \forall k\in\{0\}\cup[k^*].$$

Therefore, if the event E occurs, then by a union bound,

$$\begin{aligned}
&\mathbb{P}\left(\exists i,j\in[k_0]:|\tau^n_{i,j}-\sigma^n_{i,j}|\geq(\log N)^C\middle|\mathcal{F}_0\right)\\
&\leq\mathbb{P}\left(\exists i,j\in[k_0]:\tau^n_{i,j}\neq\tilde\tau^n_{i,j}\middle|\mathcal{F}_0\right)\\
&\quad+\sum_{k=0}^{\lfloor Nn^{-1}t_1^{-1}\log N\rfloor}\mathbb{P}\left(\{\exists i\neq j\in P_k:B_{i,j,k}\neq\mathbb{1}_{\tilde\tau^n_{i,j}\in(t_k,t_{k+1}]}\}\cap A_k\middle|\mathcal{F}_0\right)\\
&\quad+\mathbb{P}\left(\bigcup_{k=0}^{\lfloor Nn^{-1}t_1^{-1}\log N\rfloor}(A_k)^c\middle|\mathcal{F}_0\right)+\mathbb{P}\left(\exists i,j\in[k_0]:\sigma^n_{i,j}>t_{\lfloor Nn^{-1}t_1^{-1}\log N\rfloor}\middle|\mathcal{F}_0\right)\\
&\leq 2\binom{k_0}{2}(\log N)^{-2}+\sum_{k=0}^{\lfloor Nn^{-1}t_1^{-1}\log N\rfloor}\mathcal{O}(\beta_n(\log N)^{-2})+\binom{k_0}{2}(1-\beta_n)^{\lfloor Nn^{-1}t_1^{-1}\log N\rfloor}\\
&=\mathcal{O}((\log N)^{-1}),
\end{aligned}$$

where the second inequality follows for n sufficiently large by Proposition 2.4, (2.21) and (2.22), and the last inequality follows by the definition of β_n in (2.18). The result follows easily by Proposition 2.1 and then by a coupling between $(\beta_n t_1^{-1}\sigma^n_{i,j})_{i,j\in[k_0]}$ and $(\tau_{i,j})_{i,j\in[k_0]}$. □

2.1 Proof of Propositions 2.2, 2.3 and 2.4

The next five results will be used in the proofs of Propositions 2.2, 2.3 and 2.4. The first three results will also be used in Section 7 in the proof of Theorem 1.1. The first result says that a pair of lineages are unlikely to be far ahead of the front, and will be proved in Section 2.2.

Proposition 2.5. *Suppose for some $a_1>1$, $N\geq n^{a_1}$ for n sufficiently large. For n sufficiently large, on the event $E_1\cap E_2'\cap E_4$, for $i,j\in[k_0]$, $s\leq t\in\delta_n\mathbb{N}_0\cap[0,T_n^-]$ and*

$\ell_1, \ell_2 \in \mathbb{N} \cap [K, D_n^+]$, the following holds. If $t - s \geq K \log N$ then

$$\mathbb{P}\left(\tilde{\zeta}_t^{n,i} \geq \ell_1, \tilde{\zeta}_t^{n,j} \geq \ell_2, \tau_{i,j}^n > t \middle| \mathcal{F}_s\right) \leq (\log N)^7 e^{-(1+\frac{1}{4}(1-\alpha))\kappa(\ell_1+\ell_2)} \tag{2.23}$$

$$\text{and} \quad \mathbb{P}\left(\tilde{\zeta}_t^{n,i} \geq \ell_1 \middle| \mathcal{F}_s\right) \leq (\log N)^3 e^{-(1+\frac{1}{4}(1-\alpha))\kappa \ell_1}. \tag{2.24}$$

If instead $t - s \in t^* \mathbb{N}_0 \cap [0, K \log N)$ then

$$\mathbb{P}\left(\tilde{\zeta}_t^{n,i} \geq \ell_1, \tilde{\zeta}_t^{n,j} \geq \ell_2, \tau_{i,j}^n > t \middle| \mathcal{F}_s\right) \leq (\log N)^4 e^{(1+\frac{1}{4}(1-\alpha))\kappa(\tilde{\zeta}_s^{n,i} \vee 0 - \ell_1 + \tilde{\zeta}_s^{n,j} \vee 0 - \ell_2)} \tag{2.25}$$

$$\text{and} \quad \mathbb{P}\left(\tilde{\zeta}_t^{n,i} \geq \ell_1 \middle| \mathcal{F}_s\right) \leq (\log N)^2 e^{(1+\frac{1}{4}(1-\alpha))\kappa(\tilde{\zeta}_s^{n,i} \vee 0 - \ell_1)}. \tag{2.26}$$

The next result says that lineages are unlikely to be far behind the front, and will be proved in Section 2.3.

Proposition 2.6. *Suppose for some $a_1 > 1$, $N \geq n^{a_1}$ for n sufficiently large. For n sufficiently large, on the event $E_1 \cap E_2'$ the following holds. For $i \in [k_0]$,*

$$\mathbb{P}\left(\exists t \in \delta_n \mathbb{N}_0 \cap [0, T_n^-] : \tilde{\zeta}_t^{n,i} \leq D_n^- \middle| \mathcal{F}_0\right) \leq N^{-1}. \tag{2.27}$$

For $i \in [k_0]$ and $s \leq t \in \delta_n \mathbb{N}_0 \cap [0, T_n^-]$ with $t - s \geq K \log N$, if $\tilde{\zeta}_s^{n,i} \geq D_n^-$ then

$$\mathbb{P}\left(\tilde{\zeta}_t^{n,i} \leq -d_n \middle| \mathcal{F}_s\right) \leq (\log N)^{2-\frac{1}{8}\alpha C} \quad \text{and} \quad \mathbb{P}\left(\tilde{\zeta}_t^{n,i} \leq -\tfrac{1}{64}\alpha d_n + 2 \middle| \mathcal{F}_s\right) \leq (\log N)^{2-2^{-9}\alpha^2 C}. \tag{2.28}$$

For $i \in [k_0]$ and $t \in t^* \mathbb{N}_0 \cap [0, T_n^-]$,

$$\mathbb{P}\left(\tilde{\zeta}_t^{n,i} \leq -d_n \middle| \mathcal{F}_0\right) \leq (\log N)^{-\frac{1}{8}\alpha C}. \tag{2.29}$$

The next lemma gives estimates on the probability that a pair of lineages are at a particular pair of sites, and gives bounds on the increments of $\zeta^{n,i}$.

Lemma 2.7. *Suppose for some $a_1 > 1$, $N \geq n^{a_1}$ for n sufficiently large. For n sufficiently large, the following holds. Suppose the event E occurs. Take $t \in \delta_n \mathbb{N}_0 \cap [0, T_n^-]$, $i, j \in [k_0]$ and $x_i, x_j \in \frac{1}{n}\mathbb{Z}$. If $x_i, x_j \in i_{T_n - t - \gamma_n}^n$, $\zeta_t^{n,i}, \zeta_t^{n,j} \in i_{T_n - t}^n$ and $\tau_{i,j}^n > t$ then*

$$\mathbb{P}\left(\zeta_{t+\gamma_n}^{n,i} = x_i, \zeta_{t+\gamma_n}^{n,j} = x_j \middle| \mathcal{F}_t\right) = n^{-2} \pi(x_i - \mu_{T_n - t - \gamma_n}^n) \pi(x_j - \mu_{T_n - t - \gamma_n}^n)(1 + \mathcal{O}((\log N)^{-C})). \tag{2.30}$$

If $x_i, x_j \in I_{T_n - t - \epsilon_n}^n$ and $\tau_{i,j}^n > t$ then

$$\mathbb{P}\left(\zeta_{t+\epsilon_n}^{n,i} = x_i, \zeta_{t+\epsilon_n}^{n,j} = x_j \middle| \mathcal{F}_t\right) \leq 2n^{-2} \epsilon_n^{-2}. \tag{2.31}$$

Suppose instead the event $E_1 \cap E_2'$ occurs. For $t \in \delta_n \mathbb{N}_0 \cap [0, T_n^-]$, $i \in [k_0]$ and $t' \in \delta_n \mathbb{N}_0 \cap [t, t + t^*]$,

$$|\zeta_t^{n,i} - \zeta_{t'}^{n,i}| \leq (\log N)^{2/3}, \quad |\zeta_t^{n,i}| \vee |\tilde{\zeta}_t^{n,i}| \leq N^3 \quad \text{and} \quad |\zeta_t^{n,i} - \zeta_{t+\epsilon_n}^{n,i}| \leq 1. \tag{2.32}$$

Proof. Suppose the event E occurs and $\tau_{i,j}^n > t$. Then for $s \in \delta_n \mathbb{N}_0 \cap [0, T_n - t]$,

$$\begin{aligned}
&\mathbb{P}\left(\zeta_{t+s}^{n,i} = x_i, \zeta_{t+s}^{n,j} = x_j \middle| \mathcal{F}_t\right) \\
&\quad = \frac{q_{T_n - t - s, T_n - t}^n(x_i, \zeta_t^{n,i})}{p_{T_n - t}^n(\zeta_t^{n,i})} \frac{q_{T_n - t - s, T_n - t}^n(x_j, \zeta_t^{n,j}) - N^{-1} \mathbb{1}_{\zeta_t^{n,i} = \zeta_t^{n,j}, x_i = x_j}}{p_{T_n - t}^n(\zeta_t^{n,j}) - N^{-1} \mathbb{1}_{\zeta_t^{n,i} = \zeta_t^{n,j}}}.
\end{aligned} \tag{2.33}$$

If $x_i, x_j \in i_{T_n - t - \gamma_n}^n$ and $\zeta_t^{n,i}, \zeta_t^{n,j} \in i_{T_n - t}^n$ then by the definition of the event E_2 in (2.10), the events $A_{T_n - t - \gamma_n}^{(1)}(x_i, \zeta_t^{n,i})$ and $A_{T_n - t - \gamma_n}^{(1)}(x_j, \zeta_t^{n,j})$ occur. Moreover, $p_{T_n - t}^n(\zeta_t^{n,j}) \geq$

$\frac{1}{5}g(d_n) \geq \frac{1}{10}(\log N)^{-C}$ by the definition of the event E_1 in (2.10) and the definition of d_n in (2.4), and so

$$\begin{aligned}&\mathbb{P}\left(\zeta^{n,i}_{t+\gamma_n} = x_i, \zeta^{n,j}_{t+\gamma_n} = x_j \middle| \mathcal{F}_t\right)\\ &= (n^{-1}\pi(x_i - \mu^n_{T_n-t-\gamma_n}) + \mathcal{O}(n^{-1}(\log N)^{-3C})) \cdot (1 + \mathcal{O}(N^{-1}(\log N)^C))\\ &\quad \cdot (n^{-1}\pi(x_j - \mu^n_{T_n-t-\gamma_n}) + \mathcal{O}(n^{-1}(\log N)^{-3C}) + \mathcal{O}(N^{-1}(\log N)^C)).\end{aligned}$$

Since $\pi(x_i - \mu^n_{T_n-t-\gamma_n})^{-1} \vee \pi(x_j - \mu^n_{T_n-t-\gamma_n})^{-1} \leq \pi(d_n)^{-1} \vee \pi(-d_n)^{-1} = \mathcal{O}((\log N)^{2C})$, the first statement (2.30) follows.

If $x_i, x_j \in I^n_{T_n-t-\epsilon_n}$ then by the definition of the event E_2 in (2.10), the events $A^{(4)}_{T_n-t-\epsilon_n}(x_i)$ and $A^{(4)}_{T_n-t-\epsilon_n}(x_j)$ occur. If $\zeta^{n,i}_t = \zeta^{n,j}_t$ then $p^n_{T_n-t}(\zeta^{n,j}_t) - N^{-1} \geq \frac{1}{2}p^n_{T_n-t}(\zeta^{n,j}_t)$, and so (2.31) follows from (2.33).

Suppose now that the event $E_1 \cap E_2'$ occurs, and suppose for some $s \in \delta_n\mathbb{N}_0 \cap [0, T_n^-]$ that $|\zeta^{n,i}_s| \leq N^3$. Then the events $A^{(5)}_{T_n-s-\epsilon_n}(\zeta^{n,i}_s)$ and $\cap_{k\in[t^*\delta_n^{-1}]}A^{(6)}_{T_n-s-k\delta_n}(\zeta^{n,i}_s)$ occur, and so $|\zeta^{n,i}_{s+\epsilon_n} - \zeta^{n,i}_s| \leq 1$ and $|\zeta^{n,i}_s - \zeta^{n,i}_{s'}| \leq (\log N)^{2/3}$ $\forall s' \in \delta_n\mathbb{N}_0 \cap [s, s+t^*]$. Since $|\tilde{\zeta}^{n,i}_0| \leq K_0$ and $|\zeta^{n,i}_0| \leq K_0 + |\mu^n_{T_n}| \leq 2\nu N^2$ for n sufficiently large, it follows by an inductive argument that $|\zeta^{n,i}_t| \vee |\tilde{\zeta}^{n,i}_t| \leq N^3$ $\forall t \in \delta_n\mathbb{N}_0 \cap [0, T_n^-]$, which completes the proof. □

From now on in Section 2.1, we will assume for some $a_2 > 3$, $N \geq n^{a_2}$ for n sufficiently large. We will also need an estimate for the probability that a pair of lineages coalesce in a very short time interval of length δ_n.

Proposition 2.8. *Suppose the event E occurs. Take $t \in \delta_n\mathbb{N}_0 \cap [0, T_n^-]$, $i, j \in [k_0]$ and $x, y \in \frac{1}{n}\mathbb{Z}$ with $|x - y| > n^{-1}$ and $x \in I^n_{T_n-t}$. If $\zeta^{n,i}_t = x = \zeta^{n,j}_t$ and $\tau^n_{i,j} > t$ then*

$$\mathbb{P}\left(\tau^n_{i,j} \in (t, t+\delta_n] \middle| \mathcal{F}_t\right) = \begin{cases} n^2N^{-1}\delta_n g(x - \mu^n_{T_n-t})^{-1}\big(1 + \mathcal{O}((\log N)^{-C})\big) & \text{if } x \in i^n_{T_n-t},\\ \mathcal{O}(n^2N^{-1}\delta_n g(x - \mu^n_{T_n-t})^{-1}) & \text{otherwise.}\end{cases}$$

If instead $\zeta^{n,i}_t = x$, $\zeta^{n,j}_t = x + n^{-1}$ and $\tau^n_{i,j} > t$ then

$$\mathbb{P}\left(\tau^n_{i,j} \in (t, t+\delta_n] \middle| \mathcal{F}_t\right) = \begin{cases} mn^2N^{-1}\delta_n g(x - \mu^n_{T_n-t})^{-1}\big(1 + \mathcal{O}((\log N)^{-C})\big) & \text{if } x \in i^n_{T_n-t},\\ \mathcal{O}(n^2N^{-1}\delta_n g(x - \mu^n_{T_n-t})^{-1}) & \text{otherwise.}\end{cases}$$

If instead $\zeta^{n,i}_t = x, \zeta^{n,j}_t = y$ and $\tau^n_{i,j} > t$ then

$$\mathbb{P}\left(\tau^n_{i,j} \in (t, t+\delta_n] \middle| \mathcal{F}_t\right) = \mathcal{O}(n^{9/5}N^{-1}\delta_n g(x - \mu^n_{T_n-t})^{-1}\mathbb{1}_{|x-y|<Kn^{-1}}).$$

Proof. For $t \in \delta_n\mathbb{N}_0 \cap [0, T_n^-]$ and $x, x' \in \frac{1}{n}\mathbb{Z}$, if $\zeta^{n,i}_t = x$, $\zeta^{n,j}_t = x'$ and $\tau^n_{i,j} > t$, then by the definition of $\mathcal{C}^n_{T_n-t-\delta_n}(x, x')$ in (2.5),

$$\mathbb{P}\left(\tau^n_{i,j} \in (t, t+\delta_n] \middle| \mathcal{F}_t\right) = \begin{cases} \frac{|\mathcal{C}^n_{T_n-t-\delta_n}(x,x')|}{Np^n_{T_n-t}(x)\cdot Np^n_{T_n-t}(x')} & \text{if } x \neq x',\\ \frac{|\mathcal{C}^n_{T_n-t-\delta_n}(x,x)|}{Np^n_{T_n-t}(x)(Np^n_{T_n-t}(x)-1)} & \text{if } x = x'.\end{cases}$$

If $x \in I^n_{T_n-t}$ and E occurs, then by the definition of the event E_3 in (2.12), $\cap_{j=1}^3 B^{(j)}_{T_n-t-\delta_n}(x)$ occurs. Hence

$$\begin{aligned}|\mathcal{C}^n_{T_n-t-\delta_n}(x,x)| &= n^2N\delta_n p^n_{T_n-t-\delta_n}(x)(1 + \mathcal{O}(n^{-1/5})),\\ |\mathcal{C}^n_{T_n-t-\delta_n}(x,x+n^{-1})| &= \tfrac{1}{2}mn^2N\delta_n(p^n_{T_n-t-\delta_n}(x) + p^n_{T_n-t-\delta_n}(x+n^{-1}))(1 + \mathcal{O}(n^{-1/5})),\\ \text{and} \quad |\mathcal{C}^n_{T_n-t-\delta_n}(x,y)| &= \mathcal{O}(n^{9/5}N\delta_n)p^n_{T_n-t-\delta_n}(x)\mathbb{1}_{|x-y|<Kn^{-1}} \ \forall y \in \tfrac{1}{n}\mathbb{Z} \text{ with } |y-x| > n^{-1}.\end{aligned}$$

The result follows by the definition of the event E_1 in (2.10), and since $n^{-1/5} = o((\log N)^{-C})$, $Np^n_{T_n-t}(x) \geq \frac{1}{5}Ng(D_n^+) \geq \frac{1}{10}n^{1/2}N^{1/2}$ for $x \in I^n_{T_n-t}$ and $g(d_n + n^{-1})^{-1} = \mathcal{O}((\log N)^C)$. □

Finally, we need a bound on the probability that two pairs of lineages coalesce in the same time interval of length δ_n.

Proposition 2.9. *Suppose the event E occurs. For $t \in \delta_n \mathbb{N}_0 \cap [0, T_n^-]$, $x_1 \in i^n_{T_n-t}$, $x_2, x_3 \in \frac{1}{n}\mathbb{Z}$, and $i_1, i_2, i_3 \in [k_0]$, if $\zeta_t^{n,i_k} = x_k$ for $k \in \{1,2,3\}$ and $\tau^n_{i_k,i_\ell} > t \, \forall k \neq \ell \in \{1,2,3\}$ then*

$$\mathbb{P}\left(\tau^n_{i_1,i_2}, \tau^n_{i_1,i_3} \in (t, t+\delta_n] \middle| \mathcal{F}_t\right) = \mathcal{O}(n^{9/5} N^{-2} \delta_n (\log N)^{2C} \mathbb{1}_{|x_1-x_2| \vee |x_1-x_3| < Kn^{-1}}). \quad (2.34)$$

For $x_1, x_3 \in i^n_{T_n-t}$, $x_2, x_4 \in \frac{1}{n}\mathbb{Z}$ and $i_1, i_2, i_3, i_4 \in [k_0]$, if $\zeta_t^{n,i_k} = x_k$ for $k \in \{1,2,3,4\}$ and $\tau^n_{i_k,i_\ell} > t \, \forall k \neq \ell \in \{1,2,3,4\}$ then

$$\mathbb{P}\left(\tau^n_{i_1,i_2}, \tau^n_{i_3,i_4} \in (t, t+\delta_n] \middle| \mathcal{F}_t\right) = \mathcal{O}(n^4 N^{-2} \delta_n^2 (\log N)^{2C} \mathbb{1}_{|x_1-x_2| \vee |x_3-x_4| < Kn^{-1}}). \quad (2.35)$$

Proof. For the first statement, since $B^{(4)}_{T_n-t-\delta_n}(x_1)$ occurs by the definition of the event E_3 in (2.12),

$$\begin{aligned}
&\mathbb{P}\left(\tau^n_{i_1,i_2}, \tau^n_{i_1,i_3} \in (t, t+\delta_n] \middle| \mathcal{F}_t\right) \\
&= \frac{|\mathcal{C}^n_{T_n-t-\delta_n}(x_1,x_2,x_3)|}{Np^n_{T_n-t}(x_1)(Np^n_{T_n-t}(x_2) - \mathbb{1}_{x_1=x_2})(Np^n_{T_n-t}(x_3) - \mathbb{1}_{x_1=x_3} - \mathbb{1}_{x_2=x_3})} \\
&\leq \mathbb{1}_{|x_1-x_2| \vee |x_1-x_3| < Kn^{-1}} \frac{6n^{9/5} N^{-2} \delta_n p^n_{T_n-t-\delta_n}(x_1)}{p^n_{T_n-t}(x_1) p^n_{T_n-t}(x_2) p^n_{T_n-t}(x_3)}.
\end{aligned}$$

By the definition of the event E_1 in (2.10) and since $x_1 - \mu^n_{T_n-t} \leq d_n$ and $g(d_n + Kn^{-1})^{-1} = \mathcal{O}((\log N)^C)$, (2.34) follows. For the second statement, since $B^{(3)}_{T_n-t-\delta_n}(x_1)$ and $B^{(3)}_{T_n-t-\delta_n}(x_3)$ occur, letting $p(x) := p^n_{T_n-t}(x)$,

$$\begin{aligned}
&\mathbb{P}\left(\tau^n_{i_1,i_2}, \tau^n_{i_3,i_4} \in (t, t+\delta_n] \middle| \mathcal{F}_t\right) \\
&\leq \frac{|\mathcal{C}^n_{T_n-t-\delta_n}(x_1,x_2)||\mathcal{C}^n_{T_n-t-\delta_n}(x_3,x_4)|}{Np(x_1)(Np(x_2) - \mathbb{1}_{x_1=x_2})(Np(x_3) - \sum_{j=1}^2 \mathbb{1}_{x_j=x_3})(Np(x_4) - \sum_{j=1}^3 \mathbb{1}_{x_j=x_4})} \\
&\leq \mathbb{1}_{|x_1-x_2| \vee |x_3-x_4| < Kn^{-1}} \frac{24|\mathcal{C}^n_{T_n-t-\delta_n}(x_1,x_2)||\mathcal{C}^n_{T_n-t-\delta_n}(x_3,x_4)|}{N^4 p^n_{T_n-t}(x_1) p^n_{T_n-t}(x_2) p^n_{T_n-t}(x_3) p^n_{T_n-t}(x_4)}.
\end{aligned}$$

Since $\cap_{j=1}^3 B^{(j)}_{T_n-t-\delta_n}(x_1)$ and $\cap_{j=1}^3 B^{(j)}_{T_n-t-\delta_n}(x_3)$ occur, and $(x_1 - \mu^n_{T_n-t}) \vee (x_3 - \mu^n_{T_n-t}) \leq d_n$, (2.35) follows by the definition of the event E_1 in (2.10). □

We are now ready to prove Propositions 2.2-2.4.

Proof of Proposition 2.2. Suppose n is sufficiently large that $\gamma_n \leq K \log N - \delta_n$. Suppose the event E occurs. Take $t \in \delta_n \mathbb{N} \cap [t_k + 2K \log N - \delta_n, t_{k+1})$, and take $x \in \frac{1}{n}\mathbb{Z}$ such that $|x - \mu^n_{T_n-t}| \leq \frac{1}{64}\alpha d_n + 1$. By conditioning on $\mathcal{F}_t$, and then by Proposition 2.8 and the definition of $\tilde{\tau}^n_{i,j}$,

$$\begin{aligned}
&\mathbb{P}\left(\tilde{\tau}^n_{i,j} \in (t, t+\delta_n], \zeta_t^{n,i} = x \middle| \mathcal{F}_{t_k}\right) \\
&= \mathbb{E}\left[\mathbb{P}\left(\tilde{\tau}^n_{i,j} \in (t, t+\delta_n] \middle| \mathcal{F}_t\right) \mathbb{1}_{\zeta_t^{n,i}=x} \mathbb{1}_{\tau^n_{i,j} > t} \middle| \mathcal{F}_{t_k}\right] \\
&\leq \mathbb{E}\Big[n^2 N^{-1} \delta_n g(x - \mu^n_{T_n-t})^{-1}(1 + \mathcal{O}((\log N)^{-C})) \\
&\qquad \left(\mathbb{1}_{\zeta_t^{n,j}=x} + m\mathbb{1}_{|\zeta_t^{n,j}-x|=n^{-1}} + \mathcal{O}(n^{-1/5})\mathbb{1}_{|\zeta_t^{n,j}-x|<Kn^{-1}}\right) \mathbb{1}_{\zeta_t^{n,i}=x} \mathbb{1}_{\tau^n_{i,j}>t} \Big| \mathcal{F}_{t_k}\Big] \\
&= n^2 N^{-1} \delta_n g(x - \mu^n_{T_n-t})^{-1}(1 + \mathcal{O}((\log N)^{-C}))
\end{aligned}$$

$$\Big(\mathbb{P}\left(\zeta_t^{n,i}=x=\zeta_t^{n,j},\tau_{i,j}^n>t\Big|\mathcal{F}_{t_k}\right)+m\mathbb{P}\left(\zeta_t^{n,i}=x,|\zeta_t^{n,j}-x|=n^{-1},\tau_{i,j}^n>t\Big|\mathcal{F}_{t_k}\right)$$
$$+\mathcal{O}(n^{-1/5})\mathbb{P}\left(\zeta_t^{n,i}=x,|\zeta_t^{n,j}-x|<Kn^{-1},\tau_{i,j}^n>t\Big|\mathcal{F}_{t_k}\right)\Big). \tag{2.36}$$

By conditioning on $\mathcal{F}_{t-\gamma_n}$ and then on $\mathcal{F}_{t-\epsilon_n}$,

$$\begin{aligned}
&\mathbb{P}\left(\zeta_t^{n,i}=x=\zeta_t^{n,j},\tau_{i,j}^n>t\Big|\mathcal{F}_{t_k}\right)\\
&=\mathbb{E}\Big[\mathbb{P}\left(\zeta_t^{n,i}=x=\zeta_t^{n,j},\tau_{i,j}^n>t\Big|\mathcal{F}_{t-\gamma_n}\right)\mathbb{1}_{\tau_{i,j}^n>t-\gamma_n}\mathbb{1}_{|\tilde\zeta_{t-\gamma_n}^{n,i}|\vee|\tilde\zeta_{t-\gamma_n}^{n,j}|\le d_n}\Big|\mathcal{F}_{t_k}\Big]\\
&\quad+\mathbb{E}\Big[\mathbb{P}\left(\zeta_t^{n,i}=x=\zeta_t^{n,j},\tau_{i,j}^n>t\Big|\mathcal{F}_{t-\epsilon_n}\right)\mathbb{1}_{\tau_{i,j}^n>t-\epsilon_n}\mathbb{1}_{|\tilde\zeta_{t-\gamma_n}^{n,i}|\vee|\tilde\zeta_{t-\gamma_n}^{n,j}|>d_n}\Big|\mathcal{F}_{t_k}\Big].
\end{aligned}\tag{2.37}$$

For the second term on the right hand side, note that by a union bound, and then by (2.28) in Proposition 2.6 and (2.24) in Proposition 2.5, and since $\tilde\zeta_{t_k}^{n,i}\wedge\tilde\zeta_{t_k}^{n,j}\ge D_n^-$ by the definition of $I_{T_n-t_k}^{n,\epsilon}$ in (2.9), and $t-\gamma_n-t_k\ge K\log N$,

$$\begin{aligned}
&\mathbb{P}\left(|\tilde\zeta_{t-\gamma_n}^{n,i}|\vee|\tilde\zeta_{t-\gamma_n}^{n,j}|>d_n\Big|\mathcal{F}_{t_k}\right)\\
&\le\mathbb{P}\left(\tilde\zeta_{t-\gamma_n}^{n,i}\wedge\tilde\zeta_{t-\gamma_n}^{n,j}<-d_n\Big|\mathcal{F}_{t_k}\right)+\mathbb{P}\left(\tilde\zeta_{t-\gamma_n}^{n,i}\vee\tilde\zeta_{t-\gamma_n}^{n,j}>d_n\Big|\mathcal{F}_{t_k}\right)\\
&\le 2(\log N)^{2-\frac18\alpha C}+2(\log N)^3e^{-(1+\frac14(1-\alpha))\kappa\lfloor d_n\rfloor}\\
&=\mathcal{O}((\log N)^{3-\frac18\alpha C})
\end{aligned}\tag{2.38}$$

by the definition of d_n in (2.4). Therefore, by (2.37) and by (2.30) and (2.31) from Lemma 2.7,

$$\begin{aligned}
&\mathbb{P}\left(\zeta_t^{n,i}=x=\zeta_t^{n,j},\tau_{i,j}^n>t\Big|\mathcal{F}_{t_k}\right)\\
&\le n^{-2}\pi(x-\mu_{T_n-t}^n)^2\left(1+\mathcal{O}((\log N)^{-C})\right)+2n^{-2}\epsilon_n^{-2}\cdot\mathcal{O}((\log N)^{3-\frac18\alpha C})\\
&=n^{-2}\pi(x-\mu_{T_n-t}^n)^2(1+\mathcal{O}((\log N)^{-2})),
\end{aligned}$$

since $\epsilon_n^{-2}=\mathcal{O}((\log N)^4)$, $\pi(x-\mu_{T_n-t}^n)^{-2}=\mathcal{O}((\log N)^{\frac{1}{16}\alpha C})$ and we chose $C>2^{13}\alpha^{-2}$, so in particular $\frac{1}{16}\alpha C-7>2$. Hence using the same argument for the other terms on the right hand side of (2.36), and since $\pi(y-\mu_{T_n-t}^n)=\pi(x-\mu_{T_n-t}^n)(1+\mathcal{O}(n^{-1}))$ if $|x-y|<Kn^{-1}$,

$$\begin{aligned}
&\mathbb{P}\left(\tilde\tau_{i,j}^n\in(t,t+\delta_n],\zeta_t^{n,i}=x\Big|\mathcal{F}_{t_k}\right)\\
&\le N^{-1}\delta_n(1+2m)g(x-\mu_{T_n-t}^n)^{-1}\pi(x-\mu_{T_n-t}^n)^2\left(1+\mathcal{O}((\log N)^{-2})\right).
\end{aligned}$$

Note that if $\tilde\tau_{i,j}^n\in(t,t+\delta_n]$ then $|\tilde\zeta_t^{n,i}|\wedge|\tilde\zeta_t^{n,j}|\le\frac{1}{64}\alpha d_n$ by the definition of $\tilde\tau_{i,j}^n$ in (2.17), and $|\tilde\zeta_t^{n,i}-\tilde\zeta_t^{n,j}|<Kn^{-1}$ by Proposition 2.8, and so for n sufficiently large, we must have $|\tilde\zeta_t^{n,i}|\le\frac{1}{64}\alpha d_n+1$. Letting $\tilde i_s^n=\frac1n\mathbb{Z}\cap[\mu_s^n-\frac{1}{64}\alpha d_n-1,\mu_s^n+\frac{1}{64}\alpha d_n+1]$ for $s\ge0$, it follows that

$$\begin{aligned}
&\mathbb{P}\left(\tilde\tau_{i,j}^n\in(t_k+2K\log n,t_{k+1}]\Big|\mathcal{F}_{t_k}\right)\\
&\le N^{-1}\delta_n(1+2m)\left(1+\mathcal{O}((\log N)^{-2})\right)\\
&\qquad\cdot\sum_{t\in\delta_n\mathbb{N}\cap[t_k+2K\log N-\delta_n,t_{k+1})}\ \sum_{x\in\tilde i_{T_n-t}^n}g(x-\mu_{T_n-t}^n)^{-1}\pi(x-\mu_{T_n-t}^n)^2\\
&\le\beta_n\left(1+\mathcal{O}((\log N)^{-2})\right),
\end{aligned}\tag{2.39}$$

by the definition of β_n in (2.18).

For a lower bound, note that for $t \in \delta_n \mathbb{N} \cap [t_k + 2K \log N, t_{k+1})$,

$$\begin{aligned}
&\mathbb{P}\left(\tilde{\tau}^n_{i,j} \in (t, t+\delta_n] \middle| \mathcal{F}_{t_k}\right) \\
&\geq \sum_{x \in 2(\log N)^{-C}\mathbb{Z}, |x-\mu^n_{T_n-t}| \leq \frac{1}{64}\alpha d_n - 1} \mathbb{P}\left(\tilde{\tau}^n_{i,j} \in (t, t+\delta_n], |\zeta^{n,i}_t - x| < (\log N)^{-C} \middle| \mathcal{F}_{t_k}\right).
\end{aligned} \tag{2.40}$$

Now for $x \in 2(\log N)^{-C}\mathbb{Z}$ with $|x - \mu^n_{T_n-t}| \leq \frac{1}{64}\alpha d_n - 1$, by conditioning on $\mathcal{F}_t$, and then by Proposition 2.8,

$$\begin{aligned}
&\mathbb{P}\left(\tilde{\tau}^n_{i,j} \in (t, t+\delta_n], |\zeta^{n,i}_t - x| < (\log N)^{-C} \middle| \mathcal{F}_{t_k}\right) \\
&= \mathbb{E}\left[\mathbb{P}\left(\tilde{\tau}^n_{i,j} \in (t, t+\delta_n] \middle| \mathcal{F}_t\right) \mathbb{1}_{\tau^n_{i,j} > t} \mathbb{1}_{|\zeta^{n,i}_t - x| < (\log N)^{-C}} \middle| \mathcal{F}_{t_k}\right] \\
&\geq \mathbb{E}\Big[n^2 N^{-1} \delta_n g(\zeta^{n,i}_t - \mu^n_{T_n-t})^{-1}(1 - \mathcal{O}((\log N)^{-C}))(\mathbb{1}_{\zeta^{n,i}_t = \zeta^{n,j}_t} + m\mathbb{1}_{|\zeta^{n,i}_t - \zeta^{n,j}_t| = n^{-1}}) \\
&\qquad\qquad \mathbb{1}_{\tau^n_{i,j} > t} \mathbb{1}_{|\zeta^{n,i}_t - x| < (\log N)^{-C}} \Big| \mathcal{F}_{t_k}\Big] \\
&= n^2 N^{-1} \delta_n g(x - \mu^n_{T_n-t})^{-1}(1 - \mathcal{O}((\log N)^{-C})) \\
&\quad \Big(\mathbb{P}\left(\zeta^{n,i}_t = \zeta^{n,j}_t, |\zeta^{n,i}_t - x| < (\log N)^{-C}, \tau^n_{i,j} > t \middle| \mathcal{F}_{t_k}\right) \\
&\qquad + m\mathbb{P}\left(|\zeta^{n,i}_t - \zeta^{n,j}_t| = n^{-1}, |\zeta^{n,i}_t - x| < (\log N)^{-C}, \tau^n_{i,j} > t \middle| \mathcal{F}_{t_k}\right)\Big).
\end{aligned} \tag{2.41}$$

For the first term on the right hand side, by conditioning on $\mathcal{F}_{t-\gamma_n}$,

$$\begin{aligned}
&\mathbb{P}\left(\zeta^{n,i}_t = \zeta^{n,j}_t, |\zeta^{n,i}_t - x| < (\log N)^{-C}, \tau^n_{i,j} > t \middle| \mathcal{F}_{t_k}\right) \\
&\geq \mathbb{E}\Big[\mathbb{P}\left(\zeta^{n,i}_t = \zeta^{n,j}_t, |\zeta^{n,i}_t - x| < (\log N)^{-C}, \tau^n_{i,j} > t \middle| \mathcal{F}_{t-\gamma_n}\right) \\
&\qquad\qquad \mathbb{1}_{\tau^n_{i,j} > t-\gamma_n} \mathbb{1}_{|\tilde{\zeta}^{n,i}_{t-\gamma_n}| \vee |\tilde{\zeta}^{n,j}_{t-\gamma_n}| \leq d_n} \Big| \mathcal{F}_{t_k}\Big].
\end{aligned} \tag{2.42}$$

By a union bound, if $\tau^n_{i,j} > t - \gamma_n$ then

$$\begin{aligned}
\mathbb{P}\left(\tau^n_{i,j} \leq t \middle| \mathcal{F}_{t-\gamma_n}\right) &\leq \sum_{s \in \delta_n \mathbb{N} \cap [t-\gamma_n, t)} \mathbb{P}\left(\tau^n_{i,j} \in (s, s+\delta_n], \zeta^{n,i}_s \in I^n_{T_n-s} \text{ or } \zeta^{n,j}_s \in I^n_{T_n-s} \middle| \mathcal{F}_{t-\gamma_n}\right) \\
&\quad + \mathbb{P}\left(\exists s \in \delta_n \mathbb{N} \cap [t-\gamma_n, t) : \zeta^{n,i}_s, \zeta^{n,j}_s \notin I^n_{T_n-s},\ \tau^n_{i,j} > s \middle| \mathcal{F}_{t-\gamma_n}\right).
\end{aligned} \tag{2.43}$$

Suppose $|\tilde{\zeta}^{n,i}_{t-\gamma_n}| \vee |\tilde{\zeta}^{n,j}_{t-\gamma_n}| \leq d_n$. Take $s \in \delta_n \mathbb{N} \cap [t-\gamma_n, t)$, and let $I = 2\mathbb{Z} \cap [\mu^n_{T_n-s} + (\log N)^{2/3} + K + \nu t^* + 3, \mu^n_{T_n-s} + D^+_n]$; then by conditioning on $\mathcal{F}_s$ and using Proposition 2.8,

$$\begin{aligned}
&\mathbb{P}\left(\tau^n_{i,j} \in (s, s+\delta_n], \zeta^{n,i}_s \in I^n_{T_n-s} \middle| \mathcal{F}_{t-\gamma_n}\right) \\
&\leq \mathbb{E}\Big[\mathcal{O}(n^2 N^{-1} \delta_n g(\zeta^{n,i}_s - \mu^n_{T_n-s})^{-1}) \mathbb{1}_{|\zeta^{n,i}_s - \zeta^{n,j}_s| < Kn^{-1}} \mathbb{1}_{\tau^n_{i,j} > s} \mathbb{1}_{\zeta^{n,i}_s \in I^n_{T_n-s}} \Big| \mathcal{F}_{t-\gamma_n}\Big] \\
&\leq \mathcal{O}(n^2 N^{-1} \delta_n) \sum_{x' \in I} g(x' + 1 - \mu^n_{T_n-s})^{-1} \mathbb{P}\Big(|\zeta^{n,i}_s - x'| \leq 1, |\zeta^{n,j}_s - x'| \leq 2, \tau^n_{i,j} > s \Big| \mathcal{F}_{t-\gamma_n}\Big) \\
&\quad + \mathcal{O}(n^2 N^{-1} \delta_n g((\log N)^{2/3} + K + \nu t^* + 4)^{-1}).
\end{aligned} \tag{2.44}$$

Take $s' \in [s - t^*, s]$ such that $s' - (t - \gamma_n) \in t^* \mathbb{N}_0$. Then by (2.32) in Lemma 2.7, for $x' \in I$,

$$\mathbb{P}\Big(|\zeta^{n,i}_s - x'| \leq 1, |\zeta^{n,j}_s - x'| \leq 2, \tau^n_{i,j} > s \Big| \mathcal{F}_{t-\gamma_n}\Big)$$

$$\leq \mathbb{P}\Big(\zeta^{n,i}_{s'} \geq x' - 1 - (\log N)^{2/3},\, \zeta^{n,j}_{s'} \geq x' - 2 - (\log N)^{2/3}, \tau^n_{i,j} > s' \Big| \mathcal{F}_{t-\gamma_n}\Big)$$
$$\leq (\log N)^4 e^{2(1+\frac{1}{4}(1-\alpha))\kappa(d_n - (x' - 3 - (\log N)^{2/3} - \mu^n_{T_n - s'}))}$$

by (2.25) in Proposition 2.5 (since $s' - (t - \gamma_n) \leq \gamma_n \leq K \log N$ and we are assuming $\tilde{\zeta}^{n,i}_{t-\gamma_n} \vee \tilde{\zeta}^{n,j}_{t-\gamma_n} \leq d_n$). Therefore, by (2.44),

$$\begin{aligned}
&\mathbb{P}\left(\tau^n_{i,j} \in (s, s+\delta_n], \zeta^{n,i}_s \in I^n_{T_n-s} \Big| \mathcal{F}_{t-\gamma_n}\right) \\
&\leq \mathcal{O}(n^2 N^{-1} \delta_n) \\
&\quad \cdot \Bigg(\sum_{x' \in I} g(x' + 1 - \mu^n_{T_n - s})^{-1} (\log N)^{4+4C} e^{4\kappa (\log N)^{2/3}} e^{-2(1+\frac{1}{4}(1-\alpha))\kappa(x' - 3 - \mu^n_{T_n - s'})} \\
&\qquad\qquad + 2 e^{\kappa((\log N)^{2/3} + K + \nu t^* + 4)} \Bigg) \\
&= \mathcal{O}(n^2 N^{-1} \delta_n (\log N)^{4+4C} e^{4\kappa(\log N)^{2/3}})
\end{aligned} \tag{2.45}$$

since $g(y)^{-1} \leq 2e^{\kappa y}$ for $y \geq 0$, and by the definition of the event E_1 in (2.10). For the second term on the right hand side of (2.43), first note that for n sufficiently large, by the definition of the event E_1, for $s, s' > 0$ with $s' \leq s < t < t_{k+1} \leq T^-_n$ and $|s - s'| \leq t^*$ we have $|\mu^n_{T_n - s} - \mu^n_{T_n - s'}| \leq 2\nu t^*$. Hence, since we are assuming the event $E_1 \cap E'_2$ occurs, by (2.32) in Lemma 2.7 we have

$$\begin{aligned}
&\mathbb{P}\left(\exists s \in \delta_n \mathbb{N} \cap [t - \gamma_n, t) : \zeta^{n,i}_s, \zeta^{n,j}_s \notin I^n_{T_n - s},\, \tau^n_{i,j} > s \Big| \mathcal{F}_{t-\gamma_n}\right) \\
&\leq \mathbb{P}\Big(\exists s' \in [t-\gamma_n, t) : s' - (t - \gamma_n) \in t^* \mathbb{N}_0, \\
&\qquad\qquad \tilde{\zeta}^{n,i}_{s'} \wedge \tilde{\zeta}^{n,j}_{s'} \geq D^+_n - (\log N)^{2/3} - 2\nu t^*, \tau^n_{i,j} > s' \Big| \mathcal{F}_{t-\gamma_n}\Big) \\
&\leq ((t^*)^{-1} + 1)\gamma_n (\log N)^4 e^{2(1+\frac{1}{4}(1-\alpha))\kappa(d_n - (D^+_n - (\log N)^{2/3} - 2\nu t^* - 1))}
\end{aligned}$$

by (2.25) in Proposition 2.5 and since $\tilde{\zeta}^{n,i}_{t-\gamma_n} \vee \tilde{\zeta}^{n,j}_{t-\gamma_n} \leq d_n$. Note that $e^{-2(1+\frac{1}{4}(1-\alpha))\kappa D^+_n} = \left(\frac{n}{N}\right)^{(1+\frac{1}{4}(1-\alpha))(1-2c_0)} \leq \frac{n}{N}$ by (2.8) and our choice of c_0. Hence, by (2.45), substituting into (2.43),

$$\begin{aligned}
&\mathbb{P}\left(\tau^n_{i,j} \leq t \Big| \mathcal{F}_{t-\gamma_n}\right) \\
&\leq \mathcal{O}(n^2 N^{-1} \gamma_n (\log N)^{4+4C} e^{4\kappa(\log N)^{2/3}}) + \mathcal{O}(\gamma_n (\log N)^{4+4C} e^{4\kappa(\log N)^{2/3}} n N^{-1}) \\
&= \mathcal{O}(n^{-1-\frac{1}{2}(a_2 - 3)}),
\end{aligned}$$

since $N \geq n^{a_2}$ for n sufficiently large, with $a_2 > 3$. Therefore, returning to (2.42), if $|\tilde{\zeta}^{n,i}_{t-\gamma_n}| \vee |\tilde{\zeta}^{n,j}_{t-\gamma_n}| \leq d_n$ and $\tau^n_{i,j} > t - \gamma_n$,

$$\begin{aligned}
&\mathbb{P}\left(\zeta^{n,i}_t = \zeta^{n,j}_t, |\zeta^{n,i}_t - x| < (\log N)^{-C}, \tau^n_{i,j} > t \Big| \mathcal{F}_{t-\gamma_n}\right) \\
&\geq \mathbb{P}\left(\zeta^{n,i}_t = \zeta^{n,j}_t, |\zeta^{n,i}_t - x| < (\log N)^{-C} \Big| \mathcal{F}_{t-\gamma_n}\right) - \mathbb{P}\left(\tau^n_{i,j} \leq t \Big| \mathcal{F}_{t-\gamma_n}\right) \\
&\geq \pi(x - \mu^n_{T_n - t})^2 \cdot 2(\log N)^{-C} n^{-1} \left(1 - \mathcal{O}((\log N)^{-C})\right) - \mathcal{O}(n^{-1-\frac{1}{2}(a_2-3)})
\end{aligned} \tag{2.46}$$

by (2.30) in Lemma 2.7 and since $\pi(y - \mu^n_{T_n - t}) = \pi(x - \mu^n_{T_n-t})(1 + \mathcal{O}((\log N)^{-C}))$ if $|y - x| < (\log N)^{-C}$. To bound the other terms in (2.42), note first that by a union bound,

$$\begin{aligned}
&\mathbb{P}\left(\tau^n_{i,j} \leq t - \gamma_n \Big| \mathcal{F}_{t_k}\right) \\
&\leq \sum_{s \in \delta_n \mathbb{N}_0 \cap [t_k, t - \gamma_n)} \mathbb{P}\left(\tau^n_{i,j} \in (s, s+\delta_n], \zeta^{n,i}_s \in I^n_{T_n - s} \text{ or } \zeta^{n,j}_s \in I^n_{T_n - s} \Big| \mathcal{F}_{t_k}\right)
\end{aligned}$$

$$+\,\mathbb{P}\left(\exists s'\in\delta_n\mathbb{N}_0\cap[t_k,t-\gamma_n):\zeta^{n,i}_{s'}\wedge\zeta^{n,j}_{s'}\notin I^n_{T_n-s'}\Big|\mathcal{F}_{t_k}\right). \tag{2.47}$$

By Proposition 2.8, for $s\in\delta_n\mathbb{N}_0\cap[t_k,t-\gamma_n)$,

$$\begin{aligned}\mathbb{P}\left(\tau^n_{i,j}\in(s,s+\delta_n],\zeta^{n,i}_s\in I^n_{T_n-s}\Big|\mathcal{F}_{t_k}\right)&=\mathbb{E}\left[\mathbb{P}\left(\tau^n_{i,j}\in(s,s+\delta_n]\Big|\mathcal{F}_s\right)\mathbb{1}_{\zeta^{n,i}_s\in I^n_{T_n-s}}\Big|\mathcal{F}_{t_k}\right]\\&=\mathcal{O}(n^2N^{-1}\delta_n g(D^+_n)^{-1})\\&=\mathcal{O}(n^{3/2}N^{-1/2}\delta_n)\end{aligned}\tag{2.48}$$

since $\kappa D^+_n\le\frac12\log(N/n)$ by (2.8). For the second term on the right hand side of (2.47), by (2.32) in Lemma 2.7 and by the definition of the event E_1 in (2.10),

$$\begin{aligned}&\mathbb{P}\left(\exists s'\in\delta_n\mathbb{N}_0\cap[t_k,t-\gamma_n):\zeta^{n,i}_{s'}\wedge\zeta^{n,j}_{s'}\notin I^n_{T_n-s'}\Big|\mathcal{F}_{t_k}\right)\\&\le\mathbb{P}\left(\exists s'\in[t_k,t-\gamma_n):s'-t_k\in t^*\mathbb{N}_0,\tilde\zeta^{n,i}_{s'}\wedge\tilde\zeta^{n,j}_{s'}\ge D^+_n-(\log N)^{2/3}-2\nu t^*\Big|\mathcal{F}_{t_k}\right)\\&\le((t^*)^{-1}t_1+1)(\log N)^3e^{(1+\frac14(1-\alpha))\kappa((1-\epsilon)D^+_n-(D^+_n-(\log N)^{2/3}-2\nu t^*-1))}\end{aligned}$$

by (2.24) and (2.26) in Proposition 2.5 and since $\tilde\zeta^{n,i}_{t_k}\wedge\tilde\zeta^{n,j}_{t_k}\le(1-\epsilon)D^+_n$. Hence by (2.47) and (2.48), and since $\kappa(1+\frac14(1-\alpha))D^+_n\ge\frac12\log(N/n)$ by the definition of D^+_n in (2.8),

$$\begin{aligned}\mathbb{P}\left(\tau^n_{i,j}\le t-\gamma_n\Big|\mathcal{F}_{t_k}\right)&\le\mathcal{O}(t_1n^{3/2}N^{-1/2})+\mathcal{O}(t_1(\log N)^3e^{2\kappa(\log N)^{2/3}}n^{\epsilon/2}N^{-\epsilon/2})\\&=\mathcal{O}(n^{-(\frac13(a_2-3)\wedge\epsilon)}).\end{aligned}\tag{2.49}$$

Therefore, substituting into (2.42) and using (2.38) and (2.46),

$$\begin{aligned}&\mathbb{P}\left(\zeta^{n,i}_t=\zeta^{n,j}_t,|\zeta^{n,i}_t-x|<(\log N)^{-C},\tau^n_{i,j}>t\Big|\mathcal{F}_{t_k}\right)\\&\ge2\pi(x-\mu^n_{T_n-t})^2(\log N)^{-C}n^{-1}\left(1-\mathcal{O}((\log N)^{-C})\right)\\&\qquad\cdot(1-\mathcal{O}(n^{-(\frac13(a_2-3)\wedge\epsilon)})-\mathcal{O}((\log N)^{3-\frac18\alpha C})).\end{aligned}$$

Since we chose $C>2^{13}\alpha^{-2}$, we have $\frac18\alpha C-3>2$. Hence by the same argument for the second term on the right hand side of (2.41), and then substituting into (2.40),

$$\begin{aligned}&\mathbb{P}\left(\tilde\tau^n_{i,j}\in(t,t+\delta_n]\Big|\mathcal{F}_{t_k}\right)\\&\ge\sum_{x\in2(\log N)^{-C}\mathbb{Z},|x-\mu^n_{T_n-t}|\le\frac{1}{64}\alpha d_n-1}2(\log N)^{-C}nN^{-1}\delta_n(1+2m)\\&\qquad\cdot\frac{\pi(x-\mu^n_{T_n-t})^2}{g(x-\mu^n_{T_n-t})}\left(1-\mathcal{O}((\log N)^{-2})\right)\\&=\beta_nt_1^{-1}\delta_n(1-\mathcal{O}((\log N)^{-2})),\end{aligned}$$

since $\frac{1}{32}\alpha^2C>2$ and $\frac{1}{64}\alpha C>2$, which, together with (2.39), completes the proof. □

Proof of Proposition 2.3. Suppose n is sufficiently large that $2K\log N-\delta_n\ge\epsilon_n$. Suppose the event E occurs. We begin by proving the first statement (2.19). Take $s,t\in\delta_n\mathbb{N}\cap[t_k+2K\log N-\delta_n,t_{k+1})$ with $s<t$. Note that if for some $\ell,\ell'\in[k_0]$, $\tilde\tau^n_{\ell,\ell'}\in(t,t+\delta_n]$ then $|\tilde\zeta^{n,\ell}_t|\wedge|\tilde\zeta^{n,\ell'}_t|\le\frac{1}{64}\alpha d_n$ by the definition of $\tilde\tau^n_{\ell,\ell'}$ in (2.17), and $|\tilde\zeta^{n,\ell}_t-\tilde\zeta^{n,\ell'}_t|<Kn^{-1}$ by Proposition 2.8, so in particular $|\tilde\zeta^{n,\ell}_t|\le d_n$. Hence by conditioning on $\mathcal{F}_t$ and applying Proposition 2.8,

$$\begin{aligned}&\mathbb{P}\left(\tilde\tau^n_{i,j_1}\in(s,s+\delta_n],\tilde\tau^n_{i,j_2}\in(t,t+\delta_n]\Big|\mathcal{F}_{t_k}\right)\\&\le\mathbb{E}\left[\mathcal{O}(n^2N^{-1}\delta_ng(\tilde\zeta^{n,i}_t)^{-1})\mathbb{1}_{|\tilde\zeta^{n,i}_t|\le d_n}\mathbb{1}_{\tilde\tau^n_{i,j_1}\in(s,s+\delta_n]}\Big|\mathcal{F}_{t_k}\right]\\&\le\mathcal{O}(n^2N^{-1}\delta_n(\log N)^C)\mathbb{P}\left(\tilde\tau^n_{i,j_1}\in(s,s+\delta_n]\Big|\mathcal{F}_{t_k}\right).\end{aligned}\tag{2.50}$$

By conditioning on $\mathcal{F}_s$ and applying Proposition 2.8,

$$\begin{aligned}
&\mathbb{P}\left(\tilde{\tau}^n_{i,j_1} \in (s, s+\delta_n] \Big| \mathcal{F}_{t_k}\right) \\
&\leq \mathbb{E}\Big[\mathcal{O}(n^2 N^{-1}\delta_n g(\tilde{\zeta}^{n,i}_s)^{-1})\mathbb{1}_{\tau^n_{i,j_1} > s}\mathbb{1}_{|\tilde{\zeta}^{n,i}_s| \leq d_n}\mathbb{1}_{|\zeta^{n,i}_s - \zeta^{n,j_1}_s| < Kn^{-1}}\Big|\mathcal{F}_{t_k}\Big] \\
&= \mathcal{O}(n^2 N^{-1}\delta_n (\log N)^C)\mathbb{P}\left(|\tilde{\zeta}^{n,i}_s| \leq d_n, |\zeta^{n,i}_s - \zeta^{n,j_1}_s| < Kn^{-1}, \tau^n_{i,j_1} > s \Big| \mathcal{F}_{t_k}\right).
\end{aligned}$$

Then since $s - t_k \geq \epsilon_n$, by conditioning on $\mathcal{F}_{s-\epsilon_n}$,

$$\begin{aligned}
&\mathbb{P}\left(|\tilde{\zeta}^{n,i}_s| \leq d_n, |\zeta^{n,i}_s - \zeta^{n,j_1}_s| < Kn^{-1}, \tau^n_{i,j_1} > s \Big| \mathcal{F}_{t_k}\right) \\
&\leq \mathbb{E}\Big[\mathbb{P}\left(|\tilde{\zeta}^{n,i}_s| \leq d_n, |\zeta^{n,i}_s - \zeta^{n,j_1}_s| < Kn^{-1} \Big| \mathcal{F}_{s-\epsilon_n}\right)\mathbb{1}_{\tau^n_{i,j_1} > s-\epsilon_n}\Big|\mathcal{F}_{t_k}\Big] \\
&\leq \mathbb{E}\left[\sum_{x \in i^n_{T_n - s}, y \in \frac{1}{n}\mathbb{Z}, |x-y| < Kn^{-1}} \mathbb{P}\left(\zeta^{n,i}_s = x, \zeta^{n,j}_s = y \Big| \mathcal{F}_{s-\epsilon_n}\right)\mathbb{1}_{\tau^n_{i,j_1} > s-\epsilon_n}\Bigg|\mathcal{F}_{t_k}\right] \\
&\leq (2nd_n + 1)2K \cdot 2n^{-2}\epsilon_n^{-2}
\end{aligned} \tag{2.51}$$

by (2.31) in Lemma 2.7. Hence, by (2.50), and by the same argument for the case $s > t$, if $s, t \in \delta_n \mathbb{N} \cap [t_k + 2K \log N - \delta_n, t_{k+1})$ with $s \neq t$,

$$\mathbb{P}\left(\tilde{\tau}^n_{i,j_1} \in (s, s+\delta_n], \tilde{\tau}^n_{i,j_2} \in (t, t+\delta_n] \Big| \mathcal{F}_{t_k}\right) = \mathcal{O}(n^3 N^{-2}\delta_n^2 (\log N)^{2C+5}). \tag{2.52}$$

By Proposition 2.9, for $t \in \delta_n \mathbb{N} \cap [t_k + 2K \log N - \delta_n, t_{k+1})$,

$$\begin{aligned}
&\mathbb{P}\left(\tilde{\tau}^n_{i,j_1}, \tilde{\tau}^n_{i,j_2} \in (t, t+\delta_n] \Big| \mathcal{F}_{t_k}\right) \\
&= \mathcal{O}(n^{9/5} N^{-2}\delta_n (\log N)^{2C}) + \mathbb{P}\left(\tilde{\tau}^n_{i,j_1} \in (t, t+\delta_n], \tau^n_{j_1,j_2} \leq t \Big| \mathcal{F}_{t_k}\right).
\end{aligned} \tag{2.53}$$

By a union bound, and then by conditioning on $\mathcal{F}_t$ and using Proposition 2.8,

$$\begin{aligned}
&\mathbb{P}\left(\tilde{\tau}^n_{i,j_1} \in (t, t+\delta_n], \tau^n_{j_1,j_2} \in (t-\epsilon_n, t] \Big| \mathcal{F}_{t_k}\right) \\
&= \sum_{t' \in \delta_n \mathbb{N} \cap [t-\epsilon_n, t)} \mathbb{P}\left(\tilde{\tau}^n_{i,j_1} \in (t, t+\delta_n], \tau^n_{j_1,j_2} \in (t', t'+\delta_n] \Big| \mathcal{F}_{t_k}\right) \\
&\leq \sum_{t' \in \delta_n \mathbb{N} \cap [t-\epsilon_n, t)} \mathbb{E}\left[\mathcal{O}(n^2 N^{-1}\delta_n g(\tilde{\zeta}^{n,j_1}_t)^{-1})\mathbb{1}_{|\tilde{\zeta}^{n,j_1}_t| \leq d_n}\mathbb{1}_{\tau^n_{j_1,j_2} \in (t', t'+\delta_n]}\Big|\mathcal{F}_{t_k}\right] \\
&\leq \sum_{t' \in \delta_n \mathbb{N} \cap [t-\epsilon_n, t)} \mathcal{O}(n^2 N^{-1}\delta_n (\log N)^C) \\
&\qquad \cdot \mathbb{P}\left(\tau^n_{j_1,j_2} \in (t', t'+\delta_n], |\tilde{\zeta}^{n,j_1}_{t'}| \leq d_n + (\log N)^{2/3} + 1 \Big| \mathcal{F}_{t_k}\right)
\end{aligned}$$

by (2.32) in Lemma 2.7 and the definition of the event E_1 in (2.10). Then by Proposition 2.8 again, for $t' \in \delta_n \mathbb{N} \cap [t - \epsilon_n, t)$, by conditioning on $\mathcal{F}_{t'}$,

$$\begin{aligned}
&\mathbb{P}\left(\tau^n_{j_1,j_2} \in (t', t'+\delta_n], |\tilde{\zeta}^{n,j_1}_{t'}| \leq d_n + (\log N)^{2/3} + 1 \Big| \mathcal{F}_{t_k}\right) \\
&= \mathcal{O}(n^2 N^{-1}\delta_n g(d_n + (\log N)^{2/3} + 1)^{-1}).
\end{aligned}$$

Hence

$$\begin{aligned}
\mathbb{P}\left(\tilde{\tau}^n_{i,j_1} \in (t, t+\delta_n], \tau^n_{j_1,j_2} \in (t-\epsilon_n, t] \Big| \mathcal{F}_{t_k}\right) &= \mathcal{O}(n^4 N^{-2}\delta_n \epsilon_n (\log N)^C e^{2\kappa(\log N)^{2/3}}) \\
&= \mathcal{O}(n^{1-\frac{1}{2}(a_2-3)} N^{-1}\delta_n).
\end{aligned} \tag{2.54}$$

Moreover, by Proposition 2.8, conditioning on $\mathcal{F}_t$, and then conditioning on $\mathcal{F}_{t-\epsilon_n}$,

$$\begin{aligned}
&\mathbb{P}\left(\tilde{\tau}^n_{i,j_1}\in(t,t+\delta_n],\tau^n_{j_1,j_2}\le t-\epsilon_n\middle|\mathcal{F}_{t_k}\right)\\
&=\mathbb{E}\left[\mathcal{O}(n^2N^{-1}\delta_n g(\tilde{\zeta}^{n,i}_t)^{-1})\mathbb{1}_{\tau^n_{i,j_1}>t}\mathbb{1}_{|\tilde{\zeta}^{n,i}_t|\le d_n}\mathbb{1}_{|\zeta^{n,i}_t-\zeta^{n,j_1}_t|<Kn^{-1}}\mathbb{1}_{\tau^n_{j_1,j_2}\le t-\epsilon_n}\middle|\mathcal{F}_{t_k}\right]\\
&\le\mathcal{O}(n^2N^{-1}\delta_n(\log N)^C)\\
&\qquad\cdot\mathbb{E}\Big[\mathbb{P}\left(|\zeta^{n,i}_t-\zeta^{n,j_1}_t|<Kn^{-1},|\tilde{\zeta}^{n,i}_t|\le d_n\middle|\mathcal{F}_{t-\epsilon_n}\right)\mathbb{1}_{\tau^n_{i,j_1}>t-\epsilon_n}\mathbb{1}_{\tau^n_{j_1,j_2}\le t-\epsilon_n}\Big|\mathcal{F}_{t_k}\Big].
\end{aligned}\tag{2.55}$$

By the same argument as in (2.51), if $\tau^n_{i,j_1}>t-\epsilon_n$ then

$$\mathbb{P}\left(|\zeta^{n,i}_t-\zeta^{n,j_1}_t|<Kn^{-1},|\tilde{\zeta}^{n,i}_t|\le d_n\middle|\mathcal{F}_{t-\epsilon_n}\right)\le(2nd_n+1)2K\cdot 2n^{-2}\epsilon_n^{-2}=\mathcal{O}(n^{-1}(\log N)^5).$$

By the same argument as in (2.49) in the proof of Proposition 2.2,

$$\mathbb{P}\left(\tau^n_{j_1,j_2}\le t-\epsilon_n\middle|\mathcal{F}_{t_k}\right)=\mathcal{O}(n^{-(\frac{1}{3}(a_2-3)\wedge\epsilon)}).$$

Hence by (2.55),

$$\mathbb{P}\left(\tilde{\tau}^n_{i,j_1}\in(t,t+\delta_n],\tau^n_{j_1,j_2}\le t-\epsilon_n\middle|\mathcal{F}_{t_k}\right)=\mathcal{O}(n^{1-(\frac{1}{3}(a_2-3)\wedge\epsilon)}N^{-1}\delta_n(\log N)^{C+5}).\tag{2.56}$$

Therefore, by (2.53), (2.54) and (2.56),

$$\begin{aligned}
&\mathbb{P}\left(\tilde{\tau}^n_{i,j_1},\tilde{\tau}^n_{i,j_2}\in(t,t+\delta_n]\middle|\mathcal{F}_{t_k}\right)\\
&=\mathcal{O}(n^{9/5}N^{-2}\delta_n(\log N)^{2C})+\mathcal{O}(n^{1-\frac{1}{2}(a_2-3)}N^{-1}\delta_n)+\mathcal{O}(n^{1-(\frac{1}{3}(a_2-3)\wedge\epsilon)}N^{-1}\delta_n(\log N)^{C+5})\\
&=\mathcal{O}(n^{1-\frac{1}{2}(\frac{1}{3}(a_2-3)\wedge\epsilon)}N^{-1}\delta_n).
\end{aligned}$$

Hence, by (2.52) and a union bound, and since $N\ge n^3$,

$$\mathbb{P}\left(\tilde{\tau}^n_{i,j_1},\tilde{\tau}^n_{i,j_2}\in(t_k,t_{k+1}]\middle|\mathcal{F}_{t_k}\right)=\mathcal{O}(N^{-1}(\log N)^{2C+5}t_1^2)+\mathcal{O}(n^{1-\frac{1}{2}(\frac{1}{3}(a_2-3)\wedge\epsilon)}N^{-1}t_1),$$

which completes the proof of the first statement (2.19).

