



HAL
open science

Some aspects of the Fisher-KPP equation and the branching Brownian motion

Éric Brunet

► **To cite this version:**

Éric Brunet. Some aspects of the Fisher-KPP equation and the branching Brownian motion. Statistical Mechanics [cond-mat.stat-mech]. UPMC, 2016. tel-01417420

HAL Id: tel-01417420

<https://theses.hal.science/tel-01417420>

Submitted on 15 Dec 2016

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Copyright

Some aspects of the Fisher-KPP equation and the branching Brownian motion

HDR dissertation by Éric Brunet, June 2016

Defended on November the 3rd, 2016 at the ENS

Examiners

M. Jean Bérard	Université de Strasbourg
M. Giulio Biroli	CEA Saclay
M. Bernard Derrida	Collège de France
M. Victor Dotsenko	UPMC
M. Robi Peschanski	CEA Saclay
M. Zhan Shi	UPMC

Reviewers

M. Jean Bérard	Université de Strasbourg
M. Oskar Hallatschek	University of California, Berkeley
M. Robi Peschanski	CEA Saclay

Foreword

The first part of this memoir is the dissertation itself: after a chapter of introduction on the FKPP equation and the BBM, I present three independent