For the second statement (2.20), by Proposition 2.9, for $t\in\delta_n\mathbb{N}\cap[t_k+2K\log N-\delta_n,t_{k+1})$,

$$\begin{aligned}
&\mathbb{P}\left(\tilde{\tau}^n_{i_1,j_1},\tilde{\tau}^n_{i_2,j_2}\in(t,t+\delta_n]\middle|\mathcal{F}_{t_k}\right)\\
&\le\mathcal{O}(n^4N^{-2}\delta_n^2(\log N)^{2C})+\sum_{i,j\in\{i_1,i_2,j_1,j_2\},i\ne j}\mathbb{P}\left(\tilde{\tau}^n_{i_1,j_1},\tilde{\tau}^n_{i_2,j_2}\in(t,t+\delta_n],\tau^n_{i,j}\le t\middle|\mathcal{F}_{t_k}\right).
\end{aligned}$$

The second statement (2.20) then follows by the same argument as for (2.19). □

Proof of Proposition 2.4. Suppose the event E occurs. By the definition of c_0 before (2.8), we can take $\epsilon>0$ sufficiently small that $2(1+\frac{1}{4}(1-\alpha))(1-2\epsilon)(\frac{1}{2}-c_0)>1$. For $t\in\delta_n\mathbb{N}_0\cap[0,T_n^-]$ and $x\in I^{n,\epsilon}_{T_n-t}$, by conditioning on $\mathcal{F}_t$, and then by Proposition 2.8,

$$\begin{aligned}
&\mathbb{P}\left(\tau^n_{i,j}\in(t,t+\delta_n],\zeta^{n,i}_t=x\middle|\mathcal{F}_0\right)\\
&=\mathbb{E}\left[\mathbb{P}\left(\tau^n_{i,j}\in(t,t+\delta_n]\middle|\mathcal{F}_t\right)\mathbb{1}_{\tau^n_{i,j}>t}\mathbb{1}_{\zeta^{n,i}_t=x}\middle|\mathcal{F}_0\right]\\
&=\mathbb{E}\left[\mathcal{O}(n^2N^{-1}\delta_n g(x-\mu^n_{T_n-t})^{-1})\mathbb{1}_{\tau^n_{i,j}>t}\mathbb{1}_{|\zeta^{n,j}_t-x|<Kn^{-1}}\mathbb{1}_{\zeta^{n,i}_t=x}\middle|\mathcal{F}_0\right]\\
&=\mathcal{O}(n^2N^{-1}\delta_n g(x-\mu^n_{T_n-t})^{-1})\mathbb{P}\left(|\zeta^{n,j}_t-x|<Kn^{-1},\zeta^{n,i}_t=x,\tau^n_{i,j}>t\middle|\mathcal{F}_0\right).
\end{aligned}\tag{2.57}$$

If $t \geq \epsilon_n$, then for $y \in \frac{1}{n}\mathbb{Z}$ with $|y-x| < Kn^{-1}$, by conditioning on $\mathcal{F}_{t-\epsilon_n}$, and by (2.32) in Lemma 2.7,

$$\begin{aligned}
&\mathbb{P}\left(\zeta^{n,j}_t = y, \zeta^{n,i}_t = x, \tau^n_{i,j} > t \Big| \mathcal{F}_0\right) \\
&= \mathbb{E}\left[\mathbb{P}\left(\zeta^{n,j}_t = y, \zeta^{n,i}_t = x, \tau^n_{i,j} > t \Big| \mathcal{F}_{t-\epsilon_n}\right) \mathbb{1}_{\tau^n_{i,j} > t-\epsilon_n} \mathbb{1}_{|\zeta^{n,j}_{t-\epsilon_n} - y| \leq 1} \mathbb{1}_{|\zeta^{n,i}_{t-\epsilon_n} - x| \leq 1} \Big| \mathcal{F}_0\right] \\
&\leq 2n^{-2}\epsilon_n^{-2} \mathbb{P}\left(|\zeta^{n,j}_{t-\epsilon_n} - x| \leq 2, |\zeta^{n,i}_{t-\epsilon_n} - x| \leq 1, \tau^n_{i,j} > t - \epsilon_n \Big| \mathcal{F}_0\right),
\end{aligned} \tag{2.58}$$

for n sufficiently large, by (2.31) in Lemma 2.7. For $s \geq 0$, let

$$i^{n,-}_s = \tfrac{1}{n}\mathbb{Z} \cap [\mu^n_s + D^-_n, \mu^n_s - \tfrac{1}{64}\alpha d_n] \quad \text{and} \quad i^{n,+}_s = \tfrac{1}{n}\mathbb{Z} \cap [\mu^n_s + \tfrac{1}{64}\alpha d_n, \mu^n_s - (1-\epsilon)D^+_n].$$

Suppose $x \in i^{n,+}_{T_n - t}$. Since $x \leq \mu^n_{T_n-t} + (1-\epsilon)D^+_n$, if $t \geq K \log N + \epsilon_n$ then by (2.23) in Proposition 2.5, and the definition of the event E_1 in (2.10), for n sufficiently large,

$$\mathbb{P}\left(\zeta^{n,j}_{t-\epsilon_n} \geq x-2, \zeta^{n,i}_{t-\epsilon_n} \geq x-1, \tau^n_{i,j} > t-\epsilon_n \Big| \mathcal{F}_0\right) \leq (\log N)^7 e^{-2(1+\frac{1}{4}(1-\alpha))\kappa(x-3-\mu^n_{T_n-t+\epsilon_n})}.$$

Therefore, by (2.57) and (2.58), if $t \geq K \log N + \epsilon_n$ and $x \in i^{n,+}_{T_n-t}$,

$$\begin{aligned}
&\mathbb{P}\left(\tau^n_{i,j} \in (t, t+\delta_n], \zeta^{n,i}_t = x \Big| \mathcal{F}_0\right) \\
&\leq \mathcal{O}(n^2 N^{-1} \delta_n g(x - \mu^n_{T_n-t})^{-1}) \cdot 4Kn^{-2}\epsilon_n^{-2} \cdot (\log N)^7 e^{-2(1+\frac{1}{4}(1-\alpha))\kappa(x-3-\mu^n_{T_n-t+\epsilon_n})} \\
&= \mathcal{O}\left((\log N)^{11} N^{-1} \delta_n e^{-(1+\frac{1}{2}(1-\alpha))\kappa(x-\mu^n_{T_n-t})}\right)
\end{aligned} \tag{2.59}$$

by the definition of the event E_1 in (2.10), and since $g(z)^{-1} \leq 2e^{\kappa z}$ for $z \geq 0$. By (2.57) and (2.58), if $t \geq \epsilon_n$ and $x \in i^{n,-}_{T_n-t}$,

$$\mathbb{P}\left(\tau^n_{i,j} \in (t, t+\delta_n], \zeta^{n,i}_t = x \Big| \mathcal{F}_0\right) = \mathcal{O}(n^2 N^{-1} \delta_n) \cdot 4Kn^{-2}\epsilon_n^{-2} \mathbb{P}\left(|\zeta^{n,i}_{t-\epsilon_n} - x| \leq 1 \Big| \mathcal{F}_0\right).$$

Therefore, if $t \geq K \log N + \epsilon_n$,

$$\begin{aligned}
\mathbb{P}\left(\tau^n_{i,j} \in (t, t+\delta_n], \zeta^{n,i}_t \in i^{n,-}_{T_n-t} \Big| \mathcal{F}_0\right) &\leq \mathcal{O}(N^{-1}\delta_n \epsilon_n^{-2}) \sum_{x \in i^{n,-}_{T_n-t}} \mathbb{P}\left(|\zeta^{n,i}_{t-\epsilon_n} - x| \leq 1 \Big| \mathcal{F}_0\right) \\
&= \mathcal{O}(nN^{-1}\delta_n \epsilon_n^{-2} (\log N)^{2-2^{-9}\alpha^2 C})
\end{aligned}$$

by (2.28) in Proposition 2.6 and by the definition of the event E_1. By (2.59), we now have that for $t \in \delta_n \mathbb{N} \cap [K \log N + \epsilon_n, T^-_n]$,

$$\begin{aligned}
&\mathbb{P}\left(\tau^n_{i,j} \in (t, t+\delta_n], |\tilde{\zeta}^{n,i}_t| \geq \tfrac{1}{64}\alpha d_n, \zeta^{n,i}_t \in I^{n,\epsilon}_{T_n-t} \Big| \mathcal{F}_0\right) \\
&= \mathcal{O}(nN^{-1}\delta_n (\log N)^{6-2^{-9}\alpha^2 C}) + \mathcal{O}(N^{-1}\delta_n (\log N)^{11}) \sum_{x \in i^{n,+}_{T_n-t}} e^{-(1+\frac{1}{2}(1-\alpha))\kappa(x-\mu^n_{T_n-t})} \\
&= \mathcal{O}(nN^{-1}\delta_n (\log N)^{11-2^{-9}\alpha^2 C}).
\end{aligned} \tag{2.60}$$

For $t \in \delta_n \mathbb{N} \cap [\epsilon_n, T^-_n]$ and $x \in \frac{1}{n}\mathbb{Z}$ with $|x - \mu^n_{T_n-t}| \leq \frac{1}{64}\alpha d_n$, by (2.57) and (2.58),

$$\begin{aligned}
\mathbb{P}\left(\tau^n_{i,j} \in (t, t+\delta_n], \zeta^{n,i}_t = x \Big| \mathcal{F}_0\right) &\leq \mathcal{O}(n^2 N^{-1} \delta_n g(\tfrac{1}{64}\alpha d_n)^{-1}) \cdot 4K\epsilon_n^{-2} n^{-2} \\
&= \mathcal{O}(N^{-1}\delta_n (\log N)^{4+\frac{1}{64}\alpha C}).
\end{aligned}$$

Therefore, by (2.60) and since we chose $C > 2^{13}\alpha^{-2}$, for $t \in \delta_n \mathbb{N} \cap [K \log N + \epsilon_n, T^-_n]$,

$$\mathbb{P}\left(\tau^n_{i,j} \in (t, t+\delta_n], \zeta^{n,i}_t \in I^{n,\epsilon}_{T_n-t} \Big| \mathcal{F}_0\right) = \mathcal{O}(nN^{-1}\delta_n d_n (\log N)^{4+\frac{1}{64}\alpha C}). \tag{2.61}$$

Now note that for any $t \in \delta_n \mathbb{N}_0 \cap [0, T_n^-]$,

$$\begin{aligned}\mathbb{P}\left(\tau^n_{i,j} \in (t, t+\delta_n], \zeta^{n,i}_t \in I^{n,\epsilon}_{T_n - t} \middle| \mathcal{F}_0\right) &= \mathbb{E}\left[\mathbb{P}\left(\tau^n_{i,j} \in (t, t+\delta_n] \middle| \mathcal{F}_t\right) \mathbb{1}_{\zeta^{n,i}_t \in I^{n,\epsilon}_{T_n-t}} \middle| \mathcal{F}_0\right] \\ &= \mathcal{O}(n^2 N^{-1} \delta_n g(D_n^+)^{-1}) \end{aligned} \tag{2.62}$$

by Proposition 2.8. Finally, by (2.32) in Lemma 2.7 and the definition of the event E_1 in (2.10), and then by (2.23) and (2.25) in Proposition 2.5 and (2.27) in Proposition 2.6, for n sufficiently large,

$$\begin{aligned}&\mathbb{P}\left(\exists t \in \delta_n \mathbb{N}_0 \cap [0, Nn^{-1}\log N] : \zeta^{n,i}_t \wedge \zeta^{n,j}_t \notin I^{n,\epsilon}_{T_n - t}, \tau^n_{i,j} > t \middle| \mathcal{F}_0\right) \\ &\le \mathbb{P}\left(\exists t \in t^* \mathbb{N}_0 \cap [0, Nn^{-1}\log N] : \tilde\zeta^{n,i}_t \wedge \tilde\zeta^{n,j}_t \ge (1-2\epsilon) D_n^+, \tau^n_{i,j} > t \middle| \mathcal{F}_0\right) \\ &\quad + \mathbb{P}\left(\exists t \in \delta_n \mathbb{N}_0 \cap [0, Nn^{-1}\log N] : \tilde\zeta^{n,i}_t \wedge \tilde\zeta^{n,j}_t \le D_n^- \middle| \mathcal{F}_0\right) \\ &\le ((t^*)^{-1} Nn^{-1} \log N + 1)(\log N)^7 e^{2(1+\frac14(1-\alpha))\kappa(K_0 - (1-2\epsilon)D_n^+ - 1)} + 2N^{-1} \\ &\le N^{-\epsilon'} \end{aligned} \tag{2.63}$$

for some $\epsilon' > 0$, where the last inequality follows since we chose $\epsilon > 0$ sufficiently small that $2(1+\frac14(1-\alpha))(1-2\epsilon)(\frac12 - c_0) > 1$ and since $\kappa D_n^+ = (1/2 - c_0)\log(N/n)$. Hence by a union bound, and then by (2.63), (2.62), (2.61) and (2.60),

$$\begin{aligned}&\mathbb{P}\left(\{\tau^n_{i,j} \neq \tilde\tau^n_{i,j}\} \cap \{\tau^n_{i,j} \le Nn^{-1}\log N\} \middle| \mathcal{F}_0\right) \\ &\le \mathbb{P}\left(\exists t \in \delta_n \mathbb{N}_0 \cap [0, Nn^{-1}\log N] : \zeta^{n,i}_t \wedge \zeta^{n,j}_t \notin I^{n,\epsilon}_{T_n - t}, \tau^n_{i,j} > t \middle| \mathcal{F}_0\right) \\ &\quad + \sum_{\{k \in \mathbb{N}_0 : t_k \le Nn^{-1}\log N\}} \sum_{t \in \delta_n \mathbb{N}_0 \cap [t_k, t_k + 2K\log N), i' \in \{i,j\}} \mathbb{P}\left(\tau^n_{i,j} \in (t, t+\delta_n], \zeta^{n,i'}_t \in I^{n,\epsilon}_{T_n - t} \middle| \mathcal{F}_0\right) \\ &\quad + \sum_{t \in \delta_n \mathbb{N} \cap [2K\log N, Nn^{-1}\log N], i' \in \{i,j\}} \mathbb{P}\left(\tau^n_{i,j} \in (t, t+\delta_n], |\tilde\zeta^{n,i'}_t| \ge \tfrac{1}{64}\alpha d_n, \zeta^{n,i'}_t \in I^{n,\epsilon}_{T_n - t} \middle| \mathcal{F}_0\right) \\ &\le N^{-\epsilon'} + \mathcal{O}(n^2 N^{-1} g(D_n^+)^{-1} \log N) + \mathcal{O}(nN^{-1} d_n (\log N)^{4 + \frac{1}{64}\alpha C} \cdot Nn^{-1} (\log N)^{2-C}) \\ &\qquad + \mathcal{O}(nN^{-1}(\log N)^{11 - 2^{-9}\alpha^2 C} \cdot Nn^{-1}\log N) \\ &\le \tfrac12 (\log N)^{-2} \end{aligned} \tag{2.64}$$

for n sufficiently large, where the last inequality follows since we chose $C > 2^{13}\alpha^{-2}$ and so $2^{-9}\alpha^2 C - 12 > 2$ and $\frac12 C - 6 > 2$, and since $g(D_n^+)^{-1} \le 2e^{\kappa D_n^+} = \mathcal{O}\left((\frac{N}{n})^{1/2 - c_0}\right)$ and $N \ge n^3$. By a union bound and Proposition 2.2, for n sufficiently large,

$$\begin{aligned}&\mathbb{P}\left(\tau^n_{i,j} > Nn^{-1}\log N \middle| \mathcal{F}_0\right) \\ &\le \mathbb{P}\left(\exists t \in \delta_n \mathbb{N}_0 \cap [0, Nn^{-1}\log N] : \zeta^{n,i}_t \wedge \zeta^{n,j}_t \notin I^{n,\epsilon}_{T_n - t}, \tau^n_{i,j} > t \middle| \mathcal{F}_0\right) \\ &\quad + (1 - \tfrac12 \beta_n)^{\lfloor (t_1)^{-1} Nn^{-1}\log N \rfloor} \\ &\le \tfrac12 (\log N)^{-2}, \end{aligned}$$

for n sufficiently large, by (2.63) and the definition of β_n in (2.18). By (2.63) and (2.64), this completes the proof. □

2.2 Proof of Proposition 2.5

Throughout the rest of Section 2, we assume for some $a_1 > 1$, $N \ge n^{a_1}$ for n sufficiently large. We need two preliminary lemmas for the proof of Proposition 2.5. The first is an easy consequence of the definition of the event E_2'.

Lemma 2.10. *For n sufficiently large, on the event $E_1 \cap E_2'$, for $t \in \delta_n \mathbb{N}_0 \cap [0, T_n^-]$, $i,j \in [k_0]$ and $\ell_1, \ell_2 \in \frac{1}{n}\mathbb{Z} \cap [K, D_n^+]$, if $\zeta_t^{n,i}, \zeta_t^{n,j} \in I_{T_n - t}^n$,*

$$\mathbb{P}\left(\tilde{\zeta}_{t+t^*}^{n,i} \geq \ell_1, \tilde{\zeta}_{t+t^*}^{n,j} \geq \ell_2 \middle| \mathcal{F}_t \right) \mathbb{1}_{\tau_{i,j}^n > t} \leq c_1 e^{-(1+\frac{1}{2}(1-\alpha))\kappa(\ell_1 + 1 - (\tilde{\zeta}_t^{n,i} \vee K) + \ell_2 + 1 - (\tilde{\zeta}_t^{n,j} \vee K))}$$

$$\textit{and} \qquad \mathbb{P}\left(\tilde{\zeta}_{t+t^*}^{n,i} \geq \ell_1 \middle| \mathcal{F}_t \right) \leq c_1 e^{-(1+\frac{1}{2}(1-\alpha))\kappa(\ell_1 + 1 - (\tilde{\zeta}_t^{n,i} \vee K))}.$$

Proof. Write $t' = T_n - (t + t^*)$. By the definition of $q^{n,+}$ in (2.3), and the definition of $\tilde{\zeta}^{n,i}$ and $\tilde{\zeta}^{n,j}$ in (2.15), for $\ell_1, \ell_2 \in \frac{1}{n}\mathbb{Z}$, if $\tau_{i,j}^n > t$,

$$\mathbb{P}\left(\tilde{\zeta}_{t+t^*}^{n,i} \geq \ell_1, \tilde{\zeta}_{t+t^*}^{n,j} \geq \ell_2 \middle| \mathcal{F}_t \right) \leq \frac{q_{t',t'+t^*}^{n,+}(\ell_1 + \mu_{t'}^n, \zeta_t^{n,i})}{p_{t'+t^*}^n(\zeta_t^{n,i})} \frac{q_{t',t'+t^*}^{n,+}(\ell_2 + \mu_{t'}^n, \zeta_t^{n,j})}{p_{t'+t^*}^n(\zeta_t^{n,j}) - N^{-1} \mathbb{1}_{\zeta_t^{n,j} = \zeta_t^{n,i}}}. \tag{2.65}$$

By the definition of the event E_2' in (2.11), for $\ell \in I_{t'}^n$ and $z \in I_{t'+t^*}^n$ with $\ell - \mu_{t'}^n \geq K$, the event $A_{t'}^{(2)}(\ell, z)$ occurs, and so

$$\frac{q_{t',t'+t^*}^{n,+}(\ell, z)}{p_{t'+t^*}^n(z)} \leq c_1 e^{-(1+\frac{1}{2}(1-\alpha))\kappa(\ell - (z - \nu t^*) \vee (\mu_{t'}^n + K) + 2)}.$$

Note that by the definition of the event E_1 in (2.10), if $\zeta_t^{n,j} \in I_{t'+t^*}^n$ then $p_{t'+t^*}^n(\zeta_t^{n,j}) \geq \frac{1}{10}\left(\frac{n}{N}\right)^{1/2}$. Therefore by (2.65), if $\tau_{i,j}^n > t$ and $\zeta_t^{n,i}, \zeta_t^{n,j} \in I_{T_n - t}^n$, for $\ell_1, \ell_2 \in \frac{1}{n}\mathbb{Z} \cap [K, D_n^+]$,

$$\begin{aligned}
&\mathbb{P}\left(\tilde{\zeta}_{t+t^*}^{n,i} \geq \ell_1, \tilde{\zeta}_{t+t^*}^{n,j} \geq \ell_2 \middle| \mathcal{F}_t \right) \\
&\leq (1 + \mathcal{O}(N^{-1/2})) \\
&\quad \cdot c_1^2 e^{-(1+\frac{1}{2}(1-\alpha))\kappa((\ell_1 + \mu_{t'}^n) - (\zeta_t^{n,i} - \nu t^*) \vee (\mu_{t'}^n + K) + 2 + (\ell_2 + \mu_{t'}^n) - (\zeta_t^{n,j} - \nu t^*) \vee (\mu_{t'}^n + K) + 2)} \\
&\leq (1 + \mathcal{O}(N^{-1/2})) c_1^2 e^{-(1+\frac{1}{2}(1-\alpha))\kappa((\ell_1 - \tilde{\zeta}_t^{n,i} \vee K) - t^* e^{-(\log N)^{c_2}} + 2 + (\ell_2 - \tilde{\zeta}_t^{n,j} \vee K) - t^* e^{-(\log N)^{c_2}} + 2)},
\end{aligned} \tag{2.66}$$

since, by the definition of the event E_1 in (2.10), $|(\mu_{t'}^n + \nu t^*) - \mu_{T_n - t}^n| \leq t^* e^{-(\log N)^{c_2}}$. Since $c_1 < 1$ (by our choice of c_1 in (2.14)), the first statement follows by taking n sufficiently large. The second statement follows by the same argument. □

We now use Lemma 2.10 and an inductive argument to prove the following result.

Lemma 2.11. *For n sufficiently large, the following holds. For $t \in \delta_n \mathbb{N}_0 \cap [0, T_n^-]$ and $k \in [k_0]$, let*

$$\tau_t^{+,k} = \inf\left\{ s \geq t : s - t \in t^* \mathbb{N}_0, \tilde{\zeta}_s^{n,k} \geq D_n^+ \right\}. \tag{2.67}$$

Take $i,j \in [k_0]$ and let $\tau_t^+ = \tau_t^{+,i} \wedge \tau_t^{+,j} \wedge \tau_{i,j}^n$. On the event $E_1 \cap E_2'$, for $s \in [0, T_n^-]$ with $s - t \in t^* \mathbb{N}_0$, for $\ell_1, \ell_2 \in \mathbb{N} \cap [K, D_n^+]$,

$$\mathbb{P}\left(\tilde{\zeta}_s^{n,i} \geq \ell_1, \tilde{\zeta}_s^{n,j} \geq \ell_2, \tau_t^+ \geq s \middle| \mathcal{F}_t \right) \leq e^{(1+\frac{1}{4}(1-\alpha))\kappa(\tilde{\zeta}_t^{n,i} \vee K - \ell_1 + \tilde{\zeta}_t^{n,j} \vee K - \ell_2)} \tag{2.68}$$

$$\textit{and for } i' \in \{i,j\}, \quad \mathbb{P}\left(\tilde{\zeta}_s^{n,i'} \geq \ell_1, \tau_t^{+,i'} \geq s \middle| \mathcal{F}_t \right) \leq e^{(1+\frac{1}{4}(1-\alpha))\kappa(\tilde{\zeta}_t^{n,i'} \vee K - \ell_1)}. \tag{2.69}$$

Proof. Let $\lambda = \frac{1}{4}(1-\alpha)$, and recall from (2.14) that we chose $c_1 > 0$ sufficiently small that

$$\begin{aligned}
c_1((e^{\lambda\kappa} - 1)^{-1} e^{\lambda\kappa} + e^{-(1+\lambda)\kappa}(1 - e^{-(1+\lambda)\kappa})^{-1})^2 + e^{-2(1+\lambda)\kappa} < 1 \\
\text{and} \quad c_1(e^{\lambda\kappa} - 1)^{-1} e^{\lambda\kappa} + e^{-(1+\lambda)\kappa} < 1.
\end{aligned} \tag{2.70}$$

The proof is by induction. Take $t' \in [0, T_n^-]$ with $t'-t \in t^*\mathbb{N}_0$, and suppose (2.68) and (2.69) hold for $s = t'$. Let $A = e^{(1+\lambda)\kappa(\tilde{\zeta}_t^{n,i} \vee K + \tilde{\zeta}_t^{n,j} \vee K)}$. Note that by (2.32) in Lemma 2.7, if $\tau_t^+ > t'$ then $\zeta_{t'}^{n,i}, \zeta_{t'}^{n,j} \in I_{T_n - t'}^n$. For $\ell_1, \ell_2 \in \mathbb{N} \cap [K, D_n^+]$, let $J_{\ell_1,\ell_2} = \{(k_1,k_2) : k_1, k_2 \in \mathbb{N} \cap (K, D_n^+], k_1 \le \ell_1 \text{ or } k_2 \le \ell_2\}$. Then by conditioning on $\mathcal{F}_{t'}$ and applying Lemma 2.10 and a union bound,

$$
\begin{aligned}
&\mathbb{P}\left(\tilde{\zeta}_{t'+t^*}^{n,i} \ge \ell_1, \tilde{\zeta}_{t'+t^*}^{n,j} \ge \ell_2, \tau_t^+ \ge t'+t^* \middle| \mathcal{F}_t\right) \\
&\le \sum_{(k_1,k_2)\in J_{\ell_1,\ell_2}} c_1 e^{-(1+2\lambda)\kappa((\ell_1-k_1)\vee 0 + (\ell_2-k_2)\vee 0)} \\
&\qquad\qquad\qquad \cdot \mathbb{P}\left(\tilde{\zeta}_{t'}^{n,i} \in [k_1, k_1+1), \tilde{\zeta}_{t'}^{n,j} \in [k_2, k_2+1), \tau_t^+ > t' \middle| \mathcal{F}_t\right) \\
&\quad + \sum_{k \in \mathbb{N}\cap(K,D_n^+]} \Big(c_1 e^{-(1+2\lambda)\kappa((\ell_1-k)\vee 0 + \ell_2 - K)} \mathbb{P}\left(\tilde{\zeta}_{t'}^{n,i} \in [k,k+1), \tilde{\zeta}_{t'}^{n,j} \le K+1, \tau_t^{+,i} > t' \middle| \mathcal{F}_t\right) \\
&\qquad\qquad + c_1 e^{-(1+2\lambda)\kappa((\ell_2-k)\vee 0 + \ell_1 - K)} \mathbb{P}\left(\tilde{\zeta}_{t'}^{n,j} \in [k,k+1), \tilde{\zeta}_{t'}^{n,i} \le K+1, \tau_t^{+,j} > t' \middle| \mathcal{F}_t\right)\Big) \\
&\quad + c_1 e^{-(1+2\lambda)\kappa(\ell_1 - K + \ell_2 - K)} + \mathbb{P}\left(\tilde{\zeta}_{t'}^{n,i} \ge \ell_1 + 1, \tilde{\zeta}_{t'}^{n,j} \ge \ell_2+1, \tau_t^+ > t' \middle| \mathcal{F}_t\right) \\
&\le \sum_{k_1,k_2 \in \mathbb{N}\cap[K,D_n^+]} A e^{-(1+\lambda)\kappa(k_1+k_2)} c_1 e^{-(1+2\lambda)\kappa((\ell_1-k_1)\vee 0 + (\ell_2-k_2)\vee 0)} + A e^{-(1+\lambda)\kappa(\ell_1+\ell_2+2)}
\end{aligned}
$$

by the induction hypothesis and since by the definition of A, $e^{(1+\lambda)\kappa(\tilde{\zeta}_t^{n,i'} \vee K)} \le A e^{-(1+\lambda)\kappa K}$ for $i' \in \{i,j\}$ and $Ae^{-(1+\lambda)2\kappa K} \ge 1$. Therefore

$$
\begin{aligned}
&\mathbb{P}\left(\tilde{\zeta}_{t'+t^*}^{n,i} \ge \ell_1, \tilde{\zeta}_{t'+t^*}^{n,j} \ge \ell_2, \tau_t^+ \ge t'+t^* \middle| \mathcal{F}_t\right) \\
&\le Ac_1 \left(\sum_{k_1=K}^{\ell_1} e^{-(1+\lambda)\kappa k_1} e^{-(1+2\lambda)\kappa(\ell_1-k_1)} + \sum_{k_1=\ell_1+1}^{\lfloor D_n^+ \rfloor} e^{-(1+\lambda)\kappa k_1} \right) \\
&\qquad \cdot \left(\sum_{k_2=K}^{\ell_2} e^{-(1+\lambda)\kappa k_2} e^{-(1+2\lambda)\kappa(\ell_2-k_2)} + \sum_{k_2=\ell_2+1}^{\lfloor D_n^+ \rfloor} e^{-(1+\lambda)\kappa k_2} \right) + Ae^{-(1+\lambda)\kappa(\ell_1+\ell_2+2)}.
\end{aligned}
\tag{2.71}
$$

Note that

$$
\begin{aligned}
\sum_{k_1=K}^{\ell_1} e^{-(1+\lambda)\kappa k_1} e^{-(1+2\lambda)\kappa(\ell_1-k_1)} < \sum_{k_1=0}^{\ell_1} e^{-(1+2\lambda)\kappa\ell_1} e^{\lambda\kappa k_1} &< e^{-(1+2\lambda)\kappa\ell_1}(e^{\lambda\kappa}-1)^{-1} e^{\lambda\kappa(\ell_1+1)} \\
&= (e^{\lambda\kappa}-1)^{-1} e^{\lambda\kappa} e^{-(1+\lambda)\kappa\ell_1}.
\end{aligned}
$$

Hence, since $\sum_{k_1=\ell_1+1}^{\lfloor D_n^+ \rfloor} e^{-(1+\lambda)\kappa k_1} < (1-e^{-(1+\lambda)\kappa})^{-1} e^{-(1+\lambda)\kappa(\ell_1+1)}$, substituting into (2.71),

$$
\begin{aligned}
&\mathbb{P}\left(\tilde{\zeta}_{t'+t^*}^{n,i} \ge \ell_1, \tilde{\zeta}_{t'+t^*}^{n,j} \ge \ell_2, \tau_t^+ \ge t'+t^* \middle| \mathcal{F}_t\right) \\
&\le Ae^{-(1+\lambda)\kappa(\ell_1+\ell_2)} \left(c_1((e^{\lambda\kappa}-1)^{-1} e^{\lambda\kappa} + e^{-(1+\lambda)\kappa}(1-e^{-(1+\lambda)\kappa})^{-1})^2 + e^{-2(1+\lambda)\kappa} \right) \\
&\le Ae^{-(1+\lambda)\kappa(\ell_1+\ell_2)}
\end{aligned}
$$

by (2.70). Similarly, letting $A_1 = e^{(1+\lambda)\kappa(\tilde{\zeta}_t^{n,i} \vee K)}$, for $\ell \in \mathbb{N} \cap [K, D_n^+]$, by Lemma 2.10 and a union bound,

$$\begin{aligned}
&\mathbb{P}\left(\tilde{\zeta}^{n,i}_{t'+t^*} \geq \ell, \tau^{+,i}_t \geq t'+t^* \Big| \mathcal{F}_t\right) \\
&\leq \sum_{k \in \mathbb{N}\cap(K,\ell]} c_1 e^{-(1+2\lambda)\kappa(\ell-k)} \mathbb{P}\left(\tilde{\zeta}^{n,i}_{t'} \in [k,k+1), \tau^{+,i}_t > t' \Big| \mathcal{F}_t\right) \\
&\qquad + c_1 e^{-(1+2\lambda)\kappa(\ell-K)} + \mathbb{P}\left(\tilde{\zeta}^{n,i}_{t'} \geq \ell+1, \tau^{+,i}_t > t' \Big| \mathcal{F}_t\right) \\
&\leq \sum_{k \in \mathbb{N}\cap[K,\ell]} c_1 e^{-(1+2\lambda)\kappa(\ell-k)} A_1 e^{-(1+\lambda)\kappa k} + A_1 e^{-(1+\lambda)\kappa(\ell+1)}
\end{aligned}$$

by the induction hypothesis and since $A_1 e^{-(1+\lambda)\kappa K} \geq 1$. Hence

$$\begin{aligned}
\mathbb{P}\left(\tilde{\zeta}^{n,i}_{t'+t^*} \geq \ell, \tau^{+,i}_t \geq t'+t^* \Big| \mathcal{F}_t\right) &\leq A_1 \left(c_1 e^{-(1+2\lambda)\kappa\ell}(e^{\lambda\kappa}-1)^{-1} e^{\lambda\kappa(\ell+1)} + e^{-(1+\lambda)\kappa(\ell+1)}\right) \\
&= A_1 e^{-(1+\lambda)\kappa\ell}(c_1(e^{\lambda\kappa}-1)^{-1}e^{\lambda\kappa} + e^{-(1+\lambda)\kappa}) \\
&\leq A_1 e^{-(1+\lambda)\kappa\ell}
\end{aligned}$$

by (2.70). By the same argument, $\mathbb{P}\left(\tilde{\zeta}^{n,j}_{t'+t^*} \geq \ell, \tau^{+,j}_t \geq t'+t^* \Big| \mathcal{F}_t\right) \leq e^{(1+\lambda)\kappa(\tilde{\zeta}^{n,j}_t \vee K-\ell)}$. The result follows by induction. □

Proof of Proposition 2.5. If $t-s \geq K\log N$, for $i' \in \{i,j\}$, let

$$\sigma_{i'} = \inf\{s' : s' - (t - t^*\lfloor (t^*)^{-1} K \log N \rfloor) \in t^*\mathbb{N}_0, \tilde{\zeta}^{n,i'}_{s'} \leq K\}.$$

If instead $t-s < K\log N$ with $t-s \in t^*\mathbb{N}_0$, then let $\sigma_{i'} = s$ for $i' \in \{i,j\}$. Note that in both cases $t - \sigma_{i'} \leq K\log N$. Let $\lambda = \frac{1}{4}(1-\alpha)$.

Condition on $\mathcal{F}_{\sigma_i \vee \sigma_j}$ and suppose $\sigma_i \leq \sigma_j \leq t$. Recall the definition of $\tau^{+,i}_{\sigma_j}$ and $\tau^{+,j}_{\sigma_j}$ in (2.67). Then for n sufficiently large, for $\ell_1, \ell_2 \in \mathbb{N}\cap[K, D^+_n]$, by a union bound and Lemma 2.11,

$$\begin{aligned}
&\mathbb{P}\left(\tilde{\zeta}^{n,i}_t \geq \ell_1, \tilde{\zeta}^{n,j}_t \geq \ell_2, \tau^n_{i,j} > t \Big| \mathcal{F}_{\sigma_i \vee \sigma_j}\right) \\
&\leq e^{(1+\lambda)\kappa(\tilde{\zeta}^{n,i}_{\sigma_j} \vee K - \ell_1 + \tilde{\zeta}^{n,j}_{\sigma_j} \vee K - \ell_2)} + \mathbb{P}\left(\tilde{\zeta}^{n,i}_t \geq \ell_1, \tau^n_{i,j} > t, \tau^{+,i}_{\sigma_j} \geq t, \tau^{+,j}_{\sigma_j} < t \Big| \mathcal{F}_{\sigma_i \vee \sigma_j}\right) \\
&\quad + \mathbb{P}\left(\tilde{\zeta}^{n,j}_t \geq \ell_2, \tau^n_{i,j} > t, \tau^{+,j}_{\sigma_j} \geq t, \tau^{+,i}_{\sigma_j} < t \Big| \mathcal{F}_{\sigma_i \vee \sigma_j}\right) \\
&\quad\quad + \mathbb{P}\left(\tau^n_{i,j} > t, \tau^{+,i}_{\sigma_j} < t, \tau^{+,j}_{\sigma_j} < t \Big| \mathcal{F}_{\sigma_i \vee \sigma_j}\right).
\end{aligned} \tag{2.72}$$

We now bound the last three terms on the right hand side. Recall that we let $\tau^+_{\sigma_j} = \tau^{+,i}_{\sigma_j} \wedge \tau^{+,j}_{\sigma_j} \wedge \tau^n_{i,j}$. For $s' \in [\sigma_j, t]$ with $s' - \sigma_j \in t^*\mathbb{N}_0$, by conditioning on $\mathcal{F}_{s'}$,

$$\begin{aligned}
&\mathbb{P}\left(\tilde{\zeta}^{n,i}_t \geq \ell_1, \tau^n_{i,j} > t, \tau^{+,i}_{\sigma_j} \geq t, \tau^{+,j}_{\sigma_j} = s' \Big| \mathcal{F}_{\sigma_i \vee \sigma_j}\right) \\
&\leq \mathbb{E}\left[\mathbb{P}\left(\tilde{\zeta}^{n,i}_t \geq \ell_1, \tau^{+,i}_{s'} \geq t \Big| \mathcal{F}_{s'}\right) \mathbb{1}_{\tilde{\zeta}^{n,j}_{s'} \geq D^+_n, \tau^+_{\sigma_j} = s'} \Big| \mathcal{F}_{\sigma_i \vee \sigma_j}\right] \\
&\leq \sum_{\ell_1' = K}^{\ell_1 - 1} \mathbb{P}\left(\tilde{\zeta}^{n,i}_{s'} \in [\ell_1', \ell_1'+1), \tilde{\zeta}^{n,j}_{s'} \geq D^+_n, \tau^+_{\sigma_j} \geq s' \Big| \mathcal{F}_{\sigma_i \vee \sigma_j}\right) \cdot e^{(1+\lambda)\kappa(\ell_1'+1-\ell_1)} \\
&\quad + \mathbb{P}\left(\tilde{\zeta}^{n,i}_{s'} \leq K, \tilde{\zeta}^{n,j}_{s'} \geq D^+_n, \tau^+_{\sigma_j} \geq s' \Big| \mathcal{F}_{\sigma_i \vee \sigma_j}\right) \cdot e^{(1+\lambda)\kappa(K-\ell_1)} \\
&\quad + \mathbb{P}\left(\tilde{\zeta}^{n,i}_{s'} \geq \ell_1, \tilde{\zeta}^{n,j}_{s'} \geq D^+_n, \tau^+_{\sigma_j} \geq s' \Big| \mathcal{F}_{\sigma_i \vee \sigma_j}\right)
\end{aligned}$$

by (2.69) in Lemma 2.11. Therefore, by Lemma 2.11 again,

$$\begin{aligned}
&\mathbb{P}\left(\tilde{\zeta}^{n,i}_t \geq \ell_1, \tau^n_{i,j} > t, \tau^{+,i}_{\sigma_j} \geq t, \tau^{+,j}_{\sigma_j} = s' \middle| \mathcal{F}_{\sigma_i \vee \sigma_j}\right) \\
&\leq \sum_{\ell_1'=K}^{\ell_1} e^{(1+\lambda)\kappa(\tilde{\zeta}^{n,i}_{\sigma_j} \vee K - \ell_1' + \tilde{\zeta}^{n,j}_{\sigma_j} \vee K - \lfloor D^+_n \rfloor)} \cdot e^{(1+\lambda)\kappa(\ell_1'+1-\ell_1)} \\
&\qquad + e^{(1+\lambda)\kappa(\tilde{\zeta}^{n,j}_{\sigma_j} \vee K - \lfloor D^+_n \rfloor)} \cdot e^{(1+\lambda)\kappa(K-\ell_1)} \\
&\leq e^{(1+\lambda)\kappa(\tilde{\zeta}^{n,i}_{\sigma_j} \vee K + \tilde{\zeta}^{n,j}_{\sigma_j} \vee K)} (\ell_1 e^{-(1+\lambda)\kappa(\ell_1 + \lfloor D^+_n \rfloor - 1)} + e^{-(1+\lambda)\kappa(\ell_1 + \lfloor D^+_n \rfloor)}) \\
&\leq e^{(1+\lambda)\kappa(\tilde{\zeta}^{n,i}_{\sigma_j} \vee K + \tilde{\zeta}^{n,j}_{\sigma_j} \vee K + 1)} e^{-(1+\lambda)\kappa(\ell_1 + \lfloor D^+_n \rfloor)} (D^+_n + 1),
\end{aligned} \tag{2.73}$$

since $\ell_1 \leq D^+_n$. Therefore, for n sufficiently large, since $t - \sigma_j \leq K \log N$,

$$\begin{aligned}
&\mathbb{P}\left(\tilde{\zeta}^{n,i}_t \geq \ell_1, \tau^n_{i,j} > t, \tau^{+,i}_{\sigma_j} \geq t, \tau^{+,j}_{\sigma_j} < t \middle| \mathcal{F}_{\sigma_i \vee \sigma_j}\right) \\
&\leq e^{(1+\lambda)\kappa(\tilde{\zeta}^{n,i}_{\sigma_j} \vee K - \ell_1 + \tilde{\zeta}^{n,j}_{\sigma_j} \vee K - \lfloor D^+_n \rfloor + 1)} K\kappa^{-1} (\log N)^2,
\end{aligned} \tag{2.74}$$

and by the same argument,

$$\begin{aligned}
&\mathbb{P}\left(\tilde{\zeta}^{n,j}_t \geq \ell_2, \tau^n_{i,j} > t, \tau^{+,j}_{\sigma_j} \geq t, \tau^{+,i}_{\sigma_j} < t \middle| \mathcal{F}_{\sigma_i \vee \sigma_j}\right) \\
&\leq e^{(1+\lambda)\kappa(\tilde{\zeta}^{n,i}_{\sigma_j} \vee K - \lfloor D^+_n \rfloor + \tilde{\zeta}^{n,j}_{\sigma_j} \vee K - \ell_2 + 1)} K\kappa^{-1} (\log N)^2.
\end{aligned} \tag{2.75}$$

For the last term on the right hand side of (2.72), note that for $\sigma_j \leq s_1 \leq s_2 \leq t$ with $s_1 - \sigma_j, s_2 - \sigma_j \in t^* \mathbb{N}_0$, by the same argument as for (2.73),

$$\begin{aligned}
&\mathbb{P}\left(\tau^n_{i,j} > t, \tau^{+,i}_{\sigma_j} = s_1, \tau^{+,j}_{\sigma_j} = s_2 \middle| \mathcal{F}_{\sigma_i \vee \sigma_j}\right) \\
&\leq \mathbb{P}\left(\tau^n_{i,j} > s_2, \tau^{+,i}_{\sigma_j} = s_1, \tau^{+,j}_{\sigma_j} \geq s_2, \tilde{\zeta}^{n,j}_{s_2} \geq \lfloor D^+_n \rfloor \middle| \mathcal{F}_{\sigma_i \vee \sigma_j}\right) \\
&\leq e^{(1+\lambda)\kappa(\tilde{\zeta}^{n,i}_{\sigma_j} \vee K - \lfloor D^+_n \rfloor + \tilde{\zeta}^{n,j}_{\sigma_j} \vee K - \lfloor D^+_n \rfloor + 1)} (D^+_n + 1),
\end{aligned} \tag{2.76}$$

and by the same argument (2.76) also holds for $s_1 \geq s_2$. Hence by (2.72), (2.74) and (2.75), for n sufficiently large, if $\sigma_i \leq \sigma_j \leq t$ then for $\ell_1, \ell_2 \in \mathbb{N} \cap [K, D^+_n]$,

$$\mathbb{P}\left(\tilde{\zeta}^{n,i}_t \geq \ell_1, \tilde{\zeta}^{n,j}_t \geq \ell_2, \tau^n_{i,j} > t \middle| \mathcal{F}_{\sigma_i \vee \sigma_j}\right) \leq e^{(1+\lambda)\kappa(\tilde{\zeta}^{n,i}_{\sigma_j} \vee 0 - \ell_1 + \tilde{\zeta}^{n,j}_{\sigma_j} \vee 0 - \ell_2)} (\log N)^4. \tag{2.77}$$

By a simpler version of the same argument, for $i' \in \{i, j\}$ and $\ell \in \mathbb{N} \cap [K, D^+_n]$, if $\sigma_i \leq \sigma_j \leq t$ then

$$\begin{aligned}
&\mathbb{P}\left(\tilde{\zeta}^{n,i'}_t \geq \ell \middle| \mathcal{F}_{\sigma_i \vee \sigma_j}\right) \\
&\leq \mathbb{P}\left(\tilde{\zeta}^{n,i'}_t \geq \ell, \tau^{+,i'}_{\sigma_j} \geq t \middle| \mathcal{F}_{\sigma_i \vee \sigma_j}\right) + \sum_{s' \in [\sigma_j, t), s' - \sigma_j \in t^* \mathbb{N}_0} \mathbb{P}\left(\tilde{\zeta}^{n,i'}_{s'} \geq D^+_n, \tau^{+,i'}_{\sigma_j} \geq s' \middle| \mathcal{F}_{\sigma_i \vee \sigma_j}\right) \\
&\leq (\log N)^2 e^{(1+\lambda)\kappa(\tilde{\zeta}^{n,i'}_{\sigma_j} \vee 0 - \ell)}
\end{aligned} \tag{2.78}$$

for n sufficiently large, by (2.69) in Lemma 2.11. Since we let $\sigma_i = \sigma_j = s$ in the case $t - s < K \log N$, this completes the proof of (2.25) and (2.26).

From now on, assume $t - s \geq K \log N$. Condition on $\mathcal{F}_{\sigma_i \wedge \sigma_j}$ and suppose $\sigma_i \wedge \sigma_j =$

$\sigma_i \le t$; then

$$\begin{aligned}
&\mathbb{E}\left[e^{(1+\lambda)\kappa(\tilde{\zeta}^{n,i}_{\sigma_j}\vee 0)}\mathbb{1}_{\tau^{+,i}_{\sigma_i}>\sigma_j}\mathbb{1}_{\sigma_j\le t}\middle|\mathcal{F}_{\sigma_i\wedge\sigma_j}\right]\\
&\le e^{(1+\lambda)\kappa K}+\sum_{\ell=K}^{\lfloor D_n^+\rfloor}e^{(1+\lambda)\kappa(\ell+1)}\sum_{s'-\sigma_i\in t^*\mathbb{N}_0,\, s'\le t}\mathbb{P}\left(\tilde{\zeta}^{n,i}_{s'}\in[\ell,\ell+1),\tau^{+,i}_{\sigma_i}\ge s'\middle|\mathcal{F}_{\sigma_i\wedge\sigma_j}\right)\\
&\le e^{(1+\lambda)\kappa K}+\sum_{\ell=K}^{\lfloor D_n^+\rfloor}e^{(1+\lambda)\kappa(\ell+1)}((t^*)^{-1}K\log N+1)e^{(1+\lambda)\kappa(\tilde{\zeta}^{n,i}_{\sigma_i}\vee K-\ell)}\\
&\le e^{(1+\lambda)\kappa(1+K)}K\kappa^{-1}(\log N)^2 \qquad (2.79)
\end{aligned}$$

for n sufficiently large, where the second inequality follows by (2.69) in Lemma 2.11 and since $t-\sigma_i\le K\log N$, and the last inequality since $\tilde{\zeta}^{n,i}_{\sigma_i}\le K$. Therefore, if $\sigma_i\wedge\sigma_j=\sigma_i\le t$, by conditioning on $\mathcal{F}_{\sigma_i\vee\sigma_j}$, and then by (2.77), (2.78) and (2.79), and since $\tilde{\zeta}^{n,j}_{\sigma_j}\le K$ if $\sigma_j\le t$,

$$\begin{aligned}
&\mathbb{P}\left(\tilde{\zeta}^{n,i}_t\ge\ell_1,\tilde{\zeta}^{n,j}_t\ge\ell_2,\tau^n_{i,j}>t\middle|\mathcal{F}_{\sigma_i\wedge\sigma_j}\right)\\
&\le\mathbb{E}\left[\mathbb{P}\left(\tilde{\zeta}^{n,i}_t\ge\ell_1,\tilde{\zeta}^{n,j}_t\ge\ell_2,\tau^n_{i,j}>t\middle|\mathcal{F}_{\sigma_i\vee\sigma_j}\right)\mathbb{1}_{\sigma_j\le t}(\mathbb{1}_{\tau^{+,i}_{\sigma_i}>\sigma_j}+\mathbb{1}_{\tau^{+,i}_{\sigma_i}\le\sigma_j})\middle|\mathcal{F}_{\sigma_i\wedge\sigma_j}\right]\\
&\quad+\mathbb{P}\left(\sigma_j>t\middle|\mathcal{F}_{\sigma_i\wedge\sigma_j}\right)\\
&\le e^{(1+\lambda)\kappa(1+2K)}K\kappa^{-1}(\log N)^2\cdot(\log N)^4e^{-(1+\lambda)\kappa(\ell_1+\ell_2)}\\
&\quad+\mathbb{E}\left[(\log N)^2e^{(1+\lambda)\kappa(K-\ell_2)}\mathbb{1}_{\sigma_j\le t}\mathbb{1}_{\tau^{+,i}_{\sigma_i}\le\sigma_j}\middle|\mathcal{F}_{\sigma_i\wedge\sigma_j}\right]+\mathbb{P}\left(\sigma_j>t\middle|\mathcal{F}_{\sigma_i\wedge\sigma_j}\right). \qquad (2.80)
\end{aligned}$$

By (2.69) in Lemma 2.11, if $\sigma_i\wedge\sigma_j=\sigma_i\le t$, then since $\tilde{\zeta}^{n,i}_{\sigma_i}\le K$,

$$\begin{aligned}
\mathbb{P}\left(\tau^{+,i}_{\sigma_i}\le t\middle|\mathcal{F}_{\sigma_i\wedge\sigma_j}\right)&\le\sum_{s'\le t,\, s'-\sigma_i\in t^*\mathbb{N}_0}\mathbb{P}\left(\tau^{+,i}_{\sigma_i}\ge s',\tilde{\zeta}^{n,i}_{s'}\ge D_n^+\middle|\mathcal{F}_{\sigma_i\wedge\sigma_j}\right)\\
&\le((t^*)^{-1}K\log N+1)e^{(1+\lambda)\kappa(K-\lfloor D_n^+\rfloor)}. \qquad (2.81)
\end{aligned}$$

Hence, for n sufficiently large, by a union bound and then by (2.80) and (2.81) (using the same argument for the case $\sigma_j\le\sigma_i$),

$$\begin{aligned}
&\mathbb{P}\left(\tilde{\zeta}^{n,i}_t\ge\ell_1,\tilde{\zeta}^{n,j}_t\ge\ell_2,\tau^n_{i,j}>t\middle|\mathcal{F}_s\right)\\
&\le\mathbb{P}\left(\sigma_i\wedge\sigma_j>t\middle|\mathcal{F}_s\right)+\mathbb{E}\left[\mathbb{P}\left(\tilde{\zeta}^{n,i}_t\ge\ell_1,\tilde{\zeta}^{n,j}_t\ge\ell_2,\tau^n_{i,j}>t\middle|\mathcal{F}_{\sigma_i\wedge\sigma_j}\right)\mathbb{1}_{\sigma_i\wedge\sigma_j\le t}\middle|\mathcal{F}_s\right]\\
&\le\mathbb{P}\left(\sigma_i\wedge\sigma_j>t\middle|\mathcal{F}_s\right)+\mathbb{P}\left(\sigma_i\vee\sigma_j>t\middle|\mathcal{F}_s\right)+\tfrac{1}{2}(\log N)^7e^{-(1+\lambda)\kappa(\ell_1+\ell_2)} \qquad (2.82)
\end{aligned}$$

for n sufficiently large. Finally, let $t'=t-t^*\lfloor(t^*)^{-1}K\log N\rfloor\in\delta_n\mathbb{N}_0\cap[0,T_n^-]$ with $t'\ge s$, and recall the definition of $r^{n,y,\ell}_{s',s''}(\cdot)$ in (2.6). Since $(r^{n,K,t^*}_{K\log N,T_n-t'}(x))_{x\in\frac{1}{n}\mathbb{Z}}$ only depends on the Poisson processes $(\mathcal{P}^{x,i,j})_{x,i,j}$, $(\mathcal{S}^{x,i,j})_{x,i,j}$, $(\mathcal{Q}^{x,i,j,k})_{x,i,j,k}$ and $(\mathcal{R}^{x,i,y,j})_{x,y,i,j}$ in the time interval $[0,T_n-t']$, and by (2.16),

$$\mathbb{P}\left(r^{n,K,t^*}_{K\log N,T_n-t'}(x)=0\ \forall x\in\tfrac{1}{n}\mathbb{Z}\middle|\mathcal{F}_s\right)=\mathbb{P}\left(r^{n,K,t^*}_{K\log N,T_n-t'}(x)=0\ \forall x\in\tfrac{1}{n}\mathbb{Z}\middle|\mathcal{F}\right)\ge1-\left(\frac{n}{N}\right)^2$$

by the definition of the event E_4 in (2.13). By the definition of $r^{n,K,t^*}_{K\log N,T_n-t'}(x)$ in (2.6), it follows that $\mathbb{P}\left(\sigma_i\vee\sigma_j>t\middle|\mathcal{F}_s\right)\le\left(\frac{n}{N}\right)^2$. By (2.82), and since $(1+\lambda)\kappa(\ell_1+\ell_2)\le4\kappa D_n^+\le 4(1/2-c_0)\log(N/n)$ by (2.8), this completes the proof of (2.23). By a union bound and then by the same argument as in (2.78) and since $\tilde{\zeta}^{n,i}_{\sigma_i} \le K$ if $\sigma_i \le t$,

$$\begin{aligned}\mathbb{P}\left(\tilde{\zeta}^{n,i}_t \ge \ell_1 \middle| \mathcal{F}_s\right) &\le \mathbb{P}\left(\sigma_i > t \middle| \mathcal{F}_s\right) + \mathbb{E}\left[\mathbb{P}\left(\tilde{\zeta}^{n,i}_t \ge \ell_1 \middle| \mathcal{F}_{\sigma_i}\right) \mathbb{1}_{\sigma_i \le t} \middle| \mathcal{F}_s\right] \\ &\le \left(\frac{n}{N}\right)^2 + (\log N)^2 e^{(1+\lambda)\kappa(K-\ell_1)},\end{aligned}$$

which completes the proof. □

2.3 Proof of Proposition 2.6

We first prove two preliminary lemmas, similar to the lemmas in Section 2.2. Write $d_n' = \frac{1}{64}\alpha d_n$.

Lemma 2.12. *For n sufficiently large, on the event $E_1 \cap E_2'$, for $t \in \delta_n \mathbb{N}_0 \cap [0, T_n^-]$, $i \in [k_0]$ and $y, y' \le -\frac{1}{2} d_n'$, if $\tilde{\zeta}^{n,i}_t \ge y$ then*

$$\mathbb{P}\left(\tilde{\zeta}^{n,i}_{t+t^*} \le y' \middle| \mathcal{F}_t\right) \le c_1 e^{-\frac{1}{2}\alpha\kappa(y-y')}.$$

Proof. Suppose first that $y' \ge -N^3$. For n sufficiently large, by the definition of the event E_1 in (2.10), if $\tilde{\zeta}^{n,i}_t \ge y$ and $\zeta^{n,i}_t \in I^n_{T_n-t}$,

$$\begin{aligned}\mathbb{P}\left(\tilde{\zeta}^{n,i}_{t+t^*} \le y' \middle| \mathcal{F}_t\right) &\le \mathbb{P}\left(\zeta^{n,i}_{t+t^*} \le \mu^n_{T_n-t} - \nu t^* + 1 + y' \middle| \mathcal{F}_t\right) \\ &= \frac{q^{n,-}_{T_n-t-t^*, T_n-t}(\mu^n_{T_n-t} - \nu t^* + 1 + y', \tilde{\zeta}^{n,i}_t + \mu^n_{T_n-t})}{p^n_{T_n-t}(\tilde{\zeta}^{n,i}_t + \mu^n_{T_n-t})} \\ &\le c_1 e^{-\frac{1}{2}\alpha\kappa(y-y')}\end{aligned}$$

since the event $A^{(3)}_{T_n-t-t^*}(n^{-1}\lfloor n(\mu^n_{T_n-t} - \nu t^* + 1 + y')\rfloor, \zeta^{n,i}_t)$ occurs by the definition of the event E_2' in (2.11). If instead $y' < -N^3$ or $\zeta^{n,i}_t \notin I^n_{T_n-t}$ then by (2.32) in Lemma 2.7, $\mathbb{P}\left(\tilde{\zeta}^{n,i}_{t+t^*} \le y' \middle| \mathcal{F}_t\right) = 0$ almost surely. □

We now use Lemma 2.12 and an induction argument to prove the following result.

Lemma 2.13. *On the event $E_1 \cap E_2'$, for $t \in \delta_n \mathbb{N}_0 \cap [0, T_n^-]$, $i \in [k_0]$, $k \in \mathbb{N}_0$ and $t' \in [0, T_n^-]$ with $t' - t \in t^* \mathbb{N}_0$,*

$$\mathbb{P}\left(\tilde{\zeta}^{n,i}_{t'} \le -\tfrac{1}{2} d_n' - k \middle| \mathcal{F}_t\right) \le e^{-\frac{1}{4}\alpha\kappa((\frac{1}{2}d_n' + \tilde{\zeta}^{n,i}_t)\wedge 0 + k)}. \tag{2.83}$$

Proof. Recall from (2.14) that we chose $c_1 > 0$ sufficiently small that

$$c_1 + c_1 e^{3\alpha\kappa/4}(e^{\alpha\kappa/4} - 1)^{-1} + e^{-\alpha\kappa/4} < 1. \tag{2.84}$$

Let $A = e^{-\frac{1}{4}\alpha\kappa((\frac{1}{2}d_n' + \tilde{\zeta}^{n,i}_t)\wedge 0)}$. Suppose, for an induction argument, that for some $t' \ge t$ with $t' \in [0, T_n^-]$ and $t' - t \in t^* \mathbb{N}_0$, (2.83) holds for all $k \in \mathbb{N}_0$. Then by Lemma 2.12, for $k \in \mathbb{N}_0$,

$$\begin{aligned}\mathbb{P}\left(\tilde{\zeta}^{n,i}_{t'+t^*} \le -\tfrac{1}{2} d_n' - k \middle| \mathcal{F}_t\right) &\le \sum_{k'=0}^{k} \mathbb{P}\left(\tilde{\zeta}^{n,i}_{t'} \in (-\tfrac{1}{2}d_n' - k' - 1, -\tfrac{1}{2}d_n' - k'] \middle| \mathcal{F}_t\right) c_1 e^{-\frac{1}{2}\alpha\kappa(k-k'-1)} \\ &\quad + \mathbb{P}\left(\tilde{\zeta}^{n,i}_{t'} \le -\tfrac{1}{2}d_n' - k - 1 \middle| \mathcal{F}_t\right) + c_1 e^{-\frac{1}{2}\alpha\kappa k} \\ &\le \sum_{k'=0}^{k} A e^{-\frac{1}{4}\alpha\kappa k'} c_1 e^{-\frac{1}{2}\alpha\kappa(k-k'-1)} + A e^{-\frac{1}{4}\alpha\kappa(k+1)} + c_1 e^{-\frac{1}{2}\alpha\kappa k}\end{aligned}$$

by our induction hypothesis. Therefore, since $A \geq 1$,

$$\begin{aligned}\mathbb{P}\left(\tilde{\zeta}^{n,i}_{t'+t^*} \leq -\tfrac{1}{2}d'_n - k \Big| \mathcal{F}_t\right) &\leq A\left(c_1 e^{-\frac{1}{2}\alpha\kappa(k-1)} \sum_{k'=0}^{k} e^{\frac{1}{4}\alpha\kappa k'} + e^{-\frac{1}{4}\alpha\kappa(k+1)} + c_1 e^{-\frac{1}{2}\alpha\kappa k}\right)\\ &= A\left(c_1 e^{-\frac{1}{2}\alpha\kappa(k-1)} \frac{e^{\frac{1}{4}\alpha\kappa(k+1)}-1}{e^{\frac{1}{4}\alpha\kappa}-1} + e^{-\frac{1}{4}\alpha\kappa(k+1)} + c_1 e^{-\frac{1}{2}\alpha\kappa k}\right)\\ &< Ae^{-\frac{1}{4}\alpha\kappa k}\left(c_1 e^{\frac{3}{4}\alpha\kappa}(e^{\frac{1}{4}\alpha\kappa}-1)^{-1} + e^{-\frac{1}{4}\alpha\kappa} + c_1\right)\\ &\leq Ae^{-\frac{1}{4}\alpha\kappa k}\end{aligned}$$

by (2.84). The result follows by induction. □

Proof of Proposition 2.6. We begin by proving (2.27). For n sufficiently large, by (2.32) in Lemma 2.7 and then by a union bound and Lemma 2.13, and since $\tilde{\zeta}^{n,i}_0 \geq -K_0$,

$$\begin{aligned}\mathbb{P}\left(\exists t \in \delta_n \mathbb{N}_0 \cap [0, T_n^-] : \tilde{\zeta}^{n,i}_t \leq D_n^- \Big| \mathcal{F}_0\right) &\leq \mathbb{P}\left(\exists t \in t^* \mathbb{N}_0 \cap [0, T_n^-] : \tilde{\zeta}^{n,i}_t \leq \tfrac{1}{2} D_n^- \Big| \mathcal{F}_0\right)\\ &\leq ((t^*)^{-1} T_n^- + 1) e^{-\frac{1}{4}\alpha\kappa \lfloor -\frac{1}{2} D_n^- - \frac{1}{2} d'_n \rfloor}\\ &\leq N^{-1}\end{aligned}$$

for n sufficiently large, since, by (2.8), $\frac{1}{8}\alpha\kappa D_n^- = -\frac{13}{4}\log N$ and since $T_n^- \leq N^2$.

Note that the last statement (2.29) follows directly from Lemma 2.13 (since $\tilde{\zeta}^{n,i}_0 \geq -K_0$ and $\lfloor d_n - \frac{1}{2}d'_n \rfloor > \frac{1}{2} d_n$ for n sufficiently large, and by (2.4)). We now prove (2.28). Recall from (2.14) that we chose $c_1 > 0$ sufficiently small that

$$e^{-\alpha\kappa/4} + c_1(1 - e^{-\alpha\kappa/4})^{-1} < e^{-\alpha\kappa/5}. \tag{2.85}$$

Let A be a Bernoulli random variable with mean c_1 and let G be an independent geometric random variable with parameter $1 - e^{-\alpha\kappa/2}$ (with $\mathbb{P}(G \geq k) = e^{-\alpha\kappa k/2}$ for $k \in \mathbb{N}_0$). For $t' \in \delta_n \mathbb{N}_0 \cap [0, T_n^-]$, if $\tilde{\zeta}^{n,i}_{t'} \leq -\frac{1}{2}d'_n$ then by Lemma 2.12, for $k \in \mathbb{N}_0$,

$$\mathbb{P}\left(\tilde{\zeta}^{n,i}_{t'} - \tilde{\zeta}^{n,i}_{t'+t^*} \geq k \Big| \mathcal{F}_{t'}\right) \leq c_1 e^{-\frac{1}{2}\alpha\kappa k} = \mathbb{P}\left(AG - (1-A) \geq k\right).$$

Since $AG - (1-A) \geq -1$, it follows that for each $k \in \mathbb{Z}$, if $\tilde{\zeta}^{n,i}_{t'} \leq -\frac{1}{2}d'_n$ then

$$\mathbb{P}\left(\tilde{\zeta}^{n,i}_{t'} - \tilde{\zeta}^{n,i}_{t'+t^*} \geq k \Big| \mathcal{F}_{t'}\right) \leq \mathbb{P}\left(AG - (1-A) \geq k\right). \tag{2.86}$$

Let $(A_j)_{j=1}^\infty$ and $(G_j)_{j=1}^\infty$ be independent families of i.i.d. random variables with $A_1 \stackrel{d}{=} A$ and $G_1 \stackrel{d}{=} G$. Suppose $\tilde{\zeta}^{n,i}_s \geq D_n^-$ and $t - s \geq K \log N$, and take $s' \in [s, s+t^*]$ such that $t - s' \in t^* \mathbb{N}_0$. For n sufficiently large, by (2.32) in Lemma 2.7, we have $\tilde{\zeta}^{n,i}_{s'} \geq 2D_n^-$. Hence

$$\begin{aligned}\{\tilde{\zeta}^{n,i}_{s'+4\lfloor |D_n^-| \rfloor t^*} \leq -\tfrac{1}{2}d'_n\} \subseteq \{\tilde{\zeta}^{n,i}_{s'+4\lfloor |D_n^-| \rfloor t^*} \leq 0\} &\subseteq \{\tilde{\zeta}^{n,i}_{s'} - \tilde{\zeta}^{n,i}_{s'+4\lfloor |D_n^-| \rfloor t^*} \geq 2D_n^-\}\\ &= \left\{\sum_{j=1}^{4\lfloor |D_n^-| \rfloor} (\tilde{\zeta}^{n,i}_{s'+(j-1)t^*} - \tilde{\zeta}^{n,i}_{s'+jt^*}) \geq 2D_n^-\right\}.\end{aligned} \tag{2.87}$$

Then using (2.87) in the first inequality and (2.86) in the second inequality,

$$
\begin{aligned}
&\mathbb{P}\left(\tilde{\zeta}^{n,i}_{s'+\ell t^*} \leq -\tfrac{1}{2}d'_n \ \forall \ell \in \{0\}\cup[4\lfloor |D^-_n|\rfloor]\Big|\mathcal{F}_{s'}\right)\\
&\leq \mathbb{P}\Big(\tilde{\zeta}^{n,i}_{s'+\ell t^*} \leq -\tfrac{1}{2}d'_n \ \forall \ell \in \{0\}\cup[4\lfloor |D^-_n|\rfloor - 1], \sum_{j=1}^{4\lfloor |D^-_n|\rfloor} (\tilde{\zeta}^{n,i}_{s'+(j-1)t^*} - \tilde{\zeta}^{n,i}_{s'+jt^*}) \geq 2D^-_n\Big|\mathcal{F}_{s'}\Big)\\
&\leq \mathbb{P}\left(\sum_{j=1}^{4\lfloor |D^-_n|\rfloor}(A_jG_j - (1-A_j)) \geq 2D^-_n\right).
\end{aligned}
$$

By Markov's inequality,

$$
\begin{aligned}
\mathbb{P}\left(\sum_{j=1}^{4\lfloor |D^-_n|\rfloor}(A_jG_j - (1-A_j)) \geq 2D^-_n\right) &\leq e^{\frac{1}{4}\alpha\kappa\cdot 2|D^-_n|}\mathbb{E}\left[e^{\frac{1}{4}\alpha\kappa(A_1G_1-(1-A_1))}\right]^{4\lfloor |D^-_n|\rfloor}\\
&\leq e^{\frac{1}{2}\alpha\kappa|D^-_n|}\left((1-c_1)e^{-\frac{1}{4}\alpha\kappa} + c_1\frac{1-e^{-\alpha\kappa/2}}{1-e^{-\alpha\kappa/4}}\right)^{4\lfloor |D^-_n|\rfloor}\\
&\leq e^{\frac{4}{5}\alpha\kappa}e^{-\frac{3}{10}\alpha\kappa|D^-_n|}
\end{aligned}
$$

by (2.85). Therefore, since $\alpha\kappa|D^-_n| = 26\log N$ by (2.8), and since $K\log N > (4|D^-_n|+1)t^*$ for n sufficiently large, by our choice of K in Proposition 2.1,

$$
\begin{aligned}
\mathbb{P}\left(\tilde{\zeta}^{n,i}_t \leq -d_n\Big|\mathcal{F}_s\right) &\leq N^{-7} + \sum_{\ell=0}^{4\lfloor |D^-_n|\rfloor}\mathbb{E}\left[\mathbb{P}\left(\tilde{\zeta}^{n,i}_{s'+\ell t^*} \geq -\tfrac{1}{2}d'_n, \tilde{\zeta}^{n,i}_t \leq -d_n\Big|\mathcal{F}_{s'}\right)\Big|\mathcal{F}_s\right]\\
&\leq N^{-7} + \sum_{\ell=0}^{4\lfloor |D^-_n|\rfloor} e^{-\frac{1}{4}\alpha\kappa\cdot\frac{1}{2}d_n}\\
&\leq (\log N)^{2-\frac{1}{8}\alpha C}
\end{aligned}
$$

for n sufficiently large, where the second inequality follows by Lemma 2.13 and since $\lfloor d_n - \frac{1}{2}d'_n\rfloor > \frac{1}{2}d_n$, and the last inequality follows by (2.4). Since $d'_n = 2^{-6}\alpha d_n$, by the same argument, for n sufficiently large, $\mathbb{P}\left(\tilde{\zeta}^{n,i}_t \leq -d'_n + 2\Big|\mathcal{F}_s\right) \leq (\log N)^{2-2^{-9}\alpha^2 C}$. □

3 Event E_1 occurs with high probability

In this section and the following three sections, we will prove Proposition 2.1. The main result of this section (Proposition 3.1) will also imply Theorem 1.3. We begin with some notation which will be used throughout the rest of the article. For $h : \frac{1}{n}\mathbb{Z} \to \mathbb{R}$ and $x \in \frac{1}{n}\mathbb{Z}$, let

$$\nabla_n h(x) = n\left(h(x+n^{-1}) - h(x)\right)$$

and let

$$\Delta_n h(x) = n^2\left(h(x+n^{-1}) - 2h(x) + h(x-n^{-1})\right).$$

Define $f : \mathbb{R} \to \mathbb{R}$ by letting

$$f(u) = u(1-u)(2u-1+\alpha). \tag{3.1}$$

Recall the definition of the event E_1 in (2.10). In this section, we will prove the following result (along with some technical lemmas which will be used in later sections).

Proposition 3.1. *For $t \geq 0$, let $(u^n_{t,t+s})_{s\geq 0}$ denote the solution of*

$$\begin{cases}\partial_s u^n_{t,t+s} = \frac{1}{2}m\Delta_n u^n_{t,t+s} + s_0 f(u^n_{t,t+s}) & \text{for } s>0,\\ u^n_{t,t} = p^n_t.\end{cases} \tag{3.2}$$

For $c_2>0$, define the event

$$E_1' = E_1 \cap \left\{ \sup_{s\in[0,\gamma_n], x\in\frac{1}{n}\mathbb{Z}} |u^n_{t,t+s}(x) - g(x-\mu^n_t - \nu s)| \leq e^{-(\log N)^{c_2}} \ \forall t \in [\log N, N^2] \right\}. \tag{3.3}$$

Suppose for some $a_1>1$, $N\geq n^{a_1}$ for n sufficiently large. For $\ell\in\mathbb{N}$, for $b_1,c_2>0$ sufficiently small and $b_2>0$, if condition (A) *holds then for n sufficiently large,*

$$\mathbb{P}\left((E_1')^c\right) \leq \left(\frac{n}{N}\right)^{\ell}.$$

Before proving Proposition 3.1, we note that Theorem 1.3 is a trivial consequence of this result.

Proof of Theorem 1.3. By the definition of the events E_1 and E_1' in (2.10) and (3.3) respectively, on the event E_1' we have

$$\sup_{x\in\frac{1}{n}\mathbb{Z},\, t\in[\log N, N^2]} |p^n_t(x) - g(x-\mu^n_t)| \leq e^{-(\log N)^{c_2}}$$
$$\text{and}\quad |\mu^n_{t+s} - \mu^n_s - \nu s| \leq e^{-(\log N)^{c_2}} \ \forall t\in[\log N, N^2], s\in[0,1\wedge(N^2-t)].$$

Hence the result follows directly from Proposition 3.1. □

From now on in this section, we will assume for some $a_1>1$, $N\geq n^{a_1}$ for n sufficiently large. We will need some more notation; we use notation similar to [14]. For $f_1,f_2 : \frac{1}{n}\mathbb{Z}\to\mathbb{R}$, write

$$\langle f_1, f_2\rangle_n := n^{-1}\sum_{w\in\frac{1}{n}\mathbb{Z}} f_1(w)f_2(w).$$

Let $(X^n_t)_{t\geq 0}$ denote a continuous-time simple symmetric random walk on $\frac{1}{n}\mathbb{Z}$ with jump rate n^2. For $z\in\frac{1}{n}\mathbb{Z}$, let $\mathbf{P}_z(\cdot) := \mathbb{P}\left(\cdot\,|X^n_0 = z\right)$ and $\mathbf{E}_z[\cdot] := \mathbb{E}\left[\cdot\,|X^n_0 = z\right]$. Then for $z,w\in\frac{1}{n}\mathbb{Z}$ and $0\leq s\leq t$, let

$$\phi^{t,z}_s(w) := n\mathbf{P}_z\left(X^n_{m(t-s)} = w\right). \tag{3.4}$$

For $a\in\mathbb{R}$, $z,w\in\frac{1}{n}\mathbb{Z}$ and $0\leq s\leq t$, let

$$\phi^{t,z,a}_s(w) = e^{-a(t-s)}\phi^{t,z}_s(w). \tag{3.5}$$

Let $(u^n_t)_{t\geq 0}$ denote the solution of

$$\begin{cases}\partial_t u^n_t &= \frac{1}{2}m\Delta_n u^n_t + s_0 f(u^n_t) \quad \text{for } t>0,\\ u^n_0 &= p^n_0.\end{cases} \tag{3.6}$$

We will prove in Proposition 3.2 below that if t is not too large, p^n_t and u^n_t are close with high probability. By the comparison principle, $u^n_t\in[0,1]$. Since $\partial_s\phi^{t,z}_s + \frac{1}{2}m\Delta_n\phi^{t,z}_s = 0$ for $s\in(0,t)$, we have that for $a\in\mathbb{R}$, $z\in\frac{1}{n}\mathbb{Z}$ and $t\geq 0$, by integration by parts,

$$\begin{aligned}&\langle u^n_t, \phi^{t,z,a}_t\rangle_n\\ &= \langle u^n_0, \phi^{t,z,a}_0\rangle_n + \int_0^t \langle u^n_s, \partial_s\phi^{t,z,a}_s\rangle_n ds + \int_0^t \langle u^n_s, \tfrac{1}{2}m\Delta_n\phi^{t,z,a}_s\rangle_n ds + s_0\int_0^t \langle f(u^n_s), \phi^{t,z,a}_s\rangle_n ds\\ &= e^{-at}\langle p^n_0, \phi^{t,z}_0\rangle_n + \int_0^t e^{-a(t-s)}\langle s_0 f(u^n_s) + a u^n_s, \phi^{t,z}_s\rangle_n ds.\end{aligned}$$

Therefore, since $\langle u^n_t, \phi^{t,z,a}_t\rangle_n = u^n_t(z)$, it follows that for $a \in \mathbb{R}$, $z \in \frac{1}{n}\mathbb{Z}$ and $t \geq 0$,

$$u^n_t(z) = e^{-at}\langle p^n_0, \phi^{t,z}_0\rangle_n + \int_0^t e^{-a(t-s)}\langle s_0 f(u^n_s) + a u^n_s, \phi^{t,z}_s\rangle_n ds. \tag{3.7}$$

Note that by (3.7) with $a = -(1+\alpha)s_0$, since $f(u) \leq (1+\alpha)u$ for $u \in [0,1]$,

$$u^n_t(z) \leq e^{(1+\alpha)s_0 t}\langle p^n_0, \phi^{t,z}_0\rangle_n. \tag{3.8}$$

In this section, alongside proving Proposition 3.1, we will prove some preliminary tracer dynamics results which will be used in later sections, so we need some notation for tracer dynamics with an arbitrary initial set of 'tracer' type A individuals. Take $\mathcal{I}_0 \subseteq \{(x,i) : \xi^n_0(x,i) = 1\}$. Then for $t \geq 0$, let

$$\eta^n_t(x,i) = \mathbb{1}_{(\zeta^{n,t}_t(x,i), \theta^{n,t}_t(x,i)) \in \mathcal{I}_0} \quad \text{for } x \in \tfrac{1}{n}\mathbb{Z},\ i \in [N], \tag{3.9}$$

i.e. $\eta^n_t(x,i) = 1$ if and only if the i^{th} individual at x at time t is descended from an individual in $\mathcal{I}_0$ at time 0. For $t \geq 0$ and $x \in \frac{1}{n}\mathbb{Z}$, let

$$q^n_t(x) = \frac{1}{N}\sum_{i=1}^N \eta^n_t(x,i), \tag{3.10}$$

i.e. the proportion of individuals at x at time t which are descended from individuals in $\mathcal{I}_0$ at time 0. Let $(v^n_t)_{t\geq 0}$ denote the solution of

$$\begin{cases} \partial_t v^n_t &= \frac{1}{2}m\Delta_n v^n_t + s_0 v^n_t(1-u^n_t)(2u^n_t - 1 + \alpha) \quad \text{for } t > 0, \\ v^n_0 &= q^n_0. \end{cases} \tag{3.11}$$

We will prove in Proposition 3.2 below that if t is not too large, q^n_t and v^n_t are close with high probability. Note that by the comparison principle, $0 \leq v^n_t \leq u^n_t$. Moreover, for $a \in \mathbb{R}$, $t \geq 0$ and $z \in \frac{1}{n}\mathbb{Z}$, by the same argument as for (3.7),

$$v^n_t(z) = e^{-at}\langle q^n_0, \phi^{t,z}_0\rangle_n + \int_0^t e^{-a(t-s)}\langle v^n_s(s_0(1-u^n_s)(2u^n_s - 1 + \alpha) + a), \phi^{t,z}_s\rangle_n ds. \tag{3.12}$$

For $t \geq 0$ and $z \in \frac{1}{n}\mathbb{Z}$, by (3.12) with $a = -(1+\alpha)s_0$ and since $(1-u)(2u-1+\alpha) \leq 1+\alpha$ for $u \in [0,1]$,

$$v^n_t(z) \leq e^{(1+\alpha)s_0 t}\langle q^n_0, \phi^{t,z}_0\rangle_n. \tag{3.13}$$

The following result says that if t is not too large, $|p^n_t - u^n_t|$ and $|q^n_t - v^n_t|$ are small with high probability; the proof is postponed to Section 3.1.

Proposition 3.2. *Suppose $c_3 > 0$ and $\ell \in \mathbb{N}$. Then there exists $c_4 = c_4(c_3, \ell) \in (0, 1/2)$ such that for n sufficiently large, for $T \leq 2(\log N)^{c_4}$,*

$$\mathbb{P}\left(\sup_{x \in \frac{1}{n}\mathbb{Z}, |x| \leq N^5}\ \sup_{t \in [0,T]} |p^n_t(x) - u^n_t(x)| \geq \left(\frac{n}{N}\right)^{1/2 - c_3}\right) \leq \left(\frac{n}{N}\right)^{\ell}$$

and for $t \leq 2(\log N)^{c_4}$,

$$\mathbb{P}\left(\sup_{x \in \frac{1}{n}\mathbb{Z}, |x| \leq N^5} |q^n_t(x) - v^n_t(x)| \geq \left(\frac{n}{N}\right)^{1/2 - c_3}\right) \leq \left(\frac{n}{N}\right)^{\ell}.$$

For $k \in \mathbb{N}$ with $k \geq 2$, there exists a constant $C_1 = C_1(k) < \infty$ such that for $t \geq 0$,

$$\sup_{x \in \frac{1}{n}\mathbb{Z}} \mathbb{E}\left[|p^n_t(x) - u^n_t(x)|^k\right] \leq C_1\left(\frac{n^{k/2}t^{k/4}}{N^{k/2}} + N^{-k}\right)e^{C_1 t^k}. \tag{3.14}$$

We also need to control $p_t^n(x)$ when x is not in the interval $[-N^5, N^5]$ covered by Proposition 3.2.

Lemma 3.3. *For n sufficiently large, if $p_0^n(x) = 0 \ \forall x \geq N$ and $p_0^n(x) = 1 \ \forall x \leq -N$ then*

$$\mathbb{P}\left(\exists t \in [0, 2N^2],\, x \in \tfrac{1}{n}\mathbb{Z} \cap [N^5, \infty) : p_t^n(x) > 0\right) \leq e^{-N^5}$$
$$\textit{and} \qquad \mathbb{P}\left(\exists t \in [0, 2N^2],\, x \in \tfrac{1}{n}\mathbb{Z} \cap (-\infty, -N^5] : p_t^n(x) < 1\right) \leq e^{-N^5}.$$

Proof. For $x \in \frac{1}{n}\mathbb{Z}$, let

$$\tau_x := \inf\{t \geq 0 : p_t^n(x) > 0\}.$$

Let $(T_\ell)_{\ell=1}^\infty$ be a sequence of i.i.d. random variables with $T_1 \sim \text{Exp}(mr_n N^2)$. For $x > N$, τ_x occurs after time $\tau_{x-n^{-1}}$ and at a jump time in $\mathcal{R}^{x,i,x-n^{-1},j}$ for some $i, j \in [N]$. Therefore we can couple the process $(\xi_t^n(x,i))_{x \in \frac{1}{n}\mathbb{Z},\, i \in [N], t \geq 0}$ with $(T_\ell)_{\ell=1}^\infty$ in such a way that for each $\ell \in \mathbb{N}$,

$$\tau_{N+\ell n^{-1}} - \tau_{N+(\ell-1)n^{-1}} \geq T_\ell.$$

It follows that

$$\tau_{N^5} \geq \sum_{\ell=1}^{n(N^5-N)} T_\ell.$$

Therefore, letting Y_n denote a Poisson random variable with mean $2mr_n N^4$, we have that

$$\begin{aligned}\mathbb{P}\left(\tau_{N^5} \leq 2N^2\right) &\leq \mathbb{P}\left(\sum_{\ell=1}^{n(N^5-N)} T_\ell \leq 2N^2\right) \\ &= \mathbb{P}\left(Y_n \geq n(N^5 - N)\right).\end{aligned}$$

By Markov's inequality, and then since $r_n = \frac{1}{2}n^2 N^{-1}$,

$$\mathbb{P}\left(Y_n \geq n(N^5 - N)\right) \leq e^{-n(N^5-N)}\mathbb{E}\left[e^{Y_n}\right] = e^{-n(N^5-N)} e^{mn^2 N^3 (e-1)} \leq e^{-N^5}$$

for n sufficiently large, since $N \geq n$. Therefore for n sufficiently large,

$$\mathbb{P}\left(\tau_{N^5} \leq 2N^2\right) \leq e^{-N^5}.$$

Letting $\sigma_x := \inf\{t \geq 0 : p_t^n(x) < 1\}$ for $x \in \frac{1}{n}\mathbb{Z}$, by the same argument we have that

$$\mathbb{P}\left(\sigma_{-N^5} \leq 2N^2\right) \leq e^{-N^5}$$

for n sufficiently large, which completes the proof. □

Recall from (1.12) and (2.1) that $g(x) = (1 + e^{\kappa x})^{-1}$, and recall the definition of f in (3.1). Note that $u(t, x) := g(x - \nu t)$ is a travelling wave solution of the partial differential equation

$$\partial_t u = \tfrac{1}{2} m \Delta u + s_0 f(u).$$

Since $\alpha \in (0, 1)$, we have that $f(0) = f(1) = 0$, $f(u) < 0$ for $u \in (0, \frac{1}{2}(1 - \alpha))$, $f(u) > 0$ for $u \in (\frac{1}{2}(1 - \alpha), 1)$, $f'(0) < 0$ and $f'(1) < 0$. This allows us to apply results from [16] as follows. For an initial condition $u_0 : \mathbb{R} \to [0, 1]$, let $u(t, x)$ denote the solution of

$$\begin{cases} \partial_t u &= \frac{1}{2} m \Delta u + s_0 f(u) \quad \text{for } t > 0, \\ u(0, \cdot) &= u_0. \end{cases} \tag{3.15}$$

Lemma 3.4. *There exist constants $C_2 < \infty$ and $c_5 > 0$ such that for $\epsilon \leq c_5$, if u_0 is piecewise continuous with $0 \leq u_0 \leq 1$ and, for some $z_0 \in \mathbb{R}$, $|u_0(z) - g(z - z_0)| \leq \epsilon \ \forall z \in \mathbb{R}$, then*

$$|u(t, x) - g(x - \nu t - z_0)| \leq C_2 \epsilon \quad \forall x \in \mathbb{R},\ t > 0.$$

Proof. The result follows directly from Lemma 4.2 in [16] and its proof. □

Proposition 3.5. *There exist constants $c_6 > 0$ and $C_3 < \infty$ such that if u_0 is piecewise continuous with $0 \le u_0 \le 1$ and $|u_0(z) - g(z)| \le c_6 \ \forall z \in \mathbb{R}$, then for some $z_0 \in \mathbb{R}$ with $|z_0| \le 1$,*

$$|u(t,x) - g(x - \nu t - z_0)| \le C_3 e^{-c_6 t} \quad \forall x \in \mathbb{R},\ t > 0.$$

This is a slight modification of Theorem 3.1 in [16] (to ensure that C_3 and c_6 do not depend on the initial condition u_0, as long as $\|u_0 - g\|_\infty$ is sufficiently small); we postpone the proof to Appendix A. The next lemma says that if the initial condition p_0^n is not too rough, then u_t^n is close to a solution of (3.15).

Lemma 3.6. *Let $(u_t)_{t\ge 0}$ denote the solution of*

$$\begin{cases} \partial_t u_t &= \frac{1}{2} m \Delta u_t + s_0 f(u_t) \quad \text{for } t > 0, \\ u_0 &= \bar{p}_0^n, \end{cases} \tag{3.16}$$

for some $\bar{p}_0^n : \mathbb{R} \to [0,1]$ with $\bar{p}_0^n(y) = p_0^n(y) \ \forall y \in \frac{1}{n}\mathbb{Z}$. There exists a constant $C_4 < \infty$ such that for $T \ge 1$,

$$\begin{aligned} &\sup_{t\in[0,T],\, x\in\frac{1}{n}\mathbb{Z}} |u_t^n(x) - u_t(x)| \\ &\le \left(C_4 n^{-1/3} + \sup_{z_1,z_2\in\mathbb{R}, |z_1 - z_2| \le n^{-1/3}} |\bar{p}_0^n(z_1) - \bar{p}_0^n(z_2)| \right) T^2 e^{(1+\alpha)s_0 T}. \end{aligned}$$

Proof. For $t \ge 0$ and $z \in \frac{1}{n}\mathbb{Z}$, by (3.7) and since $p_0^n(y) = \bar{p}_0^n(y) \ \forall y \in \frac{1}{n}\mathbb{Z}$,

$$u_t^n(z) = \langle \bar{p}_0^n, \phi_0^{t,z} \rangle_n + s_0 \int_0^t \langle f(u_s^n), \phi_s^{t,z} \rangle_n ds.$$

Let $G_t(x) = \frac{1}{\sqrt{2\pi t}} e^{-x^2/(2t)}$; then since G is the fundamental solution of the heat equation, and using Duhamel's principle (see for example (17) and (18) in Section 2.3 on page 51 of [15] and Theorem 4.8 on page 147 of [18]), for $z \in \mathbb{R}$ and $t > 0$,

$$u_t(z) = G_{mt} * \bar{p}_0^n(z) + s_0 \int_0^t G_{m(t-s)} * f(u_s)(z) ds. \tag{3.17}$$

Letting $(B_t)_{t\ge 0}$ denote a Brownian motion, and by the definition of $\phi_s^{t,z}$ in (3.4), it follows that for $z \in \frac{1}{n}\mathbb{Z}$ and $t > 0$,

$$\begin{aligned} &|u_t^n(z) - u_t(z)| \\ &\le |\mathbf{E}_z [\bar{p}_0^n(X_{mt}^n)] - \mathbb{E}_z [\bar{p}_0^n(B_{mt})]| + s_0 \int_0^t \left| \mathbf{E}_z \left[f(u_s^n(X_{m(t-s)}^n)) \right] - \mathbb{E}_z \left[f(u_s(B_{m(t-s)})) \right] \right| ds. \end{aligned} \tag{3.18}$$

By a Skorokhod embedding argument (see e.g. Theorem 3.3.3 in [24]), for n sufficiently large, $(X_t^n)_{t\ge 0}$ and $(B_t)_{t\ge 0}$ can be coupled in such a way that $X_0^n = B_0$ and for $t \ge 0$,

$$\mathbb{P}\left(|X_{mt}^n - B_{mt}| \ge n^{-1/3} \right) \le (t+1) n^{-1/2}. \tag{3.19}$$

Since $\bar{p}_0^n \in [0,1]$, it follows that

$$|\mathbf{E}_z [\bar{p}_0^n(X_{mt}^n)] - \mathbb{E}_z [\bar{p}_0^n(B_{mt})]| \le (t+1) n^{-1/2} + \sup_{z_1,z_2\in\mathbb{R}, |z_1 - z_2| \le n^{-1/3}} |\bar{p}_0^n(z_1) - \bar{p}_0^n(z_2)|. \tag{3.20}$$

For the second term on the right hand side of (3.18), note that $\sup_{v\in[0,1]}|f(v)|<1$ and, since $f'(u)=6u(1-u)-1+\alpha(1-2u)$, we have $\sup_{v\in[0,1]}|f'(v)|=1+\alpha$. Therefore, using the triangle inequality and then by the same coupling argument as for (3.20), for $s\in[0,t]$,

$$\begin{aligned}
&\left|\mathbf{E}_z\left[f(u_s^n(X^n_{m(t-s)}))\right]-\mathbb{E}_z\left[f(u_s(B_{m(t-s)}))\right]\right|\\
&\le\left|\mathbf{E}_z\left[f(u_s^n(X^n_{m(t-s)}))\right]-\mathbf{E}_z\left[f(u_s(X^n_{m(t-s)}))\right]\right|\\
&\qquad+\left|\mathbf{E}_z\left[f(u_s(X^n_{m(t-s)}))\right]-\mathbb{E}_z\left[f(u_s(B_{m(t-s)}))\right]\right|\\
&\le(1+\alpha)\sup_{x\in\frac{1}{n}\mathbb{Z}}|u_s^n(x)-u_s(x)|+2(t+1)n^{-1/2}+(1+\alpha)\|\nabla u_s\|_\infty n^{-1/3}. \qquad (3.21)
\end{aligned}$$

We now bound $\|\nabla u_s\|_\infty$. For $t>0$ and $x\in\mathbb{R}$, by differentiating both sides of (3.17),

$$\nabla u_t(x)=G'_{mt}*\bar{p}_0^n(x)+s_0\int_0^t G'_{m(t-s)}*f(u_s)(x)ds. \qquad (3.22)$$

For the first term on the right hand side, since $\bar{p}_0^n\in[0,1]$,

$$|G'_{mt}*\bar{p}_0^n(x)|\le\int_{-\infty}^{\infty}|G'_{mt}(z)|dz=2G_{mt}(0)=2(2\pi mt)^{-1/2}.$$

For the second term on the right hand side of (3.22), since $\sup_{v\in[0,1]}|f(v)|<1$,

$$\left|\int_0^t G'_{m(t-s)}*f(u_s)(x)ds\right|\le\int_0^t\int_{-\infty}^{\infty}|G'_{m(t-s)}(z)|dzds=4(2\pi m)^{-1/2}t^{1/2}.$$

Hence by (3.22), for $t>0$,

$$\|\nabla u_t\|_\infty\le(2\pi m)^{-1/2}(2t^{-1/2}+4s_0t^{1/2}).$$

Substituting into (3.21) and then into (3.18), and using (3.20), we now have that for $t>0$ and $z\in\frac{1}{n}\mathbb{Z}$,

$$\begin{aligned}
&|u_t^n(z)-u_t(z)|\\
&\le(t+1)\,n^{-1/2}+\sup_{z_1,z_2\in\mathbb{R},|z_1-z_2|\le n^{-1/3}}|\bar{p}_0^n(z_1)-\bar{p}_0^n(z_2)|\\
&\quad+s_0\int_0^t\Big((1+\alpha)\sup_{x\in\frac{1}{n}\mathbb{Z}}|u_s^n(x)-u_s(x)|+2(t+1)n^{-1/2}\\
&\qquad\qquad+2(2\pi m)^{-1/2}(2s^{-1/2}+4s_0s^{1/2})n^{-1/3}\Big)ds.
\end{aligned}$$

Hence there exists a constant $C_4<\infty$ such that for $T\ge1$, for $t\in[0,T]$,

$$\begin{aligned}
&\sup_{x\in\frac{1}{n}\mathbb{Z}}|u_t^n(x)-u_t(x)|\\
&\le\Big(C_4n^{-1/3}+\sup_{z_1,z_2\in\mathbb{R},|z_1-z_2|\le n^{-1/3}}|\bar{p}_0^n(z_1)-\bar{p}_0^n(z_2)|\Big)T^2\\
&\qquad+(1+\alpha)s_0\int_0^t\sup_{x\in\frac{1}{n}\mathbb{Z}}|u_s^n(x)-u_s(x)|ds.
\end{aligned}$$

The result follows by Gronwall's inequality. □

The following lemma will be used in the proof of Proposition 3.1 to show that with high probability, $\sup_{|z_1-z_2|\le n^{-1/3}}|p_t^n(z_1)-p_t^n(z_2)|$ is small at large times t, which will allow us to use Lemma 3.6.

Lemma 3.7. *There exists a constant $C_5 < \infty$ such that*

$$n\langle 1, |\phi_0^{t,z+n^{-1}} - \phi_0^{t,z}|\rangle_n \le C_5 t^{-1/2} \quad \forall\, t>0, z\in \tfrac{1}{n}\mathbb{Z}, \tag{3.23}$$

and $\sup_{t\ge 1, x\in\frac{1}{n}\mathbb{Z}} |\nabla_n u_t^n(x)| \le C_5$.

Proof. For $t>0$, $z\in\frac{1}{n}\mathbb{Z}$ and $t_0\in(0,t]$, by (3.7),

$$\nabla_n u_t^n(z) = n\langle u^n_{t-t_0}, \phi_0^{t_0,z+n^{-1}} - \phi_0^{t_0,z}\rangle_n + ns_0\int_0^{t_0} \langle f(u^n_{t-t_0+s}), \phi_s^{t_0,z+n^{-1}} - \phi_s^{t_0,z}\rangle_n ds. \tag{3.24}$$

Since $u^n_{t-t_0}\in[0,1]$, we have that

$$|n\langle u^n_{t-t_0}, \phi_0^{t_0,z+n^{-1}} - \phi_0^{t_0,z}\rangle_n| \le n\langle 1, |\phi_0^{t_0,z+n^{-1}} - \phi_0^{t_0,z}|\rangle_n. \tag{3.25}$$

Let $(S_j)_{j=0}^\infty$ be a discrete-time simple symmetric random walk on $\mathbb{Z}$ with $S_0=0$. By Proposition 2.4.1 in [24] (which follows from the local central limit theorem), there exists a constant $K_1<\infty$ such that for $j\in\mathbb{N}$,

$$\sum_{y\in\mathbb{Z}} |\mathbb{P}(S_j=y-1) - \mathbb{P}(S_j=y)| \le K_1 j^{-1/2}.$$

Let $(R_s)_{s\ge 0}$ denote a Poisson process with rate 1. Then by the definition of $\phi_s^{t,z}$ in (3.4), and since $(X_s^n)_{s\ge0}$ jumps at rate n^2,

$$\begin{aligned} n\langle 1, |\phi_0^{t_0,z+n^{-1}} - \phi_0^{t_0,z}|\rangle_n &= n\sum_{y\in\frac{1}{n}\mathbb{Z}} \left|\mathbf{P}_0\left(X^n_{mt_0}=y-n^{-1}\right) - \mathbf{P}_0\left(X^n_{mt_0}=y\right)\right| \\ &\le n\sum_{y\in\frac{1}{n}\mathbb{Z}}\sum_{j=0}^\infty \mathbb{P}\left(R_{mn^2t_0}=j\right)\left|\mathbb{P}\left(S_j=ny-1\right) - \mathbb{P}\left(S_j=ny\right)\right| \\ &\le n\sum_{j=1}^\infty \mathbb{P}\left(R_{mn^2t_0}=j\right)K_1 j^{-1/2} + 2n\mathbb{P}\left(R_{mn^2t_0}=0\right). \end{aligned} \tag{3.26}$$

By Markov's inequality, and since $R_{mn^2t_0}\sim \text{Poisson}(mn^2t_0)$,

$$\begin{aligned} \mathbb{P}\left(R_{mn^2t_0}\le \tfrac12 mn^2t_0\right) = \mathbb{P}\left(e^{-R_{mn^2t_0}\log 2} \ge e^{-\frac12 mn^2t_0\log 2}\right) &\le e^{\frac12 mn^2t_0\log 2} e^{mn^2t_0(e^{-\log 2}-1)} \\ &= e^{-\frac12 mn^2t_0(1-\log 2)}. \end{aligned}$$

Therefore, by substituting into (3.26),

$$\begin{aligned} n\langle 1, |\phi_0^{t_0,z+n^{-1}} - \phi_0^{t_0,z}|\rangle_n &\le n\left((K_1+2)\mathbb{P}\left(R_{mn^2t_0}\le \tfrac12 mn^2t_0\right) + K_1(\tfrac12 mn^2t_0)^{-1/2}\right) \\ &\le t_0^{-1/2}\left((K_1+2)(n^2t_0)^{1/2}e^{-\frac12 mn^2t_0(1-\log 2)} + \sqrt{2}m^{-1/2}K_1\right) \\ &\le K_2 t_0^{-1/2}, \end{aligned} \tag{3.27}$$

where $K_2 := (K_1+2)\sup_{s\ge0}(s^{1/2}e^{-\frac12 m(1-\log 2)s}) + \sqrt{2}m^{-1/2}K_1<\infty$. This completes the proof of (3.23). For the second term on the right hand side of (3.24), since $|f(u^n_{t-t_0+s})|\le 1$ for $s\in[0,t_0]$, and then by (3.27),

$$\begin{aligned} \left|ns_0\int_0^{t_0}\langle f(u^n_{t-t_0+s}), \phi_s^{t_0,z+n^{-1}} - \phi_s^{t_0,z}\rangle_n ds\right| &\le s_0\int_0^{t_0} n\langle 1, |\phi_0^{t_0-s,z+n^{-1}} - \phi_0^{t_0-s,z}|\rangle_n ds \\ &\le 2s_0K_2t_0^{1/2}. \end{aligned}$$

Therefore, by (3.24), (3.25) and (3.27), for $t \geq 1$ and $t_0 \in (0, t]$ we have

$$\sup_{x \in \frac{1}{n}\mathbb{Z}} |\nabla_n u_t^n(x)| \leq K_2(t_0^{-1/2} + 2s_0 t_0^{1/2}),$$

and the result follows by taking $t_0 = 1$. ☐

We will use the following easy lemma repeatedly in the rest of this section, and in Section 4.

Lemma 3.8. *For $a \in \mathbb{R}$ with $|a| \leq n$ and $t \geq 0$,*

$$\mathbf{E}_0\left[e^{aX_{mt}^n}\right] = e^{\frac{1}{2}ma^2t + \mathcal{O}(ta^3n^{-1})}.$$

Proof. Let $(R_s^+)_{s\geq 0}$ and $(R_s^-)_{s\geq 0}$ be independent Poisson processes with rate 1. For $a \in \mathbb{R}$, since $(X_t^n)_{t\geq 0}$ is a continuous-time simple symmetric random walk on $\frac{1}{n}\mathbb{Z}$ with jump rate n^2, and then since $R^+_{mn^2t/2}$ and $R^-_{mn^2t/2}$ are both Poisson distributed with mean $\frac{1}{2}mn^2t$,

$$\begin{aligned}
\mathbf{E}_0\left[e^{aX_{mt}^n}\right] &= \mathbb{E}\left[e^{an^{-1}(R^+_{mn^2t/2} - R^-_{mn^2t/2})}\right] \\
&= \exp(\tfrac{1}{2}mn^2t(e^{an^{-1}} - 1))\exp(\tfrac{1}{2}mn^2t(e^{-an^{-1}} - 1)) \\
&= \exp\left(\tfrac{1}{2}mn^2t\left(an^{-1} + \tfrac{1}{2}a^2n^{-2} + \mathcal{O}\left(a^3n^{-3}\right) - an^{-1} + \tfrac{1}{2}a^2n^{-2} + \mathcal{O}\left(a^3n^{-3}\right)\right)\right) \\
&= e^{\frac{1}{2}ma^2t + \mathcal{O}(ta^3n^{-1})},
\end{aligned}$$

which completes the proof. ☐

The following two lemmas will allow us to control $p_t^n(x)$ for large x. The first lemma gives us an upper bound that we will use inductively in the proof of Proposition 3.1.

Lemma 3.9. *There exists a constant $c_7 \in (0,1)$ such that for n sufficiently large, the following holds. Suppose that $p_0^n(x) = 0\ \forall x \geq N^6$. Take $c \in (0, 1/2)$. Suppose for some $R > 0$ with $R\left(\frac{n}{N}\right)^{1/2-c} \leq c_7$ that*

$$p_0^n(x) \leq 3e^{-\kappa(1-(\log N)^{-2})x} + R\left(\frac{n}{N}\right)^{1/2-c} \quad \forall x \in \tfrac{1}{n}\mathbb{Z}, \tag{3.28}$$

and that for some $T \in (1, \log N]$, $\sup_{y \in \frac{1}{n}\mathbb{Z}, |y| \leq N, t \in [0,T]} |u_t^n(y) - g(y - \nu t)| \leq c_7(\log N)^{-2}$. Then for $t \in [0, T]$,

$$u_t^n(x) \leq \tfrac{4}{3}\left(3e^{-\kappa(1-(\log N)^{-2})(x-\nu t)} + R\left(\frac{n}{N}\right)^{1/2-c}\right) \quad \forall x \in \tfrac{1}{n}\mathbb{Z},$$

and for $t \in [1, T]$,

$$u_t^n(x) \leq (1 - c_7(\log N)^{-2})3e^{-\kappa(1-(\log N)^{-2})(x-\nu t)} + (1-c_7)R\left(\frac{n}{N}\right)^{1/2-c} \quad \forall x \in \tfrac{1}{n}\mathbb{Z}.$$

Proof. Take $d \in (0, 1/3)$ such that

$$d < \min\left(\tfrac{1}{10}(2-\alpha)s_0, \tfrac{1}{4}e^{-(1-\alpha)s_0}(1-\alpha)s_0\right). \tag{3.29}$$

Suppose that

$$R\left(\frac{n}{N}\right)^{1/2-c} < \tfrac{1}{12}(1+d)^{-1}e^{-(1-\alpha)s_0}(1-\alpha), \tag{3.30}$$

and that $T \in (1, \log N]$ with

$$\sup_{y \in \frac{1}{n}\mathbb{Z}, |y| \le N,\, t \in [0,T]} |u_t^n(y) - g(y - \nu t)| < \tfrac{1}{73} e^{-5s_0}(2-\alpha)(\log N)^{-2}. \tag{3.31}$$

Let $\theta_N = (1 - (\log N)^{-2})\kappa$, and let

$$\tau = T \wedge \inf \Big\{ t \ge 0 : \exists\, x \in \tfrac{1}{n}\mathbb{Z} \text{ s.t. } u_t^n(x) \ge (1 + d(\log N)^{-2}) 3 e^{-\theta_N (x - \nu t)} + (1+d) R \left(\frac{n}{N}\right)^{1/2 - c} \Big\}.$$

By (3.8), and then since $p_0^n(x) = 0 \; \forall x \ge N^6$, for $t \ge 0$ and $z \in \frac{1}{n}\mathbb{Z}$,

$$\begin{aligned} u_t^n(z) \le e^{(1+\alpha)s_0 t} \langle p_0^n, \phi_0^{t,z} \rangle_n &\le e^{(1+\alpha)s_0 t} \mathbf{P}_z \left(X_{mt}^n \le N^6 \right) \\ &= e^{(1+\alpha)s_0 t} \mathbf{P}_0 \left(X_{mt}^n \ge z - N^6 \right) \\ &\le e^{(1+\alpha)s_0 t} \mathbf{E}_0 \left[e^{2\theta_N X_{mt}^n} \right] e^{-2\theta_N z + 2\theta_N N^6} \\ &\le e^{(2s_0 + 3m\theta_N^2)t} e^{-2\theta_N z + 2\theta_N N^6} \end{aligned} \tag{3.32}$$

for n sufficiently large, by Markov's inequality and Lemma 3.8. Therefore, since $u_t^n(x) \in [0,1]$, there exists $N' < \infty$ such that

$$\tau = T \wedge \min_{x \in \frac{1}{n}\mathbb{Z} \cap [0, N']} \inf \Big\{ t \ge 0 : u_t^n(x) \ge (1 + d(\log N)^{-2}) 3 e^{-\theta_N(x - \nu t)} + (1+d) R \left(\frac{n}{N}\right)^{1/2-c} \Big\}.$$

Hence (by continuity of $u_t^n(x)$ for each $x \in \frac{1}{n}\mathbb{Z}$ and by our assumption on the initial condition in (3.28)) we have that $\tau > 0$. Moreover, if $\tau < T$ then there exists $x \in \frac{1}{n}\mathbb{Z} \cap [0, N']$ such that

$$u_\tau^n(x) \ge (1 + d(\log N)^{-2}) 3 e^{-\theta_N(x - \nu\tau)} + (1+d) R \left(\frac{n}{N}\right)^{1/2-c}. \tag{3.33}$$

Note that for $u \in [0,1]$,

$$f(u) + (1-\alpha)u = -2u^3 + (3-\alpha)u^2 \le (3-\alpha)u^2. \tag{3.34}$$

Now by (3.7), for $0 < t \le \tau$ and $x \in \frac{1}{n}\mathbb{Z}$, for $0 < t_0 \le t \wedge 1$,

$$\begin{aligned} u_t^n(x) &= e^{-(1-\alpha)s_0 t_0} \langle u_{t-t_0}^n, \phi_0^{t_0,x} \rangle_n \\ &\qquad + s_0 \int_0^{t_0} e^{-(1-\alpha)s_0(t_0 - s)} \langle f(u_{t-t_0+s}^n) + (1-\alpha) u_{t-t_0+s}^n, \phi_s^{t_0,x} \rangle_n ds \\ &\le e^{-(1-\alpha)s_0 t_0} \langle u_{t-t_0}^n, \phi_0^{t_0,x} \rangle_n + 3 s_0 \int_0^{t_0} e^{-(1-\alpha)s_0(t_0-s)} \langle (u_{t-t_0+s}^n)^2, \phi_s^{t_0,x} \rangle_n ds, \end{aligned} \tag{3.35}$$

where the second line follows by (3.34). Since $t \le \tau$, we have

$$\begin{aligned} \langle u_{t-t_0}^n, \phi_0^{t_0,x} \rangle_n &\le (1 + d(\log N)^{-2}) \mathbf{E}_x \left[3 e^{-\theta_N (X_{mt_0}^n - \nu(t - t_0))} \right] + (1+d) R \left(\frac{n}{N}\right)^{1/2-c} \\ &\le (1 + d(\log N)^{-2}) 3 e^{-\theta_N(x - \nu(t-t_0))} e^{\frac{1}{2} m \theta_N^2 t_0 + \mathcal{O}(t_0 n^{-1})} + (1+d) R \left(\frac{n}{N}\right)^{1/2-c}, \end{aligned}$$

by Lemma 3.8. For the second term on the right hand side of (3.35), we have that for $s \in [0, t_0)$,

$$\begin{aligned}
&\langle (u^n_{t-t_0+s})^2, \phi^{t_0,x}_s \rangle_n \\
&\le 2\left((1+d(\log N)^{-2})^2 \mathbf{E}_x \left[9e^{-2\theta_N (X^n_{m(t_0-s)} - \nu(t-t_0+s))} \right] + (1+d)^2 R^2 \left(\frac{n}{N}\right)^{1-2c} \right) \\
&\le 2\left((1+d(\log N)^{-2})^2 9e^{-2\theta_N (x-\nu(t-t_0+s))} e^{2m\theta_N^2 (t_0-s) + \mathcal{O}(t_0 n^{-1})} + (1+d)^2 R^2 \left(\frac{n}{N}\right)^{1-2c} \right) \\
&= 2(1+d(\log N)^{-2})^2 \cdot 9e^{-2\theta_N (x-\nu t)} e^{(2m\theta_N^2 - 2\theta_N \nu)(t_0-s) + \mathcal{O}(t_0 n^{-1})} + 2(1+d)^2 R^2 \left(\frac{n}{N}\right)^{1-2c},
\end{aligned}$$

where the second inequality follows by Lemma 3.8. Note that by (2.1), $(1-\alpha)s_0 + \theta_N \nu - \frac{1}{2} m\theta_N^2 = (2-\alpha-(\log N)^{-2}) s_0 (\log N)^{-2}$ and $2m\theta_N^2 - 2\theta_N \nu - (1-\alpha)s_0 \le 2m\theta_N^2 \le 2m\kappa^2 = 4s_0$. Hence for n sufficiently large, substituting into (3.35),

$$\begin{aligned}
&u^n_t(x) \\
&\le e^{-((1-\alpha)s_0 + \theta_N \nu - \frac{1}{2} m\theta_N^2)t_0 + \mathcal{O}(t_0 n^{-1})} (1+d(\log N)^{-2}) 3e^{-\theta_N (x-\nu t)} \\
&\quad + e^{-(1-\alpha)s_0 t_0} (1+d) R \left(\frac{n}{N}\right)^{1/2-c} + 6s_0 (1+d(\log N)^{-2})^2 9e^{-2\theta_N (x-\nu t)} e^{5s_0 t_0} t_0 \\
&\quad + 6s_0 (1+d)^2 R^2 \left(\frac{n}{N}\right)^{1-2c} t_0 \\
&\le (1+d(\log N)^{-2}) 3e^{-\theta_N (x-\nu t)} + (1+d) R \left(\frac{n}{N}\right)^{1/2-c} \\
&\quad + t_0 (1+d(\log N)^{-2}) 3e^{-\theta_N (x-\nu t)} \Big(18 s_0 (1+d(\log N)^{-2}) e^{-\theta_N (x-\nu t)} e^{5s_0 t_0} \\
&\qquad\qquad - e^{-\frac{1}{2}(2-\alpha)s_0 (\log N)^{-2} t_0} \tfrac{1}{2} s_0 (2-\alpha)(\log N)^{-2} \Big) \\
&\quad + t_0 (1+d) R \left(\frac{n}{N}\right)^{1/2-c} \left(6s_0 (1+d) R \left(\frac{n}{N}\right)^{1/2-c} - e^{-(1-\alpha)s_0 t_0} (1-\alpha) s_0 \right),
\end{aligned}$$

where the second inequality holds since for $y \ge 0$, $e^{-y} = 1-(1-e^{-y}) \le 1 - ye^{-y}$. Suppose x is such that

$$18(1+d(\log N)^{-2}) e^{-\theta_N (x-\nu t)} e^{5s_0 t_0} - \tfrac{1}{4} e^{-\frac{1}{2}(2-\alpha)s_0 (\log N)^{-2} t_0} (2-\alpha)(\log N)^{-2} \le 0.$$

Then since $t_0 \in (0,1]$, and by (3.30) and the definition of d in (3.29), if n is sufficiently large we have that

$$u^n_t(x) < (1+(d-2t_0 d)(\log N)^{-2}) 3e^{-\theta_N (x-\nu t)} + (1+d-2t_0 d) R \left(\frac{n}{N}\right)^{1/2-c}. \tag{3.36}$$

If instead $x \ge \nu t$ and

$$18(1+d(\log N)^{-2}) e^{-\theta_N (x-\nu t)} e^{5s_0 t_0} > \tfrac{1}{4} e^{-\frac{1}{2}(2-\alpha)s_0 (\log N)^{-2} t_0} (2-\alpha)(\log N)^{-2}, \tag{3.37}$$

then since $T \le \log N$, for n sufficiently large we have $|x| \le N$. Since $d < 1/3$ and $t_0 \le 1$, we have that for n sufficiently large,

$$\begin{aligned}
(1+(d-2t_0 d)(\log N)^{-2}) 3e^{-\theta_N (x-\nu t)} &\ge e^{-\kappa(x-\nu t)} + e^{-\theta_N (x-\nu t)} \\
&> g(x-\nu t) + \sup_{y \in \frac{1}{n}\mathbb{Z}, |y| \le N, s \in [0,T]} |u^n_s(y) - g(y-\nu s)|
\end{aligned}$$

by (3.37) and our assumption in (3.31). Therefore for n sufficiently large, in this case we also have that (3.36) holds. Finally, for n sufficiently large, if $x < \nu t$ then since $d < 1/3$, $t_0 \le 1$ and $u^n_t(x) \le 1$ we have that (3.36) holds. Hence (3.36) holds for every $x \in \frac{1}{n}\mathbb{Z}$.

Suppose that $\tau < T$; then (3.33) holds, and by setting $t = \tau$ and $t_0 = 1 \wedge \tau$, we have a contradiction by (3.36). It follows that $\tau = T$, and so the first statement of the lemma holds. The second statement follows by taking $t \geq 1$ and setting $t_0 = 1$ in (3.36). □

The next lemma will give us a corresponding lower bound on $p_t^n(x)$ for large x.

Lemma 3.10. *There exists a constant $c_8 \in (0,1)$ such that the following holds for n sufficiently large. Take $c \in (0,1/2)$. Suppose for some $R > 0$ that*

$$p_0^n(x) \geq \tfrac{1}{3} e^{-\kappa(1+(\log N)^{-2})x} \mathbb{1}_{x\geq 0} - R\left(\frac{n}{N}\right)^{1/2-c} \quad \forall x \in \tfrac{1}{n}\mathbb{Z}, \tag{3.38}$$

and that for some $T \in (1, \log N]$, $\sup_{y\in\frac{1}{n}\mathbb{Z}, |y|\leq N, t\in[0,T]} |u_t^n(y) - g(y-\nu t)| \leq c_8 (\log N)^{-2}$. Then for $t \in [0,T]$,

$$u_t^n(x) \geq \tfrac{1}{4} e^{-\kappa(1+(\log N)^{-2})(x-\nu t)} \mathbb{1}_{x\geq \nu t} - R\left(\frac{n}{N}\right)^{1/2-c} \quad \forall x \in \tfrac{1}{n}\mathbb{Z},$$

and for $t \in [1,T]$, $\forall x \in \frac{1}{n}\mathbb{Z}$,

$$u_t^n(x) \geq (1 + c_8(\log N)^{-2}) \tfrac{1}{3} e^{-\kappa(1+(\log N)^{-2})(x-\nu t)} \mathbb{1}_{x\geq \nu t - c_8} - (1-c_8) R\left(\frac{n}{N}\right)^{1/2-c}.$$

Proof. Note that for $u \in [0,1]$,

$$f(u) + (1-\alpha)u = -2u^3 + (3-\alpha)u^2 \geq 0. \tag{3.39}$$

Take $d \in \left(0, \min\left(\frac{1}{100} e^{-4(\kappa+2s_0)}(1-e^{-\kappa})(2-\alpha)s_0, \log(10/9)\kappa^{-1}\right)\right)$, and suppose

$$\sup_{y\in\frac{1}{n}\mathbb{Z}, |y|\leq N, t\in[0,T]} |u_t^n(y) - g(y-\nu t)| \leq d(\log N)^{-2}. \tag{3.40}$$

Let $\theta_N' = (1 + (\log N)^{-2})\kappa$. For some $t_1 \in [0,T]$, suppose

$$u_{t_1}^n(x) \geq \tfrac{1}{3} e^{-\theta_N'(x-\nu t_1)} \mathbb{1}_{x\geq \nu t_1} - R\left(\frac{n}{N}\right)^{1/2-c} \quad \forall x \in \tfrac{1}{n}\mathbb{Z}. \tag{3.41}$$

Take $t \in (t_1, t_1+1]$ and let $t_0 = t - t_1$. Then for $x \in \frac{1}{n}\mathbb{Z}$, by (3.7),

$$\begin{aligned} u_t^n(x) &= e^{-(1-\alpha)s_0 t_0} \langle u_{t_1}^n, \phi_0^{t_0,x}\rangle_n + s_0 \int_0^{t_0} e^{-(1-\alpha)s_0(t_0-s)} \langle f(u_{t_1+s}^n) + (1-\alpha)u_{t_1+s}^n, \phi_s^{t_0,x}\rangle_n ds \\ &\geq e^{-(1-\alpha)s_0 t_0} \langle u_{t_1}^n, \phi_0^{t_0,x}\rangle_n \end{aligned}$$

by (3.39). Hence by (3.41),

$$u_t^n(x) \geq e^{-(1-\alpha)s_0 t_0} \left(\mathbf{E}_x\left[\tfrac{1}{3} e^{-\theta_N'(X_{mt_0}^n - \nu t_1)} \mathbb{1}_{X_{mt_0}^n \geq \nu t_1}\right] - R\left(\frac{n}{N}\right)^{1/2-c}\right). \tag{3.42}$$

Note that

$$\begin{aligned} &\mathbf{E}_x\left[e^{-\theta_N'(X_{mt_0}^n - \nu t_1)} \mathbb{1}_{X_{mt_0}^n \geq \nu t_1}\right] \\ &= \mathbf{E}_x\left[e^{-\theta_N'(X_{mt_0}^n - \nu t_1)}\right] - \mathbf{E}_x\left[e^{-\theta_N'(X_{mt_0}^n - \nu t_1)} \mathbb{1}_{X_{mt_0}^n < \nu t_1}\right] \\ &= e^{-\theta_N'(x-\nu t_1)} e^{\frac{1}{2}m(\theta_N')^2 t_0 + \mathcal{O}(n^{-1}t_0)} - e^{\theta_N' \nu t_1} \mathbf{E}_x\left[e^{-\theta_N' X_{mt_0}^n} \mathbb{1}_{X_{mt_0}^n < \nu t_1}\right] \end{aligned} \tag{3.43}$$

by Lemma 3.8. For the second term on the right hand side, using Markov's inequality and Lemma 3.8 in the second inequality,

$$\begin{aligned}\mathbf{E}_x\left[e^{-\theta'_N X^n_{mt_0}}\mathbb{1}_{X^n_{mt_0}<\nu t_1}\right] &\leq \sum_{k=\lfloor x-\nu t_1\rfloor}^{\infty} e^{-\theta'_N(x-k-1)}\mathbf{P}_x\left(X^n_{mt_0}\leq x-k\right)\\ &\leq e^{-\theta'_N x}\sum_{k=\lfloor x-\nu t_1\rfloor}^{\infty} e^{\theta'_N(k+1)}e^{-2\theta'_N k}e^{2m(\theta'_N)^2 t_0+\mathcal{O}(t_0 n^{-1})}\\ &\leq e^{-\theta'_N x}e^{\theta'_N+2m(\theta'_N)^2t_0+\mathcal{O}(t_0n^{-1})}e^{-\theta'_N\lfloor x-\nu t_1\rfloor}(1-e^{-\theta'_N})^{-1}.\end{aligned}$$

Suppose $x\geq \nu t_1$ with

$$e^{-\theta'_N(x-\nu t_1)}\leq e^{-3(\theta'_N+m(\theta'_N)^2)}(1-e^{-\theta'_N})\tfrac{1}{5}(2-\alpha)s_0(\log N)^{-2}. \tag{3.44}$$

Then by (3.43) and since $t_0\leq 1$, for n sufficiently large,

$$\begin{aligned}&e^{-(1-\alpha)s_0t_0}\mathbf{E}_x\left[\tfrac{1}{3}e^{-\theta'_N(X^n_{mt_0}-\nu t_1)}\mathbb{1}_{X^n_{mt_0}\geq \nu t_1}\right]\\ &\geq e^{-(1-\alpha)s_0t_0}\tfrac{1}{3}e^{-\theta'_N(x-\nu t_1)}\big(e^{\frac{1}{2}m(\theta'_N)^2t_0+\mathcal{O}(t_0n^{-1})}-e^{3(\theta'_N+m(\theta'_N)^2)}e^{-\theta'_N(x-\nu t_1)}(1-e^{-\theta'_N})^{-1}\big)\\ &\geq \tfrac{1}{3}e^{-\theta'_N(x-\nu t)}e^{((-1+\alpha)s_0-\theta'_N\nu+\frac{1}{2}m(\theta'_N)^2+\mathcal{O}(n^{-1}))t_0}\\ &\qquad\qquad\cdot(1-e^{3(\theta'_N+m(\theta'_N)^2)}e^{-\theta'_N(x-\nu t_1)}(1-e^{-\theta'_N})^{-1})\\ &\geq \tfrac{1}{3}e^{-\theta'_N(x-\nu t)}e^{\frac{1}{2}(2-\alpha)s_0(\log N)^{-2}t_0}(1-\tfrac{1}{5}(2-\alpha)s_0(\log N)^{-2})\end{aligned}$$

for n sufficiently large, where the second inequality holds since $t_1=t-t_0$ and the last inequality follows since by (2.1) we have $(-1+\alpha)s_0-\theta'_N\nu+\frac{1}{2}m(\theta'_N)^2\geq(2-\alpha)s_0(\log N)^{-2}$ and by our assumption (3.44) on x.

By (3.42), it follows that for n sufficiently large, if $x\geq\nu t_1$ and (3.44) holds, then for $t\in(t_1,t_1+1]$,

$$\begin{aligned}u^n_t(x)\geq \tfrac{1}{3}e^{-\theta'_N(x-\nu t)}e^{\frac{1}{2}(2-\alpha)s_0(\log N)^{-2}(t-t_1)}(1-\tfrac{1}{5}(2-\alpha)s_0(\log N)^{-2})&\\ -e^{-(1-\alpha)s_0(t-t_1)}R\left(\frac{n}{N}\right)^{1/2-c}.&\end{aligned} \tag{3.45}$$

If instead $t\in(t_1,(t_1+1)\wedge T]$ and $x\geq\nu t$ with $e^{-\theta'_N(x-\nu t_1)}>e^{-3(\theta'_N+m(\theta'_N)^2)}(1-e^{-\theta'_N})\frac{1}{5}(2-\alpha)s_0(\log N)^{-2}$, then if n is sufficiently large, we have $|x|\leq N$ and so by (3.40),

$$u^n_t(x)\geq g(x-\nu t)-d(\log N)^{-2}\geq \tfrac{1}{2}e^{-\kappa(x-\nu t)}-\tfrac{1}{20}e^{-\theta'_N(x-\nu t_1)}\geq\tfrac{9}{20}e^{-\theta'_N(x-\nu t)}, \tag{3.46}$$

where the second inequality follows since $g(y)\geq\frac{1}{2}e^{-\kappa y}\ \forall y\geq 0$ and by (2.1), the definition of d and our assumption on x. For $x\in[\nu t-d,\nu t]$, by (3.40),

$$u^n_t(x)\geq\tfrac{1}{2}-d(\log N)^{-2}\geq\tfrac{2}{5}e^{\theta'_N d}\geq\tfrac{2}{5}e^{-\theta'_N(x-\nu t)} \tag{3.47}$$

for n sufficiently large, since $e^{\kappa d}\leq 10/9$ by the definition of d. Since (3.41) holds for $t_1=0$ by our assumption in (3.38), for n sufficiently large that $e^{\frac{9}{40}(2-\alpha)s_0(\log N)^{-2}}(1-\frac{1}{5}(2-\alpha)s_0(\log N)^{-2})\geq 1$, (3.41) holds for each $t_1\in\frac{1}{2}\mathbb{N}_0\cap[0,T]$ by (3.45) and (3.46) and by induction. Then for $t\in[1,T]$, there exists $t_1\in[0,T]$ such that (3.41) holds and with $t-t_1\in[1/2,1]$, and the result follows by (3.45), (3.46) and (3.47). $\square$

The following result will allow us to show that $|u^n_{t,t+s}(x)-g(x-\mu^n_t-\nu s)|$ is small in the proof of Proposition 3.1.

Lemma 3.11. *Suppose* $(u_t^{n,1})_{t\geq 0}$ *and* $(u_t^{n,2})_{t\geq 0}$ *solve* (3.6) *with initial conditions* $p_0^{n,1}$ *and* $p_0^{n,2}$ *respectively. Then for* $t \geq 0$*,*

$$\sup_{x\in\frac{1}{n}\mathbb{Z}} |u_t^{n,1}(x) - u_t^{n,2}(x)| \leq e^{(1+\alpha)s_0 t} \sup_{y\in\frac{1}{n}\mathbb{Z}} |p_0^{n,1}(y) - p_0^{n,2}(y)|.$$

Proof. By (3.7), for $x \in \frac{1}{n}\mathbb{Z}$ and $t \geq 0$,

$$\begin{aligned} |u_t^{n,1}(x) - u_t^{n,2}(x)| &\leq \langle |p_0^{n,1} - p_0^{n,2}|, \phi_0^{t,x}\rangle_n + s_0 \int_0^t \langle |f(u_s^{n,1}) - f(u_s^{n,2})|, \phi_s^{t,x}\rangle_n ds \\ &\leq \sup_{y\in\frac{1}{n}\mathbb{Z}} |p_0^{n,1}(y) - p_0^{n,2}(y)| + (1+\alpha)s_0 \int_0^t \sup_{y\in\frac{1}{n}\mathbb{Z}} |u_s^{n,1}(y) - u_s^{n,2}(y)| ds \end{aligned}$$

since $\sup_{u\in[0,1]} |f'(u)| = 1+\alpha$. The result follows by Gronwall's inequality. □

We are now ready to prove Proposition 3.1.

Proof of Proposition 3.1. Without loss of generality, assume $b_2 \in (0, 1/3)$ is sufficiently small that $\left(\frac{n}{N}\right)^{1/3} \leq n^{-b_2}$ for n sufficiently large. Take $c_5, c_6 > 0$ as defined in Lemma 3.4 and Proposition 3.5. Let $b_1 = \frac{1}{2}(c_5 \wedge c_6)$, and suppose condition (A) holds. Define the event

$$A = \left\{p_t^n(x) = 0 \ \forall t \in [0, 2N^2], x \geq N^5\right\} \cap \left\{p_t^n(x) = 1 \ \forall t \in [0, 2N^2], x \leq -N^5\right\}.$$

Recall from (2.8) that $D_n^+ = (1/2 - c_0)\kappa^{-1}\log(N/n)$. Take $c_3 \in (0, c_0 \wedge 1/6)$, and take $\ell' \in \mathbb{N}$ sufficiently large that $N^2\left(\frac{n}{N}\right)^{\ell'} \leq \left(\frac{n}{N}\right)^{\ell+1}$ for n sufficiently large. Take $c_4 = c_4(c_3, \ell') \in (0, 1/2)$ as defined in Proposition 3.2, and let $T_0 = (\log N)^{c_4}$. By making c_4 smaller if necessary, we can assume $c_4 < a_0$ (recall from Section 1.2 that $(\log N)^{a_0} \leq \log n$ for n sufficiently large). For $k \in \mathbb{Z}$, let $t_k = (k+1)T_0$, and for $k \in \mathbb{N}_0$, let $(u_t^{n,k})_{t\geq 0}$ denote the solution of

$$\begin{cases} \partial_t u_t^{n,k} &= \frac{1}{2}m\Delta_n u_t^{n,k} + s_0 f(u_t^{n,k}) \quad \text{for } t > 0, \\ u_0^{n,k} &= p_{t_{k-1}}^n. \end{cases}$$

For $k \in \mathbb{N}_0$, define the event

$$A_k = \left\{ \sup_{x\in\frac{1}{n}\mathbb{Z}, |x|\leq N^5} \sup_{t\in[0,2T_0]} |p_{t+t_{k-1}}^n(x) - u_t^{n,k}(x)| \leq \left(\frac{n}{N}\right)^{1/2-c_3} \right\}.$$

Let $j_0 = \lfloor N^2 T_0^{-1} \rfloor$. Note that by a union bound, and then by Proposition 3.2 and Lemma 3.3, for n sufficiently large,

$$\mathbb{P}\left(A^c \cup \bigcup_{j=0}^{j_0+1} A_j^c\right) \leq 2e^{-N^5} + (j_0+2)\left(\frac{n}{N}\right)^{\ell'} \leq \left(\frac{n}{N}\right)^{\ell} \tag{3.48}$$

by our choice of ℓ'. From now on, suppose that the event $A \cap \bigcap_{j=0}^{j_0+1} A_j$ occurs.

For $k \in \mathbb{N}_0$, let $(u_t^k)_{t\geq 0}$ denote the solution of

$$\begin{cases} \partial_t u_t^k &= \frac{1}{2}m\Delta u_t^k + s_0 f(u_t^k) \quad \text{for } t > 0, \\ u_0^k &= \bar{p}_{t_{k-1}}^n, \end{cases}$$

where $\bar{p}_{t_{k-1}}^n : \mathbb{R} \to [0,1]$ is the linear interpolation of $p_{t_{k-1}}^n : \frac{1}{n}\mathbb{Z} \to [0,1]$.

Now for an induction argument, for $k \in \mathbb{N}_0$ with $k \le j_0 + 1$, suppose there exists $z_{k-1} \in \mathbb{R}$ with $|z_{k-1}| \le k$ such that

$$D_k := \sup_{x \in \frac{1}{n}\mathbb{Z}} |p^n_{t_{k-1}}(x) - g(x - \nu t_{k-1} - z_{k-1})| \le \tfrac{1}{2}(c_5 \wedge c_6) = b_1 \tag{3.49}$$

$$\text{and} \quad \sup_{x_1, x_2 \in \frac{1}{n}\mathbb{Z}, |x_1 - x_2| \le n^{-1/3}} |p^n_{t_{k-1}}(x_1) - p^n_{t_{k-1}}(x_2)| \le n^{-b_2}. \tag{3.50}$$

(Note that (3.49) and (3.50) hold for $k = 0$, by condition (A).) Then by the triangle inequality,

$$\begin{aligned} \|\bar{p}^n_{t_{k-1}} - g(\cdot - \nu t_{k-1} - z_{k-1})\|_\infty &\le D_k + n^{-1}\|\nabla g\|_\infty + n^{-b_2} \\ &\le c_5 \wedge c_6 \end{aligned} \tag{3.51}$$

for n sufficiently large. Hence by Proposition 3.5, there exists $z_k \in \mathbb{R}$ with $|z_k| \le k+1$ such that

$$|u^k_t(x) - g(x - \nu(t_{k-1} + t) - z_k)| \le C_3 e^{-c_6 t} \quad \forall x \in \mathbb{R},\ t > 0. \tag{3.52}$$

Therefore by Lemma 3.6, for $t \in [0, 2T_0]$,

$$\sup_{x \in \frac{1}{n}\mathbb{Z}} |u^{n,k}_t(x) - g(x - \nu(t_{k-1} + t) - z_k)| \le (C_4 n^{-1/3} + 2n^{-b_2}) 4T_0^2 e^{2(1+\alpha)s_0 T_0} + C_3 e^{-c_6 t}. \tag{3.53}$$

Then by the definition of the event A_k, for $t \in [T_0, 2T_0]$,

$$\begin{aligned} &\sup_{x \in \frac{1}{n}\mathbb{Z}, |x| \le N^5} |p^n_{t_{k-1}+t}(x) - g(x - \nu(t_{k-1} + t) - z_k)| \\ &\le \left(\frac{n}{N}\right)^{1/2 - c_3} + (C_4 n^{-1/3} + 2n^{-b_2}) 4T_0^2 e^{2(1+\alpha)s_0 T_0} + C_3 e^{-c_6 T_0} \\ &\le e^{-\frac{1}{2}c_6 T_0} \end{aligned}$$

for n sufficiently large (since $c_4 < a_0$). Therefore, for n sufficiently large, since $k \le j_0 + 1$ and $|z_k| \le k+1$, and by the definition of the event A, we have that for $t \in [T_0, 2T_0]$,

$$\begin{aligned} &\sup_{x \in \frac{1}{n}\mathbb{Z}} |p^n_{t_{k-1}+t}(x) - g(x - \nu(t_{k-1} + t) - z_k)| \\ &\quad \le \max\left(e^{-\frac{1}{2}c_6 T_0}, \sup_{y \ge N^5 - N^3} g(y), \sup_{y \le -N^5 + N^2} (1 - g(y)) \right) = e^{-\frac{1}{2}c_6 T_0}. \end{aligned} \tag{3.54}$$

By the definitions of the events A_k and A, and then by Lemma 3.7 and our choice of b_2 and c_3, we have that

$$\begin{aligned} \sup_{x_1, x_2 \in \frac{1}{n}\mathbb{Z}, |x_1 - x_2| \le n^{-1/3}} |p^n_{t_k}(x_1) - p^n_{t_k}(x_2)| &\le n^{-1} \lfloor n^{2/3} \rfloor \sup_{x \in \frac{1}{n}\mathbb{Z}} |\nabla_n u^{n,k}_{T_0}(x)| + 2\left(\frac{n}{N}\right)^{1/2 - c_3} \\ &\le n^{-b_2} \end{aligned}$$

for n sufficiently large. By induction, we now have that for n sufficiently large, for $k \in \mathbb{N}$ with $k \le j_0 + 1$, there exists $z_{k-1} \in \mathbb{R}$ with $|z_{k-1}| \le k$ such that (3.49) and (3.50) hold with $D_k \le e^{-\frac{1}{2}c_6 T_0}$. By Lemma 3.4 and (3.51), if n is sufficiently large then for $t \ge 0$ and $x \in \mathbb{R}$,

$$|u^k_t(x) - g(x - \nu(t_{k-1} + t) - z_{k-1})| \le C_2(D_k + 2n^{-b_2})$$

and so by (3.52), $\|g(\cdot - z_k) - g(\cdot - z_{k-1})\|_\infty \le C_2(D_k + 2n^{-b_2})$. For n sufficiently large, since $\nabla g(0) = -\kappa/4$, it follows that

$$|z_{k-1} - z_k| \le 5\kappa^{-1} C_2 (D_k + 2n^{-b_2}) \le e^{-\frac{1}{3}c_6 T_0}.$$

Therefore, by (3.54), for n sufficiently large, for $k \in \mathbb{N}_0$ with $k \le j_0$,

$$|z_{k+1} - z_k| \le e^{-\frac{1}{3}c_6 T_0} \quad \text{and} \quad \sup_{t \in [t_k, t_{k+1}],\, x \in \frac{1}{n}\mathbb{Z}} |p^n_t(x) - g(x - \nu t - z_k)| \le e^{-\frac{1}{2}c_6 T_0}. \tag{3.55}$$

Note that for $k \in \mathbb{N}_0$ with $k \le j_0$, by (3.55),

$$\begin{aligned} &\sup_{x \in \frac{1}{n}\mathbb{Z}, |x - (z_k + \nu t_k)| \le N,\, t \in [0, T_0]} |u^{n,k+1}_t(x) - g(x - \nu(t + t_k) - z_k)| \\ &\le e^{-\frac{1}{2}c_6 T_0} + \sup_{|x| \le N^5,\, t \in [0, T_0]} |u^{n,k+1}_t(x) - p^n_{t+t_k}(x)| \\ &\le e^{-\frac{1}{2}c_6 T_0} + \left(\frac{n}{N}\right)^{1/2 - c_3} \end{aligned} \tag{3.56}$$

by the definition of the event A_{k+1}.

We now use Lemma 3.9 to prove an upper bound on $p^n_t(x)$ for large x. Let $c_9 = c_7 \wedge c_8 \in (0,1)$ and $R_0 = e^{-\frac{1}{2}c_6 T_0} \left(\frac{n}{N}\right)^{-(1/2 - c_3)}$. Define $(R_k)_{k=1}^\infty$ inductively by letting $R_k = (1 - c_9) R_{k-1} + 1$ for $k \ge 1$. Let

$$k^* = \frac{\log(2c_9^{-1}) - \log R_0}{\log(1 - c_9/2)}.$$

Then since $R_k \le (1 - c_9/2) R_{k-1}$ if $R_{k-1} \ge 2c_9^{-1}$ and $R_k \le 2c_9^{-1} - 1$ if $R_{k-1} \le 2c_9^{-1}$, we have $R_k \le 2c_9^{-1}$ for $k \ge k^*$. Suppose n is sufficiently large that $e^{-\frac{1}{2}c_6 T_0} \le c_9$ and $e^{-\frac{1}{2}c_6 T_0} + \left(\frac{n}{N}\right)^{1/2 - c_3} \le c_9 (\log N)^{-2}$. Then by Lemma 3.9, (3.56) and the definition of the event A, for $k \in \mathbb{N}_0$ with $k \le j_0$, if

$$p^n_{t_k}(x) \le 3e^{-\kappa(1 - (\log N)^{-2})(x - \nu t_k - z_k)} + R_k \left(\frac{n}{N}\right)^{1/2 - c_3} \quad \forall x \in \tfrac{1}{n}\mathbb{Z}, \tag{3.57}$$

then for $t \in [0, T_0]$,

$$u^{n,k+1}_t(x) \le \tfrac{4}{3}\left(3e^{-\kappa(1 - (\log N)^{-2})(x - \nu(t + t_k) - z_k)} + R_k \left(\frac{n}{N}\right)^{1/2 - c_3}\right) \quad \forall x \in \tfrac{1}{n}\mathbb{Z}.$$

Therefore, by the definition of the events A_{k+1} and A, for $t \in [t_k, t_{k+1}]$ and $x \in \frac{1}{n}\mathbb{Z}$,

$$p^n_t(x) \le 4e^{-\kappa(1 - (\log N)^{-2})(x - \nu t - z_k)} + (1 + \tfrac{4}{3}R_k)\left(\frac{n}{N}\right)^{1/2 - c_3}. \tag{3.58}$$

Moreover, by Lemma 3.9 and (3.56), for $t \in [1, T_0]$ and $x \in \frac{1}{n}\mathbb{Z}$,

$$u^{n,k+1}_t(x) \le (1 - c_7(\log N)^{-2}) 3e^{-\kappa(1 - (\log N)^{-2})(x - \nu(t + t_k) - z_k)} + (1 - c_7) R_k \left(\frac{n}{N}\right)^{1/2 - c_3},$$

and so by the definition of the events A_{k+1} and A, for $x \in \frac{1}{n}\mathbb{Z}$,

$$\begin{aligned} p^n_{t_{k+1}}(x) &\le (1 - c_7(\log N)^{-2}) 3e^{-\kappa(1 - (\log N)^{-2})(x - \nu t_{k+1} - z_k)} + (1 + (1 - c_7) R_k)\left(\frac{n}{N}\right)^{1/2 - c_3} \\ &\le 3e^{-\kappa(1 - (\log N)^{-2})(x - \nu t_{k+1} - z_{k+1})} + R_{k+1}\left(\frac{n}{N}\right)^{1/2 - c_3} \end{aligned}$$

for n sufficiently large, by the definition of R_{k+1} and since $|z_k - z_{k+1}| \le e^{-\frac{1}{3}c_6 T_0}$ by (3.55). Note that (3.57) holds for $k = 0$ by (3.55) and the definition of R_0, and since $g(y) \le e^{-\kappa y} \wedge 1$ $\forall y \in \mathbb{R}$. Hence by induction, (3.57) holds for each $0 \le k \le j_0$. Therefore, by (3.58), for $k \ge k^*$, for $t \in [t_k, t_{k+1}]$ and $x \in \frac{1}{n}\mathbb{Z}$,

$$p^n_t(x) \le 4e^{-\kappa(1 - (\log N)^{-2})(x - \nu t - z_k)} + (1 + \tfrac{8}{3}c_9^{-1})\left(\frac{n}{N}\right)^{1/2 - c_3}. \tag{3.59}$$

We now use Lemma 3.10 to establish a corresponding lower bound. By Lemma 3.10 and (3.56), if for some $k \in \mathbb{N}_0$ with $k \le j_0$

$$p^n_{t_k}(x) \ge \tfrac{1}{3} e^{-\kappa(1+(\log N)^{-2})(x-\nu t_k - z_k)} \mathbb{1}_{x \ge \nu t_k + z_k} - R_k \left(\frac{n}{N}\right)^{1/2-c_3} \quad \forall x \in \tfrac{1}{n}\mathbb{Z}, \tag{3.60}$$

then for $t \in [0, T_0]$,

$$u^{n,k+1}_t(x) \ge \tfrac{1}{4} e^{-\kappa(1+(\log N)^{-2})(x-\nu(t+t_k)-z_k)} \mathbb{1}_{x \ge \nu(t_k+t)+z_k} - R_k \left(\frac{n}{N}\right)^{1/2-c_3} \quad \forall x \in \tfrac{1}{n}\mathbb{Z}.$$

Hence by the definition of the event A_{k+1} and since $p^n_t \ge 0$, for $t \in [t_k, t_{k+1}]$ and $x \in \frac{1}{n}\mathbb{Z}$,

$$p^n_t(x) \ge \tfrac{1}{4} e^{-\kappa(1+(\log N)^{-2})(x-\nu t - z_k)} \mathbb{1}_{x \ge \nu t + z_k} - (1+R_k) \left(\frac{n}{N}\right)^{1/2-c_3}. \tag{3.61}$$

Moreover, by Lemma 3.10 and (3.56), for $t \in [1, T_0]$ and $x \in \frac{1}{n}\mathbb{Z}$,

$$\begin{aligned} u^{n,k+1}_t(x) \ge (1+c_8(\log N)^{-2}) \tfrac{1}{3} e^{-\kappa(1+(\log N)^{-2})(x-\nu(t+t_k)-z_k)} \mathbb{1}_{x \ge \nu(t_k+t)+z_k-c_8} \\ - (1-c_8) R_k \left(\frac{n}{N}\right)^{1/2-c_3}, \end{aligned}$$

and so by the definition of the event A_{k+1} and since $p^n_t \ge 0$, for $x \in \frac{1}{n}\mathbb{Z}$,

$$\begin{aligned} p^n_{t_{k+1}}(x) &\ge (1+c_8(\log N)^{-2}) \tfrac{1}{3} e^{-\kappa(1+(\log N)^{-2})(x-\nu t_{k+1}-z_k)} \mathbb{1}_{x \ge \nu t_{k+1}+z_k-c_8} \\ &\qquad - ((1-c_8)R_k + 1) \left(\frac{n}{N}\right)^{1/2-c_3} \\ &\ge \tfrac{1}{3} e^{-\kappa(1+(\log N)^{-2})(x-\nu t_{k+1}-z_{k+1})} \mathbb{1}_{x \ge \nu t_{k+1}+z_{k+1}} - R_{k+1} \left(\frac{n}{N}\right)^{1/2-c_3} \end{aligned}$$

for n sufficiently large, by the definition of R_{k+1} and since $|z_k - z_{k+1}| \le e^{-\frac{1}{3}c_6 T_0}$. By (3.55) and the definition of R_0, and since $g(z) \ge \frac{1}{2} e^{-\kappa z}$ for $z \ge 0$, (3.60) holds for $k = 0$. Hence by induction, (3.60) holds for each $0 \le k \le j_0$. Then by (3.61), for $k \ge k^*$, for $t \in [t_k, t_{k+1}]$ and $x \in \frac{1}{n}\mathbb{Z}$,

$$p^n_t(x) \ge \tfrac{1}{4} e^{-\kappa(1+(\log N)^{-2})(x-\nu t - z_k)} \mathbb{1}_{x \ge \nu t + z_k} - (1+2c_9^{-1}) \left(\frac{n}{N}\right)^{1/2-c_3}. \tag{3.62}$$

We are now ready to complete the proof. Take $c_2 \in (0, c_4)$. Recall from (1.13) that for $t \ge 0$, $\mu^n_t = \sup\{x \in \frac{1}{n}\mathbb{Z} : p^n_t(x) \ge 1/2\}$. By (3.55) and since $\nabla g(0) = -\kappa/4$, for n sufficiently large, for $k \in \mathbb{N}_0$ with $k \le j_0$, for $t \in [t_k, t_{k+1}]$,

$$|(\nu t + z_k) - \mu^n_t| \le 5\kappa^{-1} e^{-\frac{1}{2}c_6 T_0}. \tag{3.63}$$

Therefore, for n sufficiently large, by (3.55),

$$\sup_{x \in \frac{1}{n}\mathbb{Z}, t \in [T_0, N^2]} |p^n_t(x) - g(x - \mu^n_t)| \le e^{-\frac{1}{2}c_6 T_0} + 5\kappa^{-1} e^{-\frac{1}{2}c_6 T_0} \|\nabla g\|_\infty \le e^{-2(\log N)^{c_2}} \tag{3.64}$$

since $c_2 < c_4$. By (3.63) and since $|z_0| \le 1$ and $|z_k - z_{k-1}| \le e^{-\frac{1}{3}c_6 T_0}$ $\forall k \in \mathbb{N}$ with $k \le j_0$, if n is sufficiently large we have $|\mu^n_{\log N}| \le 2\nu \log N$ and for $t \in [\log N, N^2]$ and $s \in [0,1]$ with $t + s \le N^2$,

$$|\mu^n_{t+s} - \mu^n_t - \nu s| \le 10\kappa^{-1} e^{-\frac{1}{2}c_6 T_0} + e^{-\frac{1}{3}c_6 T_0} \le e^{-(\log N)^{c_2}}.$$

Now for $t \in [\frac{1}{2}(\log N)^2, N^2]$, take $x \in \frac{1}{n}\mathbb{Z}$ such that $g(x - \mu^n_t) \le 2e^{-(\log N)^{c_2}}$. Then for n sufficiently large that $k^* \le \frac{1}{2}(\log N)^{3/2}$, by (3.59) and (3.63),

$$p^n_t(x) \le 4e^{-\kappa(1-(\log N)^{-2})(x-\mu^n_t - 5\kappa^{-1}e^{-\frac{1}{2}c_6 T_0})} + (1+\tfrac{8}{3}c_9^{-1}) \left(\frac{n}{N}\right)^{1/2-c_3} \le 5g((x-\mu^n_t) \wedge (D^+_n + 2)))$$

for n sufficiently large, since $\kappa D_n^+(\log N)^{-1} \leq 1/2$, $c_3 < c_0$ and $g(y) \sim e^{-\kappa y}$ as $y \to \infty$. Similarly, for n sufficiently large, by (3.62) and (3.63), if $x - \mu_t^n \leq D_n^+ + 2$ then

$$p_t^n(x) \geq \tfrac{1}{4} e^{-\kappa(1+(\log N)^{-2})(x-\mu_t^n+5\kappa^{-1}e^{-\frac{1}{2}c_6 T_0})} - (1+2c_9^{-1})\left(\frac{n}{N}\right)^{1/2-c_3} \geq \tfrac{1}{5} g(x-\mu_t^n).$$

If instead $g(x-\mu_t^n) \geq 2e^{-(\log N)^{c_2}}$, then $p_t^n(x) \in [\frac{1}{2}g(x-\mu_t^n), \frac{3}{2}g(x-\mu_t^n)]$ by (3.64).

Finally, for $t \in [\log N, N^2]$, let $(\tilde{u}^n_{t,t+s})_{s\geq 0}$ solve (3.2) with $\tilde{u}^n_{t,t}(x) = g(x-\mu_t^n)$ for $x \in \frac{1}{n}\mathbb{Z}$. Recall the definition of γ_n in (2.4). Then for $s \in [0,\gamma_n]$, by Lemma 3.11 and (3.64),

$$\begin{aligned}
&\sup_{x\in\frac{1}{n}\mathbb{Z}} |u^n_{t,t+s}(x) - g(x-\mu_t^n - \nu s)| \\
&\leq e^{(1+\alpha)s_0\gamma_n} e^{-2(\log N)^{c_2}} + \sup_{x\in\frac{1}{n}\mathbb{Z}} |\tilde{u}^n_{t,t+s}(x) - g(x-\mu_t^n-\nu s)| \\
&\leq e^{(1+\alpha)s_0\gamma_n} e^{-2(\log N)^{c_2}} + (C_4 + \|\nabla g\|_\infty) n^{-1/3}\gamma_n^2 e^{(1+\alpha)s_0\gamma_n} \\
&\leq e^{-(\log N)^{c_2}}
\end{aligned}$$

for n sufficiently large, where the second inequality follows by Lemma 3.6 and since $(g(\cdot - \mu_t^n - \nu s))_{s\geq 0}$ solves (3.16). The result follows by (3.48) and by the definitions of E_1 in (2.10) and E_1' in (3.3). □

3.1 Proof of Proposition 3.2

The proof of Proposition 3.2 uses similar arguments to those in [14]. The following lemma is the main step in the proof.

Lemma 3.12. *Suppose $\phi : [0,\infty) \times \frac{1}{n}\mathbb{Z} \to \mathbb{R}$ is continuously differentiable in t, and write $\phi_t(x) := \phi(t,x)$. Suppose that for any $t > 0$,*

$$\sup_{s\in[0,t]} \langle |\phi_s|, 1\rangle_n < \infty \quad \text{and} \quad \sup_{s\in[0,t]} \langle |\partial_s \phi_s|, 1\rangle_n < \infty.$$

Then for $t \geq 0$,

$$\begin{aligned}
&\langle q_t^n, \phi_t\rangle_n - \langle q_0^n, \phi_0\rangle_n - \int_0^t \langle q_s^n, \partial_s\phi_s\rangle_n ds \\
&\quad = s_0 \int_0^t \langle q_s^n(1-p_s^n)(2p_s^n - 1 + \alpha), \phi_s\rangle_n ds + \tfrac{1}{2} m \int_0^t \langle q_s^n, \Delta_n \phi_s\rangle_n ds + M_t^n(\phi), \qquad (3.65)
\end{aligned}$$

where $(M_t^n(\phi))_{t\geq 0}$ is a martingale with $M_0^n(\phi) = 0$ and

$$\langle M^n(\phi)\rangle_t \leq \frac{n}{N} \int_0^t \langle (1+m)q_s^n(\cdot) + \tfrac{1}{2}m(q_s^n(\cdot - n^{-1}) + q_s^n(\cdot + n^{-1})), \phi_s^2\rangle_n ds.$$

Before proving Lemma 3.12, we prove the following useful consequence.

Corollary 3.13. *For $a \in \mathbb{R}$, $t \geq 0$ and $z \in \frac{1}{n}\mathbb{Z}$,*

$$\begin{aligned}
q_t^n(z) &= e^{-at}\langle q_0^n, \phi_0^{t,z}\rangle_n \\
&\quad + \int_0^t e^{-a(t-s)} \langle q_s^n(s_0(1-p_s^n)(2p_s^n - 1 + \alpha) + a), \phi_s^{t,z}\rangle_n ds + M_t^n(\phi^{t,z,a}). \qquad (3.66)
\end{aligned}$$

Proof. Recall the definitions of $\phi^{t,z}$ and $\phi^{t,z,a}$ in (3.4) and (3.5). Note that $\partial_s \phi_s^{t,z} + \frac{1}{2} m \Delta_n \phi_s^{t,z} = 0$ for $s \in (0,t)$. Hence

$$\partial_s \phi_s^{t,z,a} + \tfrac{1}{2} m \Delta_n \phi_s^{t,z,a} = a\phi_s^{t,z,a}.$$

Therefore, by substituting $\phi_s(x) := \phi_s^{t,z,a}(x)$ into (3.65) in Lemma 3.12 we have

$$\langle q_t^n, \phi_t^{t,z,a}\rangle_n = \langle q_0^n, \phi_0^{t,z,a}\rangle_n + \int_0^t \langle q_s^n(s_0(1-p_s^n)(2p_s^n - 1 + \alpha) + a), \phi_s^{t,z,a}\rangle_n ds + M_t^n(\phi^{t,z,a}).$$

Since $\phi_t^{t,z,a}(w) = n\mathbb{1}_{w=z}$, the result follows. □

Proof of Lemma 3.12. For $t \geq 0$, $x \in \frac{1}{n}\mathbb{Z}$ and $i \in [N]$, by the definition of η^n in (3.9) we have that

$$\begin{aligned}
\eta_t^n(x,i) = \eta_0^n(x,i) &+ \sum_{j\in[N]\setminus\{i\}} \int_0^t (\eta_{s-}^n(x,j) - \eta_{s-}^n(x,i)) d\mathcal{P}_s^{x,i,j} \\
&+ \sum_{j\in[N]\setminus\{i\}} \int_0^t \xi_{s-}^n(x,j)(\eta_{s-}^n(x,j) - \eta_{s-}^n(x,i)) d\mathcal{S}_s^{x,i,j} \\
&+ \sum_{j\neq k\in[N]\setminus\{i\}} \int_0^t \mathbb{1}_{\xi_{s-}^n(x,j)=\xi_{s-}^n(x,k)}(\eta_{s-}^n(x,j) - \eta_{s-}^n(x,i)) d\mathcal{Q}_s^{x,i,j,k} \\
&+ \sum_{j\in[N], y\in\{x-n^{-1},\, x+n^{-1}\}} \int_0^t (\eta_{s-}^n(y,j) - \eta_{s-}^n(x,i)) d\mathcal{R}_s^{x,i,y,j}.
\end{aligned}$$

Recall from (3.10) that $q_s^n(y) = N^{-1}\sum_{j\in[N]} \eta_s^n(y,j)$ for $y \in \frac{1}{n}\mathbb{Z}$ and $s \geq 0$. By integration by parts applied to $\eta_t^n(x,i)\phi_t(x)$, and then summing over i and x, using our assumptions on ϕ,

$$\begin{aligned}
&\langle q_t^n, \phi_t\rangle_n - \langle q_0^n, \phi_0\rangle_n - \int_0^t \langle q_s^n, \partial_s\phi_s\rangle_n ds \\
&\quad= \frac{1}{Nn} \sum_{x\in\frac{1}{n}\mathbb{Z}} \sum_{i=1}^N \sum_{j\in[N]\setminus\{i\}} \int_0^t (\eta_{s-}^n(x,j) - \eta_{s-}^n(x,i))\phi_s(x) d\mathcal{P}_s^{x,i,j} \\
&\qquad+ \frac{1}{Nn} \sum_{x\in\frac{1}{n}\mathbb{Z}} \sum_{i=1}^N \sum_{j\in[N]\setminus\{i\}} \int_0^t \xi_{s-}^n(x,j)(\eta_{s-}^n(x,j) - \eta_{s-}^n(x,i))\phi_s(x) d\mathcal{S}_s^{x,i,j} \\
&\qquad+ \frac{1}{Nn} \sum_{x\in\frac{1}{n}\mathbb{Z}} \sum_{i=1}^N \sum_{j\neq k\in[N]\setminus\{i\}} \int_0^t \mathbb{1}_{\xi_{s-}^n(x,j)=\xi_{s-}^n(x,k)}(\eta_{s-}^n(x,j) - \eta_{s-}^n(x,i))\phi_s(x) d\mathcal{Q}_s^{x,i,j,k} \\
&\qquad+ \frac{1}{Nn} \sum_{x\in\frac{1}{n}\mathbb{Z}} \sum_{i=1}^N \sum_{j\in[N], y\in\{x-n^{-1},\, x+n^{-1}\}} \int_0^t (\eta_{s-}^n(y,j) - \eta_{s-}^n(x,i))\phi_s(x) d\mathcal{R}_s^{x,i,y,j}.
\end{aligned} \tag{3.67}$$

We shall consider each line on the right hand side of (3.67) separately. For the first line,

$$\begin{aligned}
A_t^1 &:= \frac{1}{Nn} \sum_{x\in\frac{1}{n}\mathbb{Z}} \sum_{i=1}^N \sum_{j\in[N]\setminus\{i\}} \int_0^t (\eta_{s-}^n(x,j) - \eta_{s-}^n(x,i))\phi_s(x) d\mathcal{P}_s^{x,i,j} \\
&= \frac{1}{Nn} \sum_{x\in\frac{1}{n}\mathbb{Z}} \sum_{i=1}^N \sum_{j\in[N]\setminus\{i\}} \int_0^t (\eta_{s-}^n(x,j) - \eta_{s-}^n(x,i))\phi_s(x)(d\mathcal{P}_s^{x,i,j} - r_n(1-(\alpha+1)s_n)ds) \\
&\quad+ \frac{1}{Nn} \sum_{x\in\frac{1}{n}\mathbb{Z}} \sum_{i=1}^N \sum_{j\in[N]\setminus\{i\}} \int_0^t (\eta_{s-}^n(x,j) - \eta_{s-}^n(x,i))\phi_s(x) r_n(1-(\alpha+1)s_n) ds.
\end{aligned}$$

Now for $x \in \frac{1}{n}\mathbb{Z}$ and $s \in [0,t]$,

$$\sum_{i=1}^{N} \sum_{j\in[N]\setminus\{i\}} (\eta^n_{s-}(x,j) - \eta^n_{s-}(x,i)) = 0.$$

Hence

$$\begin{aligned} A^1_t &= M^{n,1}_t(\phi) \\ &:= \frac{1}{Nn} \sum_{x\in\frac{1}{n}\mathbb{Z}} \sum_{i=1}^{N} \sum_{j\in[N]\setminus\{i\}} \int_0^t (\eta^n_{s-}(x,j) - \eta^n_{s-}(x,i))\phi_s(x)(d\mathcal{P}^{x,i,j}_s - r_n(1-(\alpha+1)s_n)ds), \end{aligned} \tag{3.68}$$

which is a martingale (since we assumed $\sup_{s\in[0,t']}\langle|\phi_s|,1\rangle_n < \infty$ for any $t'>0$). For the second line on the right hand side of (3.67),

$$\begin{aligned} A^2_t &:= \frac{1}{Nn} \sum_{x\in\frac{1}{n}\mathbb{Z}} \sum_{i=1}^{N} \sum_{j\in[N]\setminus\{i\}} \int_0^t \xi^n_{s-}(x,j)(\eta^n_{s-}(x,j) - \eta^n_{s-}(x,i))\phi_s(x)d\mathcal{S}^{x,i,j}_s \\ &= \frac{1}{Nn} \sum_{x\in\frac{1}{n}\mathbb{Z}} \sum_{i=1}^{N} \sum_{j\in[N]\setminus\{i\}} \int_0^t \xi^n_{s-}(x,j)(\eta^n_{s-}(x,j) - \eta^n_{s-}(x,i))\phi_s(x)(d\mathcal{S}^{x,i,j}_s - r_n\alpha s_n ds) \\ &\quad + \frac{1}{Nn} \sum_{x\in\frac{1}{n}\mathbb{Z}} \sum_{i=1}^{N} \sum_{j\in[N]\setminus\{i\}} \int_0^t \xi^n_{s-}(x,j)(\eta^n_{s-}(x,j) - \eta^n_{s-}(x,i))\phi_s(x)r_n\alpha s_n ds. \end{aligned}$$

For the expression on the last line, for $x \in \frac{1}{n}\mathbb{Z}$ and $s \in [0,t]$, since $\xi^n_{s-}(x,j) = 1$ if $\eta^n_{s-}(x,j) = 1$,

$$\begin{aligned} &\sum_{i=1}^{N} \sum_{j\in[N]\setminus\{i\}} \xi^n_{s-}(x,j)(\eta^n_{s-}(x,j) - \eta^n_{s-}(x,i)) \\ &\quad = \sum_{i=1}^{N} \sum_{j\in[N]\setminus\{i\}} \eta^n_{s-}(x,j) - \sum_{i=1}^{N} \eta^n_{s-}(x,i)\left(\sum_{j=1}^{N} \xi^n_{s-}(x,j) - 1\right) \\ &\quad = (N-1)Nq^n_{s-}(x) - Nq^n_{s-}(x)(Np^n_{s-}(x)-1) \\ &\quad = N^2 q^n_{s-}(x)(1-p^n_{s-}(x)). \end{aligned}$$

Therefore we can write

$$\begin{aligned} &\frac{1}{Nn} \sum_{x\in\frac{1}{n}\mathbb{Z}} \sum_{i=1}^{N} \sum_{j\in[N]\setminus\{i\}} \int_0^t \xi^n_{s-}(x,j)(\eta^n_{s-}(x,j) - \eta^n_{s-}(x,i))\phi_s(x)r_n\alpha s_n ds \\ &= \alpha N r_n s_n \int_0^t \langle q^n_{s-}(1-p^n_{s-}), \phi_s\rangle_n ds. \end{aligned}$$

Hence, since $Nr_ns_n = s_0$ by (1.11),

$$A^2_t = \alpha s_0 \int_0^t \langle q^n_s(1-p^n_s), \phi_s\rangle_n ds + M^{n,2}_t(\phi), \tag{3.69}$$

where

$$M^{n,2}_t(\phi) := \frac{1}{Nn} \sum_{x\in\frac{1}{n}\mathbb{Z}} \sum_{i=1}^{N} \sum_{j\in[N]\setminus\{i\}} \int_0^t \xi^n_{s-}(x,j)(\eta^n_{s-}(x,j) - \eta^n_{s-}(x,i))\phi_s(x)(d\mathcal{S}^{x,i,j}_s - r_n\alpha s_n ds) \tag{3.70}$$

is a martingale. For the third line on the right hand side of (3.67),

$$
\begin{aligned}
A_t^3 &:= \frac{1}{Nn} \sum_{x \in \frac{1}{n}\mathbb{Z}} \sum_{i=1}^{N} \sum_{j \neq k \in [N] \setminus \{i\}} \int_0^t \mathbb{1}_{\xi_{s-}^n(x,j) = \xi_{s-}^n(x,k)} (\eta_{s-}^n(x,j) - \eta_{s-}^n(x,i)) \phi_s(x) d\mathcal{Q}_s^{x,i,j,k} \\
&= \frac{1}{Nn} \sum_{x \in \frac{1}{n}\mathbb{Z}} \sum_{i=1}^{N} \sum_{j \neq k \in [N] \setminus \{i\}} \int_0^t \mathbb{1}_{\xi_{s-}^n(x,j) = \xi_{s-}^n(x,k)} (\eta_{s-}^n(x,j) - \eta_{s-}^n(x,i)) \phi_s(x) \\
&\qquad \cdot (d\mathcal{Q}_s^{x,i,j,k} - \tfrac{1}{N} r_n s_n ds) \\
&\quad + \frac{1}{Nn} \sum_{x \in \frac{1}{n}\mathbb{Z}} \sum_{i=1}^{N} \sum_{j \neq k \in [N] \setminus \{i\}} \int_0^t \mathbb{1}_{\xi_{s-}^n(x,j) = \xi_{s-}^n(x,k)} (\eta_{s-}^n(x,j) - \eta_{s-}^n(x,i)) \phi_s(x) \tfrac{1}{N} r_n s_n ds.
\end{aligned}
$$

For $x \in \frac{1}{n}\mathbb{Z}$ and $s \in [0,t]$, since $\eta_{s-}^n(x,j) = 0$ if $\xi_{s-}^n(x,j) = 0$,

$$
\begin{aligned}
&\sum_{i=1}^{N} \sum_{j \neq k \in [N] \setminus \{i\}} \mathbb{1}_{\xi_{s-}^n(x,j) = \xi_{s-}^n(x,k)} (\eta_{s-}^n(x,j) - \eta_{s-}^n(x,i)) \\
&\quad = \sum_{i,j,k \in [N] \text{ distinct}} \Big(\mathbb{1}_{\eta_{s-}^n(x,j) = \xi_{s-}^n(x,k) = 1} - \mathbb{1}_{\xi_{s-}^n(x,j) = \xi_{s-}^n(x,k) = \eta_{s-}^n(x,i) = 1} \\
&\qquad\qquad - \mathbb{1}_{\xi_{s-}^n(x,j) = \xi_{s-}^n(x,k) = 0,\, \eta_{s-}^n(x,i) = 1} \Big) \\
&\quad = (N-2) N q_{s-}^n(x) (N p_{s-}^n(x) - 1) - N q_{s-}^n(x) (N p_{s-}^n(x) - 1)(N p_{s-}^n(x) - 2) \\
&\qquad - N q_{s-}^n(x) (N - N p_{s-}^n(x))(N - N p_{s-}^n(x) - 1) \\
&\quad = N^3 q_{s-}^n(x) (1 - p_{s-}^n(x))(2 p_{s-}^n(x) - 1).
\end{aligned}
$$

Therefore, since $N r_n s_n = s_0$,

$$
A_t^3 = s_0 \int_0^t \langle q_s^n (1 - p_s^n)(2 p_s^n - 1), \phi_s \rangle_n ds + M_t^{n,3}(\phi), \tag{3.71}
$$

where

$$
\begin{aligned}
&M_t^{n,3}(\phi) \\
&:= \frac{1}{Nn} \sum_{x \in \frac{1}{n}\mathbb{Z}} \sum_{i=1}^{N} \sum_{j \neq k \in [N] \setminus \{i\}} \int_0^t \mathbb{1}_{\xi_{s-}^n(x,j) = \xi_{s-}^n(x,k)} (\eta_{s-}^n(x,j) - \eta_{s-}^n(x,i)) \phi_s(x) \\
&\qquad \cdot (d\mathcal{Q}_s^{x,i,j,k} - \tfrac{1}{N} r_n s_n ds)
\end{aligned} \tag{3.72}
$$

is a martingale. Finally, for the fourth line on the right hand side of (3.67),

$$
\begin{aligned}
A_t^4 &:= \frac{1}{Nn} \sum_{x \in \frac{1}{n}\mathbb{Z}} \sum_{i=1}^{N} \sum_{j \in [N], y \in \{x - n^{-1},\, x + n^{-1}\}} \int_0^t (\eta_{s-}^n(y,j) - \eta_{s-}^n(x,i)) \phi_s(x) d\mathcal{R}_s^{x,i,y,j} \\
&= \frac{1}{Nn} \sum_{x \in \frac{1}{n}\mathbb{Z}} \sum_{i=1}^{N} \sum_{j \in [N], y \in \{x - n^{-1},\, x + n^{-1}\}} \int_0^t (\eta_{s-}^n(y,j) - \eta_{s-}^n(x,i)) \phi_s(x) (d\mathcal{R}_s^{x,i,y,j} - m r_n ds) \\
&\quad + \frac{1}{Nn} \sum_{x \in \frac{1}{n}\mathbb{Z}} \sum_{i=1}^{N} \sum_{j \in [N], y \in \{x - n^{-1},\, x + n^{-1}\}} \int_0^t (\eta_{s-}^n(y,j) - \eta_{s-}^n(x,i)) \phi_s(x) m r_n ds.
\end{aligned}
$$

For $x \in \frac{1}{n}\mathbb{Z}$ and $s \in [0,t]$,

$$
\begin{aligned}
&\sum_{i,j \in [N], y \in \{x - n^{-1},\, x + n^{-1}\}} (\eta_{s-}^n(y,j) - \eta_{s-}^n(x,i)) \\
&= N^2 (q_{s-}^n(x - n^{-1}) + q_{s-}^n(x + n^{-1})) - 2 N^2 q_{s-}^n(x).
\end{aligned}
$$

Therefore we can write

$$\begin{aligned}
&\frac{1}{Nn}\sum_{x\in\frac{1}{n}\mathbb{Z}}\sum_{i=1}^{N}\sum_{j\in[N],y\in\{x-n^{-1},\,x+n^{-1}\}}\int_0^t(\eta^n_{s-}(y,j)-\eta^n_{s-}(x,i))\phi_s(x)mr_n ds\\
&=\frac{mr_n}{Nn}\sum_{x\in\frac{1}{n}\mathbb{Z}}\int_0^t(N^2(q^n_{s-}(x-n^{-1})+q^n_{s-}(x+n^{-1}))-2N^2q^n_{s-}(x))\phi_s(x)ds\\
&=\frac{Nmr_n}{n}\sum_{x\in\frac{1}{n}\mathbb{Z}}\int_0^t q^n_{s-}(x)(\phi_s(x+n^{-1})+\phi_s(x-n^{-1})-2\phi_s(x))ds\\
&=\frac{Nmr_n}{n^2}\int_0^t\langle q^n_s,\Delta_n\phi_s\rangle_n ds,
\end{aligned}$$

where the second equality follows by summation by parts. Hence, since $Nr_nn^{-2}=\frac{1}{2}$,

$$A^4_t=\tfrac{1}{2}m\int_0^t\langle q^n_s,\Delta_n\phi_s\rangle_n ds+M^{n,4}_t(\phi), \tag{3.73}$$

where

$$\begin{aligned}
M^{n,4}_t(\phi):=\frac{1}{Nn}\sum_{x\in\frac{1}{n}\mathbb{Z}}\sum_{i=1}^{N}\sum_{j\in[N],y\in\{x-n^{-1},\,x+n^{-1}\}}\int_0^t(\eta^n_{s-}(y,j)-\eta^n_{s-}(x,i))\phi_s(x)\\
\cdot(d\mathcal{R}^{x,i,y,j}_s-mr_n ds)
\end{aligned} \tag{3.74}$$

is a martingale. Combining (3.68), (3.69), (3.71) and (3.73) with (3.67), we have that

$$\begin{aligned}
&\langle q^n_t,\phi_t\rangle_n-\langle q^n_0,\phi_0\rangle_n-\int_0^t\langle q^n_s,\partial_s\phi_s\rangle_n ds\\
&=s_0\int_0^t\langle q^n_s(1-p^n_s)(2p^n_s-1+\alpha),\phi_s\rangle_n ds+\tfrac{1}{2}m\int_0^t\langle q^n_s,\Delta_n\phi_s\rangle_n ds+M^n_t(\phi),
\end{aligned}$$

where $M^n_t(\phi):=\sum_{i=1}^4 M^{n,i}_t(\phi)$ is a martingale with $M^n_0(\phi)=0$.

It remains to bound $\langle M^n(\phi)\rangle_t$. Since $(\mathcal{P}^{x,i,j})$, $(\mathcal{S}^{x,i,j})$, $(\mathcal{Q}^{x,i,j,k})$ and $(\mathcal{R}^{x,i,y,j})$ are independent families of Poisson processes,

$$\langle M^n(\phi)\rangle_t=\sum_{i=1}^4\langle M^{n,i}(\phi)\rangle_t. \tag{3.75}$$

By the definition of $M^{n,1}(\phi)$ in (3.68), we have that for $t\ge 0$,

$$\begin{aligned}
\langle M^{n,1}(\phi)\rangle_t&=\frac{1}{N^2n^2}r_n(1-(\alpha+1)s_n)\sum_{x\in\frac{1}{n}\mathbb{Z}}\sum_{i=1}^{N}\sum_{j\in[N]\setminus\{i\}}\int_0^t(\eta^n_{s-}(x,j)-\eta^n_{s-}(x,i))^2\phi_s(x)^2ds\\
&=\frac{r_n}{n^2}(1-(\alpha+1)s_n)\int_0^t\sum_{x\in\frac{1}{n}\mathbb{Z}}2q^n_{s-}(x)(1-q^n_{s-}(x))\phi_s(x)^2ds\\
&\le\frac{r_n}{n}(1-(\alpha+1)s_n)\int_0^t\langle 2q^n_s,\phi^2_s\rangle_n ds.
\end{aligned} \tag{3.76}$$

By the same argument, by the definition of $M^{n,2}(\phi)$ in (3.70),

$$\langle M^{n,2}(\phi)\rangle_t\le\frac{r_n}{n}\alpha s_n\int_0^t\langle 2q^n_s,\phi^2_s\rangle_n ds.$$

Then by the definition of $M^{n,3}(\phi)$ in (3.72),

$$\begin{aligned}
&\langle M^{n,3}(\phi)\rangle_t \\
&= \frac{1}{N^2n^2}\frac{r_n s_n}{N}\sum_{x\in\frac{1}{n}\mathbb{Z}}\sum_{i=1}^{N}\sum_{j\neq k\in[N]\setminus\{i\}}\int_0^t \mathbb{1}_{\xi^n_{s-}(x,j)=\xi^n_{s-}(x,k)}(\eta^n_{s-}(x,j)-\eta^n_{s-}(x,i))^2\phi_s(x)^2ds \\
&\leq \frac{1}{N^2n^2}\frac{r_n s_n}{N}\sum_{x\in\frac{1}{n}\mathbb{Z}} N^3\int_0^t 2q^n_{s-}(x)(1-q^n_{s-}(x))\phi_s(x)^2ds \\
&\leq \frac{r_n}{n}s_n\int_0^t\langle 2q^n_s,\phi_s^2\rangle_n ds.
\end{aligned}$$

Finally, by the definition of $M^{n,4}(\phi)$ in (3.74),

$$\begin{aligned}
\langle M^{n,4}(\phi)\rangle_t &\leq \frac{1}{N^2n^2}mr_n\sum_{x\in\frac{1}{n}\mathbb{Z}}N^2\int_0^t(q^n_{s-}(x-n^{-1})+2q^n_{s-}(x)+q^n_{s-}(x+n^{-1}))\phi_s(x)^2ds \\
&= \frac{mr_n}{n}\int_0^t\langle q^n_s(\cdot-n^{-1})+2q^n_s(\cdot)+q^n_s(\cdot+n^{-1}),\phi_s^2\rangle_n ds.
\end{aligned}$$

By (3.75), and since $r_n n^{-1}=\frac{1}{2}nN^{-1}$ by (1.11), the result follows. □

The following result, which is a version of the local central limit theorem in [24], will be used several times in the rest of the article. Recall that we let $(X^n_t)_{t\geq 0}$ denote a simple symmetric random walk on $\frac{1}{n}\mathbb{Z}$ with jump rate n^2.

Lemma 3.14 (Theorem 2.5.6 in [24])**.** *For $x\in\frac{1}{n}\mathbb{Z}$ and $t>0$ with $|x|\leq\frac{1}{2}nt$,*

$$\mathbf{P}_0\left(X^n_t=x\right)=\frac{1}{n}\frac{1}{\sqrt{2\pi t}}e^{-\frac{x^2}{2t}}e^{\mathcal{O}(n^{-1}t^{-1/2}+n^{-1}|x|^3t^{-2})}.$$

The next lemma gives us useful bounds on $\langle M^n(\phi^{t,z})\rangle_t$.

Lemma 3.15. *There exists a constant $C_6<\infty$ such that for $t\geq 0$, $s\in[0,t]$ and $z\in\frac{1}{n}\mathbb{Z}$,*

$$\langle 1,(\phi^{t,z}_s)^2\rangle_n = n\mathbf{P}_0\left(X^n_{2m(t-s)}=0\right),\qquad \int_0^t\langle 1,(\phi^{t,z}_s)^2\rangle_n ds\leq C_6t^{1/2} \tag{3.77}$$

$$\text{and}\qquad \langle M^n(\phi^{t,z})\rangle_t\leq C_6t^{1/2}\frac{n}{N}. \tag{3.78}$$

Proof. For $s\in[0,t]$, by the definition of $\phi^{t,z}_s$ in (3.4) and by translational invariance,

$$\begin{aligned}
\sum_{x\in\frac{1}{n}\mathbb{Z}}\phi^{t,z}_s(x)^2 &= n^2\sum_{x\in\frac{1}{n}\mathbb{Z}}\mathbf{P}_0\left(X^n_{m(t-s)}=x\right)^2 \\
&= n^2\sum_{x\in\frac{1}{n}\mathbb{Z}}\mathbf{P}_0\left(X^n_{m(t-s)}=-x\right)\mathbf{P}_0\left(X^n_{m(t-s)}=x\right) \\
&= n^2\mathbf{P}_0\left(X^n_{2m(t-s)}=0\right),
\end{aligned} \tag{3.79}$$

where the second line follows by symmetry. (This argument is used in (54) of [14].) By Lemma 3.14, for $t_0>0$,

$$\int_0^{t_0}n\mathbf{P}_0\left(X^n_s=0\right)ds\leq\min(nt_0,n^{-1})+\int_{t_0\wedge n^{-2}}^{t_0}(2\pi s)^{-1/2}e^{\mathcal{O}(1)}ds\leq K_3t_0^{1/2},$$

for some constant K_3. By (3.79), the first statement (3.77) follows, and the second statement (3.78) follows by Lemma 3.12 and since $q^n_s\in[0,1]$. □

We will use the following lemma in the proof of Proposition 3.2, and also later on in Section 4.

Lemma 3.16. *For $k \in \mathbb{N}$, $t \geq 0$ and $z \in \frac{1}{n}\mathbb{Z}$,*

$$\begin{aligned}
&|q_t^n(z) - v_t^n(z)|^k \\
&\leq 3^{2k-1} s_0^k t^{k-1} \left(\int_0^t \langle |q_s^n - v_s^n|^k, \phi_s^{t,z} \rangle_n ds + \int_0^t \sup_{x \in \frac{1}{n}\mathbb{Z}} v_s^n(x)^k \langle |p_s^n - u_s^n|^k, \phi_s^{t,z} \rangle_n ds \right) \\
&\quad + 3^{k-1} |M_t^n(\phi^{t,z})|^k.
\end{aligned}$$

Proof. By Corollary 3.13 and (3.12) with $a = 0$, for $t \geq 0$ and $z \in \frac{1}{n}\mathbb{Z}$,

$$\begin{aligned}
&|q_t^n(z) - v_t^n(z)| \\
&\leq s_0 \int_0^t |\langle (q_s^n - v_s^n)(1 - p_s^n)(2p_s^n - 1 + \alpha), \phi_s^{t,z} \rangle_n| ds \\
&\quad + s_0 \int_0^t |\langle v_s^n((1 - p_s^n)(2p_s^n - 1 + \alpha) - (1 - u_s^n)(2u_s^n - 1 + \alpha)), \phi_s^{t,z} \rangle_n| ds + |M_t^n(\phi^{t,z})|.
\end{aligned}$$

Therefore, since $|(1-u)(2u-1+\alpha)| \leq 1+\alpha$ for $u \in [0,1]$, and since $|(1-x)(2x-1+\alpha) - (1-y)(2y-1+\alpha)| \leq 3|x-y|$ for $x, y \in [0,1]$, for $k \in \mathbb{N}$,

$$\begin{aligned}
|q_t^n(z) - v_t^n(z)|^k &\leq 3^{k-1} s_0^k \left(\int_0^t \langle (1+\alpha)|q_s^n - v_s^n|, \phi_s^{t,z} \rangle_n ds \right)^k \\
&\quad + 3^{k-1} s_0^k \left(\int_0^t \langle v_s^n \cdot 3|p_s^n - u_s^n|, \phi_s^{t,z} \rangle_n ds \right)^k + 3^{k-1} |M_t^n(\phi^{t,z})|^k. \qquad (3.80)
\end{aligned}$$

Note that by the definition of $\phi^{t,z}$ in (3.4), for $s \in [0,t]$, $\langle 1, \phi_s^{t,z} \rangle_n = 1$. Hence by two applications of Jensen's inequality,

$$\begin{aligned}
\left(\int_0^t \langle (1+\alpha)|q_s^n - v_s^n|, \phi_s^{t,z} \rangle_n ds \right)^k &\leq t^{k-1} (1+\alpha)^k \int_0^t \langle |q_s^n - v_s^n|, \phi_s^{t,z} \rangle_n^k ds \\
&\leq t^{k-1} (1+\alpha)^k \int_0^t \langle |q_s^n - v_s^n|^k, \phi_s^{t,z} \rangle_n ds.
\end{aligned}$$

Similarly,

$$\left(\int_0^t \langle 3v_s^n |p_s^n - u_s^n|, \phi_s^{t,z} \rangle_n ds \right)^k \leq t^{k-1} 3^k \int_0^t \sup_{x \in \frac{1}{n}\mathbb{Z}} v_s^n(x)^k \langle |p_s^n - u_s^n|^k, \phi_s^{t,z} \rangle_n ds.$$

The result follows by (3.80). □

We will use the following form of the Burkholder-Davis-Gundy inequality (see the proof of Lemma 4 in [28]) in the proof of Proposition 3.2 and also later in Section 4.

Lemma 3.17 (Burkholder-Davis-Gundy inequality)**.** *For $k \in \mathbb{N}$ with $k \geq 2$ there exists $C(k) < \infty$ such that for $(M_t)_{t \geq 0}$ a càdlàg martingale with $M_0 = 0$, for $t \geq 0$,*

$$\mathbb{E}\left[\sup_{s \in [0,t]} |M_s|^k \right] \leq C(k) \mathbb{E}\left[\langle M \rangle_t^{k/2} + \sup_{s \in [0,t]} |M_s - M_{s-}|^k \right].$$

We are now ready to finish this section by proving Proposition 3.2.

Proof of Proposition 3.2. For $t > 0$ and $z \in \frac{1}{n}\mathbb{Z}$, by Lemma 3.12 we have that almost surely

$$\sup_{s\in[0,t]} |M^n_s(\phi^{t,z}) - M^n_{s-}(\phi^{t,z})| = \sup_{s\in[0,t]} |\langle q^n_s, \phi^{t,z}_s\rangle_n - \langle q^n_{s-}, \phi^{t,z}_s\rangle_n| \le N^{-1}.$$

It follows by Lemma 3.15 and Lemma 3.17 that for $k \ge 2$,

$$\mathbb{E}\left[\sup_{s\in[0,t]} |M^n_s(\phi^{t,z})|^k\right] \le C(k)\left(\left(C_6 t^{1/2}\frac{n}{N}\right)^{k/2} + N^{-k}\right).$$

By Lemma 3.16, and since $\langle 1, \phi^{t,z}_s\rangle_n = 1$ and $v^n_s \in [0,1]$ for $s \in [0,t]$,

$$\begin{aligned}
&\mathbb{E}\left[|q^n_t(z) - v^n_t(z)|^k\right] \\
&\le 3^{2k-1}s_0^k t^{k-1}\left(\int_0^t \sup_{x\in\frac{1}{n}\mathbb{Z}} \mathbb{E}\left[|q^n_s(x) - v^n_s(x)|^k\right] ds + \int_0^t \sup_{x\in\frac{1}{n}\mathbb{Z}} \mathbb{E}\left[|p^n_s(x) - u^n_s(x)|^k\right] ds\right) \\
&\quad + 3^{k-1}C(k)\left(\left(C_6 t^{1/2}\frac{n}{N}\right)^{k/2} + N^{-k}\right).
\end{aligned} \tag{3.81}$$

Temporarily setting $\eta^n_0 = \xi^n_0$ and so $q^n_0 = p^n_0$, we have $p^n_s = q^n_s$ and $v^n_s = u^n_s$ $\forall s \ge 0$, and by Gronwall's inequality, for $t \ge 0$,

$$\sup_{x\in\frac{1}{n}\mathbb{Z}} \mathbb{E}\left[|p^n_t(x) - u^n_t(x)|^k\right] \le 3^{k-1}C(k)\left(\left(C_6 t^{1/2}\frac{n}{N}\right)^{k/2} + N^{-k}\right) e^{3^{2k-1}2s_0^k t^k}.$$

It follows that there exists a constant $C_1 = C_1(k) < \infty$ such that for $t \ge 0$,

$$\sup_{x\in\frac{1}{n}\mathbb{Z}} \mathbb{E}\left[|p^n_t(x) - u^n_t(x)|^k\right] \le C_1\left(\frac{n^{k/2}t^{k/4}}{N^{k/2}} + N^{-k}\right) e^{C_1 t^k}, \tag{3.82}$$

which establishes (3.14). Then substituting into (3.81),

$$\begin{aligned}
\mathbb{E}\left[|q^n_t(z) - v^n_t(z)|^k\right] &\le 3^{2k-1}s_0^k t^{k-1}\int_0^t \sup_{x\in\frac{1}{n}\mathbb{Z}} \mathbb{E}\left[|q^n_s(x) - v^n_s(x)|^k\right] ds \\
&\quad + 3^{2k-1}s_0^k t^{k-1}\int_0^t C_1\left(\frac{n^{k/2}s^{k/4}}{N^{k/2}} + N^{-k}\right) e^{C_1 s^k} ds \\
&\quad + 3^{k-1}C(k)\left(\left(C_6 t^{1/2}\frac{n}{N}\right)^{k/2} + N^{-k}\right).
\end{aligned}$$

Hence by Gronwall's inequality, there exists a constant $K_4 = K_4(k) < \infty$ such that for $t \ge 0$,

$$\sup_{x\in\frac{1}{n}\mathbb{Z}} \mathbb{E}\left[|q^n_t(x) - v^n_t(x)|^k\right] \le K_4(t^{5k/4}+1)e^{C_1 t^k}\left(\frac{n}{N}\right)^{k/2} e^{3^{2k-1}s_0^k t^k}. \tag{3.83}$$

Note that for $x \in \frac{1}{n}\mathbb{Z}$, the rate at which $(p^n_t(x))_{t\ge 0}$ jumps is bounded above by

$$N^2 r_n(1-(\alpha+1)s_n) + N^2 r_n \alpha s_n + N^3 \cdot \tfrac{1}{N} r_n s_n + 2N^2 m r_n = N^2 r_n(1+2m) = \tfrac{1}{2}Nn^2(1+2m)$$

by (1.11). Therefore, for $t \ge 0$ and $x \in \frac{1}{n}\mathbb{Z}$, letting $Z \sim \text{Poisson}(\frac{1}{2}(1+2m))$ and then using Markov's inequality,

$$\mathbb{P}\left(\sup_{s\in[0,n^{-2}N^{-1}]} |p^n_{t+s}(x) - p^n_t(x)| \ge N^{-1/2}\right) \le \mathbb{P}\left(Z \ge N^{1/2}\right) \le e^{-2N^{1/2}}\mathbb{E}\left[e^{2Z}\right] \le e^{-N^{1/2}}$$

for n sufficiently large. Suppose $T \le N$. Then by a union bound,

$$\begin{aligned}
&\mathbb{P}\left(\exists t \in n^{-2}N^{-1}\mathbb{N}_0 \cap [0,T], x \in \tfrac{1}{n}\mathbb{Z} \cap [-N^5, N^5] : \sup_{s \in [0, n^{-2}N^{-1}]} |p^n_{t+s}(x) - p^n_t(x)| \ge N^{-1/2}\right) \\
&\le \sum_{t \in n^{-2}N^{-1}\mathbb{N}_0 \cap [0,T]} \sum_{x \in \frac{1}{n}\mathbb{Z} \cap [-N^5, N^5]} \mathbb{P}\left(\sup_{s \in [0, n^{-2}N^{-1}]} |p^n_{t+s}(x) - p^n_t(x)| \ge N^{-1/2}\right) \\
&\le (n^2NT+1)(2N^5 n+1)e^{-N^{1/2}} \\
&\le e^{-N^{1/2}/2} \qquad (3.84)
\end{aligned}$$

for n sufficiently large. For $t_1, t_2 \ge 0$ and $x \in \frac{1}{n}\mathbb{Z}$, since $\sup_{u \in [0,1]} |f(u)| < 1$,

$$\begin{aligned}
|u^n_{t_1}(x) - u^n_{t_2}(x)| &\le \tfrac{1}{2}m \sup_{s \ge 0, y \in \frac{1}{n}\mathbb{Z}} |\Delta_n u^n_s(y)||t_1 - t_2| + s_0|t_1 - t_2| \\
&\le (mn^2 + s_0)|t_1 - t_2|.
\end{aligned}$$

Therefore for n sufficiently large, for $t \ge 0$ and $x \in \frac{1}{n}\mathbb{Z}$,

$$\sup_{s \in [0, n^{-2}N^{-1}]} |u^n_{t+s}(x) - u^n_t(x)| \le 2mN^{-1}. \qquad (3.85)$$

Then by (3.84), (3.85) and a union bound, for $c_3 \in (0, 1/2)$, for n sufficiently large that $2mN^{-1} + N^{-1/2} \le \frac{1}{2}\left(\frac{n}{N}\right)^{1/2-c_3}$,

$$\begin{aligned}
&\mathbb{P}\left(\sup_{x \in \frac{1}{n}\mathbb{Z},\, |x| \le N^5} \sup_{t \in [0,T]} |p^n_t(x) - u^n_t(x)| \ge \left(\frac{n}{N}\right)^{1/2-c_3}\right) \\
&\le \sum_{t \in n^{-2}N^{-1}\mathbb{N}_0 \cap [0,T]} \sum_{x \in \frac{1}{n}\mathbb{Z},\, |x| \le N^5} \mathbb{P}\left(|p^n_t(x) - u^n_t(x)| \ge \tfrac{1}{2}\left(\frac{n}{N}\right)^{1/2-c_3}\right) + e^{-N^{1/2}/2}.
\end{aligned}$$

Hence for $k \in \mathbb{N}$ with $k \ge 2$, by Markov's inequality, and then by (3.82),

$$\begin{aligned}
&\mathbb{P}\left(\sup_{x \in \frac{1}{n}\mathbb{Z},\, |x| \le N^5} \sup_{t \in [0,T]} |p^n_t(x) - u^n_t(x)| \ge \left(\frac{n}{N}\right)^{1/2-c_3}\right) \\
&\le \sum_{t \in n^{-2}N^{-1}\mathbb{N}_0 \cap [0,T]} \sum_{x \in \frac{1}{n}\mathbb{Z},\, |x| \le N^5} \mathbb{E}\left[|p^n_t(x) - u^n_t(x)|^k\right] 2^k \left(\frac{n}{N}\right)^{-k(1/2-c_3)} + e^{-N^{1/2}/2} \\
&\le \sum_{t \in n^{-2}N^{-1}\mathbb{N}_0 \cap [0,T]} \sum_{x \in \frac{1}{n}\mathbb{Z},\, |x| \le N^5} C_1 \left(\frac{n^{k/2}t^{k/4}}{N^{k/2}} + N^{-k}\right) e^{C_1 t^k} 2^k \left(\frac{n}{N}\right)^{-k(1/2-c_3)} + e^{-N^{1/2}/2} \\
&\le (n^2NT+1)(2nN^5+1)C_1 \left(\frac{n^{k/2}T^{k/4}}{N^{k/2}} + N^{-k}\right) e^{C_1 T^k} 2^k \left(\frac{n}{N}\right)^{-k(1/2-c_3)} + e^{-N^{1/2}/2}.
\end{aligned}$$

Take $\ell' \in \mathbb{N}$ sufficiently large that $n^4N^7 e^{2^k(C_1 + 3^{2k-1}s_0^k)(\log N)^{1/2}} \left(\frac{n}{N}\right)^{\ell'} \le 1$ for n sufficiently large. For $\ell \in \mathbb{N}$, take $c_4 \in (0, \frac{1}{2}c_3(\ell + \ell' + 1)^{-1})$. Since $1/(2c_4) > (\ell + \ell' + 1)/c_3$ and $c_3 < 1/2$, we can take $k \in \mathbb{N} \cap ((\ell + \ell')/c_3, 1/(2c_4))$ with $k \ge 2$. Therefore for $T \le 2(\log N)^{c_4}$, for n sufficiently large,

$$\begin{aligned}
&\mathbb{P}\left(\sup_{x \in \frac{1}{n}\mathbb{Z},\, |x| \le N^5} \sup_{t \in [0,T]} |p^n_t(x) - u^n_t(x)| \ge \left(\frac{n}{N}\right)^{1/2-c_3}\right) \\
&\le n^4N^7 \left(\frac{n}{N}\right)^{k/2} e^{C_1 2^k (\log N)^{c_4 k}} \left(\frac{n}{N}\right)^{-k(1/2-c_3)} + e^{-N^{1/2}/2} \\
&\le \left(\frac{n}{N}\right)^{\ell}
\end{aligned}$$

for n sufficiently large, since $kc_3 > \ell + \ell'$ and $c_4 k < 1/2$. Similarly, by a union bound and Markov's inequality, and then by (3.83), for $t \leq 2(\log N)^{c_4}$,

$$\begin{aligned}
&\mathbb{P}\left(\sup_{x\in \frac{1}{n}\mathbb{Z},\, |x|\leq N^5} |q_t^n(x) - v_t^n(x)| \geq \left(\frac{n}{N}\right)^{1/2-c_3} \right) \\
&\leq \sum_{x\in \frac{1}{n}\mathbb{Z}, |x|\leq N^5} \mathbb{E}\left[|q_t^n(x) - v_t^n(x)|^k \right] \left(\frac{n}{N}\right)^{-k(1/2-c_3)} \\
&\leq (2nN^5+1)K_4(t^{5k/4}+1)e^{C_1 t^k} e^{3^{2k-1} s_0^k t^k} \left(\frac{n}{N}\right)^{kc_3} \\
&\leq \left(\frac{n}{N}\right)^{\ell}
\end{aligned}$$

for n sufficiently large, which completes the proof. □

4 Event E_2 occurs with high probability

Recall the definitions of the events E_2 and E_2' in (2.10) and (2.11). In this section, we will prove the following result.

Proposition 4.1. *For $c_1, c_2 > 0$, for $t^* \in \mathbb{N}$ sufficiently large and $K \in \mathbb{N}$ sufficiently large (depending on t^*), the following holds. If $a_1 > 1$ and $N \geq n^{a_1}$ for n sufficiently large, then for n sufficiently large,*

$$\mathbb{P}\left((E_2')^c \cap E_1' \right) \leq \left(\frac{n}{N}\right)^2 .$$

Moreover, if $a_2 > 3$ and $N \geq n^{a_2}$ for n sufficiently large, then for n sufficiently large,

$$\mathbb{P}\left((E_2)^c \cap E_1' \right) \leq \left(\frac{n}{N}\right)^2 .$$

Suppose from now on in this section that for some $a_1 > 1$, $N \geq n^{a_1}$ for n sufficiently large, and fix $c_1, c_2 > 0$. We begin by proving that for t, x_1 and x_2 such that x_1 and x_2 are not too far from the front, the event $A_t^{(1)}(x_1, x_2)$ occurs with high probability. Recall the definition of $(v_t^n)_{t\geq 0}$ in (3.11). We begin by showing that the solution of a PDE closely related to (3.11) can be written in terms of a diffusion $(Z_t)_{t\geq 0}$.

Lemma 4.2. *Suppose $h : \mathbb{R} \to [0,1]$ is measurable, and take $t_0 \geq 0$. For $x \in \mathbb{R}$ and $t \geq t_0$, let*

$$v_t(x) = g(x-\nu t)\mathbb{E}_{x-\nu t}\left[\frac{h(Z_{t-t_0}+\nu t_0)}{g(Z_{t-t_0})} \right],$$

where under $\mathbb{P}_{x_0}$, $(Z_t)_{t\geq 0}$ solves the SDE

$$dZ_t = \nu\, dt + \frac{m\nabla g(Z_t)}{g(Z_t)}\, dt + \sqrt{m}\, dB_t, \quad Z_0 = x_0, \tag{4.1}$$

and $(B_t)_{t\geq 0}$ is a Brownian motion. Then $v_{t_0} = h$ and

$$\partial_t v_t(x) = \tfrac{1}{2} m\Delta v_t(x) + s_0 v_t(x)(1-g(x-\nu t))(2g(x-\nu t)-1+\alpha) \quad \text{for } t > t_0,\ x \in \mathbb{R}.$$

Proof. For $t \geq t_0$ and $x \in \mathbb{R}$, let

$$v_t^{(1)}(x) = \mathbb{E}_{x-\nu t}\left[\frac{h(Z_{t-t_0}+\nu t_0)}{g(Z_{t-t_0})} \right] = v_t(x) g(x-\nu t)^{-1}.$$

Recall (4.1). Since $\mathcal{A}f(x) := \frac{1}{2}m\Delta f(x) + \left(\nu + \frac{m\nabla g(x)}{g(x)}\right)\nabla f(x)$ is the generator of the diffusion $(Z_t)_{t\geq 0}$, for $t > t_0$ and $x \in \mathbb{R}$,

$$\partial_t v_t^{(1)}(x) = \tfrac{1}{2}m\Delta v_t^{(1)}(x) + \left(\nu + \frac{m\nabla g(x-\nu t)}{g(x-\nu t)}\right)\nabla v_t^{(1)}(x) - \nu\nabla v_t^{(1)}(x)$$

(see for example Theorem 7.1.5 in [13]). Therefore

$$\begin{aligned}\partial_t v_t(x) &= -\nu\nabla g(x-\nu t)v_t^{(1)}(x) + \tfrac{1}{2}mg(x-\nu t)\Delta v_t^{(1)}(x) + m\nabla g(x-\nu t)\nabla v_t^{(1)}(x)\\ &= \tfrac{1}{2}m\Delta v_t(x) - \tfrac{1}{2}m\frac{\Delta g(x-\nu t)}{g(x-\nu t)}v_t(x) - \nu\frac{\nabla g(x-\nu t)}{g(x-\nu t)}v_t(x).\end{aligned}$$

Since $\Delta g = -\kappa^2 g(1-g)(2g-1)$ and $\nabla g = -\kappa g(1-g)$, the result follows by (2.1). □

We now show that for $(u_t^n)_{t\geq 0}$ and $(v_t^n)_{t\geq 0}$ defined as in (3.6) and (3.11), if we have that $\sup_{s\in[0,t],\, x\in\frac{1}{n}\mathbb{Z}}|u_s^n(x) - g(x-\nu s)|$ is small then v_t^n is approximately given by an expectation of a function of Z_t. The proof is similar to the proof of Lemma 3.6.

Lemma 4.3. *Take $\delta, \epsilon \in (0,1)$. For $t \geq 0$ and $x \in \mathbb{R}$, let*

$$v_t(x) = g(x-\nu t)\mathbb{E}_{x-\nu t}\left[\bar{q}_0^n(Z_t)g(Z_t)^{-1}\right],$$

where $\bar{q}_0^n : \mathbb{R} \to [0,1]$ is the linear interpolation of $q_0^n : \frac{1}{n}\mathbb{Z} \to [0,1]$, and $(Z_t)_{t\geq 0}$ is defined in (4.1). Suppose that $T \geq 1$, $\sup_{x\in\frac{1}{n}\mathbb{Z}, s\in[0,T]}|u_s^n(x) - g(x-\nu s)| \leq \delta$ and $\sup_{x_1,x_2\in\frac{1}{n}\mathbb{Z}, |x_1-x_2|\leq n^{-1/3}}|q_0^n(x_1) - q_0^n(x_2)| \leq \epsilon$. There exists a constant $C_7 < \infty$ such that for n sufficiently large, for $t \in [0,T]$,

$$\sup_{x\in\frac{1}{n}\mathbb{Z}}|v_t^n(x) - v_t(x)| \leq \left(C_7(n^{-1/3}+\delta)\sup_{x\in\frac{1}{n}\mathbb{Z}} q_0^n(x) + 2\epsilon\right)e^{5s_0T}T^2.$$

Proof. For $t > 0$ and $x \in \mathbb{R}$, let $G_t(x) = \frac{1}{\sqrt{2\pi t}}e^{-x^2/(2t)}$. For $s \geq 0$ and $x \in \mathbb{R}$, let $f_s(x) = v_s(x)(1-g(x-\nu s))(2g(x-\nu s) - 1 + \alpha)$. By Lemma 4.2, for any fixed $a \in \mathbb{R}$, $v_t(x)$ solves the equation

$$\partial_t v_t(x) = (\tfrac{1}{2}m\Delta v_t(x) - av_t) + s_0 f_t + av_t \quad \text{for } t > 0,\ x \in \mathbb{R}.$$

Since $e^{-at}G_{mt}(x)$ is the fundamental solution of the equation $\partial_t v = \frac{1}{2}m\Delta v - av$, Duhamel's principle (see for example (17) and (18) in Section 2.3 on page 51 of [15] and Theorem 4.8 on page 147 of [18]) implies that for $a \in \mathbb{R}$, $z \in \mathbb{R}$ and $t > 0$,

$$v_t(z) = e^{-at}G_{mt} * v_0(z) + \int_0^t e^{-a(t-s)}G_{m(t-s)} * (s_0 f_s + av_s)(z)ds. \tag{4.2}$$

Therefore, by (4.2) with $a = -(1+\alpha)s_0$, and since $(1-u)(2u-1+\alpha) \leq 1+\alpha$ for $u \in [0,1]$,

$$v_t(z) \leq e^{(1+\alpha)s_0 t}G_{mt} * v_0(z). \tag{4.3}$$

Letting $(B_t)_{t\geq 0}$ denote a Brownian motion, it follows from (3.12) and (4.2) with $a = 0$ that for $z \in \frac{1}{n}\mathbb{Z}$ and $t \geq 0$,

$$\begin{aligned}|v_t^n(z) - v_t(z)| &\leq |\mathbf{E}_z[q_0^n(X_{mt}^n)] - \mathbb{E}_z[v_0(B_{mt})]|\\ &\quad + s_0\int_0^t \Big|\mathbf{E}_z\left[v_s^n(1-u_s^n)(2u_s^n - 1 + \alpha)(X_{m(t-s)}^n)\right] - \mathbb{E}_z\left[f_s(B_{m(t-s)})\right]\Big|ds.\end{aligned} \tag{4.4}$$

Recall from (3.19) in the proof of Lemma 3.6 that for n sufficiently large, $(X_t^n)_{t\geq 0}$ and $(B_t)_{t\geq 0}$ can be coupled in such a way that $X_0^n = B_0$ and for $t\geq 0$,

$$\mathbb{P}\left(|X_{mt}^n - B_{mt}| \geq n^{-1/3}\right) \leq (t+1)n^{-1/2}. \tag{4.5}$$

Since $v_0 = \bar{q}_0^n$, which is the linear interpolation of q_0^n, it follows that for $z\in \frac{1}{n}\mathbb{Z}$ and $t\geq 0$,

$$\begin{aligned}
&|\mathbf{E}_z\left[q_0^n(X_{mt}^n)\right] - \mathbb{E}_z\left[v_0(B_{mt})\right]| \\
&\leq (t+1)n^{-1/2}\sup_{x\in\frac{1}{n}\mathbb{Z}} q_0^n(x) + \sup_{x_1,x_2\in\mathbb{R},|x_1-x_2|\leq n^{-1/3}} |\bar{q}_0^n(x_1)-\bar{q}_0^n(x_2)| \\
&\leq (t+1)n^{-1/2}\sup_{x\in\frac{1}{n}\mathbb{Z}} q_0^n(x) + 2\epsilon
\end{aligned} \tag{4.6}$$

for n sufficiently large. For the second term on the right hand side of (4.4), note that if $t\leq T$ then for $s\in[0,t]$ and $y\in\frac{1}{n}\mathbb{Z}$,

$$|(1-u_s^n(y))(2u_s^n(y)-1+\alpha) - (1-g(y-\nu s))(2g(y-\nu s)-1+\alpha)| \leq 3\delta.$$

Hence by the triangle inequality and then by (4.5), for $s\in[0,t]$,

$$\begin{aligned}
&\left|\mathbf{E}_z\left[v_s^n(1-u_s^n)(2u_s^n-1+\alpha)(X_{m(t-s)}^n)\right] - \mathbb{E}_z\left[f_s(B_{m(t-s)})\right]\right| \\
&\quad\leq \mathbf{E}_z\left[(|(v_s^n - v_s)(1-u_s^n)(2u_s^n-1+\alpha)| + 3\delta v_s)(X_{m(t-s)}^n)\right] \\
&\qquad + \left|\mathbf{E}_z\left[f_s(X_{m(t-s)}^n)\right] - \mathbb{E}_z\left[f_s(B_{m(t-s)})\right]\right| \\
&\quad\leq 3\left(\sup_{x\in\frac{1}{n}\mathbb{Z}}|v_s^n(x)-v_s(x)| + \delta\sup_{x\in\mathbb{R}} v_s(x)\right) + 2(t+1)n^{-1/2}\sup_{x\in\mathbb{R}}|f_s(x)| + n^{-1/3}\sup_{x\in\mathbb{R}}|\nabla f_s(x)| \\
&\quad\leq 3\Big(\sup_{x\in\frac{1}{n}\mathbb{Z}}|v_s^n(x)-v_s(x)| + (\delta + 2(t+1)n^{-1/2})e^{(1+\alpha)s_0 s}\|v_0\|_\infty \\
&\qquad\qquad + n^{-1/3}(\|\nabla v_s\|_\infty + e^{(1+\alpha)s_0 s}\|v_0\|_\infty\|\nabla g\|_\infty)\Big)
\end{aligned} \tag{4.7}$$

by (4.3). It remains to bound $\|\nabla v_s\|_\infty$. For $t>0$ and $x\in\mathbb{R}$, by differentiating both sides of (4.2),

$$\nabla v_t(x) = G'_{mt} * v_0(x) + s_0\int_0^t G'_{m(t-s)} * f_s(x)ds. \tag{4.8}$$

For the first term on the right hand side,

$$|G'_{mt} * v_0(x)| \leq \|v_0\|_\infty \int_{-\infty}^{\infty} |G'_{mt}(z)|dz = 2\|v_0\|_\infty G_{mt}(0) = 2\|v_0\|_\infty (2\pi m t)^{-1/2}.$$

For the second term on the right hand side of (4.8), since $|f_s(\cdot)|\leq (1+\alpha)e^{(1+\alpha)s_0 s}\|v_0\|_\infty$ by (4.3),

$$\left|\int_0^t G'_{m(t-s)} * f_s(x)ds\right| \leq (1+\alpha)e^{(1+\alpha)s_0 t}\|v_0\|_\infty \int_0^t 2G_{m(t-s)}(0)ds,$$

and so by (4.8), for $t>0$,

$$\|\nabla v_t\|_\infty \leq (2t^{-1/2} + 4s_0(1+\alpha)e^{(1+\alpha)s_0 t}t^{1/2})(2\pi m)^{-1/2}\|v_0\|_\infty.$$

Substituting into (4.7) and then into (4.4), using (4.6), we now have that for $t \in [0,T]$ and $z \in \frac{1}{n}\mathbb{Z}$,

$$\begin{aligned}
&|v_t^n(z) - v_t(z)| \\
&\le (t+1)n^{-1/2} \sup_{x \in \frac{1}{n}\mathbb{Z}} q_0^n(x) + 2\epsilon \\
&\quad + 3s_0 \int_0^t \Bigg(\sup_{x \in \frac{1}{n}\mathbb{Z}} |v_s^n(x) - v_s(x)| + e^{(1+\alpha)s_0 t} \|v_0\|_\infty (\delta + 2(t+1)n^{-1/2} + n^{-1/3}\|\nabla g\|_\infty) \\
&\qquad\qquad + (s^{-1/2} + 2s_0(1+\alpha)e^{(1+\alpha)s_0 t} s^{1/2}) m^{-1/2} \|v_0\|_\infty n^{-1/3} \Bigg) ds.
\end{aligned}$$

The result follows by Gronwall's inequality and since $\|v_0\|_\infty = \sup_{x \in \frac{1}{n}\mathbb{Z}} q_0^n(x)$. $\square$

By the theory of speed and scale (see for example [21]), $(Z_t)_{t\ge 0}$ as defined in (4.1) has scale function S and speed measure density M given by

$$S(x) = \int_0^x \tfrac{1}{4} e^{-\alpha\kappa y} g(y)^{-2} dy \quad \text{and} \quad M(x) = \frac{4}{m} e^{\alpha\kappa x} g(x)^2. \tag{4.9}$$

Therefore $(Z_t)_{t\ge 0}$ has a stationary distribution with density π as defined in (1.15). We now establish some useful upper bounds on the total variation distance between π and the law of Z_t at a large time t. Recall the definitions of γ_n and d_n in (2.4).

Lemma 4.4. *Take $z_0 \in \mathbb{R}$ and suppose $(Z_t^{(1)})_{t\ge 0}$ and $(Z_t^{(2)})_{t\ge 0}$ solve the SDEs*

$$\begin{aligned}
& dZ_t^{(1)} = \nu dt + \frac{m\nabla g(Z_t^{(1)})}{g(Z_t^{(1)})} dt + \sqrt{m} dB_t^{(1)}, \quad Z_0^{(1)} = z_0 \\
\text{and} \quad & dZ_t^{(2)} = \nu dt + \frac{m\nabla g(Z_t^{(2)})}{g(Z_t^{(2)})} dt + \sqrt{m} dB_t^{(2)}, \quad Z_0^{(2)} = Z,
\end{aligned}$$

where $(B_t^{(1)})_{t\ge 0}$ and $(B_t^{(2)})_{t\ge 0}$ are independent Brownian motions and Z is an independent random variable with density π. Let

$$T^Z = \inf\{t \ge 0 : Z_t^{(1)} = Z_t^{(2)}\}.$$

Then for n sufficiently large, if $|z_0| \le d_n + 1$,

$$\mathbb{P}\left(T^Z \ge \tfrac{1}{2}\gamma_n\right) \le (\log N)^{-12C}. \tag{4.10}$$

For $A < \infty$, for $t \ge 0$ sufficiently large (depending on A), if $|z_0| \le A$,

$$\mathbb{P}\left(T^Z \ge t\right) \le 2m^{-1/2} t^{-1/4}. \tag{4.11}$$

Remark 4.5. The first bound (4.10) will be used in the proof of Proposition 4.1, and the weaker bound in (4.11) will be used in Section 7 in the proof of Theorem 1.1.

Proof. Suppose first that $|z_0| \le d_n + 1$. Since $g(x) \le \min(e^{-\kappa x}, 1)\ \forall x \in \mathbb{R}$, for $y_0 > 0$ we have

$$\begin{aligned}
&\int_{y_0}^\infty g(y)^2 e^{\alpha\kappa y} dy \le (2-\alpha)^{-1} \kappa^{-1} e^{-(2-\alpha)\kappa y_0} \\
\text{and} \quad &\int_{-\infty}^{-y_0} g(y)^2 e^{\alpha\kappa y} dy \le \alpha^{-1} \kappa^{-1} e^{-\alpha\kappa y_0}.
\end{aligned} \tag{4.12}$$

It follows that since $d_n = \kappa^{-1} C \log \log N$,

$$\mathbb{P}\left(|Z_0^{(2)}| \geq 13\alpha^{-1} d_n\right) \leq 2\alpha^{-1}\kappa^{-1}\left(\int_{-\infty}^{\infty} g(y)^2 e^{\alpha\kappa y} dy\right)^{-1} (\log N)^{-13C}. \tag{4.13}$$

Take $(Z_t)_{t\geq 0}$ as defined in (4.1), and for $a \in \mathbb{R}$, let

$$\tau^a = \inf\{t \geq 0 : Z_t = a\}.$$

By (4.9) and the theory of speed and scale (see for example [21]), and then since $g(y) \in [\frac{1}{2}e^{-\kappa y}, e^{-\kappa y}]$ $\forall y \geq 0$, for $x > 0$,

$$\begin{aligned}\mathbb{P}_{x/2}\left(\tau^x < \tau^0\right) = \frac{S(0) - S(x/2)}{S(0) - S(x)} \leq \frac{\int_0^{x/2} 4e^{-\alpha\kappa y} e^{2\kappa y} dy}{\int_0^x e^{-\alpha\kappa y} e^{2\kappa y} dy} &= 4\frac{e^{(2-\alpha)\kappa x/2} - 1}{e^{(2-\alpha)\kappa x} - 1} \\ &\leq 8e^{-(2-\alpha)\kappa x/2}\end{aligned}$$

for $x \geq \kappa^{-1}\log 2$. Similarly, since $g(y) \in [1/2, 1]$ $\forall y \leq 0$,

$$\mathbb{P}_{-x/2}\left(\tau^{-x} < \tau^0\right) = \frac{S(0) - S(-x/2)}{S(0) - S(-x)} \leq \frac{\int_{-x/2}^0 4e^{-\alpha\kappa y} dy}{\int_{-x}^0 e^{-\alpha\kappa y} dy} = 4\frac{e^{\alpha\kappa x/2} - 1}{e^{\alpha\kappa x} - 1} \leq 8e^{-\alpha\kappa x/2}$$

for $x \geq \alpha^{-1}\kappa^{-1}\log 2$. Hence for n sufficiently large,

$$\max\left(\mathbb{P}_{13\alpha^{-1}d_n}\left(\tau^{26\alpha^{-1}d_n} < \tau^0\right), \mathbb{P}_{-13\alpha^{-1}d_n}\left(\tau^{-26\alpha^{-1}d_n} < \tau^0\right)\right) \leq 8(\log N)^{-13C}. \tag{4.14}$$

Let $(B_t)_{t\geq 0}$ denote a Brownian motion. Note that $\frac{\nabla g(y)}{g(y)} \in [-\kappa, 0]$ $\forall y \in \mathbb{R}$, and so $|\nu + \frac{m\nabla g(y)}{g(y)}| < \sqrt{2s_0 m}$ by (2.1). Hence for $x \in \mathbb{R}$ with $|x| \geq 13\alpha^{-1}d_n$,

$$\mathbb{P}_x\left(\tau^0 < 1\right) \leq \mathbb{P}\left(\sup_{t\in[0,1]} \sqrt{m}B_t \geq 13\alpha^{-1}d_n - \sqrt{2ms_0}\right) \leq 2e^{-\frac{1}{2m}(13\alpha^{-1}d_n - \sqrt{2ms_0})^2} \tag{4.15}$$

by the reflection principle and a Gaussian tail bound. Therefore by a union bound,

$$\begin{aligned}&\mathbb{P}\left(\exists j \in \{1,2\}, t \in [0, \gamma_n] : |Z_t^{(j)}| \geq 26\alpha^{-1}d_n\right) \\ &\leq \mathbb{P}\left(|Z_0^{(2)}| \geq 13\alpha^{-1}d_n\right) \\ &\quad + 2\lceil\gamma_n\rceil \max\left(\mathbb{P}_{13\alpha^{-1}d_n}\left(\tau^{26\alpha^{-1}d_n} < \tau^0\right), \mathbb{P}_{-13\alpha^{-1}d_n}\left(\tau^{-26\alpha^{-1}d_n} < \tau^0\right)\right) \\ &\quad + 2\lceil\gamma_n\rceil \max\left(\mathbb{P}_{13\alpha^{-1}d_n}\left(\tau^0 < 1\right), \mathbb{P}_{-13\alpha^{-1}d_n}\left(\tau^0 < 1\right)\right) \\ &\leq \tfrac{1}{2}(\log N)^{-12C}\end{aligned} \tag{4.16}$$

for n sufficiently large, by (4.13), (4.14) and (4.15).

For $t \geq 0$, define the σ-algebra $\mathcal{F}_t^Z = \sigma((Z_s^{(1)})_{s\leq t}, (Z_s^{(2)})_{s\leq t})$. Note that if $Z_t^{(1)} \leq Z_t^{(2)}$ then for $s \in [t, T^Z \vee t]$,

$$\begin{aligned}&Z_s^{(2)} - Z_s^{(1)} \\ &= (Z_t^{(2)} - Z_t^{(1)}) + m\int_t^s \left(\frac{\nabla g(Z_u^{(2)})}{g(Z_u^{(2)})} - \frac{\nabla g(Z_u^{(1)})}{g(Z_u^{(1)})}\right) du + \sqrt{m}((B_s^{(2)} - B_t^{(2)}) - (B_s^{(1)} - B_t^{(1)})) \\ &\leq (Z_t^{(2)} - Z_t^{(1)}) + \sqrt{m}((B_s^{(2)} - B_t^{(2)}) - (B_s^{(1)} - B_t^{(1)})),\end{aligned} \tag{4.17}$$

since $y \mapsto \frac{\nabla g(y)}{g(y)}$ is decreasing. Therefore, for n sufficiently large, for $t \geq 0$, if $|Z_t^{(1)}| \vee |Z_t^{(2)}| \leq 26\alpha^{-1} d_n$ then

$$\begin{aligned}\mathbb{P}\left(T^Z > t + \gamma_n^{1/2} \middle| \mathcal{F}_t^Z\right) &\leq \mathbb{P}_{52\alpha^{-1}d_n}\left(\sqrt{2m}B_s \geq 0\ \forall s \in [0, \gamma_n^{1/2}]\right) \\ &\leq \mathbb{P}_{52\alpha^{-1}\kappa^{-1}C+1}\left(\sqrt{2m}B_s \geq 0\ \forall s \in [0,1]\right) := p > 0 \qquad (4.18)\end{aligned}$$

by Brownian scaling and since $d_n = \kappa^{-1} C \log\log N$ and $\gamma_n = \lfloor (\log\log N)^4 \rfloor$. Therefore by (4.16) and a union bound, for n sufficiently large,

$$\begin{aligned}&\mathbb{P}\left(T^Z \geq \tfrac{1}{2}\gamma_n\right) \\ &\leq \tfrac{1}{2}(\log N)^{-12C} + \mathbb{P}\left(T^Z \geq \tfrac{1}{2}\gamma_n, |Z^{(1)}_{k\gamma_n^{1/2}}| \vee |Z^{(2)}_{k\gamma_n^{1/2}}| \leq 26\alpha^{-1} d_n\ \forall k \in \mathbb{N}_0 \cap [0, \tfrac{1}{2}\gamma_n^{1/2}]\right) \\ &\leq \tfrac{1}{2}(\log N)^{-12C} + p^{\lfloor \gamma_n^{1/2}/2 \rfloor}\end{aligned}$$

by (4.18), which completes the proof of (4.10).

Now take $A < \infty$ and suppose $|z_0| \leq A$. Then for $t \geq A^4$, by a union bound and (4.17),

$$\begin{aligned}\mathbb{P}\left(T^Z \geq t\right) &\leq \mathbb{P}\left(|Z_0^{(2)}| \geq t^{1/4}\right) + \mathbb{P}_{2t^{1/4}}\left(\sqrt{2m}B_s \geq 0\ \forall s \in [0,t]\right) \\ &\leq 2\alpha^{-1}\kappa^{-1}\left(\int_{-\infty}^{\infty} g(y)^2 e^{\alpha\kappa y} dy\right)^{-1} e^{-\alpha\kappa t^{1/4}} + \mathbb{P}_0\left(|B_{2mt}| \leq 2t^{1/4}\right)\end{aligned}$$

by (4.12) and the reflection principle. Since $\mathbb{P}_0\left(|B_{2mt}| \leq 2t^{1/4}\right) \leq \frac{4t^{1/4}}{(4\pi m t)^{1/2}}$, the result follows by taking t sufficiently large. □

Fix $x_0 \in \frac{1}{n}\mathbb{Z}$, and take $(v_t^n)_{t\geq 0}$ as in (3.11) with $v_0^n(x) = p_0^n(x_0)\mathbb{1}_{x=x_0}$, and where $(u_t^n)_{t\geq 0}$ is defined in (3.6). The following result will be combined with a bound on $|q^n_{\gamma_n} - v^n_{\gamma_n}|$ to show that the event $A_t^{(1)}(x_1, x_2)$ occurs with high probability for suitable t, x_1 and x_2. Recall that we fixed $c_2 > 0$ at the start of Section 4.

Lemma 4.6. *Suppose* $\sup_{x\in\frac{1}{n}\mathbb{Z}, s\in[0,\gamma_n]} |u_s^n(x) - g(x-\nu s)| \leq e^{-(\log N)^{c_2}}$. *For n sufficiently large, if $|x_0| \leq d_n$ and $|x - \nu\gamma_n| \leq d_n + 1$,*

$$\frac{v^n_{\gamma_n}(x)}{g(x-\nu\gamma_n)} = \frac{\pi(x_0)}{g(x_0)} p_0^n(x_0) n^{-1}(1 + \mathcal{O}((\log N)^{-4C})).$$

Proof. Let $t_0 = (\log N)^{-12C}$. For $x \in \frac{1}{n}\mathbb{Z}$, let $P^n_{t_0,x_0}(x) = \mathbf{P}_x\left(X^n_{mt_0} = x_0\right)$, and let $\bar{P}^n_{t_0,x_0} : \mathbb{R} \to [0,1]$ denote the linear interpolation of $P^n_{t_0,x_0}$. Let $\bar{v}^n_{t_0}$ denote the linear interpolation of $v^n_{t_0}$. For $t \geq t_0$ and $x \in \mathbb{R}$, let

$$v_t(x) = g(x-\nu t)\mathbb{E}_{x-\nu t}\left[\frac{\bar{v}^n_{t_0}(Z_{t-t_0} + \nu t_0)}{g(Z_{t-t_0})}\right], \qquad (4.19)$$

where $(Z_t)_{t\geq 0}$ is defined in (4.1). By (3.13), for $t \geq 0$ and $y \in \frac{1}{n}\mathbb{Z}$,

$$v_t^n(y) \leq e^{(1+\alpha)s_0 t} p_0^n(x_0)\mathbf{P}_y\left(X^n_{mt} = x_0\right), \qquad (4.20)$$

and so for $t \geq t_0$ and $x \in \mathbb{R}$,

$$\begin{aligned}&v_t(x) \\ &\leq g(x-\nu t)p_0^n(x_0)e^{(1+\alpha)s_0 t_0}\Big(\mathbb{E}_{x-\nu t}\left[g(Z_{t-t_0})^{-1}\bar{P}^n_{t_0,x_0}(Z_{t-t_0}+\nu t_0)\mathbb{1}_{|Z_{t-t_0}+\nu t_0 - x_0| < n^{1/4}}\right] \\ &\qquad + \mathbb{E}_{x-\nu t}\left[g(Z_{t-t_0})^{-1}\bar{P}^n_{t_0,x_0}(Z_{t-t_0}+\nu t_0)\mathbb{1}_{|Z_{t-t_0}+\nu t_0 - x_0| \geq n^{1/4}}\right]\Big). \qquad (4.21)\end{aligned}$$

For the first term on the right hand side, we have that if n is sufficiently large that $n^{1/4}+n^{-1}\le \frac{1}{2}mnt_0$, then by Lemma 3.14,

$$\begin{aligned}&\mathbb{E}_{x-\nu t}\left[g(Z_{t-t_0})^{-1}\bar{P}^n_{t_0,x_0}(Z_{t-t_0}+\nu t_0)\mathbb{1}_{|Z_{t-t_0}+\nu t_0-x_0|<n^{1/4}}\right]\\&\le n^{-1}(2\pi m t_0)^{-1/2}e^{\mathcal{O}(n^{-1/5})}\mathbb{E}_{x-\nu t}\left[g(Z_{t-t_0})^{-1}e^{-(Z_{t-t_0}+\nu t_0-x_0)^2/(2mt_0)}\right].\end{aligned}$$

For the second term on the right hand side of (4.21), by the definition of $\bar{P}^n_{t_0,x_0}$ and then by Markov's inequality, for n sufficiently large,

$$\begin{aligned}&\mathbb{E}_{x-\nu t}\left[g(Z_{t-t_0})^{-1}\bar{P}^n_{t_0,x_0}(Z_{t-t_0}+\nu t_0)\mathbb{1}_{|Z_{t-t_0}+\nu t_0-x_0|\ge n^{1/4}}\right]\\&\le \mathbb{E}_{x-\nu t}\left[(1+e^{\kappa Z_{t-t_0}})\mathbf{P}_0\left(X^n_{mt_0}\ge |Z_{t-t_0}+\nu t_0-x_0|-n^{-1}\right)\mathbb{1}_{|Z_{t-t_0}+\nu t_0-x_0|\ge n^{1/4}}\right]\\&\le \mathbb{E}_{x-\nu t}\left[(1+e^{\kappa Z_{t-t_0}})e^{-3\kappa|Z_{t-t_0}+\nu t_0-x_0|}e^{3\kappa n^{-1}}\mathbf{E}_0\left[e^{3\kappa X^n_{mt_0}}\right]\mathbb{1}_{|Z_{t-t_0}+\nu t_0-x_0|\ge n^{1/4}}\right]\\&\le e^{10s_0t_0}(e^{-3\kappa n^{1/4}}+e^{\kappa|x_0|}e^{-2\kappa n^{1/4}})\end{aligned}$$

by Lemma 3.8 and since $e^{\kappa Z_{t-t_0}}e^{-3\kappa|Z_{t-t_0}+\nu t_0-x_0|}\le e^{(-\nu t_0+x_0)\kappa}e^{-2\kappa|Z_{t-t_0}+\nu t_0-x_0|}$ and $\frac{1}{2}m\kappa^2=s_0$. Substituting into (4.21), it follows that

$$\begin{aligned}v_t(x)\le{}& g(x-\nu t)p_0^n(x_0)e^{(1+\alpha)s_0t_0}n^{-1}(2\pi m t_0)^{-1/2}\\&\left(\mathcal{O}(nt_0^{1/2}e^{\kappa|x_0|}e^{-2\kappa n^{1/4}})+e^{\mathcal{O}(n^{-1/5})}\mathbb{E}_{x-\nu t}\left[g(Z_{t-t_0})^{-1}e^{-(Z_{t-t_0}+\nu t_0-x_0)^2/(2mt_0)}\right]\right).\end{aligned}\tag{4.22}$$

Note that for $y\in\mathbb{R}$,

$$\begin{aligned}g(y)^{-1}e^{-(y+\nu t_0-x_0)^2/(2mt_0)}&\le 1+e^{\kappa(x_0-\nu t_0)}e^{(\kappa-(2mt_0)^{-1}(y+\nu t_0-x_0))(y+\nu t_0-x_0)}\\&\le 1+e^{\kappa|x_0|+s_0t_0}\end{aligned}$$

since $\frac{1}{2}m\kappa^2=s_0$ and so $\sup_{z\in\mathbb{R}}(\kappa z-(2mt_0)^{-1}z^2)=s_0t_0$. Hence by Lemma 4.4, for n sufficiently large, if $t-t_0\ge\gamma_n/2$ and $|x-\nu t|\le d_n+1$, then

$$\begin{aligned}&\mathbb{E}_{x-\nu t}\left[g(Z_{t-t_0})^{-1}e^{-(Z_{t-t_0}+\nu t_0-x_0)^2/(2mt_0)}\right]\\&\le\int_{-\infty}^{\infty}\pi(y)g(y)^{-1}e^{-(y+\nu t_0-x_0)^2/(2mt_0)}dy+3e^{\kappa|x_0|}(\log N)^{-12C}.\end{aligned}\tag{4.23}$$

Note that $g(y)e^{\alpha\kappa y}\le\min(e^{\alpha\kappa y},e^{-(1-\alpha)\kappa y})\le 1\ \forall y\in\mathbb{R}$. Therefore, since $y\mapsto g(y)$ is decreasing, and letting $(B_s)_{s\ge0}$ denote a Brownian motion,

$$\begin{aligned}&\int_{-\infty}^{\infty}g(y)e^{\alpha\kappa y}e^{-(y+\nu t_0-x_0)^2/(2mt_0)}dy\\&\le g(x_0-\nu t_0-t_0^{1/3})\int_{-\infty}^{\infty}e^{\alpha\kappa y}e^{-(y+\nu t_0-x_0)^2/(2mt_0)}dy\\&\quad+\int_{-\infty}^{\infty}e^{-(y+\nu t_0-x_0)^2/(2mt_0)}\mathbb{1}_{|y+\nu t_0-x_0|>t_0^{1/3}}dy\\&\le(2\pi m t_0)^{1/2}\left(g(x_0-\nu t_0-t_0^{1/3})\mathbb{E}_{x_0-\nu t_0}\left[e^{\alpha\kappa B_{mt_0}}\right]+\mathbb{P}_0\left(|B_{mt_0}|>t_0^{1/3}\right)\right)\\&\le(2\pi m t_0)^{1/2}\left(g(x_0-\nu t_0-t_0^{1/3})e^{\alpha\kappa(x_0-\nu t_0)}e^{\frac{1}{2}m\alpha^2\kappa^2t_0}+2e^{-t_0^{-1/3}/(2m)}\right)\end{aligned}$$

by a Gaussian tail bound. Therefore if $|x_0|\le d_n$, by (4.23) and since $|\frac{\nabla g(y)}{g(y)}|\le\kappa\ \forall y\in\mathbb{R}$ and $g(y)^{-1}e^{-\alpha\kappa y}\le 2e^{\kappa|y|}\ \forall y\in\mathbb{R}$,

$$\begin{aligned}&\mathbb{E}_{x-\nu t}\left[g(Z_{t-t_0})^{-1}e^{-(Z_{t-t_0}+\nu t_0-x_0)^2/(2mt_0)}\right]\\&\le(2\pi m t_0)^{1/2}\pi(x_0)g(x_0)^{-1}(1+\mathcal{O}(t_0^{1/3})+\mathcal{O}(t_0^{-1/2}e^{2\kappa d_n}(\log N)^{-12C})).\end{aligned}$$

Substituting into (4.22), we have that if $t-t_0\geq \gamma_n/2$, $|x-\nu t|\leq d_n+1$ and $|x_0|\leq d_n$,

$$\frac{v_t(x)}{g(x-\nu t)}\leq n^{-1}p_0^n(x_0)\pi(x_0)g(x_0)^{-1}(1+\mathcal{O}((\log N)^{-4C})). \tag{4.24}$$

For a lower bound, note that by (3.12) with $a=(1-\alpha)s_0$ and since $(1-u)(2u-1+\alpha)\geq \alpha-1$ $\forall u\in[0,1]$, for $y\in\frac{1}{n}\mathbb{Z}$,

$$v_{t_0}^n(y)\geq e^{-(1-\alpha)s_0t_0}p_0^n(x_0)P_{t_0,x_0}^n(y).$$

Suppose n is sufficiently large that $t_0^{1/3}+n^{-1}\leq \frac{1}{2}mnt_0$, and then by (4.19),

$$\begin{aligned} v_t(x) &\geq g(x-\nu t)\mathbb{E}_{x-\nu t}\left[g(Z_{t-t_0})^{-1}e^{-(1-\alpha)s_0t_0}p_0^n(x_0)\bar{P}_{t_0,x_0}^n(Z_{t-t_0}+\nu t_0)\mathbb{1}_{|Z_{t-t_0}+\nu t_0-x_0|<t_0^{1/3}}\right] \\ &\geq g(x-\nu t)p_0^n(x_0)e^{-(1-\alpha)s_0t_0}g(x_0-\nu t_0-t_0^{1/3})^{-1} \\ &\quad \mathbb{E}_{x-\nu t}\left[n^{-1}(2\pi m t_0)^{-1/2}e^{-(Z_{t-t_0}+\nu t_0-x_0)^2/(2mt_0)}e^{\mathcal{O}(n^{-1}t_0^{-2})}\mathbb{1}_{|Z_{t-t_0}+\nu t_0-x_0|<t_0^{1/3}}\right] \end{aligned} \tag{4.25}$$

by Lemma 3.14. By Lemma 4.4, for n sufficiently large, if $t-t_0\geq\gamma_n/2$ and $|x-\nu t|\leq d_n+1$,

$$\begin{aligned} &\mathbb{E}_{x-\nu t}\left[e^{-(Z_{t-t_0}+\nu t_0-x_0)^2/(2mt_0)}\mathbb{1}_{|Z_{t-t_0}+\nu t_0-x_0|<t_0^{1/3}}\right] \\ &\geq \int_{-\infty}^{\infty}\pi(y)e^{-(y+\nu t_0-x_0)^2/(2mt_0)}\mathbb{1}_{|y+\nu t_0-x_0|<t_0^{1/3}}dy-(\log N)^{-12C}. \end{aligned} \tag{4.26}$$

Since $y\mapsto g(y)$ is decreasing,

$$\begin{aligned} &\int_{-\infty}^{\infty}g(y)^2e^{\alpha\kappa y}e^{-(y+\nu t_0-x_0)^2/(2mt_0)}\mathbb{1}_{|y+\nu t_0-x_0|<t_0^{1/3}}dy \\ &\geq g(x_0-\nu t_0+t_0^{1/3})^2e^{\alpha\kappa(x_0-\nu t_0-t_0^{1/3})}(2\pi m t_0)^{1/2}\left(1-\mathbb{P}_0\left(|B_{mt_0}|>t_0^{1/3}\right)\right) \\ &\geq g(x_0)^2e^{\alpha\kappa x_0}(2\pi m t_0)^{1/2}(1-\mathcal{O}(e^{-t_0^{-1/3}/(2m)})-\mathcal{O}(t_0^{1/3})) \end{aligned}$$

by a Gaussian tail bound and since $|\frac{\nabla g(y)}{g(y)}|\leq\kappa$ $\forall y\in\mathbb{R}$. Therefore if $t-t_0\geq\gamma_n/2$, $|x-\nu t|\leq d_n+1$ and $|x_0|\leq d_n$, by (4.26) and (4.25), and since $(\log N)^{-12C}t_0^{-1/2}\pi(x_0)^{-1}=\mathcal{O}((\log N)^{-4C})$,

$$\frac{v_t(x)}{g(x-\nu t)}\geq p_0^n(x_0)n^{-1}\pi(x_0)g(x_0)^{-1}(1-\mathcal{O}((\log N)^{-4C})). \tag{4.27}$$

It remains to bound $|v_{\gamma_n}^n(x)-v_{\gamma_n}(x)|$. By (4.20) and Lemma 3.14, for $z\in\frac{1}{n}\mathbb{Z}$ and $t>0$,

$$v_t^n(z)\leq e^{2s_0t}p_0^n(x_0)n^{-1}(2\pi m t)^{-1/2}e^{\mathcal{O}(n^{-1}t^{-1/2})}. \tag{4.28}$$

Therefore, by Lemma 4.3, for n sufficiently large,

$$\begin{aligned} &\sup_{x\in\frac{1}{n}\mathbb{Z}}|v_{\gamma_n}^n(x)-v_{\gamma_n}(x)| \\ &\leq\Big(C_7(n^{-1/3}+e^{-(\log N)^{c_2}})e^{2s_0t_0}p_0^n(x_0)(mt_0)^{-1/2}n^{-1}+2n^{-1/3}\sup_{z\in\frac{1}{n}\mathbb{Z}}|\nabla_n v_{t_0}^n(z)|\Big)e^{5s_0\gamma_n}\gamma_n^2. \end{aligned} \tag{4.29}$$

Let $t_1 = t_0/2$; then for $z \in \frac{1}{n}\mathbb{Z}$, by (3.12), and then using (4.28) and Lemma 3.7 in the last inequality,

$$\begin{aligned}
&|\nabla_n v^n_{t_0}(z)| \\
&= \Big| n\langle v^n_{t_1}, \phi_0^{t_1,z+n^{-1}} - \phi_0^{t_1,z}\rangle_n \\
&\qquad + ns_0 \int_0^{t_1} \langle v^n_{t_1+s}(1-u^n_{t_1+s})(2u^n_{t_1+s}-1+\alpha), \phi_s^{t_1,z+n^{-1}} - \phi_s^{t_1,z}\rangle_n ds \Big| \\
&\le \sup_{x\in\frac{1}{n}\mathbb{Z}, s\in[0,t_1]} v^n_{t_1+s}(x) \left(n\langle 1, |\phi_0^{t_1,z+n^{-1}} - \phi_0^{t_1,z}|\rangle_n + ns_0\int_0^{t_1} \langle 1+\alpha, |\phi_s^{t_1,z+n^{-1}} - \phi_s^{t_1,z}|\rangle_n ds \right) \\
&\le e^{2s_0 t_0} p_0^n(x_0) n^{-1}(mt_1)^{-1/2} \left(C_5 t_1^{-1/2} + \int_0^{t_1} 2s_0 C_5 (t_1-s)^{-1/2} ds \right)
\end{aligned}$$

for n sufficiently large. Hence

$$\sup_{z\in\frac{1}{n}\mathbb{Z}} |\nabla_n v^n_{t_0}(z)| \le e^{2s_0t_0} p_0^n(x_0) n^{-1} m^{-1/2} C_5 (2t_0^{-1} + 4s_0).$$

By (4.29) it follows that for n sufficiently large,

$$\sup_{x\in\frac{1}{n}\mathbb{Z}} |v^n_{\gamma_n}(x) - v_{\gamma_n}(x)| \le p_0^n(x_0) n^{-1}(e^{-\frac{1}{2}(\log N)^{c_2}} \vee n^{-1/6}).$$

By (4.24) and (4.27), this completes the proof. □

We now show that $|q^n_{\gamma_n} - v^n_{\gamma_n}|$ is small with high probability, which, combined with the previous lemma, will imply that $A_t^{(1)}(x_1, x_2)$ occurs with high probability for suitable x_1, x_2 and t. This result is stronger than Proposition 3.2 (but only applies when $q_0^n(x) = p_0^n(x_0)\mathbb{1}_{x=x_0}$ for some x_0), and will also be used to show that $A_t^{(4)}(x)$ occurs with high probability for suitable x and t.

Lemma 4.7. *For $c, c' \in (0, 1/2)$ and $\ell \in \mathbb{N}$, the following holds for n sufficiently large. Suppose $N \ge n^3$, and for some $x_0 \in \frac{1}{n}\mathbb{Z}$, $q_0^n(x) = p_0^n(x_0)\mathbb{1}_{x=x_0}$ and $p_0^n(x_0) \ge \left(\frac{n^2}{N}\right)^{1-c}$. For $t \le \gamma_n$ and $z \in \frac{1}{n}\mathbb{Z}$,*

$$\mathbb{P}\left(|q_t^n(z) - v_t^n(z)| \ge \left(\frac{n}{N}\right)^{1/2-c'} p_0^n(x_0)^{1/2} n^{-1/2} \right) \le \left(\frac{n}{N}\right)^{\ell},$$

where $(q_t^n)_{t\ge0}$ and $(v_t^n)_{t\ge0}$ are defined in (3.10) *and* (3.11) *respectively.*

Proof. By Lemma 3.14, there exists a constant $K_5 > 1$ such that

$$\mathbf{P}_0\left(X^n_{mt} = 0\right) \le K_5 n^{-1} t^{-1/2} \quad \forall\, n \in \mathbb{N},\, t > 0. \tag{4.30}$$

By Corollary 3.13 with $a = -(1+\alpha)s_0$, for $t \ge 0$ and $z \in \frac{1}{n}\mathbb{Z}$,

$$\begin{aligned}
q_t^n(z) &\le e^{(1+\alpha)s_0 t}\langle q_0^n, \phi_0^{t,z}\rangle_n + M_t^n(\phi^{t,z,-(1+\alpha)s_0}) \\
&\le e^{(1+\alpha)s_0 t} p_0^n(x_0) \min(K_5 n^{-1} t^{-1/2}, 1) + M_t^n(\phi^{t,z,-(1+\alpha)s_0})
\end{aligned} \tag{4.31}$$

by (4.30). Let

$$\tau = \inf\left\{ t > 0 : \sup_{x\in\frac{1}{n}\mathbb{Z}} q_t^n(x) \ge K_5 e^{2s_0\gamma_n} p_0^n(x_0) n^{-1} t^{-1/2} \right\}.$$

We will show that $\tau > \gamma_n$ with high probability. By Lemma 3.12, for $t > 0$,

$$\sup_{s\in[0,t]} |M^n_s(\phi^{t,z,-(1+\alpha)s_0}) - M^n_{s-}(\phi^{t,z,-(1+\alpha)s_0})| = \sup_{s\in[0,t]} |\langle q^n_s - q^n_{s-}, \phi^{t,z,-(1+\alpha)s_0}_s\rangle_n| \\ \le e^{(1+\alpha)s_0 t} N^{-1}.$$

Therefore, by the Burkholder-Davis-Gundy inequality as stated in Lemma 3.17, for $t \ge 0$, $z \in \frac{1}{n}\mathbb{Z}$ and $k \in \mathbb{N}$ with $k \ge 2$,

$$\mathbb{E}\left[\sup_{s\in[0,t]} |M^n_{s\wedge\tau}(\phi^{t,z,-(1+\alpha)s_0})|^k\right] \le C(k)\mathbb{E}\left[\langle M^n(\phi^{t,z,-(1+\alpha)s_0})\rangle^{k/2}_{t\wedge\tau} + e^{(1+\alpha)s_0 tk} N^{-k}\right]. \tag{4.32}$$

For $t \le \gamma_n$, by the definition of τ and by Lemma 3.12, and then by Lemma 3.15,

$$\langle M^n(\phi^{t,z,-(1+\alpha)s_0})\rangle_{t\wedge\tau} \le \frac{n}{N}\int_0^t \langle (1+2m)K_5 e^{2s_0\gamma_n} p^n_0(x_0) n^{-1} s^{-1/2}, (\phi^{t,z}_s)^2 e^{2(1+\alpha)s_0(t-s)}\rangle_n ds \\ \le \frac{n}{N}(1+2m)K_5 e^{6s_0\gamma_n} p^n_0(x_0)\int_0^t s^{-1/2}\mathbf{P}_0\left(X^n_{2m(t-s)} = 0\right) ds. \tag{4.33}$$

Then by (4.30),

$$\int_0^t s^{-1/2}\mathbf{P}_0\left(X^n_{2m(t-s)} = 0\right) ds \le \int_0^t s^{-1/2} K_5 n^{-1}(2(t-s))^{-1/2} ds \\ = K_5 n^{-1} 2^{-1/2}\cdot 2\int_0^{t/2} s^{-1/2}(t-s)^{-1/2} ds \\ \le 2^{3/2} K_5 n^{-1}.$$

Hence, by (4.33), for $t \le \gamma_n$,

$$\langle M^n(\phi^{t,z,-(1+\alpha)s_0})\rangle_{t\wedge\tau} \le \frac{1}{N}(1+2m)2^{3/2}K_5^2 e^{6s_0\gamma_n} p^n_0(x_0). \tag{4.34}$$

For $b \in (0, 1/2)$ and $\ell_1 \in \mathbb{N}$, take $k \in \mathbb{N}$ with $k > \ell_1/b$. Then for n sufficiently large, for $t \le \gamma_n$ and $z \in \frac{1}{n}\mathbb{Z}$, by Markov's inequality and (4.32), and since $p^n_0(x_0)^{1/2}N^{-1/2} \ge (\frac{n^2}{N})^{1/2}N^{-1/2} = nN^{-1}$,

$$\mathbb{P}\left(|M^n_{t\wedge\tau}(\phi^{t,z,-(1+\alpha)s_0})| \ge \left(\frac{n}{N}\right)^{1/2-b} p^n_0(x_0)^{1/2} n^{-1/2}\right) \\ \le \left(\frac{n}{N}\right)^{-k(1/2-b)} p^n_0(x_0)^{-k/2} n^{k/2} C(k)\cdot 2\left(\frac{1}{N}(1+2m)2^{3/2}K_5^2 e^{6s_0\gamma_n} p^n_0(x_0)\right)^{k/2} \\ \le \left(\frac{n}{N}\right)^{\ell_1} \tag{4.35}$$

for n sufficiently large, since $bk > \ell_1$ and $\gamma_n = \lfloor(\log\log N)^4\rfloor$. Now let $b = c/4$. Then for n sufficiently large, since $N \ge n^3$ and then since $p^n_0(x_0) \ge (\frac{n^2}{N})^{1-c}$,

$$\left(\frac{n}{N}\right)^{1/2-b} n^{-1/2} \le \left(\frac{n^2}{N}\right)^{(1-c)/2} n^{-1} \le \tfrac{1}{3}K_5 e^{2s_0\gamma_n}(\gamma_n + N^{-1})^{-1/2} p^n_0(x_0)^{1/2} n^{-1}. \tag{4.36}$$

Since $p^n_0(x_0) \ge n^2 N^{-1}$, we can take n sufficiently large that

$$N^{-1} \le \tfrac{1}{3}K_5 e^{2s_0\gamma_n}(\gamma_n + N^{-1})^{-1/2} p^n_0(x_0) n^{-1} \tag{4.37}$$

and also, since $\alpha < 1$ and $N \geq n^3$,

$$e^{(1+\alpha)s_0 t} t^{-1/2} \leq \tfrac{1}{3} e^{2s_0\gamma_n} (t+N^{-1})^{-1/2} \; \forall t \in [N^{-1}, \gamma_n] \quad \text{and} \quad \tfrac{1}{3} n^{-1} (2N^{-1})^{-1/2} \geq 1. \tag{4.38}$$

If $|M^n_{t\wedge\tau}(\phi^{t,z,-(1+\alpha)s_0})| \leq \left(\frac{n}{N}\right)^{1/2-b} p^n_0(x_0)^{1/2} n^{-1/2}$ and $t \in [0, \tau \wedge \gamma_n]$ then by (4.31), and since $K_5 > 1$,

$$\begin{aligned} q^n_t(z) &\leq K_5 e^{(1+\alpha)s_0 t} p^n_0(x_0) \min(n^{-1}t^{-1/2}, 1) + \left(\frac{n}{N}\right)^{1/2-b} p^n_0(x_0)^{1/2} n^{-1/2} \\ &\leq K_5 e^{2s_0\gamma_n} (t+N^{-1})^{-1/2} p^n_0(x_0) n^{-1} - N^{-1}, \end{aligned} \tag{4.39}$$

by (4.36), (4.37) and (4.38) (using the second equation in (4.38) for the case $t \leq N^{-1}$). Take $\ell_2 \in \mathbb{N}$ and let $Y_n \sim$ Poisson$((2m+1)N^{2-\ell_2} r_n)$. Then for $t \geq 0$ and $z \in \frac{1}{n}\mathbb{Z}$, since $(q^n_s(z))_{s\geq 0}$ jumps at rate at most $(2m+1)r_n N^2$,

$$\mathbb{P}\left(\sup_{s\in[0,N^{-\ell_2}]} |q^n_{t+s}(z) - q^n_t(z)| > N^{-1} \right) \leq \mathbb{P}(Y_n \geq 2) \leq (\tfrac{1}{2}(2m+1)N^{1-\ell_2} n^2)^2 \tag{4.40}$$

since $r_n = \frac{1}{2} n^2 N^{-1}$. Therefore, for $\ell_1, \ell_2 \in \mathbb{N}$, letting $\mathcal{A} = N^{-\ell_2}\mathbb{N}_0 \cap [0, \gamma_n]$, by a union bound and (4.39),

$$\begin{aligned} &\mathbb{P}(\tau \leq \gamma_n) \\ &\leq \mathbb{P}\left(\exists t \in \mathcal{A}, z \in \tfrac{1}{n}\mathbb{Z} : |z - x_0| \leq N^5, |M^n_{t\wedge\tau}(\phi^{t,z,-(1+\alpha)s_0})| \geq \left(\frac{n}{N}\right)^{1/2-b} p^n_0(x_0)^{1/2} n^{-1/2} \right) \\ &\quad + \mathbb{P}\left(\exists t \in \mathcal{A}, z \in \tfrac{1}{n}\mathbb{Z} : |z - x_0| \leq N^5, \sup_{s\in[0,N^{-\ell_2}]} |q^n_{t+s}(z) - q^n_t(z)| > N^{-1} \right) \\ &\quad + \mathbb{P}\left(\exists z \in \tfrac{1}{n}\mathbb{Z}, t \in [0,\gamma_n] : |z - x_0| > N^5, q^n_t(z) > 0 \right) \\ &\leq \sum_{t\in\mathcal{A}} (2nN^5+1) \left(\frac{n}{N}\right)^{\ell_1} + \sum_{t\in\mathcal{A}} (2nN^5+1)(\tfrac{1}{2}(2m+1)N^{1-\ell_2}n^2)^2 + 2e^{-N^5}, \end{aligned}$$

for n sufficiently large, by (4.35) and (4.40), and by the same argument as Lemma 3.3 for the last term. For $\ell' \in \mathbb{N}$, take ℓ_2 sufficiently large that $\gamma_n N^{\ell_2+5} n (N^{1-\ell_2}n^2)^2 = \gamma_n N^{7-\ell_2} n^5 \leq \left(\frac{n}{N}\right)^{\ell'+1}$ for n sufficiently large, and then take ℓ_1 sufficiently large that $\gamma_n N^{\ell_2+5} n \left(\frac{n}{N}\right)^{\ell_1} \leq \left(\frac{n}{N}\right)^{\ell'+1}$ for n sufficiently large. It follows that for n sufficiently large,

$$\mathbb{P}(\tau \leq \gamma_n) \leq \left(\frac{n}{N}\right)^{\ell'}. \tag{4.41}$$

Note that by (3.13) and then by (4.30), for $t \geq 0$ and $z \in \frac{1}{n}\mathbb{Z}$,

$$v^n_t(z) \leq e^{(1+\alpha)s_0 t} \langle q^n_0, \phi^{t,z}_0 \rangle_n \leq e^{(1+\alpha)s_0 t} p^n_0(x_0) \min(K_5 n^{-1} t^{-1/2}, 1). \tag{4.42}$$

Take $k \in \mathbb{N}$ with $k \geq 2$. By Lemma 3.16 and since $q^n_t, v^n_t \in [0,1]$, we have that for $t \geq 0$ and $z \in \frac{1}{n}\mathbb{Z}$,

$$\begin{aligned} &|q^n_t(z) - v^n_t(z)|^k \\ &\leq 3^{2k-1} s_0^k t^{k-1} \left(\int_0^t \langle |q^n_s - v^n_s|^k, \phi^{t,z}_s \rangle_n ds + \int_0^t \sup_{x\in\frac{1}{n}\mathbb{Z}} v^n_s(x)^k \langle |p^n_s - u^n_s|^k, \phi^{t,z}_s \rangle_n ds \right) \\ &\quad + \mathbb{1}_{\tau<t} + 3^{k-1} |M^n_{t\wedge\tau}(\phi^{t,z})|^k. \end{aligned}$$

Therefore, by (3.14) in Proposition 3.2 and by (4.42) and (4.41), for $\ell' \in \mathbb{N}$, for n sufficiently large, for $t \le \gamma_n$ and $z \in \frac{1}{n}\mathbb{Z}$,

$$\begin{aligned}
&\mathbb{E}\left[|q_t^n(z) - v_t^n(z)|^k\right] \\
&\le 3^{2k-1} s_0^k t^{k-1} \int_0^t \sup_{x \in \frac{1}{n}\mathbb{Z}} \mathbb{E}\left[|q_s^n(x) - v_s^n(x)|^k\right] ds \\
&\quad + 3^{2k-1} s_0^k t^{k-1} e^{(1+\alpha)s_0 tk} p_0^n(x_0)^k \int_0^t (K_5 n^{-1} s^{-1/2} \wedge 1)^k C_1 \left(\frac{n^{k/2} s^{k/4}}{N^{k/2}} + N^{-k}\right) e^{C_1 s^k} ds \\
&\quad + \left(\frac{n}{N}\right)^{\ell'} + 3^{k-1} \mathbb{E}\left[|M^n_{t\wedge\tau}(\phi^{t,z})|^k\right]. \qquad (4.43)
\end{aligned}$$

Take ℓ' sufficiently large that for n sufficiently large,

$$\left(\frac{n}{N}\right)^{\ell'} \le N^{-k/2} \left(\frac{n^2}{N}\right)^{k/2} \le N^{-k/2} p_0^n(x_0)^{k/2}.$$

Note that for the second term on the right hand side of (4.43),

$$\begin{aligned}
&\int_0^t (K_5 n^{-1} s^{-1/2} \wedge 1)^k C_1 \left(\frac{n^{k/2} s^{k/4}}{N^{k/2}} + N^{-k}\right) e^{C_1 s^k} ds \\
&\le C_1 \int_0^t (K_5^{k/2} N^{-k/2} + N^{-k}) e^{C_1 s^k} ds \\
&\le C_1 (K_5^{k/2} N^{-k/2} + N^{-k}) t e^{C_1 t^k}.
\end{aligned}$$

By the same argument as in (4.32) and (4.34), since $t \le \gamma_n$,

$$\mathbb{E}\left[|M^n_{t\wedge\tau}(\phi^{t,z})|^k\right] \le C(k) \left(\left(\frac{1}{N}(1+2m) 2^{3/2} K_5^2 e^{2s_0\gamma_n} p_0^n(x_0) \right)^{k/2} + N^{-k} \right).$$

Note that $N^{-1/2} p_0^n(x_0)^{1/2} \ge nN^{-1}$. Hence substituting into (4.43) and then by Gronwall's inequality, there exists a constant $K_6 = K_6(k)$ such that for n sufficiently large, for $t \in [0, \gamma_n]$,

$$\begin{aligned}
&\sup_{x \in \frac{1}{n}\mathbb{Z}} \mathbb{E}\left[|q_t^n(x) - v_t^n(x)|^k\right] \\
&\le K_6 (\gamma_n^k e^{(1+\alpha)s_0\gamma_n k} e^{C_1 \gamma_n^k} + 1 + e^{s_0 \gamma_n k}) N^{-k/2} p_0^n(x_0)^{k/2} e^{3^{2k-1} s_0^k \gamma_n^{k-1} t}. \qquad (4.44)
\end{aligned}$$

The result now follows by Markov's inequality, taking $k \in \mathbb{N}$ sufficiently large that $kc' > \ell$, and then taking n sufficiently large that (4.44) holds with this choice of k. □

We are now ready to prove that $A_t^{(1)}(x_1, x_2)$ occurs with high probability for suitable t, x_1 and x_2. For $t \ge 0$ and $x_1 \in \frac{1}{n}\mathbb{Z}$, let $(v^n_{t,t+s}(x_1, \cdot))_{s\ge 0}$ denote the solution of

$$\begin{cases} \partial_s v^n_{t,t+s}(x_1,\cdot) = \frac{1}{2} m \Delta_n v^n_{t,t+s}(x_1,\cdot) + s_0 v^n_{t,t+s}(x_1,\cdot)(1 - u^n_{t,t+s})(2u^n_{t,t+s} - 1 + \alpha) \text{ for } s > 0, \\ v^n_{t,t}(x_1, x) = p_t^n(x_1) \mathbb{1}_{x = x_1}, \end{cases} \qquad (4.45)$$

where $(u^n_{t,t+s})_{s\ge 0}$ is defined in (3.2). Recall the definition of $q^n_{t_1,t_2}(x_1,x_2)$ in (2.2).

Proposition 4.8. *Suppose $N \ge n^3$ for n sufficiently large. For $\ell \in \mathbb{N}$, the following holds for n sufficiently large. For $t \in [(\log N)^2 - \gamma_n, N^2 - \gamma_n]$ and $x_1, x_2 \in \frac{1}{n}\mathbb{Z}$,*

$$\mathbb{P}\left(A_t^{(1)}(x_1,x_2)^c \cap \{|x_1 - \mu_t^n| \vee |x_2 - \mu^n_{t+\gamma_n}| \le d_n\} \cap E_1' \right) \le \left(\frac{n}{N}\right)^{\ell}.$$

Proof. Fix $c' \in (0,1/4)$. By Lemma 4.7, for n sufficiently large,

$$\mathbb{P}\left(\left\{|q^n_{t,t+\gamma_n}(x_1,x_2)-v^n_{t,t+\gamma_n}(x_1,x_2)| \geq \left(\frac{n}{N}\right)^{1/2-c'} n^{-1/2}\right\} \cap \left\{p^n_t(x_1) \geq \left(\frac{n^2}{N}\right)^{3/4}\right\}\right) \leq \left(\frac{n}{N}\right)^{\ell}. \tag{4.46}$$

Suppose n is sufficiently large that $(\log N)^2 - \gamma_n \geq \frac{1}{2}(\log N)^2 \vee \log N$. Recall the definition of E'_1 in (3.3). By Lemma 4.6, if E'_1 occurs and $|x_1-\mu^n_t| \leq d_n$, $|x_2-\nu\gamma_n-\mu^n_t| \leq d_n+1$ then

$$\frac{v^n_{t,t+\gamma_n}(x_1,x_2)}{g(x_2-\nu\gamma_n-\mu^n_t)} = \frac{\pi(x_1-\mu^n_t)}{g(x_1-\mu^n_t)} p^n_t(x_1) n^{-1}(1+\mathcal{O}((\log N)^{-4C})).$$

Suppose $|x_1-\mu^n_t| \vee |x_2-\mu^n_{t+\gamma_n}| \leq d_n$ and E'_1 occurs. Then if n is sufficiently large, by the definition of E_1 in (2.10) we have $p^n_t(x_1) \wedge p^n_{t+\gamma_n}(x_2) \geq \frac{1}{10}(\log N)^{-C}$, $|x_2-\nu\gamma_n-\mu^n_t| \leq d_n+1$, $|p^n_t(x_1)-g(x_1-\mu^n_t)| \leq e^{-(\log N)^{c_2}}$, $|p^n_{t+\gamma_n}(x_2)-g(x_2-\mu^n_{t+\gamma_n})| \leq e^{-(\log N)^{c_2}}$ and $|\mu^n_{t+\gamma_n}-(\mu^n_t+\nu\gamma_n)| \leq \lceil \gamma_n \rceil e^{-(\log N)^{c_2}}$. Hence for n sufficiently large, if $|q^n_{t,t+\gamma_n}(x_1,x_2) - v^n_{t,t+\gamma_n}(x_1,x_2)| \leq \left(\frac{n}{N}\right)^{1/2-c'} n^{-1/2} \leq n^{-3/2+2c'}$, then $A^{(1)}_t(x_1,x_2)$ occurs. By (4.46), this completes the proof. □

The next two lemmas will be used to show that $A^{(2)}_t(x_1,x_2)$ and $A^{(3)}_t(x_1,x_2)$ occur with high probability for suitable t, x_1 and x_2. Recall that we fixed $c_1>0$ at the start of Section 4, and recall the definition of D^+_n in (2.8).

Lemma 4.9. *For $\epsilon > 0$ sufficiently small, $t^* \in \mathbb{N}$ sufficiently large and $K \in \mathbb{N}$ sufficiently large (depending on t^*), the following holds for n sufficiently large. Suppose $\sup_{s\in[0,t^*],x\in\frac{1}{n}\mathbb{Z}} |u^n_s(x)-g(x-\nu s)| < \epsilon$, and also $p^n_t(x) \in [\frac{1}{6}g(x-\nu t), 6g(x-\nu t)]$ $\forall t \in [0,t^*]$, $x \leq \nu t + D^+_n + 1$ and $p^n_t(x) \leq 6g(D^+_n)$ $\forall t \in [0,t^*], x \geq \nu t + D^+_n$. Suppose $q^n_0(z) = p^n_0(z)\mathbb{1}_{z\geq \ell}$ for some $\ell \in \frac{1}{n}\mathbb{Z} \cap [K, D^+_n]$. Then for $z \leq \nu t^* + D^+_n + 1$,*

$$\frac{v^n_{t^*}(z)}{p^n_{t^*}(z)} \leq \tfrac{1}{2} c_1 e^{-(1+\frac{1}{2}(1-\alpha))\kappa(\ell-(z-\nu t^*)\vee K+2)},$$

where $(v^n_t)_{t\geq 0}$ is defined in (3.11).

Proof. Let $\lambda = \frac{1}{2}(1-\alpha)$. Note that since $(\alpha-2)^2>1$, we have $\frac{1}{4}(1-\alpha^2) < 1-\alpha$. Take $a \in (\frac{1}{4}(1-\alpha^2), 1-\alpha)$ so that

$$\lambda^2+\lambda\alpha-a = \tfrac{1}{2}(1-\alpha)(\tfrac{1}{2}(1-\alpha)+\alpha)-a = \tfrac{1}{4}(1-\alpha^2)-a < 0.$$

Take $t^* \in \mathbb{N}$ sufficiently large that $144 e^{(\lambda^2+\lambda\alpha-a)s_0 t^*} \leq \frac{1}{3}c_1 e^{-2\kappa(1+\lambda)}$. Take $\epsilon \in (0, \frac{1}{2}(1-\alpha))$ sufficiently small that $(1-\epsilon)(2\epsilon-1+\alpha) < -a$. Then take $K \in \mathbb{N}$ sufficiently large that $\nu t^* \leq K/6$, $2s_0 t^* e^{4s_0 t^*} e^{-\lambda\kappa K/6} \leq 1$, $72 e^{5s_0 t^*} e^{-(1-\lambda)\kappa K/2} \leq \frac{1}{2}c_1 e^{-2\kappa(1+\lambda)}$, $2g(K/3)+2\epsilon < 1-\alpha$ and

$$(1-g(x)-\epsilon)(2(g(x)+\epsilon)-1+\alpha) \leq -a \qquad \text{for } x \geq K/3.$$

Then for $s \geq 0$ and $x \in \frac{1}{n}\mathbb{Z}$, if $x-\nu s \geq K/3$ and $|u^n_s(x)-g(x-\nu s)| < \epsilon$ we have

$$(1-u^n_s(x))(2u^n_s(x)-1+\alpha)+a \leq 0. \tag{4.47}$$

If instead $x-\nu s \leq K/3$, then by (3.13),

$$v^n_s(x) \leq e^{(1+\alpha)s_0 s}\mathbf{E}_x\left[p^n_0(X^n_{ms})\mathbb{1}_{X^n_{ms}\geq \ell}\right] \leq e^{(1+\alpha)s_0 s}\sup_{y\geq \ell} p^n_0(y)\mathbf{P}_0\left(X^n_{ms} \geq \ell - \tfrac{1}{3}K-\nu s\right).$$

Moreover, for $u \in [0,1]$, we have $(1-u)(2u-1+\alpha)+a \le 2$.

Suppose $\ell \in [K, D_n^+]$ and $\sup_{s\in[0,t^*],x\in\frac{1}{n}\mathbb{Z}} |u_s^n(x) - g(x-\nu s)| < \epsilon$. For $z \in \frac{1}{n}\mathbb{Z}$ and $t \in [0,t^*]$ we have by (3.12) and (4.47) that

$$\begin{aligned} v_t^n(z) &\le e^{-as_0t}\langle q_0^n, \phi_0^{t,z}\rangle_n + \int_0^t 2s_0 e^{-as_0(t-s)} \sup_{x-\nu s\le K/3} v_s^n(x)ds \\ &\le \sup_{x\ge\ell} p_0^n(x)\left(e^{-as_0t}\mathbf{P}_z\left(X_{mt}^n \ge \ell\right) + 2s_0e^{(1+\alpha)s_0t}\int_0^t \mathbf{P}_0\left(X_{ms}^n \ge \ell - \tfrac{1}{3}K - \nu s\right)ds\right). \end{aligned} \tag{4.48}$$

By Markov's inequality and Lemma 3.8, and since $\frac{1}{2}m\kappa^2 = s_0$,

$$\begin{aligned} \mathbf{P}_z\left(X_{mt}^n \ge \ell\right) = \mathbf{P}_0\left(X_{mt}^n \ge \ell - z\right) &\le e^{-\lambda\kappa(\ell-z)}\mathbf{E}_0\left[e^{\lambda\kappa X_{mt}^n}\right] \\ &= e^{-\lambda\kappa(\ell-z)}e^{(\lambda^2+\mathcal{O}(n^{-1}))s_0t}. \end{aligned}$$

Therefore, applying the same argument to the second term on the right hand side of (4.48),

$$\begin{aligned} v_t^n(z) &\le \sup_{x\ge\ell} p_0^n(x)(e^{-\lambda\kappa(\ell-z)}e^{(\lambda^2-a+\mathcal{O}(n^{-1}))s_0t} + 2s_0te^{(1+\alpha)s_0t}e^{-\lambda\kappa(\ell-\frac{1}{3}K-\nu t)}e^{(\lambda^2+\mathcal{O}(n^{-1}))s_0t}) \\ &\le \sup_{x\ge\ell} p_0^n(x)e^{-\lambda\kappa(\ell-z)}e^{(\lambda^2-a+\mathcal{O}(n^{-1}))s_0t}(1+2s_0te^{(1+\alpha+a+\lambda\alpha)s_0t}e^{-\lambda\kappa(z-\frac{1}{3}K)}), \end{aligned}$$

since $\kappa\nu = \alpha s_0$. Hence for $z \in [\frac{1}{2}K + \nu t^*, D_n^+ + 1 + \nu t^*]$, using our choice of K in the second inequality, using that $\kappa\nu = \alpha s_0$ in the third line, and using our choice of t^* in the last inequality,

$$\begin{aligned} \frac{v_{t^*}^n(z)}{p_{t^*}^n(z)} &\le \frac{6g(\ell)}{\frac{1}{6}g(z-\nu t^*)}e^{-\lambda\kappa(\ell-z)}e^{(\lambda^2-a+\mathcal{O}(n^{-1}))s_0t^*}(1+2s_0t^*e^{4s_0t^*}e^{-\lambda\kappa K/6}) \\ &\le 36e^{-\kappa\ell}\cdot 2e^{\kappa(z-\nu t^*)}e^{-\lambda\kappa(\ell-z)}e^{(\lambda^2-a+\mathcal{O}(n^{-1}))s_0t^*}\cdot 2 \\ &= 144e^{-(1+\lambda)\kappa(\ell-(z-\nu t^*))}e^{(\lambda^2+\alpha\lambda-a+\mathcal{O}(n^{-1}))s_0t^*} \\ &\le \tfrac{1}{2}c_1e^{-(1+\lambda)\kappa(\ell-(z-\nu t^*)+2)} \end{aligned} \tag{4.49}$$

for n sufficiently large. Also, for any $z \in \frac{1}{n}\mathbb{Z}$ and $t \ge 0$, by (3.13) and then by Markov's inequality and Lemma 3.8, and since $\frac{1}{2}m\kappa^2 = s_0$,

$$\begin{aligned} v_t^n(z) \le e^{(1+\alpha)s_0t}\sup_{x\ge\ell} p_0^n(x)\mathbf{P}_z\left(X_{mt}^n \ge \ell\right) &\le e^{(1+\alpha)s_0t}\sup_{x\ge\ell} p_0^n(x)e^{-\kappa(\ell-z)}\mathbf{E}_0\left[e^{\kappa X_{mt}^n}\right] \\ &\le e^{(1+\alpha)s_0t}\sup_{x\ge\ell} p_0^n(x)e^{2s_0t}e^{-\kappa(\ell-z)} \end{aligned}$$

for n sufficiently large. Therefore, for $z \le \frac{1}{2}K + \nu t^* \le \frac{2}{3}K$, using that $g(\ell) \le e^{-\kappa\ell}$, $g(K/2)^{-1} \le 2e^{\kappa K/2}$ and $\kappa\nu = \alpha s_0$ in the second inequality, using that $\ell - \frac{1}{2}K \ge \frac{1}{2}K$ in the third inequality, and using our choice of K in the last inequality,

$$\begin{aligned} \frac{v_{t^*}^n(z)}{p_{t^*}^n(z)} \le e^{(1+\alpha)s_0t^*}\frac{6g(\ell)}{\frac{1}{6}g(K/2)}e^{2s_0t^*}e^{-\kappa(\ell-\frac{1}{2}K-\nu t^*)} &\le 72e^{5s_0t^*}e^{-2\kappa(\ell-\frac{1}{2}K)} \\ &\le 72e^{5s_0t^*}e^{-(1+\lambda)\kappa(\ell-\frac{1}{2}K)}e^{-(1-\lambda)\kappa\cdot\frac{1}{2}K} \\ &\le \tfrac{1}{2}c_1e^{-(1+\lambda)\kappa(\ell-\frac{1}{2}K+2)}. \end{aligned}$$

By (4.49), this completes the proof. □

Lemma 4.10. *For $\epsilon > 0$ sufficiently small and $t^* \in \mathbb{N}$ sufficiently large, for $K \in \mathbb{N}$ sufficiently large (depending on t^*), the following holds for n sufficiently large. Suppose $\sup_{s\in[0,t^*],\, x\in\frac{1}{n}\mathbb{Z}} |u^n_s(x) - g(x-\nu s)| < \epsilon$, and $p^n_t(x) \geq \frac{1}{6}g(x-\nu t)$ $\forall t \in [0,t^*]$, $x \leq \nu t + D^+_n$. Suppose $q^n_0(z) = p^n_0(z)\mathbb{1}_{z\leq \ell}$ for some $\ell \in \frac{1}{n}\mathbb{Z}$ with $\ell \leq -K$. Then for $z \leq \nu t^* + D^+_n$,*

$$\frac{v^n_{t^*}(z)}{p^n_{t^*}(z)} \leq \tfrac{1}{2}c_1 e^{-\frac{1}{2}\alpha\kappa((z-\nu t^*)-\ell+1)}, \tag{4.50}$$

where $(v^n_t)_{t\geq 0}$ is defined in (3.11).

Proof. Take $c \in (0, \alpha^2/4)$. Take $t^* \in \mathbb{N}$ sufficiently large that $e^{(c-\alpha^2/4)s_0 t^*} < \frac{1}{48}c_1 e^{-\kappa}$. Suppose $\sup_{s\in[0,t^*],x\in\frac{1}{n}\mathbb{Z}} |u^n_s(x) - g(x-\nu s)| < c/4$. Take $K \in \mathbb{N}$ sufficiently large that $g(-K/2) \geq 1 - c/4$, $2s_0 t^* e^{13 s_0 t^*} e^{-\kappa K/2} < \frac{1}{48}c_1 e^{-\kappa}$ and $e^{7s_0t^*}e^{-\kappa K} < \frac{1}{24}c_1 e^{-\kappa}$. Then for $s \in [0,t^*]$ and $x \in \frac{1}{n}\mathbb{Z}$ with $x \leq -\frac{1}{2}K + \nu s$, we have

$$(1-u^n_s(x))(2u^n_s(x) - 1 + \alpha) \leq (\tfrac{1}{4}c + 1 - g(x-\nu s))(1+\alpha) \leq c.$$

Take $\ell \in \frac{1}{n}\mathbb{Z}$ with $\ell \leq -K$. By (3.12) with $a = -cs_0$, and since $(1-u)(2u-1+\alpha) - c \leq 2$ for $u \in [0,1]$, for $t \in [0,t^*]$ and $z \in \frac{1}{n}\mathbb{Z}$,

$$\begin{aligned} v^n_t(z) &\leq e^{cs_0 t}\langle q^n_0, \phi^{t,z}_0\rangle_n + s_0\int_0^t e^{cs_0(t-s)}\langle 2v^n_s(\cdot)\mathbb{1}_{\cdot\geq -\frac{1}{2}K+\nu s}, \phi^{t,z}_s\rangle_n ds \\ &\leq e^{cs_0 t}\mathbf{P}_z\left(X^n_{mt} \leq \ell\right) + 2s_0 e^{cs_0 t}\int_0^t \sup_{x\geq -\frac{1}{2}K+\nu s} v^n_s(x) ds. \end{aligned} \tag{4.51}$$

For $s \in [0,t]$ and $x \geq -\frac{1}{2}K + \nu s$, by (3.13),

$$\begin{aligned} v^n_s(x) \leq e^{(1+\alpha)s_0 s}\mathbf{P}_x\left(X^n_{ms} \leq \ell\right) &\leq e^{(1+\alpha)s_0 s}\mathbf{P}_0\left(X^n_{ms} \geq -\ell - \tfrac{1}{2}K + \nu s\right) \\ &\leq e^{(1+\alpha)s_0 s} e^{3\kappa(\ell + \frac{1}{2}K - \nu s)} e^{10 s_0 s}, \end{aligned}$$

for n sufficiently large, by Markov's inequality and Lemma 3.8, and since $\frac{1}{2}m\kappa^2 = s_0$. Hence by (4.51) and then by Lemma 3.8 and since $\frac{1}{2}m\kappa^2 = s_0$, $\kappa\nu = \alpha s_0$ and $\ell \leq -K$, for $z \leq \nu t^*$,

$$\begin{aligned} v^n_{t^*}(z) &\leq e^{cs_0t^*}e^{-\frac{1}{2}\alpha\kappa(z-\ell)}\mathbf{E}_0\left[e^{\frac{1}{2}\alpha\kappa X^n_{mt^*}}\right] + 2s_0t^*e^{13s_0t^*}e^{3\kappa(\ell+\frac{1}{2}K)} \\ &\leq e^{-\frac{1}{2}\alpha\kappa((z-\nu t^*)-\ell)}e^{(c-\frac{1}{4}\alpha^2+\mathcal{O}(n^{-1}))s_0t^*} + 2s_0t^*e^{13s_0t^*}e^{\kappa\ell}e^{-\kappa K/2} \\ &\leq \tfrac{1}{24}c_1 e^{-\frac{1}{2}\alpha\kappa((z-\nu t^*)-\ell+1)}, \end{aligned}$$

where the last line follows by our choice of t^* and K and since $z \leq \nu t^*$. Hence for $z \leq \nu t^*$, since $p^n_{t^*}(z) \geq \frac{1}{12}$, we have that (4.50) holds. For $z \in [\nu t^*, \nu t^* + D^+_n]$, by (3.13) and then by Markov's inequality and Lemma 3.8, and since $\ell \leq -K$, for n sufficiently large,

$$\begin{aligned} v^n_{t^*}(z) \leq e^{(1+\alpha)s_0t^*}\mathbf{P}_z\left(X^n_{mt^*} \leq \ell\right) \leq e^{(1+\alpha)s_0t^*}e^{-2\kappa(z-\ell)}e^{5s_0t^*} &\leq e^{7s_0t^*}e^{-\kappa K}e^{-\kappa z}e^{-\kappa(z-\ell)} \\ &\leq \tfrac{1}{24}c_1 e^{-\kappa z}e^{-\frac{1}{2}\alpha\kappa((z-\nu t^*)-\ell+1)} \end{aligned}$$

by our choice of K and since $z - \ell \geq 0$. The result follows since $p^n_{t^*}(z) \geq \frac{1}{12}e^{-\kappa(z-\nu t^*)} \geq \frac{1}{12}e^{-\kappa z}$. ∎

For $t \geq 0$ and $x_1 \in \frac{1}{n}\mathbb{Z}$, let $(v^{n,+}_{t,t+s}(x_1,\cdot))_{s\geq 0}$ denote the solution of

$$\begin{cases} \partial_s v^{n,+}_{t,t+s}(x_1,\cdot) &= \frac{1}{2}m\Delta_n v^{n,+}_{t,t+s}(x_1,\cdot) + s_0 v^{n,+}_{t,t+s}(x_1,\cdot)(1-u^n_{t,t+s})(2u^n_{t,t+s} - 1 + \alpha) \text{ for } s > 0, \\ v^{n,+}_{t,t}(x_1,x) &= p^n_t(x)\mathbb{1}_{x\geq x_1}, \end{cases}$$

where $(u^n_{t,t+s})_{s\geq 0}$ is defined in (3.2). Similarly, let $(v^{n,-}_{t,t+s}(x_1,\cdot))_{s\geq 0}$ denote the solution of

$$\begin{cases}\partial_s v^{n,-}_{t,t+s}(x_1,\cdot) &= \frac{1}{2}m\Delta_n v^{n,-}_{t,t+s}(x_1,\cdot)+s_0 v^{n,-}_{t,t+s}(x_1,\cdot)(1-u^n_{t,t+s})(2u^n_{t,t+s}-1+\alpha) \text{ for } s>0,\\ v^{n,-}_{t,t}(x_1,x) &= p^n_t(x)\mathbb{1}_{x\leq x_1}.\end{cases}$$

We now use Lemmas 4.9 and 4.10 to prove the following result.

Lemma 4.11. *For $t^*\in\mathbb{N}$ sufficiently large, and $K\in\mathbb{N}$ sufficiently large (depending on t^*), for $\ell\in\mathbb{N}$, the following holds for n sufficiently large. For $t\in[(\log N)^2-t^*,N^2-t^*]$ and $x_1,x_2\in\frac{1}{n}\mathbb{Z}$ with $x_1-x_2\leq(\log N)^{2/3}$,*

$$\mathbb{P}\left(A^{(2)}_t(x_1,x_2)^c\cap\{x_1-\mu^n_t\in[K,D^+_n],x_2-\mu^n_{t+t^*}\leq D^+_n\}\cap E'_1\right)\leq\left(\frac{n}{N}\right)^{\ell}. \tag{4.52}$$

For $t\in[(\log N)^2-t^*,N^2-t^*]$ and $x_1,x_2\in\frac{1}{n}\mathbb{Z}$ with $x_2-x_1\leq(\log N)^{2/3}$,

$$\mathbb{P}\left(A^{(3)}_t(x_1,x_2)^c\cap\{x_1-\mu^n_t\leq -K\}\cap E'_1\right)\leq\left(\frac{n}{N}\right)^{\ell}. \tag{4.53}$$

Proof. Take $t^*,K\in\mathbb{N}$ sufficiently large that Lemmas 4.9 and 4.10 hold. Recall the definition of E'_1 in (3.3). Suppose n is sufficiently large that $(\log N)^2-t^*\geq\frac{1}{2}(\log N)^2\vee\log N$, and E'_1 occurs. Take $t\in[(\log N)^2-t^*,N^2-t^*]$ and $x_1,x_2\in\frac{1}{n}\mathbb{Z}$ with $x_1-x_2\leq(\log N)^{2/3}$. Recall from (2.8) that $D^+_n=(1/2-c_0)\kappa^{-1}\log(N/n)$. Take $c_3\in(0,c_0)$ and suppose $|q^{n,+}_{t,t+t^*}(x_1,x_2)-v^{n,+}_{t,t+t^*}(x_1,x_2)|\leq\left(\frac{n}{N}\right)^{1/2-c_3}$. Then for n sufficiently large, by Lemma 4.9 and (3.3), and by the definition of the event E_1 in (2.10), if $x_1-\mu^n_t\in[K,D^+_n]$ and $x_2-\mu^n_{t+t^*}\leq D^+_n$,

$$\begin{aligned}\frac{q^{n,+}_{t,t+t^*}(x_1,x_2)}{p^n_{t+t^*}(x_2)}&\leq\tfrac{1}{2}c_1e^{-(1+\frac{1}{2}(1-\alpha))\kappa(x_1-(x_2-\nu t^*)\vee(\mu^n_t+K)+2)}+5g(D^+_n)^{-1}\left(\frac{n}{N}\right)^{1/2-c_3}\\&\leq c_1e^{-(1+\frac{1}{2}(1-\alpha))\kappa(x_1-(x_2-\nu t^*)\vee(\mu^n_t+K)+2)}\end{aligned}$$

for n sufficiently large, since $x_1-x_2\leq(\log N)^{2/3}$ and $g(D^+_n)^{-1}\leq 2\left(\frac{N}{n}\right)^{1/2-c_0}$ with $c_0>c_3$. By Proposition 3.2, the first statement (4.52) follows.

Now take $t\in[(\log N)^2-t^*,N^2-t^*]$ and $x_1,x_2\in\frac{1}{n}\mathbb{Z}$ with $x_2-x_1\leq(\log N)^{2/3}$. Suppose E'_1 occurs and suppose $|q^{n,-}_{t,t+t^*}(x_1,x_2)-v^{n,-}_{t,t+t^*}(x_1,x_2)|\leq\left(\frac{n}{N}\right)^{1/4}$. If $x_1-\mu^n_t\leq -K$, then for n sufficiently large, $x_2-\mu^n_{t+t^*}\leq(\log N)^{2/3}$ and so $p^n_{t+t^*}(x_2)^{-1}\leq 10e^{\kappa(\log N)^{2/3}}$. Hence by Lemma 4.10,

$$\begin{aligned}\frac{q^{n,-}_{t,t+t^*}(x_1,x_2)}{p^n_{t+t^*}(x_2)}&\leq\tfrac{1}{2}c_1e^{-\frac{1}{2}\alpha\kappa((x_2-\nu t^*)-x_1+1)}+10e^{\kappa(\log N)^{2/3}}\left(\frac{n}{N}\right)^{1/4}\\&\leq c_1e^{-\frac{1}{2}\alpha\kappa((x_2-\nu t^*)-x_1+1)}\end{aligned}$$

for n sufficiently large. By Proposition 3.2, the second statement (4.53) follows, which completes the proof. □

We now show that $A^{(4)}_t(x)$ and $A^{(5)}_t(x)$ occur with high probability for suitable x and t.

Lemma 4.12. *For $\ell\in\mathbb{N}$, the following holds for n sufficiently large. For $x\in\frac{1}{n}\mathbb{Z}$ and $t\geq 0$,*

$$\mathbb{P}\left(A^{(5)}_t(x)^c\right)\leq\left(\frac{n}{N}\right)^{\ell}. \tag{4.54}$$

If there exists $a_2>3$ such that $N\geq n^{a_2}$ for n sufficiently large, then for $t\in[(\log N)^2-\epsilon_n,N^2-\epsilon_n]$ and $x\in\frac{1}{n}\mathbb{Z}$,

$$\mathbb{P}\left(A^{(4)}_t(x)^c\cap\{x-\mu^n_t\leq D^+_n\}\cap E'_1\right)\leq\left(\frac{n}{N}\right)^{\ell}. \tag{4.55}$$

Proof. For $t \geq 0$ and $x_1, x_2 \in \frac{1}{n}\mathbb{Z}$, by Corollary 3.13 with $a = -(1+\alpha)s_0$,

$$\mathbb{E}\left[q^n_{t,t+\epsilon_n}(x_1,x_2)\right] \leq e^{(1+\alpha)s_0\epsilon_n}\mathbf{P}_{x_2}\left(X^n_{m\epsilon_n} = x_1\right) \leq e^{(1+\alpha)s_0\epsilon_n}e^{-(\log N)^{3/2}|x_1-x_2|}e^{m(\log N)^3\epsilon_n}$$

for n sufficiently large, by Markov's inequality and Lemma 3.8. Recall from (2.4) that $\epsilon_n \leq (\log N)^{-2}$. Therefore, for n sufficiently large, for $x \in \frac{1}{n}\mathbb{Z}$, by a union bound and then by Markov's inequality,

$$\begin{aligned}\mathbb{P}\left(A^{(5)}_t(x)^c\right) &\leq \sum_{x' \in \frac{1}{n}\mathbb{Z}, |x-x'|\geq 1} \mathbb{P}\left(q^n_{t,t+\epsilon_n}(x',x) \geq N^{-1}\right) \\ &\leq Ne^{(1+\alpha)s_0\epsilon_n}N^m \sum_{x' \in \frac{1}{n}\mathbb{Z}, |x-x'|\geq 1} e^{-(\log N)^{3/2}|x-x'|},\end{aligned}$$

which completes the proof of (4.54).

From now on, assume there exists $a_2 > 3$ such that $N \geq n^{a_2}$ for n sufficiently large. Suppose n is sufficiently large that $(\log N)^2 - \epsilon_n \geq \frac{1}{2}(\log N)^2 \vee \log N$, and take $t \in [(\log N)^2 - \epsilon_n, N^2 - \epsilon_n]$ and $x_1, x_2 \in \frac{1}{n}\mathbb{Z}$ with $|x_1 - x_2| \leq 1$. Recall the definition of $(v^n_{t,t+s}(x_1,\cdot))_{s\geq 0}$ in (4.45). By (3.13), and then by Lemma 3.14, there exists a constant $K_7 < \infty$ such that for n sufficiently large,

$$v^n_{t,t+\epsilon_n}(x_1,x_2) \leq e^{(1+\alpha)s_0\epsilon_n}p^n_t(x_1)\mathbf{P}_{x_2}\left(X^n_{m\epsilon_n} = x_1\right) \leq K_7 n^{-1}\epsilon_n^{-1/2}p^n_t(x_1).$$

Suppose E'_1 occurs and $x_1 \leq \mu^n_t + D^+_n$. Then for n sufficiently large, by the definition of the event E_1 in (2.10) and since $|x_1 - x_2| \leq 1$, there exists a constant $K_8 < \infty$ such that $\frac{p^n_t(x_1)}{p^n_{t+\epsilon_n}(x_2)} \leq K_8$, and so

$$\frac{v^n_{t,t+\epsilon_n}(x_1,x_2)}{p^n_{t+\epsilon_n}(x_2)} \leq K_7K_8n^{-1}\epsilon_n^{-1/2}. \tag{4.56}$$

Recall from (2.8) that $D^+_n = (1/2 - c_0)\kappa^{-1}\log(N/n)$. Take $c' \in (0, c_0/2)$ and suppose

$$|q^n_{t,t+\epsilon_n}(x_1,x_2) - v^n_{t,t+\epsilon_n}(x_1,x_2)| \leq \left(\frac{n}{N}\right)^{1/2-c'} p^n_t(x_1)^{1/2}n^{-1/2}.$$

By (4.56) and then since $x_2 \leq \mu^n_t + D^+_n + 1$ and by the definition of K_8,

$$\begin{aligned}\frac{q^n_{t,t+\epsilon_n}(x_1,x_2)}{p^n_{t+\epsilon_n}(x_2)} &\leq K_7K_8n^{-1}\epsilon_n^{-1/2} + p^n_{t+\epsilon_n}(x_2)^{-1/2}\left(\frac{n}{N}\right)^{1/2-c'}\left(\frac{p^n_t(x_1)}{p^n_{t+\epsilon_n}(x_2)}\right)^{1/2} n^{-1/2} \\ &\leq K_7K_8n^{-1}\epsilon_n^{-1/2} + 10^{1/2}e^{\frac{1}{2}\kappa(D^+_n+2)}\left(\frac{n}{N}\right)^{1/2-c'}K_8^{1/2}n^{-1/2} \\ &\leq (K_7K_8+1)n^{-1}\epsilon_n^{-1/2} \end{aligned} \tag{4.57}$$

for n sufficiently large, since $N \geq n^3$ and so $e^{\frac{1}{2}\kappa D^+_n}\left(\frac{n}{N}\right)^{1/2-c'} = \left(\frac{n}{N}\right)^{1/4+c_0/2-c'} \leq n^{-1/2}$. For $c \in (0, \frac{1}{2}(a_2-2)^{-1}(a_2-3))$, we have $3/2 - 2c < a_2(1/2 - c)$ and so since $N \geq n^{a_2}$ we have $p^n_t(x_1) \geq \frac{1}{10}e^{-\kappa D^+_n} \geq \frac{1}{10}\left(\frac{n}{N}\right)^{1/2} \geq \left(\frac{n^2}{N}\right)^{1-c}$ for n sufficiently large. Hence by Lemma 4.7, for n sufficiently large,

$$\begin{aligned}\mathbb{P}\Big(\{|q^n_{t,t+\epsilon_n}(x_1,x_2) - v^n_{t,t+\epsilon_n}(x_1,x_2)| \geq \left(\frac{n}{N}\right)^{1/2-c'} p^n_t(x_1)^{1/2}n^{-1/2}\} \\ \cap \{x_1 \leq \mu^n_t + D^+_n\} \cap E'_1\Big) \leq \left(\frac{n}{N}\right)^{\ell+1},\end{aligned}$$

and by (4.57), it follows that for n sufficiently large,

$$\mathbb{P}\left(\{q^n_{t,t+\epsilon_n}(x_1,x_2) > n^{-1}\epsilon_n^{-1}p^n_{t+\epsilon_n}(x_2)\} \cap \{x_1 - \mu^n_t \leq D^+_n\} \cap E'_1\right) \leq \left(\frac{n}{N}\right)^{\ell+1}.$$

By the same argument as for the proof of (4.54), the second statement (4.55) now follows. □

Finally we show that $A_t^{(6)}(x)$ occurs with high probability; the proof is similar to the first half of the proof of Lemma 4.12.

Lemma 4.13. *For $\ell \in \mathbb{N}$ and $t^* \in \mathbb{N}$, the following holds for n sufficiently large. For $t \geq 0$ and $x \in \frac{1}{n}\mathbb{Z}$,*

$$\mathbb{P}\left(A_t^{(6)}(x)^c\right) \leq \left(\frac{n}{N}\right)^\ell.$$

Proof. By Corollary 3.13 with $a = -(1+\alpha)s_0$, for $k \in [t^*\delta_n^{-1}]$ and $x' \in \frac{1}{n}\mathbb{Z}$,

$$\begin{aligned}\mathbb{E}\left[q^n_{t,t+k\delta_n}(x',x)\right] &\leq e^{(1+\alpha)s_0 t^*}\mathbf{P}_x\left(X^n_{mk\delta_n} = x'\right)\\ &\leq e^{(1+\alpha)s_0 t^*}e^{-(\log N)^{1/2}|x-x'|}\mathbf{E}_0\left[e^{X^n_{mk\delta_n}(\log N)^{1/2}}\right]\\ &\leq e^{(1+\alpha)s_0 t^*}e^{-(\log N)^{1/2}|x-x'|}e^{mt^*\log N}\end{aligned}$$

for n sufficiently large, where the second inequality follows by Markov's inequality, and the third by Lemma 3.8. Therefore, by a union bound and Markov's inequality,

$$\begin{aligned}&\mathbb{P}\left(\exists x' \in \tfrac{1}{n}\mathbb{Z}, k \in [t^*\delta_n^{-1}] : |x-x'| \geq (\log N)^{2/3}, q^n_{t,t+k\delta_n}(x',x) \geq N^{-1}\right)\\ &\leq t^*\delta_n^{-1}\cdot Ne^{(1+\alpha)s_0t^*}N^{mt^*}\sum_{x'\in\frac{1}{n}\mathbb{Z},|x-x'|\geq(\log N)^{2/3}} e^{-(\log N)^{1/2}|x-x'|}\\ &\leq \left(\frac{n}{N}\right)^\ell\end{aligned}$$

for n sufficiently large. □

We can now end this section by proving Proposition 4.1.

Proof of Proposition 4.1. Note that if $x_1 - x_2 > (\log N)^{2/3}$ and $A_t^{(6)}(x_2)$ occurs, then $A_t^{(2)}(x_1,x_2)$ occurs. Similarly, if $x_2 - x_1 > (\log N)^{2/3}$ and $A_t^{(6)}(x_2)$ occurs, then $A_t^{(3)}(x_1,x_2)$ occurs. The result now follows directly from Proposition 4.8 and Lemmas 4.11, 4.12 and 4.13. □

5 Event E_3 occurs with high probability

In this section, we will prove the following result.

Proposition 5.1. *For $K \in \mathbb{N}$ sufficiently large, for $c_2 > 0$, if $N \geq n^3$ for n sufficiently large, then for n sufficiently large, if $p_0^n(x) = 0 \;\forall x \geq N$,*

$$\mathbb{P}\left((E_3)^c \cap E_1\right) \leq \left(\frac{n}{N}\right)^2.$$

By the definition of the events E_1 and E_3 in (2.10) and (2.12), Proposition 5.1 follows directly from the following result.

Lemma 5.2. *For $\ell \in \mathbb{N}$, for $K \in \mathbb{N}$ sufficiently large, for $c_2 > 0$, if $N \geq n^3$ for n sufficiently large then the following holds for n sufficiently large. If $p_0^n(y) = 0 \;\forall y \geq N$ then for $t \in [(\log N)^2 - \delta_n, N^2]$, $x \in \frac{1}{n}\mathbb{Z}$ with $x \geq -N^5$ and $j \in \{1,2,3,4\}$,*

$$\mathbb{P}\left(B_t^{(j)}(x)^c \cap E_1 \cap \{x \leq \mu_t^n + D_n^+ + 1\}\right) \leq \left(\frac{n}{N}\right)^\ell. \tag{5.1}$$

Proof. We begin by proving (5.1) with $j=1$. For $x \in \frac{1}{n}\mathbb{Z}$, $i \in [N]$ and $0 \leq t_1 < t_2$, let $\mathcal{A}^{x,i}[t_1,t_2)$ denote the total number of points in the time interval $[t_1,t_2)$ in the Poisson processes $(\mathcal{P}^{x,i,i'})_{i'\in[N]\setminus\{i\}}$, $(\mathcal{S}^{x,i,i'})_{i'\in[N]\setminus\{i\}}$, $(\mathcal{Q}^{x,i,i',i''})_{i',i''\in[N]\setminus\{i\},i'\neq i''}$ and

$(\mathcal{R}^{x,i,y,i'})_{i' \in [N], y \in \{x \pm n^{-1}\}}$. (These points correspond to the times at which the individual (x,i) may be replaced by offspring of another individual.) For $t \geq 0$ and $x \in \frac{1}{n}\mathbb{Z}$, let

$$\mathcal{C}_t^{n,1}(x) = \{(i,j) : i \neq j \in [N], \mathcal{P}^{x,i,j}[t, t+\delta_n) = 1 = \mathcal{A}^{x,i}[t, t+\delta_n),\ \mathcal{A}^{x,j}[t, t+\delta_n) = 0, \\ \xi_t^n(x,j) = 1\}.$$

Recall the definition of $\mathcal{C}_t^n(x,x)$ in (2.5). If $(i,j) \in \mathcal{C}_t^{n,1}(x)$, then

$$(\zeta_{\delta_n}^{n,t+\delta_n}(x,i), \theta_{\delta_n}^{n,t+\delta_n}(x,i)) = (x,j) = (\zeta_{\delta_n}^{n,t+\delta_n}(x,j), \theta_{\delta_n}^{n,t+\delta_n}(x,j)),$$

and so $(i,j), (j,i) \in \mathcal{C}_t^n(x,x)$. Note that if $(i,j) \in \mathcal{C}_t^{n,1}(x)$ then $(j,i) \notin \mathcal{C}_t^{n,1}(x)$; therefore

$$|\mathcal{C}_t^n(x,x)| \geq 2|\mathcal{C}_t^{n,1}(x)|. \tag{5.2}$$

For $t \geq 0$, $x \in \frac{1}{n}\mathbb{Z}$ and $i \in [N]$, let

$$\mathcal{D}_t^n(x,i) = \{(y,j) \in \tfrac{1}{n}\mathbb{Z} \times [N] : (\zeta_s^{n,t+s}(y,j), \theta_s^{n,t+s}(y,j)) = (x,i) \text{ for some } s \in [0, \delta_n]\}, \tag{5.3}$$

the set of labels of individuals whose time-t ancestor at some time in $[t, t+\delta_n]$ is (x,i). Define

$$\mathcal{M}_t^n = \max_{x \in \frac{1}{n}\mathbb{Z} \cap [-2N^5, N^5],\, i \in [N]} |\mathcal{D}_t^n(x,i)|. \tag{5.4}$$

For $t \geq 0$ and $x \in \frac{1}{n}\mathbb{Z}$, let

$$\mathcal{C}_t^{n,2}(x) \\ = \Big\{(i,j) : i \neq j \in [N], \Big(\mathcal{P}^{x,i,j} + \mathcal{S}^{x,i,j} + \sum_{k \in [N] \setminus \{i,j\}} \mathcal{Q}^{x,i,j,k}\Big)[t, t+\delta_n) > 0,\ \xi_t^n(x,j) = 1\Big\}. \tag{5.5}$$

Suppose $(i,j) \in \mathcal{C}_t^n(x,x)$, and $(i,j), (j,i) \notin \mathcal{C}_t^{n,2}(x)$. Then there exist $s \in [0, \delta_n]$, $(y,k) \notin \{(x,i), (x,j)\}$ and $i' \in \{i,j\}$ such that $(\zeta_s^{n,t+\delta_n}(x,i'), \theta_s^{n,t+\delta_n}(x,i')) = (y,k)$. Then letting $(x_0, i_0) = (\zeta_{\delta_n}^{n,t+\delta_n}(x,i), \theta_{\delta_n}^{n,t+\delta_n}(x,i))$, we have (x,i), (x,j), $(y,k) \in \mathcal{D}_t^n(x_0, i_0)$. Since $\zeta^{n,t+\delta_n}(x,i)$ only jumps in increments of $\pm n^{-1}$, and $(\zeta_s^{n,t+\delta_n}(x,i), \theta_s^{n,t+\delta_n}(x,i)) \in \mathcal{D}_t^n(x_0, i_0)$ $\forall s \in [0, \delta_n]$, we have $|x - x_0| < |\mathcal{D}_t^n(x_0, i_0)| n^{-1}$. Hence if $x_0 \in [-2N^5, N^5]$ then $|x - x_0| < \mathcal{M}_t^n n^{-1}$. Therefore, by the definition of $q^{n,-}$ in (2.3), if $q_{t,t+\delta_n}^{n,-}(-2N^5, x) = 0$ and $p_t^n(y) = 0$ $\forall y \geq N^5$, then

$$|\mathcal{C}_t^n(x,x)| \leq 2|\mathcal{C}_t^{n,2}(x)| + 2\binom{\mathcal{M}_t^n}{2} |\{(x_0, i_0) \in \tfrac{1}{n}\mathbb{Z} \times [N] : |x - x_0| < \mathcal{M}_t^n n^{-1}, |\mathcal{D}_t^n(x_0, i_0)| \geq 3\}|. \tag{5.6}$$

We now use the inequalities (5.2) and (5.6) to give lower and upper bounds on $|\mathcal{C}_t^n(x,x)|$.

We begin with a lower bound. For $x \in \frac{1}{n}\mathbb{Z}$, $i \in [N]$ and $0 \leq t_1 < t_2$, let $\mathcal{A}^{1,x,i}[t_1, t_2)$ denote the total number of points in the time interval $[t_1, t_2)$ in the Poisson processes $(\mathcal{P}^{x,i,j})_{j \in [N] \setminus \{i\}, \xi_{t_1}^n(x,j) = 1}$. Let $\mathcal{A}^{2,x,i}[t_1, t_2)$ denote the total number of points in the time interval $[t_1, t_2)$ in the Poisson processes $(\mathcal{P}^{x,i,j})_{j \in [N] \setminus \{i\}, \mathcal{A}^{x,j}[t_1, t_2) > 0}$. Now fix $t \geq 0$ and $x \in \frac{1}{n}\mathbb{Z}$ and let

$$\begin{aligned} A^{(1)} &= |\{i \in [N] : \xi_t^n(x,i) = 1, \mathcal{A}^{x,i}[t, t+\delta_n) = 1 = \mathcal{A}^{1,x,i}[t, t+\delta_n)\}|, \\ A^{(2)} &= |\{i \in [N] : \xi_t^n(x,i) = 0, \mathcal{A}^{x,i}[t, t+\delta_n) = 1 = \mathcal{A}^{1,x,i}[t, t+\delta_n)\}|, \\ \text{and} \qquad B &= |\{i \in [N] : \mathcal{A}^{x,i}[t, t+\delta_n) = 1 = \mathcal{A}^{2,x,i}[t, t+\delta_n)\}|. \end{aligned}$$

Then by (5.2) and the definition of $\mathcal{C}_t^{n,1}(x)$,

$$|\mathcal{C}_t^n(x,x)| \geq 2|\mathcal{C}_t^{n,1}(x)| \geq 2(A^{(1)} + A^{(2)} - B). \tag{5.7}$$

Let $(X^n_j)_{j=1}^\infty$ be i.i.d., let $(Y^n_j)_{j=1}^\infty$ be i.i.d., and let $(Z^n_j)_{j=1}^\infty$ be i.i.d., with

$$\begin{aligned} X^n_1 &\sim \text{Poisson}\,(r_n\delta_n(1-(\alpha+1)s_n)) \\ Y^n_1 &\sim \text{Poisson}\,(r_n\delta_n(\alpha s_n + N^{-1}s_n(N-2))) \\ \text{and} \quad Z^n_1 &\sim \text{Poisson}\,(mr_n\delta_n). \end{aligned}$$

Recall from (1.11) that $r_n = \frac{1}{2}n^2N^{-1}$ and $s_n = 2s_0n^{-2}$. Then conditional on $p^n_t(x)$, $A^{(1)} \sim \text{Bin}(Np^n_t(x), p_1)$ and $A^{(2)} \sim \text{Bin}(N(1-p^n_t(x)), p_2)$ are independent, with

$$\begin{aligned} p_1 &= \mathbb{P}\left(\sum_{j=1}^{Np^n_t(x)-1} X^n_j = 1, \sum_{j=Np^n_t(x)}^{N-1} X^n_j + \sum_{j=1}^{N-1} Y^n_j + \sum_{j=1}^{2N} Z^n_j = 0\right) \\ &= \mathbb{1}_{p^n_t(x)>0}\Big(\tfrac{1}{2}n^2\delta_n(p^n_t(x)-N^{-1})(1+\mathcal{O}(n^{-2})) + \mathcal{O}((n^2\delta_n(p^n_t(x)-N^{-1}))^2)\Big) \\ &\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad \cdot \big(1-\mathcal{O}(n^2\delta_n)\big) \\ &= \mathbb{1}_{p^n_t(x)>0}\tfrac{1}{2}n^2\delta_n(p^n_t(x)-N^{-1})(1+\mathcal{O}(n^{-2}+n^2\delta_n)) \end{aligned}$$

and

$$\begin{aligned} p_2 &= \mathbb{P}\left(\sum_{j=1}^{Np^n_t(x)} X^n_j = 1, \sum_{j=Np^n_t(x)+1}^{N-1} X^n_j + \sum_{j=1}^{N-1} Y^n_j + \sum_{j=1}^{2N} Z^n_j = 0\right) \\ &= \tfrac{1}{2}n^2\delta_n p^n_t(x)(1+\mathcal{O}(n^{-2}+n^2\delta_n)). \end{aligned}$$

Hence

$$\mathbb{E}\left[A^{(1)}+A^{(2)}\middle| p^n_t(x)\right] = \tfrac{1}{2}Nn^2\delta_n p^n_t(x)(1+\mathcal{O}(n^{-2}+n^2\delta_n+N^{-1}p^n_t(x)^{-1})).$$

Recall from (2.4) that $\delta_n = \lfloor N^{1/2}n^2\rfloor^{-1}$. Suppose n is sufficiently large that $(\log N)^2 - \delta_n \geq \frac{1}{2}(\log N)^2$. Then on the event E_1, for $t \in [(\log N)^2 - \delta_n, N^2]$ and $x \leq \mu^n_t + D^+_n + 1$, by (2.10) and (2.8) we have $N^{-1}p^n_t(x)^{-1} \leq 10N^{-1}e^{\kappa(D^+_n+1)} \leq 10e^\kappa N^{-1/2}n^{-1/2}$ and

$$Nn^2\delta_n p^n_t(x) \geq \tfrac{1}{5}N^{1/2}g(x-\mu^n_t) \geq \tfrac{1}{10}N^{1/2}e^{-\kappa(D^+_n+1)} \geq 2n^{1/2} \tag{5.8}$$

for n sufficiently large. Hence for n sufficiently large, for $t \in [(\log N)^2 - \delta_n, N^2]$ and $x \in \frac{1}{n}\mathbb{Z}$, by conditioning on $p^n_t(x)$ and then applying Theorem 2.3(c) in [25],

$$\begin{aligned} \mathbb{P}\left(\left\{A^{(1)}+A^{(2)} \leq \tfrac{1}{2}Nn^2\delta_n p^n_t(x)(1-n^{-1/5})\right\} \cap \{x \leq \mu^n_t + D^+_n + 1\} \cap E_1\right) &\leq e^{-\frac{1}{3}n^{-2/5}n^{1/2}} \\ &= e^{-\frac{1}{3}n^{1/10}}. \end{aligned} \tag{5.9}$$

For an upper bound on B, first let

$$A' = |\{i \in [N] : \mathcal{A}^{x,i}[t, t+\delta_n) > 0\}|.$$

Then $A' \sim \text{Bin}(N, p)$ where

$$p = \mathbb{P}\left(\sum_{j=1}^{N-1}(X^n_j + Y^n_j) + \sum_{j=1}^{2N} Z^n_j > 0\right) = \tfrac{1}{2}n^2\delta_n(1+2m)(1+\mathcal{O}(n^2\delta_n + n^{-2})).$$

Conditional on A', we have $B \leq \text{Bin}(A', \frac{A'-1}{(1+2m)N-1})$. By Theorem 2.3(b) in [25], for n sufficiently large,

$$\mathbb{P}\left(A' \geq Nn^2\delta_n(1+2m)\right) \leq e^{-\frac{1}{8}Nn^2\delta_n(1+2m)}. \tag{5.10}$$

Moreover, since $\delta_n = \lfloor N^{1/2} n^2 \rfloor^{-1}$, letting $B' \sim \text{Bin}(\lfloor 2N^{1/2}(1+2m)\rfloor, 2N^{-1/2})$, for n sufficiently large,

$$\begin{aligned}\mathbb{P}\left(B \geq n^{1/4}, A' \leq N n^2 \delta_n (1+2m)\right) &\leq \mathbb{P}\left(B' \geq n^{1/4}\right)\\ &\leq e^{-n^{1/4}}(1+(e-1)2N^{-1/2})^{\lfloor 2N^{1/2}(1+2m)\rfloor}\\ &\leq e^{-\frac{1}{2}n^{1/4}}, \end{aligned} \tag{5.11}$$

where the second inequality follows by Markov's inequality. Therefore, by (5.7), (5.8), (5.9), (5.10) and (5.11), for n sufficiently large, for $t \in [(\log N)^2 - \delta_n, N^2]$ and $x \in \frac{1}{n}\mathbb{Z}$,

$$\begin{aligned}&\mathbb{P}\left(\left\{|\mathcal{C}_t^n(x,x)| \leq N n^2 \delta_n p_t^n(x)(1-2n^{-1/5})\right\} \cap \{x \leq \mu_t^n + D_n^+ + 1\} \cap E_1\right)\\ &\quad \leq e^{-\frac{1}{3}n^{1/10}} + e^{-\frac{1}{8}N^{1/2}} + e^{-\frac{1}{2}n^{1/4}}. \end{aligned} \tag{5.12}$$

For an upper bound on $|\mathcal{C}_t^n(x,x)|$, note that by the definition of $\mathcal{C}_t^{n,2}(x)$ in (5.5), conditional on $p_t^n(x)$,

$$|\mathcal{C}_t^{n,2}(x)| \sim \text{Bin}(N p_t^n(x)(N-1), p'),$$

where

$$\begin{aligned} p' &= \mathbb{P}\left(\left(\mathcal{P}^{x,1,2} + \mathcal{S}^{x,1,2} + \sum_{k\in[N]\setminus\{1,2\}} \mathcal{Q}^{x,1,2,k}\right)[0,\delta_n) > 0\right)\\ &= r_n \delta_n (1 + \mathcal{O}(r_n\delta_n + n^{-2}N^{-1})). \end{aligned}$$

Then $N p_t^n(x)(N-1)p' = \frac{1}{2} N n^2 \delta_n p_t^n(x)(1+\mathcal{O}(n^2 N^{-1}\delta_n + N^{-1}))$. Hence for n sufficiently large, for $t \in [(\log N)^2 - \delta_n, N^2]$ and $x \in \frac{1}{n}\mathbb{Z}$, by Theorem 2.3(b) in [25] and (5.8),

$$\begin{aligned}&\mathbb{P}\left(\left\{|\mathcal{C}_t^{n,2}(x)| \geq \tfrac{1}{2} N n^2 \delta_n p_t^n(x)(1+n^{-1/5})\right\} \cap \{x \leq \mu_t^n + D_n^+ + 1\} \cap E_1\right)\\ &\quad \leq e^{-\frac{1}{3}n^{-2/5}\cdot n^{1/2}} = e^{-\frac{1}{3}n^{1/10}}. \end{aligned} \tag{5.13}$$

We now bound the second term on the right hand side of (5.6). For $x \in \frac{1}{n}\mathbb{Z}$, $i \in [N]$ and $0 \leq t_1 < t_2$, let $\mathcal{B}^{x,i}[t_1,t_2)$ denote the total number of points in the time interval $[t_1,t_2)$ in the Poisson processes $(\mathcal{P}^{x,i',i})_{i'\in[N]\setminus\{i\}}$, $(\mathcal{S}^{x,i',i})_{i'\in[N]\setminus\{i\}}$, $(\mathcal{Q}^{x,i',i,i''})_{i',i''\in[N]\setminus\{i\}, i'\neq i''}$ and $(\mathcal{R}^{y,i',x,i})_{i'\in[N], y\in\{x\pm n^{-1}\}}$. (These points correspond to the times at which offspring of the individual (x,i) may replace another individual.) Let $\mathcal{B}^{1,x,i}[t_1,t_2)$ denote the total number of points in the interval $[t_1,t_2)$ in $(\mathcal{P}^{x,i',i})_{i'\in[N]\setminus\{i\}, \mathcal{B}^{x,i'}[t_1,t_2)>0}$, $(\mathcal{S}^{x,i',i})_{i'\in[N]\setminus\{i\}, \mathcal{B}^{x,i'}[t_1,t_2)>0}$, $(\mathcal{Q}^{x,i',i,i''})_{i',i''\in[N]\setminus\{i\}, i''\neq i', \mathcal{B}^{x,i'}[t_1,t_2)>0}$ and $(\mathcal{R}^{y,i',x,i})_{i'\in[N], y\in\{x\pm n^{-1}\}, \mathcal{B}^{y,i'}[t_1,t_2)>0}$. Then fix $x \in \frac{1}{n}\mathbb{Z}$ and $t \geq 0$, and let

$$\begin{aligned} &C^{(1)} = |\{i \in [N] : \mathcal{B}^{x,i}[t,t+\delta_n) \geq 2\}|\\ \text{and}\quad &C^{(2)} = |\{i \in [N] : \mathcal{B}^{x,i}[t,t+\delta_n) = 1 = \mathcal{B}^{1,x,i}[t,t+\delta_n)\}|. \end{aligned}$$

By the definition of $\mathcal{D}_t^n(x,i)$ in (5.3), we have that

$$|\{i \in [N] : |\mathcal{D}_t^n(x,i)| \geq 3\}| \leq C^{(1)} + C^{(2)}. \tag{5.14}$$

Then $C^{(1)} \sim \text{Bin}(N, p'')$, where

$$p'' = \mathbb{P}\left(\mathcal{B}^{x,1}[t,t+\delta_n) \geq 2\right) \leq (r_n \delta_n N(1+2m))^2 = \tfrac{1}{4} n^4 \delta_n^2 (1+2m)^2.$$

Therefore, by Markov's inequality and since $n^4\delta_n^2 \leq 2N^{-1}$ for n sufficiently large,

$$\mathbb{P}\left(C^{(1)} \geq n^{1/4}\right) \leq e^{-n^{1/4}}(1+(e-1)\tfrac{1}{4}n^4\delta_n^2(1+2m)^2)^N \leq e^{-\frac{1}{2}n^{1/4}}$$

for n sufficiently large. For $y \in \frac{1}{n}\mathbb{Z}$, let $D_y = |\{i \in [N] : \mathcal{B}^{y,i}[t, t+\delta_n) > 0\}|$. Then conditional on D_x, $D_{x-n^{-1}}$ and $D_{x+n^{-1}}$ we have

$$C^{(2)} \le \text{Bin}(D_x, \tfrac{(D_x-1)(1-2N^{-1}s_n)+m(D_{x-n^{-1}}+D_{x+n^{-1}})}{(1-2N^{-1}s_n)(N-1)+2mN}).$$

By the same argument as in (5.10) and (5.11), it follows that for n sufficiently large,

$$\mathbb{P}\left(C^{(2)} \ge n^{1/4}\right) \le 3e^{-\frac{1}{8}Nn^2\delta_n(1+2m)} + e^{-\frac{1}{2}n^{1/4}}.$$

Therefore, by (5.14), for n sufficiently large, for $x \in \frac{1}{n}\mathbb{Z}$ and $t \ge 0$,

$$\mathbb{P}\left(|\{i \in [N] : |\mathcal{D}_t^n(x,i)| \ge 3\}| \ge 2n^{1/4}\right) \le 3e^{-\frac{1}{8}Nn^2\delta_n(1+2m)} + 2e^{-\frac{1}{2}n^{1/4}}. \tag{5.15}$$

For $K \in \mathbb{N}$, let $S_n^K \sim \text{Poisson}((2m+1)Nr_n(K-1)\delta_n)$. Then since a set of k individuals produces offspring individuals at total rate at most $(2m+1)Nr_nk$, for $i \in [N]$,

$$\begin{aligned}\mathbb{P}\left(|\mathcal{D}_t^n(x,i)| \ge K\right) \le \mathbb{P}\left(S_n^K \ge K-1\right) &\le ((2m+1)Nr_n(K-1)\delta_n)^{K-1} \\ &\le ((2m+1)(K-1))^{K-1}N^{-(K-1)/2}\end{aligned}$$

for n sufficiently large. Therefore, by the definition of $\mathcal{M}_t^n$ in (5.4), for $\ell \in \mathbb{N}$, for $K \in \mathbb{N}$ sufficiently large that $7 - \frac{1}{2}(K-1) < -\ell$, for $t \ge 0$,

$$\mathbb{P}\left(\mathcal{M}_t^n \ge K\right) \le \sum_{x \in \frac{1}{n}\mathbb{Z}\cap[-2N^5, N^5], i \in [N]} \mathbb{P}\left(|\mathcal{D}_t^n(x,i)| \ge K\right) \le \tfrac{1}{3}\left(\frac{n}{N}\right)^{\ell} \tag{5.16}$$

for n sufficiently large. For $x \ge -N^5$ and $t \ge 0$, by Corollary 3.13 with $a = -(1+\alpha)s_0$, and then by Markov's inequality,

$$\begin{aligned}\mathbb{E}\left[q_{t,t+\delta_n}^{n,-}(-2N^5, x)\right] \le e^{(1+\alpha)s_0\delta_n}\langle \mathbb{1}_{\cdot \le -2N^5}, \phi_0^{\delta_n, x}\rangle_n &\le e^{(1+\alpha)s_0\delta_n}\mathbf{E}_0\left[e^{X^n_{m\delta_n}}\right]e^{-N^5} \\ &\le e^{1-N^5}\end{aligned} \tag{5.17}$$

for n sufficiently large, by Lemma 3.8. By Lemma 3.3, for $t \le N^2$, $\mathbb{P}\left(p_t^n(y) = 0\, \forall y \ge N^5\right) \ge 1 - e^{-N^5}$. By (5.6), (5.8), (5.13), (5.15) and (5.16), it now follows that for $\ell \in \mathbb{N}$, for n sufficiently large, for $x \in \frac{1}{n}\mathbb{Z}$ with $x \ge -N^5$ and $t \in [(\log N)^2 - \delta_n, N^2]$,

$$\mathbb{P}\left(\left\{|\mathcal{C}_t^n(x,x)| \ge Nn^2\delta_n p_t^n(x)(1+2n^{-1/5})\right\} \cap \{x \le \mu_t^n + D_n^+ + 1\} \cap E_1\right) \le \tfrac{1}{2}\left(\frac{n}{N}\right)^{\ell}. \tag{5.18}$$

By (5.12), we now have that (5.1) holds with $j = 1$.

For $t \ge 0$ and $x, y \in \frac{1}{n}\mathbb{Z}$ with $|x-y| = n^{-1}$, let

$$\begin{aligned}\mathcal{C}_t^{n,1}(x,y) = \{(i,j) \in [N]^2 : \mathcal{R}^{x,i,y,j}[t, t+\delta_n) = 1 = \mathcal{A}^{x,i}[t, t+\delta_n), \\ \mathcal{A}^{y,j}[t,t+\delta_n) = 0, \xi_t^n(y,j) = 1\}, \\ \mathcal{C}_t^{n,2}(x,y) = \{(i,j) \in [N]^2 : \mathcal{R}^{x,i,y,j}[t,t+\delta_n) > 0, \xi_t^n(y,j) = 1\}.\end{aligned}$$

Then $|\mathcal{C}_t^n(x, x+n^{-1})| \ge |\mathcal{C}_t^{n,1}(x, x+n^{-1})| + |\mathcal{C}_t^{n,1}(x+n^{-1}, x)|$. If $q_{t,t+\delta_n}^{n,-}(-2N^5, x) = 0$ and $p_t^n(y) = 0\ \forall y \ge N^5$, then by the same argument as for (5.6),

$$\begin{aligned}|\mathcal{C}_t^n(x, x+n^{-1})| \le |\mathcal{C}_t^{n,2}(x, x+n^{-1})| + |\mathcal{C}_t^{n,2}(x+n^{-1}, x)| \\ + \binom{\mathcal{M}_t^n}{2}|\{(x_0, i_0) \in \tfrac{1}{n}\mathbb{Z} \times [N] : |x - x_0| < \mathcal{M}_t^n n^{-1}, |\mathcal{D}_t^n(x_0, i_0)| \ge 3\}|.\end{aligned}$$

By the same argument as for (5.12) and (5.18), it follows that for n sufficiently large, for $x \in \frac{1}{n}\mathbb{Z}$ with $x \geq -N^5$ and $t \in [(\log N)^2 - \delta_n, N^2]$, (5.1) holds with $j = 2$.

Suppose for some $k > 1$ that $x, y \in \frac{1}{n}\mathbb{Z}$ with $x \geq -N^5$ and $|x - y| = kn^{-1}$. Suppose $\mathcal{C}^n_t(x, y) \neq \emptyset$. Take $(i, j) \in \mathcal{C}^n_t(x, y)$, and let $(x_0, i_0) = (\zeta^{n,t+\delta_n}_{\delta_n}(x, i), \theta^{n,t+\delta_n}_{\delta_n}(x, i))$. Since $(\zeta^{n,t+\delta_n}_s(x, i), \theta^{n,t+\delta_n}_s(x, i)) \in \mathcal{D}^n_t(x_0, i_0)$ and $(\zeta^{n,t+\delta_n}_s(y, j), \theta^{n,t+\delta_n}_s(y, j)) \in \mathcal{D}^n_t(x_0, i_0)$ $\forall s \in [0, \delta_n]$, we have $(x, i), (y, j) \in \mathcal{D}^n_t(x_0, i_0)$ and $|\mathcal{D}^n_t(x_0, i_0)| \geq \max(k, n|x_0 - x|) + 1 \geq 3$. If $p^n_t(y) = 0$ $\forall y \geq N^5$ and $q^{n,-}_{t,t+\delta_n}(-2N^5, x) = 0$, then by (5.4) it follows that $k < \mathcal{M}^n_t$ and $|x_0 - x| < \mathcal{M}^n_t n^{-1}$. Therefore

$$|\mathcal{C}^n_t(x, y)| \leq \mathbb{1}_{|x-y|<\mathcal{M}^n_t n^{-1}} \binom{\mathcal{M}^n_t}{2} |\{(x_0, i_0) \in \tfrac{1}{n}\mathbb{Z} \times [N] : |x_0 - x| < \mathcal{M}^n_t n^{-1}, |\mathcal{D}^n_t(x_0, i_0)| \geq 3\}|.$$

By Lemma 3.3, (5.17), (5.8), (5.15) and (5.16), it follows that for $K \in \mathbb{N}$ sufficiently large, for n sufficiently large, for $x \geq -N^5$ and $t \in [(\log N)^2 - \delta_n, N^2]$, (5.1) holds with $j = 3$.

Finally, suppose $x, y, y' \in \frac{1}{n}\mathbb{Z}$ with $x \geq -N^5$ and suppose $\mathcal{C}^n_t(x, y, y') \neq \emptyset$. Take $(i, j, j') \in \mathcal{C}^n_t(x, y, y')$, and let $(x_0, i_0) = (\zeta^{n,t+\delta_n}_{\delta_n}(x, i), \theta^{n,t+\delta_n}_{\delta_n}(x, i))$. Suppose that $p^n_t(y) = 0$ $\forall y \geq N^5$ and $q^{n,-}_{t,t+\delta_n}(-2N^5, x) = 0$. Then $(x, i), (y, j), (y', j') \in \mathcal{D}^n_t(x_0, i_0)$, and moreover $|x - x_0| < \mathcal{M}^n_t n^{-1}$ and $|x - y| \vee |x - y'| < \mathcal{M}^n_t n^{-1}$. Therefore

$$\begin{aligned}
&|\mathcal{C}^n_t(x, y, y')| \\
&\leq \mathbb{1}_{|x-y|\vee|x-y'|<\mathcal{M}^n_t n^{-1}} (\mathcal{M}^n_t)^3 |\{(x_0, i_0) \in \tfrac{1}{n}\mathbb{Z} \times [N] : |x_0 - x| < \mathcal{M}^n_t n^{-1}, |\mathcal{D}^n_t(x_0, i_0)| \geq 3\}|.
\end{aligned}$$

By Lemma 3.3, (5.17), (5.8), (5.15) and (5.16), it follows that for $K \in \mathbb{N}$ sufficiently large, for n sufficiently large, for $x \geq -N^5$ and $t \in [(\log N)^2 - \delta_n, N^2]$, (5.1) holds with $j = 4$. This completes the proof. □

6 Event E_4 occurs with high probability

In this section, we complete the proof of Proposition 2.1 by proving the following result.

Proposition 6.1. *Suppose for some $a_1 > 1$, $N \geq n^{a_1}$ for n sufficiently large. For $b_1 > 0$ sufficiently small, $b_2 > 0$ and $t^* \in \mathbb{N}$, for $K \in \mathbb{N}$ sufficiently large, then for n sufficiently large, if condition* (A) *holds,*

$$\mathbb{P}\left((E_4)^c\right) \leq \left(\frac{n}{N}\right)^2.$$

Proposition 2.1 now follows directly from Propositions 3.1, 4.1, 5.1 and 6.1. From now on in this section, we assume that there exists $a_1 > 1$ such that $N \geq n^{a_1}$ for n sufficiently large. We begin by proving the following lemma, which we will then use iteratively to show that with high probability no lineages consistently stay far ahead of the front. Recall the definition of q^n_t from (3.10). Fix $t^* \in \mathbb{N}$.

Lemma 6.2. *There exist $c \in (0, 1)$ and $\epsilon \in (0, 1)$ such that for $K \in \mathbb{N}$ sufficiently large, the following holds. Suppose q^n_0 and $((\mathcal{P}^{x,i,j})_{x,i,j}, (\mathcal{S}^{x,i,j})_{x,i,j}, (\mathcal{Q}^{x,i,j,k})_{x,i,j,k}, (\mathcal{R}^{x,i,y,j})_{x,i,y,j})$ are independent, and define the event*

$$A = \left\{ \sup_{t\in[0,t^*],\, x\in\frac{1}{n}\mathbb{Z}} |p^n_t(x) - g(x - \mu^n_t)| \leq \epsilon \right\} \cap \left\{ \sup_{t\in[0,t^*]} \mu^n_t \leq 2\nu t^* \right\}.$$

Then

$$\sup_{z\geq K} \mathbb{E}\left[q^n_{t^*}(z)\right] \leq c \sup_{x\in\frac{1}{n}\mathbb{Z}} \mathbb{E}\left[q^n_0(x)\right] + 4 s_0 t^* \mathbb{P}\left(A^c\right). \tag{6.1}$$

Proof. Let $\delta = \mathbb{P}(A^c)$. For $a \in \mathbb{R}$, $t \geq 0$ and $z \in \frac{1}{n}\mathbb{Z}$, by Lemma 3.12, $(M^n_s(\phi^{t,z,as_0}))_{s\geq 0}$ is a martingale with $M^n_0(\phi^{t,z,as_0}) = 0$. Hence by Corollary 3.13,

$$\begin{aligned}&\mathbb{E}\left[q^n_t(z)\right]\\ &= e^{-as_0 t}\langle \mathbb{E}\left[q^n_0\right], \phi^{t,z}_0\rangle_n + s_0 \int_0^t e^{-as_0(t-s)}\langle \mathbb{E}\left[q^n_s((1-p^n_s)(2p^n_s - 1+\alpha)+a)\right], \phi^{t,z}_s\rangle_n ds.\end{aligned} \tag{6.2}$$

Take $a \in (0, 1-\alpha)$ and then take $\epsilon \in (0, \frac{1}{2}(1-\alpha))$ sufficiently small that $(1-\epsilon)(2\epsilon - 1 + \alpha) < -a$. Take $K \in \mathbb{N}$ sufficiently large that $1 - g(K/2 - 2t^*\nu) - \epsilon > 0$, $e^{-as_0 t^*} + 2s_0 t^* e^{(2s_0+m)t^* - K/2} < 1$ and

$$(1 - g(x - 2\nu t^*) - \epsilon)(2(g(x-2\nu t^*) + \epsilon) - 1 + \alpha) \leq -a \qquad \text{for } x \geq K/2.$$

Then on the event A,

$$(1 - p^n_s(x))(2p^n_s(x) - 1 + \alpha) + a \leq 0 \qquad \forall\, x \geq K/2,\ s \in [0, t^*].$$

It follows that for $x \geq K/2$ and $s \in [0, t^*]$, since $p^n_s(x), q^n_s(x) \in [0,1]$,

$$\mathbb{E}\left[q^n_s(x)((1-p^n_s(x))(2p^n_s(x) - 1 + \alpha) + a)\right] \leq \mathbb{E}\left[q^n_s(x)(1+\alpha+a)\mathbb{1}_{A^c}\right] \leq 2\delta,$$

and for $x \leq K/2$ and $s \in [0, t^*]$,

$$\mathbb{E}\left[q^n_s(x)((1-p^n_s(x))(2p^n_s(x) - 1 + \alpha) + a)\right] \leq \mathbb{E}\left[q^n_s(x)(1+\alpha+a)\right] \leq 2\mathbb{E}\left[q^n_s(x)\right].$$

Hence for $t \in [0, t^*]$ and $z \in \frac{1}{n}\mathbb{Z}$, substituting into (6.2),

$$\begin{aligned}\mathbb{E}\left[q^n_t(z)\right] &\leq e^{-as_0 t}\langle \mathbb{E}\left[q^n_0\right], \phi^{t,z}_0\rangle_n + s_0\int_0^t e^{-as_0(t-s)}\langle 2\delta + 2\sup_{y\in\frac{1}{n}\mathbb{Z}} \mathbb{E}\left[q^n_s(y)\right]\mathbb{1}_{\cdot \leq K/2}, \phi^{t,z}_s\rangle_n ds\\ &\leq e^{-as_0 t}\sup_{x\in\frac{1}{n}\mathbb{Z}} \mathbb{E}\left[q^n_0(x)\right] + 2s_0 t^*\delta + 2s_0\int_0^t \sup_{y\in\frac{1}{n}\mathbb{Z}}\mathbb{E}\left[q^n_s(y)\right]\mathbf{P}_z\left(X^n_{m(t-s)} \leq K/2\right) ds.\end{aligned} \tag{6.3}$$

In particular, for $t \in [0, t^*]$, since $a > 0$,

$$\sup_{z\in\frac{1}{n}\mathbb{Z}}\mathbb{E}\left[q^n_t(z)\right] \leq \sup_{x\in\frac{1}{n}\mathbb{Z}}\mathbb{E}\left[q^n_0(x)\right] + 2s_0 t^*\delta + 2s_0\int_0^t \sup_{y\in\frac{1}{n}\mathbb{Z}}\mathbb{E}\left[q^n_s(y)\right] ds.$$

By Gronwall's inequality, it follows that for $t \in [0, t^*]$,

$$\sup_{z\in\frac{1}{n}\mathbb{Z}}\mathbb{E}\left[q^n_t(z)\right] \leq \left(\sup_{x\in\frac{1}{n}\mathbb{Z}}\mathbb{E}\left[q^n_0(x)\right] + 2s_0 t^*\delta\right) e^{2s_0 t}. \tag{6.4}$$

Therefore, substituting the bound in (6.4) into (6.3), for $t \in [0, t^*]$ and $z \in \frac{1}{n}\mathbb{Z}$ with $z \geq K$,

$$\begin{aligned}\mathbb{E}\left[q^n_t(z)\right] &\leq e^{-as_0 t}\sup_{x\in\frac{1}{n}\mathbb{Z}}\mathbb{E}\left[q^n_0(x)\right] + 2s_0 t^*\delta\\ &\quad + 2s_0\int_0^t e^{2s_0 s}\left(\sup_{x\in\frac{1}{n}\mathbb{Z}}\mathbb{E}\left[q^n_0(x)\right] + 2s_0 t^*\delta\right)\mathbf{P}_K\left(X^n_{m(t-s)} \leq K/2\right) ds.\end{aligned}$$

For $0 \leq s \leq t \leq t^*$, by Markov's inequality and Lemma 3.8,

$$\mathbf{P}_K\left(X^n_{m(t-s)} \leq K/2\right) = \mathbf{P}_0\left(X^n_{m(t-s)} \geq K/2\right) \leq e^{-K/2}\mathbb{E}\left[e^{X^n_{m(t-s)}}\right] \leq e^{mt^* - K/2}$$

for n sufficiently large. Hence for $z \in \frac{1}{n}\mathbb{Z}$ with $z \geq K$,

$$\mathbb{E}\left[q^n_{t^*}(z)\right] \leq (e^{-as_0 t^*} + 2s_0 t^* e^{(2s_0+m)t^*-K/2}) \sup_{x \in \frac{1}{n}\mathbb{Z}} \mathbb{E}\left[q^n_0(x)\right] + 2s_0 t^* \delta(1 + 2s_0 t^* e^{(2s_0+m)t^*-K/2}),$$

which completes the proof, since at the start of the proof we chose K sufficiently large that $e^{-as_0 t^*} + 2s_0 t^* e^{(2s_0+m)t^*-K/2} < 1$. $\square$

Take $c \in (0,1)$ and $\epsilon \in (0,1)$ as in Lemma 6.2. For $t \geq 0$, define the σ-algebra $\mathcal{F}'_t = \sigma((p^n_s(x))_{s \in [0,t], x \in \frac{1}{n}\mathbb{Z}})$. The following result will easily imply Proposition 6.1.

Proposition 6.3. *For $\ell \in \mathbb{N}$, there exists $\ell' \in \mathbb{N}$ such that for $K \in \mathbb{N}$ sufficiently large and $c_2 > 0$, the following holds for n sufficiently large. Take $t \in \delta_n \mathbb{N}_0 \cap [0, T_n^-]$ and let $t' = T_n - t - t^* \lfloor (t^*)^{-1} K \log N \rfloor$. Suppose $p^n_{t'}(x) = 0 \; \forall x \geq N^5$ and $\mathbb{P}\left((E_1)^c | \mathcal{F}'_{t'}\right) \leq \left(\frac{n}{N}\right)^{\ell'}$. Then*

$$\mathbb{P}\left(r^{n,K,t^*}_{K \log N, T_n - t}(x) = 0 \; \forall x \in \tfrac{1}{n}\mathbb{Z} \middle| \mathcal{F}'_{t'}\right) \geq 1 - \left(\frac{n}{N}\right)^{\ell}.$$

Proof. Take ℓ' sufficiently large that $nN^6 \left(\frac{n}{N}\right)^{\ell'} \leq \left(\frac{n}{N}\right)^{\ell+1}$ for n sufficiently large. Then take $c' \in (c,1)$ and take $K > t^*(\ell'+1)(-\log c')^{-1}$ sufficiently large that Lemma 6.2 holds. Suppose

$$\mathbb{P}\left((E_1)^c | \mathcal{F}'_{t'}\right) \leq \left(\frac{n}{N}\right)^{\ell'}. \tag{6.5}$$

For $k \in \mathbb{N}$ and $x \in \frac{1}{n}\mathbb{Z}$, let $r^n_k(x) = r^{n,K,t^*}_{kt^*, t'+kt^*}(x)$. Take $k \in \mathbb{N}$ with $kt^* \leq K \log N$. Then by the definition of $r^{n,y,\ell}_{s,t}$ in (2.6),

$$\begin{aligned}
\sup_{z \in \frac{1}{n}\mathbb{Z}} \mathbb{E}\left[r^n_k(z) \middle| \mathcal{F}'_{t'}\right] &= \sup_{z \in \frac{1}{n}\mathbb{Z}} \mathbb{E}\left[r^n_k(z) \mathbb{1}_{z \geq \mu^n_{t'+kt^*}+K} \left(\mathbb{1}_{E_1} + \mathbb{1}_{(E_1)^c}\right) \middle| \mathcal{F}'_{t'}\right] \\
&\leq \sup_{z \in \frac{1}{n}\mathbb{Z},\, z \geq \mu^n_{t'} + \nu k t^* + K - \nu t^*} \mathbb{E}\left[r^n_k(z) \middle| \mathcal{F}'_{t'}\right] + \mathbb{P}\left((E_1)^c | \mathcal{F}'_{t'}\right)
\end{aligned}$$

for n sufficiently large, by the definition of the event E_1 in (2.10). Therefore, by (6.5) and then by Lemma 6.2 with $q^n_0 = r^n_{k-1}(\cdot + \mu^n_{t'} + \lfloor \nu(k-1)t^* n \rfloor n^{-1})$,

$$\begin{aligned}
\sup_{z \in \frac{1}{n}\mathbb{Z}} \mathbb{E}\left[r^n_k(z) \middle| \mathcal{F}'_{t'}\right] &\leq \sup_{z \in \frac{1}{n}\mathbb{Z},\, z \geq \mu^n_{t'} + \lfloor \nu(k-1)t^* n \rfloor n^{-1} + K} \mathbb{E}\left[r^n_k(z) \middle| \mathcal{F}'_{t'}\right] + \left(\frac{n}{N}\right)^{\ell'} \\
&\leq c \sup_{x \in \frac{1}{n}\mathbb{Z}} \mathbb{E}\left[r^n_{k-1}(x) \middle| \mathcal{F}'_{t'}\right] + (1 + 4s_0 t^*) \left(\frac{n}{N}\right)^{\ell'}
\end{aligned} \tag{6.6}$$

for n sufficiently large. Recall that we chose $c' \in (c,1)$, and let

$$k^* = \min\left\{k \in \mathbb{N}_0 : \sup_{x \in \frac{1}{n}\mathbb{Z}} \mathbb{E}\left[r^n_k(x) \middle| \mathcal{F}'_{t'}\right] \leq \frac{1 + 4s_0 t^*}{c' - c} \left(\frac{n}{N}\right)^{\ell'}\right\}.$$

Then for $k \in \mathbb{N}$ with $k \leq \min(k^*, (t^*)^{-1} K \log N)$, we have $(c'-c) \sup_{x \in \frac{1}{n}\mathbb{Z}} \mathbb{E}\left[r^n_{k-1}(x) \middle| \mathcal{F}'_{t'}\right] \geq (1 + 4s_0 t^*) \left(\frac{n}{N}\right)^{\ell'}$ by the definition of k^*, and so by (6.6),

$$\sup_{x \in \frac{1}{n}\mathbb{Z}} \mathbb{E}\left[r^n_k(x) | \mathcal{F}'_{t'}\right] \leq c' \sup_{x \in \frac{1}{n}\mathbb{Z}} \mathbb{E}\left[r^n_{k-1}(x) | \mathcal{F}'_{t'}\right] \leq \ldots \leq (c')^k \sup_{x \in \frac{1}{n}\mathbb{Z}} \mathbb{E}\left[r^n_0(x) | \mathcal{F}'_{t'}\right] \leq (c')^k.$$

Hence for n sufficiently large, since $\lfloor (t^*)^{-1} K \log N \rfloor - 1 > (\ell'+1)(-\log c')^{-1} \log(N/n)$ by our choice of K, we have $k^* < (t^*)^{-1} K \log N$. For $k \in \mathbb{N} \cap [k^*+1, (t^*)^{-1} K \log N]$, if

$$\sup_{x \in \frac{1}{n}\mathbb{Z}} \mathbb{E}\left[r^n_{k-1}(x) | \mathcal{F}'_{t'}\right] \leq \frac{1 + 4s_0 t^*}{c' - c} \left(\frac{n}{N}\right)^{\ell'}$$

then by (6.6),

$$\sup_{x\in\frac{1}{n}\mathbb{Z}} \mathbb{E}\left[r^n_k(x)\middle|\mathcal{F}'_{t'}\right] \le \left(\frac{c}{c'-c}+1\right)(1+4s_0t^*)\left(\frac{n}{N}\right)^{\ell'} \le \frac{1+4s_0t^*}{c'-c}\left(\frac{n}{N}\right)^{\ell'} \tag{6.7}$$

since $c'<1$. Therefore, by induction, (6.7) holds for all $k\in\mathbb{N}\cap[k^*,(t^*)^{-1}K\log N]$. By a union bound, and then by Lemma 3.3 and since $p^n_{t'}(x)=0\ \forall x\ge N^5$, and by (6.7),

$$\begin{aligned}
&\mathbb{P}\left(\sup_{x\in\frac{1}{n}\mathbb{Z}} r^n_{\lfloor (t^*)^{-1}K\log N\rfloor}(x)>0\middle|\mathcal{F}'_{t'}\right)\\
&\le \mathbb{P}\left(\exists x\ge 2N^5 : p^n_{T_n-t}(x)>0\middle|\mathcal{F}'_{t'}\right)+\mathbb{P}\left(\mu^n_{T_n-t}\le 0\middle|\mathcal{F}'_{t'}\right)\\
&\quad+\sum_{x\in\frac{1}{n}\mathbb{Z}\cap[K,2N^5]} N\mathbb{E}\left[r^n_{\lfloor (t^*)^{-1}K\log N\rfloor}(x)\middle|\mathcal{F}'_{t'}\right]\\
&\le e^{-N^5}+\mathbb{P}\left((E_1)^c|\mathcal{F}'_{t'}\right)+2nN^5\cdot N\frac{1+4s_0t^*}{c'-c}\left(\frac{n}{N}\right)^{\ell'}\\
&\le \left(\frac{n}{N}\right)^{\ell}
\end{aligned}$$

for n sufficiently large, by (6.5) and our choice of ℓ'. □

Proof of Proposition 6.1. Take $\ell\in\mathbb{N}$ sufficiently large that $\left(\frac{n}{N}\right)^{\ell-2}N^2\delta_n^{-1}\le\left(\frac{n}{N}\right)^3$ for n sufficiently large. Take $\ell'\in\mathbb{N}$ and $K\in\mathbb{N}$ sufficiently large that Proposition 6.3 holds. By Proposition 3.1, by taking $b_1,c_2>0$ sufficiently small, $\mathbb{P}\left((E_1)^c\right)\le\left(\frac{n}{N}\right)^{\ell+\ell'}$ for n sufficiently large. For $t\in\delta_n\mathbb{N}_0\cap[0,T_n^-]$, let

$$D_t=\left\{r^{n,K,t^*}_{K\log N,T_n-t}(x)=0\ \forall x\in\tfrac{1}{n}\mathbb{Z}\right\}.$$

Then by Proposition 6.3, letting $t'=T_n-t-t^*\lfloor (t^*)^{-1}K\log N\rfloor$,

$$\mathbb{P}\left(D_t^c\middle|\mathcal{F}'_{t'}\right)\le\left(\frac{n}{N}\right)^{\ell}+\mathbb{1}_{\{\mathbb{P}\left((E_1)^c|\mathcal{F}'_{t'}\right)>\left(\frac{n}{N}\right)^{\ell'}\}}+\mathbb{1}_{\{\exists x\ge N^5:p^n_{t'}(x)>0\}}.$$

Hence by Markov's inequality and Lemma 3.3,

$$\mathbb{P}\left(D_t^c\right)\le\left(\frac{n}{N}\right)^{\ell}+\left(\frac{N}{n}\right)^{\ell'}\mathbb{P}\left((E_1)^c\right)+e^{-N^5}\le 3\left(\frac{n}{N}\right)^{\ell}$$

for n sufficiently large. Therefore, by (2.13) and a union bound, and then by Markov's inequality,

$$\mathbb{P}\left((E_4)^c\right)\le\sum_{t\in\delta_n\mathbb{N}_0\cap[0,T_n^-]}\mathbb{P}\left(\mathbb{P}\left(D_t^c|\mathcal{F}\right)\ge\left(\frac{n}{N}\right)^2\right)\le\sum_{t\in\delta_n\mathbb{N}_0\cap[0,T_n^-]}\left(\frac{N}{n}\right)^2\mathbb{P}\left(D_t^c\right)\le\left(\frac{n}{N}\right)^2$$

for n sufficiently large, by our choice of ℓ, which completes the proof. □

7 Proofs of Theorems 1.1 and 1.4

The proofs of Theorems 1.1 and 1.4 use results from Sections 2, 3, 4 and 6. We first prove Theorem 1.1, and then Theorem 1.4 will follow easily from the proof of Theorem 1.1.

Proof of Theorem 1.1. Take $T_n \in [(\log N)^2, N^2]$ and $T'_n \geq 0$ with $T_n - T'_n \geq (\log N)^2$. Recall from (2.4) that $\delta_n = \lfloor N^{1/2} n^2 \rfloor^{-1}$, and let $S_n = T_n - \delta_n \lfloor \delta_n^{-1} T'_n \rfloor$. Take $b_1, c_2 > 0$ sufficiently small and $t^*, K \in \mathbb{N}$ sufficiently large that Proposition 3.1 holds with $\ell = 1$ and Propositions 4.1 and 6.1 hold. Assume $c_2 < a_0$ (recall that $(\log N)^{a_0} \leq \log n$ for n sufficiently large). Recall (2.7), and similarly to (2.16), for $t \in [0, T_n]$ let

$$\mathcal{F}_t = \sigma(\mathcal{F}, \sigma((\zeta_s^{n,T_n}(X_0, J_0))_{s \leq t})).$$

Condition on $\mathcal{F}_0$, and suppose the event $E'_1 \cap E'_2 \cap E_4$ occurs, so in particular by (2.10) and (3.3),

$$|p^n_{S_n}(x) - g(x - \mu^n_{S_n})| \leq e^{-(\log N)^{c_2}} \ \forall x \in \tfrac{1}{n}\mathbb{Z}. \tag{7.1}$$

Fix $x_0 \in \mathbb{R}$ and take $\epsilon > 0$. Define $v_0 : \frac{1}{n}\mathbb{Z} \to [0,1]$ by letting

$$v_0(y) = \begin{cases} p^n_{S_n}(y) & \text{for } y < \mu^n_{S_n} + x_0, \\ \min(p^n_{S_n}(y), N^{-1}\lfloor N h(y) \rfloor) & \text{for } y \in [\mu^n_{S_n} + x_0, \mu^n_{S_n} + x_0 + \epsilon], \\ 0 & \text{for } y > \mu^n_{S_n} + x_0 + \epsilon, \end{cases} \tag{7.2}$$

where $h : [\mu^n_{S_n} + \lfloor x_0 n \rfloor n^{-1}, \mu^n_{S_n} + \lceil (x_0+\epsilon) n \rceil n^{-1}] \to [0,1]$ is linear with $h(\mu^n_{S_n} + \lfloor x_0 n \rfloor n^{-1}) = p^n_{S_n}(\mu^n_{S_n} + \lfloor x_0 n \rfloor n^{-1})$ and $h(\mu^n_{S_n} + \lceil (x_0+\epsilon) n \rceil n^{-1}) = 0$. For each $y \in \frac{1}{n}\mathbb{Z}$, take $I_y \subseteq \{(y,i) : \xi^n_{S_n}(y,i) = 1\}$ measurable with respect to $\sigma((\xi^n_{S_n}(x,j))_{x \in \frac{1}{n}\mathbb{Z}, j \in [N]})$ such that $|I_y| = N v_0(y)$. Then let $I = \cup_{y \in \frac{1}{n}\mathbb{Z}} I_y$. For $t \geq S_n$ and $x \in \frac{1}{n}\mathbb{Z}$, let

$$\tilde{q}^n_t(x) = N^{-1} |\{i \in [N] : (\zeta^{n,t}_{t-S_n}(x,i), \theta^{n,t}_{t-S_n}(x,i)) \in I\}|,$$

the proportion of individuals at x at time t which are descended from the set I at time S_n. Recall the definition of $q^{n,-}$ in (2.3) and note that for $t \geq S_n$ and $x \in \frac{1}{n}\mathbb{Z}$,

$$q^{n,-}_{S_n,t}(\mu^n_{S_n} + x_0, x) \leq \tilde{q}^n_t(x) \leq q^{n,-}_{S_n,t}(\mu^n_{S_n} + x_0 + \epsilon, x). \tag{7.3}$$

Let $(\tilde{v}^n_t)_{t \geq S_n}$ solve

$$\begin{cases} \partial_t \tilde{v}^n_t = \frac{1}{2} m \Delta_n \tilde{v}^n_t + s_0 \tilde{v}^n_t (1 - u^n_{S_n,t})(2u^n_{S_n,t} - 1 + \alpha) & \text{for } t > S_n, \\ \tilde{v}^n_{S_n} = v_0, \end{cases}$$

where $(u^n_{S_n,t})_{t \geq S_n}$ is defined as in (3.2). Recall the definition of γ_n in (2.4). Note that by Proposition 3.2, for n sufficiently large, for $t \leq S_n + \gamma_n$,

$$\mathbb{P}\left(\sup_{x \in \frac{1}{n}\mathbb{Z} \cap [-N^5, N^5]} |\tilde{q}^n_t(x) - \tilde{v}^n_t(x)| \geq \left(\frac{n}{N}\right)^{1/4} \right) \leq \frac{n}{N}. \tag{7.4}$$

For $t \geq 0$ and $x \in \mathbb{R}$, let

$$\tilde{v}_t(x) = g(x - \mu^n_{S_n} - \nu t) \mathbb{E}_{x - \mu^n_{S_n} - \nu t}\left[\bar{v}_0(Z_t + \mu^n_{S_n}) g(Z_t)^{-1} \right], \tag{7.5}$$

where $\bar{v}_0 : \mathbb{R} \to [0,1]$ is the linear interpolation of v_0, and $(Z_t)_{t \geq 0}$ is defined in (4.1). By Lemma 4.3 and the definition of the event E'_1 in (3.3), for n sufficiently large,

$$\begin{aligned} &\sup_{x \in \frac{1}{n}\mathbb{Z},\, t \in [0,\gamma_n]} |\tilde{v}^n_{S_n+t}(x) - \tilde{v}_t(x)| \\ &\leq (C_7(n^{-1/3} + e^{-(\log N)^{c_2}}) + 2 \sup_{x_1,x_2 \in \frac{1}{n}\mathbb{Z}, |x_1 - x_2| \leq n^{-1/3}} |v_0(x_1) - v_0(x_2)|) e^{5 s_0 \gamma_n} \gamma_n^2. \end{aligned}$$

By the definition of v_0 in (7.2) and by (7.1),

$$\sup_{x_1,x_2 \in \frac{1}{n}\mathbb{Z}, |x_1 - x_2| \leq n^{-1/3}} |v_0(x_1) - v_0(x_2)| \leq 2(2e^{-(\log N)^{c_2}} + n^{-1/3} \|\nabla g\|_\infty) + \epsilon^{-1} n^{-1/3} + N^{-1}.$$

Therefore, for n sufficiently large, for $t \in [0, \gamma_n]$ and $x \in \frac{1}{n}\mathbb{Z}$ with $|x - \mu^n_{S_n+t}| \le d_n$,

$$\Big| \frac{\tilde{v}^n_{S_n+t}(x)}{g(x-\mu^n_{S_n} - \nu t)} - \mathbb{E}_{x-\mu^n_{S_n}-\nu t}\left[\bar{v}_0(Z_t + \mu^n_{S_n}) g(Z_t)^{-1}\right] \Big| \le e^{-\frac{1}{2}(\log N)^{c_2}}. \tag{7.6}$$

From now on, we consider two different cases; suppose first that $T'_n \le \gamma_n$. Recalling (7.3) and (7.4), suppose for all $x \in \frac{1}{n}\mathbb{Z} \cap [-N^5, N^5]$ that

$$q^{n,-}_{S_n,T_n}(\mu^n_{S_n} + x_0, x) \le \tilde{v}^n_{T_n}(x) + \left(\frac{n}{N}\right)^{1/4} \quad \text{and} \quad q^{n,-}_{S_n,T_n}(\mu^n_{S_n} + x_0 + \epsilon, x) \ge \tilde{v}^n_{T_n}(x) - \left(\frac{n}{N}\right)^{1/4}.$$

By the definition of the event E_1 in (2.10), for n sufficiently large, if $x \in \frac{1}{n}\mathbb{Z}$ with $|x - \mu^n_{T_n}| \le K_0$ then since we are assuming $T'_n \le \gamma_n$ we have $|x - \mu^n_{S_n} - \nu(T_n - S_n)| \le 2K_0$, and so by (7.6),

$$\begin{aligned}
&\frac{q^{n,-}_{S_n,T_n}(\mu^n_{S_n} + x_0, x)}{g(x - \mu^n_{S_n} - \nu(T_n - S_n))} \\
&\le \mathbb{E}_{x-\mu^n_{S_n}-\nu(T_n-S_n)}\left[\bar{v}_0(Z_{T_n-S_n} + \mu^n_{S_n}) g(Z_{T_n-S_n})^{-1}\right] + e^{-\frac{1}{2}(\log N)^{c_2}} + \left(\frac{n}{N}\right)^{1/4} g(2K_0)^{-1}
\end{aligned} \tag{7.7}$$

and

$$\begin{aligned}
&\frac{q^{n,-}_{S_n,T_n}(\mu^n_{S_n} + x_0 + \epsilon, x)}{g(x - \mu^n_{S_n} - \nu(T_n - S_n))} \\
&\ge \mathbb{E}_{x-\mu^n_{S_n}-\nu(T_n-S_n)}\left[\bar{v}_0(Z_{T_n-S_n} + \mu^n_{S_n}) g(Z_{T_n-S_n})^{-1}\right] - e^{-\frac{1}{2}(\log N)^{c_2}} - \left(\frac{n}{N}\right)^{1/4} g(2K_0)^{-1}.
\end{aligned} \tag{7.8}$$

Applying (4.11) in Lemma 4.4, it follows that

$$\begin{aligned}
&\frac{q^{n,-}_{S_n,T_n}(\mu^n_{S_n} + x_0, x)}{g(x - \mu^n_{S_n} - \nu(T_n - S_n))} \\
&\le \int_{-\infty}^{\infty} \pi(y) \bar{v}_0(y + \mu^n_{S_n}) g(y)^{-1} dy + 2m^{-1/2}(T_n - S_n)^{-1/4} \sup_{z\in\mathbb{R}} |\bar{v}_0(z + \mu^n_{S_n}) g(z)^{-1}| \\
&\quad + e^{-\frac{1}{2}(\log N)^{c_2}} + \left(\frac{n}{N}\right)^{1/4} g(2K_0)^{-1} \\
&\le \int_{-\infty}^{x_0+\epsilon} \pi(y) dy + \epsilon
\end{aligned} \tag{7.9}$$

for n sufficiently large, since by (7.1) and by the definition of v_0 in (7.2), $v_0(y + \mu^n_{S_n}) \le (g(y) + e^{-(\log N)^{c_2}}) \mathbb{1}_{y \le x_0+\epsilon}$ $\forall y \in \frac{1}{n}\mathbb{Z}$, and since we are assuming that $T'_n \to \infty$ as $n \to \infty$. Similarly, since $v_0(y + \mu^n_{S_n}) \ge (g(y) - e^{-(\log N)^{c_2}}) \mathbb{1}_{y<x_0}$ $\forall y \in \frac{1}{n}\mathbb{Z}$, for n sufficiently large we have

$$\frac{q^{n,-}_{S_n,T_n}(\mu^n_{S_n} + x_0 + \epsilon, x)}{g(x - \mu^n_{S_n} - \nu(T_n - S_n))} \ge \int_{-\infty}^{x_0} \pi(y) dy - \epsilon. \tag{7.10}$$

For n sufficiently large, since $|T_n - T'_n - S_n| \le \delta_n$ we have that $|\mu^n_{T_n - T'_n} - \mu^n_{S_n}| \le \epsilon$. Recall the definition of G_{K_0,T_n} in (1.14). Then for $(X_0, J_0) \in G_{K_0,T_n}$ we have $|X_0 - \mu^n_{T_n}| \le K_0$, and so for n sufficiently large, by the definition of the event E_1 in (2.10) and by (7.10),

$$\mathbb{P}\left(\zeta^{n,T_n}_{T_n-S_n}(X_0, J_0) \le \mu^n_{T_n-T'_n} + x_0 + 2\epsilon \Big| \mathcal{F}_0\right) \ge \frac{q^{n,-}_{S_n,T_n}(\mu^n_{S_n} + x_0 + \epsilon, X_0)}{p^n_{T_n}(X_0)} \ge \int_{-\infty}^{x_0} \pi(y) dy - 2\epsilon$$

and by (7.9),

$$\mathbb{P}\left(\zeta^{n,T_n}_{T_n-S_n}(X_0,J_0)\le \mu^n_{T_n-T_n'}+x_0-\epsilon\Big|\mathcal{F}_0\right)\le \frac{q^{n,-}_{S_n,T_n}(\mu^n_{S_n}+x_0,X_0)}{p^n_{T_n}(X_0)}\le \int_{-\infty}^{x_0+\epsilon}\pi(y)dy+2\epsilon.$$

Hence letting $y_0=x_0+2\epsilon$, by (7.3) and (7.4), for n sufficiently large,

$$\begin{aligned}&\mathbb{P}\left(\zeta^{n,T_n}_{T_n-S_n}(X_0,J_0)-\mu^n_{T_n-T_n'}\le y_0\right)\\ &\ge\left(\int_{-\infty}^{y_0-2\epsilon}\pi(y)dy-2\epsilon\right)\left(1-\frac{n}{N}-\mathbb{P}\left((E_1'\cap E_2'\cap E_4)^c\right)\right)\\ &\ge\int_{-\infty}^{y_0-2\epsilon}\pi(y)dy-3\epsilon \end{aligned}\tag{7.11}$$

for n sufficiently large, by Propositions 3.1, 4.1 and 6.1. Similarly, for n sufficiently large,

$$\mathbb{P}\left(\zeta^{n,T_n}_{T_n-S_n}(X_0,J_0)-\mu^n_{T_n-T_n'}\le y_0\right)\le\int_{-\infty}^{y_0+2\epsilon}\pi(y)dy+3\epsilon.\tag{7.12}$$

By the same argument as in the proof of Lemma 4.12, by Corollary 3.13 with $a=-(1+\alpha)s_0$, and since $|T_n-T_n'-S_n|\le\delta_n$, we have that for $x_1,x_2\in\frac{1}{n}\mathbb{Z}$,

$$\begin{aligned}\mathbb{E}\left[q^n_{T_n-T_n',S_n}(x_1,x_2)\right]&\le e^{(1+\alpha)s_0\delta_n}\mathbf{P}_{x_2}\left(X^n_{m(S_n-(T_n-T_n'))}=x_1\right)\\ &\le e^{(1+\alpha)s_0\delta_n}e^{-n^{1/2}|x_1-x_2|}e^{mn\delta_n}\end{aligned}$$

for n sufficiently large, by Lemma 3.8. Therefore, by a union bound and since, on the event $E_1\cap E_2'$, $|\zeta^{n,T_n}_{T_n-S_n}(X_0,J_0)|\le N^3$ by Lemma 2.7, and then by Markov's inequality and Propositions 3.1 and 4.1,

$$\begin{aligned}&\mathbb{P}\left(|\zeta^{n,T_n}_{T_n'}(X_0,J_0)-\zeta^{n,T_n}_{T_n-S_n}(X_0,J_0)|\ge n^{-1/3}\right)\\ &\le\sum_{x_1\in\frac{1}{n}\mathbb{Z},\,x_2\in\frac{1}{n}\mathbb{Z}\cap[-N^3,N^3],|x_1-x_2|\ge n^{-1/3}}\mathbb{P}\left(q^n_{T_n-T_n',S_n}(x_1,x_2)\ge N^{-1}\right)+\mathbb{P}\left((E_1\cap E_2')^c\right)\\ &\le Ne^{(1+\alpha)s_0\delta_n}e^{mn\delta_n}\sum_{x_1\in\frac{1}{n}\mathbb{Z},\,x_2\in\frac{1}{n}\mathbb{Z}\cap[-N^3,N^3],|x_1-x_2|\ge n^{-1/3}}e^{-n^{1/2}|x_1-x_2|}+2\frac{n}{N}\\ &\le 3\frac{n}{N}\end{aligned}\tag{7.13}$$

for n sufficiently large. Since $\epsilon>0$ can be taken arbitrarily small, this, together with (7.11) and (7.12), completes the proof in the case $T_n'\le\gamma_n$.

Now suppose instead that $T_n'\ge\gamma_n$, and take $s\in t^*\mathbb{N}_0$ such that $T_n-s\in[S_n+\gamma_n-t^*,S_n+\gamma_n]$. Recall from (2.4) that $d_n=\kappa^{-1}C\log\log N$. By Propositions 2.5 and 2.6, if $(X_0,J_0)\in G_{K_0,T_n}$,

$$\mathbb{P}\left(|\zeta^{n,T_n}_s(X_0,J_0)-\mu^n_{T_n-s}|\ge d_n\Big|\mathcal{F}_0\right)=\mathcal{O}((\log N)^{3-\frac{1}{8}\alpha C})=\mathcal{O}((\log N)^{-1})\tag{7.14}$$

since we chose $C>2^{13}\alpha^{-2}$ at the start of Section 2. Suppose for all $y\in\frac{1}{n}\mathbb{Z}\cap[-N^5,N^5]$ that

$$\begin{aligned}&q^{n,-}_{S_n,T_n-s}(\mu^n_{S_n}+x_0,y)\le\tilde{v}^n_{T_n-s}(y)+\left(\frac{n}{N}\right)^{1/4}\\ \text{and}\quad &q^{n,-}_{S_n,T_n-s}(\mu^n_{S_n}+x_0+\epsilon,y)\ge\tilde{v}^n_{T_n-s}(y)-\left(\frac{n}{N}\right)^{1/4}.\end{aligned}$$

Take $x \in \frac{1}{n}\mathbb{Z}$ with $|x - \mu^n_{T_n-s}| \le d_n$. Then for n sufficiently large, by the definition of the event E_1 in (2.10), and by (7.6) and by (4.10) in Lemma 4.4,

$$\begin{aligned}
&\frac{q^{n,-}_{S_n,T_n-s}(\mu^n_{S_n}+x_0,x)}{g(x-\mu^n_{S_n}-\nu(T_n-s-S_n))}\\
&\le \int_{-\infty}^{\infty}\pi(y)\bar{v}_0(y+\mu^n_{S_n})g(y)^{-1}dy + (\log N)^{-12C}\sup_{z\in\mathbb{R}}|\bar{v}_0(z+\mu^n_{S_n})g(z)^{-1}|\\
&\quad + e^{-\frac{1}{2}(\log N)^{c_2}} + \left(\frac{n}{N}\right)^{1/4} g(d_n+1)^{-1}\\
&\le \int_{-\infty}^{x_0+\epsilon}\pi(y)dy + \epsilon
\end{aligned}$$

for n sufficiently large, as in (7.9). Hence for n sufficiently large that $|\mu^n_{T_n-T'_n} - \mu^n_{S_n}| \le \epsilon$, if $|\zeta^{n,T_n}_s(X_0,J_0) - \mu^n_{T_n-s}| \le d_n$ then

$$\begin{aligned}
\mathbb{P}\left(\zeta^{n,T_n}_{T_n-S_n}(X_0,J_0) \le \mu^n_{T_n-T'_n} + x_0 - \epsilon \middle| \mathcal{F}_s\right) &\le \frac{q^{n,-}_{S_n,T_n-s}(\mu^n_{S_n}+x_0, \zeta^{n,T_n}_s(X_0,J_0))}{p^n_{T_n-s}(\zeta^{n,T_n}_s(X_0,J_0))}\\
&\le \int_{-\infty}^{x_0+\epsilon}\pi(y)dy + 2\epsilon
\end{aligned}$$

for n sufficiently large, and similarly

$$\mathbb{P}\left(\zeta^{n,T_n}_{T_n-S_n}(X_0,J_0) \le \mu^n_{T_n-T'_n} + x_0 + 2\epsilon \middle| \mathcal{F}_s\right) \ge \int_{-\infty}^{x_0}\pi(y)dy - 2\epsilon.$$

As in (7.11) and (7.12), it follows by (7.14), (7.3), (7.4) and Propositions 3.1, 4.1 and 6.1 that for n sufficiently large,

$$\int_{-\infty}^{y_0-2\epsilon}\pi(y)dy - 3\epsilon \le \mathbb{P}\left(\zeta^{n,T_n}_{T_n-S_n}(X_0,J_0) - \mu^n_{T_n-T'_n} \le y_0\right) \le \int_{-\infty}^{y_0+2\epsilon}\pi(y)dy + 3\epsilon.$$

By (7.13) and since $\epsilon > 0$ can be taken arbitrarily small, this completes the proof. □

Proof of Theorem 1.4. We begin by proving the following claim. Let $(Z_t)_{t\ge0}$ be defined as in (4.1). For $t_* > 0$, there exists $C_* = C_*(t_*) > 0$ such that for $x_1, x_2, y_1, y_2 \in \mathbb{R}$ and $t_1, t_2 \ge t_*$ with $|t_1 - t_2| \le 1$,

$$\left|\mathbb{P}_{x_1}(Z_{t_1} \le y_1) - \mathbb{P}_{x_2}(Z_{t_2} \le y_2)\right| \le C_*(|x_1-x_2|^{1/2} + |y_1-y_2|^{1/2} + |t_1-t_2|^{1/6}). \quad (7.15)$$

To prove the claim, first let $(Z^{(1)}_t)_{t\ge0}$ and $(Z^{(2)}_t)_{t\ge0}$ solve (4.1), with $Z^{(1)}_0 = x_1$ and $Z^{(2)}_0 = x_2$. We can couple $Z^{(1)}$ and $Z^{(2)}$ with a Brownian motion $(B_t)_{t\ge0}$ in such a way that

$$\begin{aligned}
Z^{(1)}_t &= x_1 + \nu t + m\int_0^t \frac{\nabla g(Z^{(1)}_s)}{g(Z^{(1)}_s)}ds + \sqrt{m}B_t\\
\text{and}\quad Z^{(2)}_t &= x_2 + \nu t + m\int_0^t \frac{\nabla g(Z^{(2)}_s)}{g(Z^{(2)}_s)}ds + \sqrt{m}B_t
\end{aligned}$$

for $t \in [0,\tau]$, where $\tau = \inf\{t \ge 0 : Z^{(1)}_t = Z^{(2)}_t\}$, and $Z^{(1)}_t = Z^{(2)}_t$ for $t \ge \tau$. Then for $t \in [0,\tau]$ we have

$$Z^{(1)}_t - Z^{(2)}_t = x_1 - x_2 + m\int_0^t \left(\frac{\nabla g(Z^{(1)}_s)}{g(Z^{(1)}_s)} - \frac{\nabla g(Z^{(2)}_s)}{g(Z^{(2)}_s)}\right)ds.$$

Since $y \mapsto \frac{\nabla g(y)}{g(y)}$ is decreasing, it follows that $|Z_t^{(1)} - Z_t^{(2)}| \leq |x_1 - x_2| \; \forall t \geq 0$. Therefore

$$\mathbb{P}_{x_1}\left(Z_{t_1} \leq y_1\right) = \mathbb{P}\left(Z_{t_1}^{(1)} \leq y_1\right) \leq \mathbb{P}\left(Z_{t_1}^{(2)} \leq y_1 + |x_1 - x_2|\right) = \mathbb{P}_{x_2}\left(Z_{t_1} \leq y_1 + |x_1 - x_2|\right). \tag{7.16}$$

Now for any $C > 0$ we can use a union bound to write

$$\begin{aligned}&\mathbb{P}_{x_2}\left(Z_{t_1} \leq y_1 + |x_1 - x_2|\right)\\ &\quad \leq \mathbb{P}_{x_2}\left(Z_{t_2} \leq y_1 + |x_1 - x_2| + C|t_1 - t_2|^{1/3}\right) + \mathbb{P}_{x_2}\left(|Z_{t_1} - Z_{t_2}| \geq C|t_1 - t_2|^{1/3}\right).\end{aligned} \tag{7.17}$$

To bound the second term on the right hand side, note that we can write

$$|Z_{t_1} - Z_{t_2}| \leq \big(\nu + m \sup_{y\in\mathbb{R}} \big|\tfrac{\nabla g(y)}{g(y)}\big|\big)|t_1 - t_2| + \sqrt{m}|B_{|t_1 - t_2|}|,$$

where $(B_t)_{t\geq 0}$ is a Brownian motion. Therefore, since $|t_1 - t_2| \leq 1$, for $C > 0$ a sufficiently large constant, we can write

$$\mathbb{P}_{x_2}\left(|Z_{t_1} - Z_{t_2}| \geq C|t_1 - t_2|^{1/3}\right) \leq \mathbb{P}\left(|B_{|t_1 - t_2|}| \geq |t_1 - t_2|^{1/3}\right) \leq 2e^{-\frac{1}{2}|t_1 - t_2|^{-1/3}}, \tag{7.18}$$

where the last inequality follows by a Gaussian tail estimate. For the first term on the right hand side of (7.17), note that for $z \in \mathbb{R}$ and $\delta \in (0, t_2]$, by conditioning on $Z_{t_2-\delta}$, and then letting $(B_t)_{t\geq 0}$ denote a Brownian motion,

$$\begin{aligned}&\mathbb{P}_{x_2}\left(Z_{t_2} \in [z, z+\delta]\right)\\ &\leq \sup_{x\in\mathbb{R}} \mathbb{P}_x\left(Z_\delta \in [z, z+\delta]\right)\\ &\leq \sup_{x\in\mathbb{R}} \mathbb{P}_x\left(\sqrt{m}B_\delta \in \big[z - \big(\nu + m\sup_{y\in\mathbb{R}}\big|\tfrac{\nabla g(y)}{g(y)}\big|\big)\delta, z + \big(1 - \nu + m\sup_{y\in\mathbb{R}}\big|\tfrac{\nabla g(y)}{g(y)}\big|\big)\delta\big]\right)\\ &\leq \frac{\delta^{1/2}}{\sqrt{2\pi m}}\big(1 + 2m\sup_{y\in\mathbb{R}}\big|\tfrac{\nabla g(y)}{g(y)}\big|\big),\end{aligned} \tag{7.19}$$

where the last inequality follows since the density of B_δ is bounded by $(2\pi\delta)^{-1/2}$. Therefore, by a union bound and applying (7.19) with $z = y_1 - |y_1 - y_2|$ and $\delta = |y_1 - y_2| + |x_1 - x_2| + C|t_1 - t_2|^{1/3}$, if $t_2 \geq |y_1 - y_2| + |x_1 - x_2| + C|t_1 - t_2|^{1/3}$ then

$$\begin{aligned}&\mathbb{P}_{x_2}\left(Z_{t_2} \leq y_1 + |x_1 - x_2| + C|t_1 - t_2|^{1/3}\right)\\ &\leq \mathbb{P}_{x_2}\left(Z_{t_2} \leq y_2\right) + (2\pi m)^{-1/2}(|y_1 - y_2| + |x_1 - x_2| + C|t_1 - t_2|^{1/3})^{1/2}\big(1 + 2m\sup_{y\in\mathbb{R}}\big|\tfrac{\nabla g(y)}{g(y)}\big|\big).\end{aligned} \tag{7.20}$$

Hence by combining (7.16), (7.17), (7.18) and (7.20), we have that for $t_* > 0$, there exists $C_* = C_*(t_*) > 0$ such that for $t_1, t_2 \geq t_*$ with $|t_1 - t_2| \leq 1$ and $x_1, x_2, y_1, y_2 \in \mathbb{R}$,

$$\mathbb{P}_{x_1}\left(Z_{t_1} \leq y_1\right) \leq \mathbb{P}_{x_2}\left(Z_{t_2} \leq y_2\right) + C_*(|x_1 - x_2|^{1/2} + |y_1 - y_2|^{1/2} + |t_1 - t_2|^{1/6}).$$

By bounding $\mathbb{P}_{x_2}\left(Z_{t_2} \leq y_2\right)$ in the same way, the claim (7.15) follows.

We now use the claim to prove the result. First take $K > 0$ sufficiently large that for any $x \in [-K_0, K_0]$ we have

$$\mathbb{P}_x\left(|Z_{t_0}| > K\right) < \tfrac{1}{2}\delta.$$

Then note that it suffices to prove that for $y_0 \in [-K, K]$,

$$\left|\mathbb{P}\left(\zeta_{t_0}^{n,T_n}(X_0, J_0) - \mu_{T_n - t_0}^n \leq y_0\right) - \mathbb{P}_{X_0 - \mu_{T_n}^n}\left(Z_{t_0} \leq y_0\right)\right| < \tfrac{1}{2}\delta.$$

For $t \in [0, T_n]$, let $\mathcal{F}_t = \sigma(\mathcal{F}, \sigma((\zeta_s^{n,T_n}(X_0, J_0))_{s \le t}))$. Let $S_n = T_n - \delta_n \lfloor \delta_n^{-1} t_0 \rfloor$. Condition on $\mathcal{F}_0$, and suppose the event $E_1' \cap E_2' \cap E_4$ occurs, so in particular (7.1) holds. Fix $x_0 \in [-K-1, K+1]$ and $\epsilon > 0$, define v_0 as in (7.2) in the proof of Theorem 1.1, and let $\bar{v}_0$ denote the linear interpolation of v_0. Define $\tilde{v}_t(x)$ as in (7.5) in the proof of Theorem 1.1. Then by the same argument as for (7.7) and (7.8) in the proof of Theorem 1.1, for n sufficiently large, if for all $x \in \frac{1}{n}\mathbb{Z} \cap [-N^5, N^5]$ we have

$$q^{n,-}_{S_n,T_n}(\mu^n_{S_n} + x_0, x) \le \tilde{v}^n_{T_n}(x) + \left(\frac{n}{N}\right)^{1/4} \quad \text{and} \quad q^{n,-}_{S_n,T_n}(\mu^n_{S_n} + x_0 + \epsilon, x) \ge \tilde{v}^n_{T_n}(x) - \left(\frac{n}{N}\right)^{1/4},$$

then (7.7) and (7.8) hold for all $x \in \frac{1}{n}\mathbb{Z}$ with $|x - \mu^n_{T_n}| \le K_0$. By the definition of v_0 in (7.2) and since (7.1) holds, we have $v_0(y + \mu^n_{S_n}) \le (g(y) + e^{-(\log N)^{c_2}})\mathbb{1}_{y \le x_0 + \epsilon}$ $\forall y \in \frac{1}{n}\mathbb{Z}$, and so for n sufficiently large, using (7.7) we have

$$\begin{aligned}
&\frac{q^{n,-}_{S_n,T_n}(\mu^n_{S_n} + x_0, X_0)}{g(X_0 - \mu^n_{S_n} - \nu(T_n - S_n))} \\
&\le \mathbb{E}_{X_0 - \mu^n_{S_n} - \nu(T_n - S_n)}\left[(1 + \mathcal{O}(n^{-1}) + e^{-(\log N)^{c_2}} g(x_0 + \epsilon)^{-1})\mathbb{1}_{Z_{T_n - S_n} \le x_0 + \epsilon}\right] \\
&\qquad\qquad + e^{-\frac{1}{2}(\log N)^{c_2}} + \left(\frac{n}{N}\right)^{1/4} g(2K_0)^{-1} \\
&\le \mathbb{P}_{X_0 - \mu^n_{T_n}}(Z_{t_0} \le x_0) + C_*(t_0/2)\epsilon^{1/2} + \epsilon,
\end{aligned}$$

where the second inequality follows for n sufficiently large by (7.15) and since we have $|x_0| \le K + 1$, $|T_n - S_n - t_0| \le \delta_n$, and since (by the definition of the event E_1 in (2.10)) we have $|\mu^n_{S_n} + \nu(T_n - S_n) - \mu^n_{T_n}| \le (t_0 + 1)e^{-(\log N)^{c_2}}$. By the same argument, using (7.8), we have that for n sufficiently large,

$$\frac{q^{n,-}_{S_n,T_n}(\mu^n_{S_n} + x_0 + \epsilon, X_0)}{g(X_0 - \mu^n_{S_n} - \nu(T_n - S_n))} \ge \mathbb{P}_{X_0 - \mu^n_{T_n}}(Z_{t_0} \le x_0) - C_*(t_0/2)\epsilon^{1/2} - \epsilon.$$

The result now follows by exactly the same argument as in the proof of Theorem 1.1 from (7.9) and (7.10). □

8 Glossary

Here we list frequently used notation. In the second column of the table we give a brief heuristic description, and in the third column we refer to the section or equation where the notation is defined.

Notation	**Meaning**	**Defn./Sect.**
$\xi^n_t(x,i)$	type of ith individual at site x at time t	Section 1.1
$p^n_t(x)$	proportion of type A at site x at time t	Section 1.1
s_n	selection parameter	(1.11)
r_n	time scaling parameter	(1.11)
$(\mathcal{P}^{x,i,j}_t)_{t\ge 0}$	Poisson process corresponding to neutral reproduction events	Section 1.1
$(\mathcal{S}^{x,i,j}_t)_{t\ge 0}$	Poisson process corresponding to selective reproduction events giving an advantage to type A	Section 1.1
$(\mathcal{Q}^{x,i,j,k}_t)_{t\ge 0}$	Poisson process corresponding to selective reproduction events giving an advantage to the majority type	Section 1.1

$(\mathcal{R}_t^{x,i,y,j})_{t\geq 0}$	Poisson process corresponding to migration events	Section 1.1
$(\zeta_t^{n,T}(x,i), \theta_t^{n,T}(x,i))$	site and label of time-$(T-t)$ ancestor of ith individual at site x at time T	Section 1.1
g	travelling wave profile	(1.12)
μ_t^n	position of random travelling front at time t	(1.13)
$G_{R,t}$	set of (sites and labels of) type A individuals within distance R of the front at time t	(1.14)
π	density of stationary distribution for diffusion (1.6)	(1.15)
κ, ν	constants	(2.1)
$q_{t_1,t_2}^n(x_1, x_2)$	proportion of individuals at x_2 at time t_2 which are type A and whose time-t_1 ancestor was at x_1	(2.2)
$q_{t_1,t_2}^{n,+}(x_1, x_2)$ $(q_{t_1,t_2}^{n,-}(x_1, x_2))$	proportion of individuals at x_2 at time t_2 which are type A and whose time-t_1 ancestor was $\geq$ $(\leq)$ x_1	(2.3)
C	large constant	Section 2
$\delta_n, \epsilon_n, \gamma_n, d_n$	deterministic quantities depending on n	(2.4)
$\mathcal{C}_t^n(x_1, x_2, \ldots, x_\ell)$	set of ℓ-tuples of distinct type A individuals at $x_1, \ldots, x_\ell$ at time $t+\delta_n$ with common ancestor at time t	(2.5)
$r_{s,t}^{n,y,\ell}(x)$	proportion of individuals at x at time t which are type A and whose ancestors stayed distance y ahead of the front for time s	(2.6)
T_n	time at which sample of type A individuals is taken	Section 2
$\mathcal{F}$	σ-algebra generated by tracer random variables	(2.7)
$A_t^{(j)}(x_1, x_2)$, $A_t^{(j')}(x)$	'good' events that control the motion of a single ancestral lineage	Section 2
$B_t^{(j)}(x)$	'good' events that control the probability that a pair (or triple) of lineages coalesce in a time interval of length δ_n	Section 2
D_n^+, D_n^-	w.h.p., a pair of lineages in the sample are never both more than D_n^+ ahead of the front (before they coalesce), and no lineage is $\|D_n^-\|$ behind the front	(2.8)
$I_t^n, I_t^{n,\epsilon}, i_t^n$	intervals around the front location at time t	(2.9)
E_1	'good' event that says $p_t^n(\cdot) \approx g(\cdot - \mu_t^n)$ and $\mu_{t+s}^n - \mu_t^n \approx \nu s$	(2.10)
T_n^-	$T_n^- = T_n - (\log N)^2$	Section 2
E_2, E_2'	'good' events defined as an intersection of $A_t^{(j)}(x_1, x_2)$ and $A_t^{(j')}(x)$ events	(2.10), (2.11)
E_3	'good' event defined as an intersection of $B_t^{(j)}(x)$ events	(2.12)
E_4	'good' event that says (conditional on $\mathcal{F}$) w.h.p., no lineages stay far ahead of the front for a long time	(2.13)
E	$E = \cap_{j=1}^4 E_j$	Section 2

$\zeta_t^{n,i}$ ($\tilde{\zeta}_t^{n,i}$)	site (location relative to the front) of ith ancestral lineage in the sample at time $T_n - t$	(2.15)
$\tau_{i,j}^n$	time (backwards in time from T_n) when ith and jth ancestral lineages coalesce	Section 2
$\mathcal{F}_t$	σ-algebra generated by $\mathcal{F}$ and ancestral lineages in sample up to time t (backwards in time)	(2.16)
t_k	$t_k = k\lfloor (\log N)^C \rfloor$	Section 2
$\tilde{\tau}_{i,j}^n$	coalescence time $\tau_{i,j}^n$ if coalescence happens fairly near the front and not too soon after a time t_k	(2.17)
β_n	approximate probability that a given pair of lineages coalesce in a time interval of length t_1	(2.18)
∇_n	$\nabla_n h(x) = n(h(x+n^{-1}) - h(x))$	Section 3
Δ_n	$\Delta_n h(x) = n^2(h(x+n^{-1}) - 2h(x) + h(x-n^{-1}))$	Section 3
f	$f(u) = u(1-u)(2u-1+\alpha)$	(3.1)
$\langle \cdot, \cdot \rangle_n$	$\langle f_1, f_2 \rangle_n = n^{-1} \sum_{w \in \frac{1}{n}\mathbb{Z}} f_1(w) f_2(w)$	Section 3
$(X_t^n)_{t \geq 0}$	continuous-time SSRW on $\frac{1}{n}\mathbb{Z}$, jump rate n^2	Section 3
$\mathbf{P}_z$, $\mathbf{E}_z$	$\mathbf{P}_z(\cdot) := \mathbb{P}\left(\cdot \,\middle|\, X_0^n = z\right)$, $\mathbf{E}_z[\cdot] := \mathbb{E}\left[\cdot \,\middle|\, X_0^n = z\right]$	Section 3
$\phi_s^{t,z}, \phi_s^{t,z,a}$	rescaled transition probabilities for X^n	(3.4), (3.5)
$(u_t^n)_{t \geq 0}$	solution of system of ODEs, discrete approximation of (1.16)	(3.6)
$\eta_t^n(x,i)$	indicator function of the event that the ith individual at x at time t is descended from an individual in $\mathcal{I}_0$ at time 0	(3.9)
$q_t^n(x)$	proportion of individuals at x at time t descended from $\mathcal{I}_0$ at time 0	(3.10)
$(v_t^n)_{t \geq 0}$	solution of system of ODEs; $q_t^n \approx v_t^n$ w.h.p.	(3.11)

A Proof of Proposition 3.5

Proof of Proposition 3.5. By rescaling time and space, we can assume $m = 2$ and $s_0 = 1$. In this proof, we use the notation and refer to results from [16]. The only change required in the proof is in Section 5, where we need to control $\sup_z |h(z,t)|$ at large times t.

Take $\delta > 0$ and suppose $|\varphi(z) - U(z)| \leq \delta \;\forall z \in \mathbb{R}$. Then by Lemma 4.2, for some constant C_0, if δ is sufficiently small then $|u(x+ct,t) - U(x)| \leq C_0\delta \;\forall x \in \mathbb{R}, t > 0$. Therefore, by Lemma 4.5, there exists $z_0 \in \mathbb{R}$ such that $\lim_{t\to\infty} \sup_{x\in\mathbb{R}} |u(x+ct,t) - U(x-z_0)| = 0$ and so $\sup_{x\in\mathbb{R}} |U(x) - U(x-z_0)| \leq C_0\delta$. It follows that

$$|u(x+ct,t) - U(x-z_0)| \leq 2C_0\delta \quad \forall x \in \mathbb{R},\ t > 0.$$

Hence by the definition of $w(z,t)$ in the proof of Lemma 4.5, and by the estimates in Lemma 4.3, for t sufficiently large (depending on δ),

$$|w(z,t) - U(z-z_0)| \leq 3C_0\delta \quad \forall z \in \mathbb{R}. \tag{A.1}$$

By the definition of $\alpha(t)$ in (5.1), for t sufficiently large (depending on δ), it follows that

$$\begin{aligned} 0 &= \int_{-\infty}^{\infty} e^{cz} h(z,t) U'(z - z_0 - \alpha(t)) dz \\ &\geq \int_{-\infty}^{\infty} e^{cz} U'(z - z_0 - \alpha(t))(U(z-z_0) - 3C_0\delta - U(z - z_0 - \alpha(t)) dz. \end{aligned}$$

There exists a constant $a > 0$ such that if $\alpha(t) \geq \delta^{1/2}$ and if δ is sufficiently small then

$$\int_{z_0+\alpha(t)-\delta^{1/2}}^{z_0+\alpha(t)} e^{cz} U'(z - z_0 - \alpha(t))(U(z - z_0) - 3C_0\delta - U(z - z_0 - \alpha(t))dz \\ \geq a\delta e^{c(z_0+\alpha(t))}.$$

For $R < \infty$, if δ is sufficiently small and $\alpha(t) \geq \delta^{1/2}$ then for $z \in \mathbb{R}$ with $|z-(z_0+\alpha(t))| \leq R$ we have $U(z - z_0) - U(z - z_0 - \alpha(t)) \geq 3C_0\delta$. Therefore

$$0 \geq a\delta e^{c(z_0+\alpha(t))} \\ - 3C_0\delta\Big(\int_{z_0+\alpha(t)+R}^{\infty} e^{cz} U'(z - z_0 - \alpha(t))dz + \int_{-\infty}^{z_0+\alpha(t)-R} e^{cz} U'(z - z_0 - \alpha(t))dz\Big),$$

which, by the tail behaviour of U', is a contradiction for R sufficiently large. By the same argument for the case $\alpha(t) \leq -\delta^{1/2}$, it follows that if δ is sufficiently small, $|\alpha(t)| \leq \delta^{1/2}$ for t sufficiently large (depending on δ).

Hence by (A.1), for $b > 0$, if δ is sufficiently small then for t sufficiently large (depending on δ and b), $\sup_z |h(z, t)| \leq b$. Therefore, if δ is sufficiently small then the inequality

$$\frac{1}{2}\frac{d}{dt}\|y\|^2 \leq -\frac{M}{2}\|y\|^2 + \mathcal{O}(e^{-Kt})$$

(which appears before (5.3)) holds for $t \geq T$, where $T = T(\delta)$ and $K = K(\delta)$.

This is the only modification required in the proof. □

References

[1] J W Arntzen, *An amphibian species pushed out of Britain by a moving hybrid zone*, Molecular Ecology **28** (2019), no. 23, 5145–5154.

[2] N H Barton, *The dynamics of hybrid zones*, Heredity **43** (1979), no. 3, 341–359.

[3] N H Barton, A M Etheridge, and A K Sturm, *Coalescence in a random background*, Ann. Appl. Probab. **14** (2004), no. 2, 754–785. MR2052901

[4] N H Barton and G M Hewitt, *Adaptation, speciation and hybrid zones*, Nature **341** (1989), no. 6242, 497–503.

[5] J Berestycki, N Berestycki, and J Schweinsberg, *The genealogy of branching Brownian motion with absorption*, Ann. Probab. **41** (2013), no. 2, 527–618. MR3077519

[6] G Birzu, O Hallatschek, and K Korolev, *Fluctuations uncover a distinct class of traveling waves*, Proc. Nat. Acad. Sci. U.S.A. **115** (2018), no. 6, E3645–E3654. MR3796428

[7] N Biswas, A Etheridge, and A Klimek, *The spatial Lambda-Fleming-Viot process with fluctuating selection*, Electronic Journal of Probability **26** (2021). MR4235476

[8] M Bramson, *Convergence of solutions of the Kolmogorov equation to travelling waves*, vol. 285, American Mathematical Soc., 1983. MR0705746

[9] É Brunet and B Derrida, *Shift in the velocity of a front due to a cutoff*, Phys. Rev. E **56** (1997), no. 3, 2597–2604. MR1473413

[10] É Brunet and B Derrida, *Effect of microscopic noise on front propagation*, J Statist. Phys. **103** (2001), no. 1-2, 269–282. MR1828730

[11] É Brunet, B Derrida, A H Mueller, and S Munier, *Noisy travelling waves: effect of selection on genealogies*, Europhys. Lett. **76** (2006), no. 1, 1–7. MR2299937

[12] P J Donnelly and T G Kurtz, *Genealogical processes for Fleming-Viot models with selection and recombination*, Ann. Appl. Probab. **9** (1999), no. 4, 1091–1148. MR1728556

[13] R Durrett, *Stochastic calculus: a practical introduction*, vol. 6, CRC Press, 1996. MR1398879

[14] R Durrett and W-T Fan, *Genealogies in expanding populations*, The Annals of Applied Probability **26** (2016), no. 6, 3456–3490. MR3582808

[15] L C Evans, *Partial differential equations*, American Mathematical Soc., 2010. MR2597943

[16] P C Fife and J B McLeod, *The approach of solutions of nonlinear diffusion equations to travelling front solutions*, Archive for Rational Mechanics and Analysis **65** (1977), no. 4, 335–361. MR0442480

[17] R A Fisher, *The wave of advance of advantageous genes*, Ann. Eugenics **7** (1937), no. 4, 355–369.

[18] G B Folland, *Introduction to partial differential equations*, Princeton University Press, 1995. MR1357411

[19] J Garnier, T Giletti, F Hamel, and L Roques, *Inside dynamics of pulled and pushed fronts*, Journal de Mathématiques Pures et Appliquées **98** (2012), no. 4, 428–449. MR2968163

[20] O Hallatschek and D R Nelson, *Gene surfing in expanding populations*, Theoretical Population Biology **73** (2008), no. 1, 158–170. MR0602140

[21] S Karlin and H E Taylor, *A second course in stochastic processes*, Elsevier, 1981. MR0611513

[22] M Kimura, *Stepping stone model of population*, Ann. Rep. Nat. Inst. Genetics Japan **3** (1953), 62–63.

[23] A Kolmogorov, I Petrovsky, and N Piscounov, *Étude de l'equation de la diffusion avec croissance de la quantité de matière et son application à un problème biologique*, Moscow Univ. Math. Bull. **1** (1937), 1–25.

[24] G F Lawler and V Limic, *Random walk: a modern introduction*, vol. 123, Cambridge University Press, 2010. MR2677157

[25] C McDiarmid, *Concentration*, Probabilistic methods for algorithmic discrete mathematics, Springer, 1998, pp. 195–248. MR1678578

[26] C Mueller, L Mytnik, and J Quastel, *Effect of noise on front propagation in reaction-diffusion equations of KPP type*, Inv. Math. **184** (2011), no. 2, 405–453. MR2793860

[27] C Mueller and R Sowers, *Random travelling waves for the KPP equation with noise*, J. Functional Anal. **128** (1995), no. 2, 439–498. MR1319963

[28] C Mueller and R Tribe, *Stochastic pde's arising from the long range contact and long range voter processes*, Probability Theory and Related Fields **102** (1995), no. 4, 519–545. MR1346264

[29] M Nordborg and S M Krone, *Separation of timescales and convergence to the coalescent in structured populations*, Modern developments in theoretical population genetics: the legacy of Gustave Malécot (M Slatkin and M Veuille, eds.), Oxford University Press, 2002.

[30] L Roques, J Garnier, F Hamel, and E K Klein, *Allee effect promotes diversity in traveling waves of colonization*, Proceedings of the National Academy of Sciences **109** (2012), no. 23, 8828–8833. MR3033229

[31] F Rothe, *Convergence to pushed fronts*, The Rocky Mountain Journal of Mathematics **11** (1981), 617–633. MR0639447

[32] A N Stokes, *On two types of moving front in quasilinear diffusion*, Math. Biosci. **31** (1976), 307–315. MR0682241

[33] J Tourniaire, *A branching particle system as a model of semi pushed fronts*, arXiv preprint arXiv:2111.00096 (2021).

[34] K Uchiyama, *The behavior of solutions of some non-linear diffusion equations for large time*, Journal of Mathematics of Kyoto University **18** (1978), no. 3, 453–508. MR0509494

[35] W van Saarloos, *Front propagation into unstable states*, Phys. Rep. **386** (2003), no. 2-6, 29–222.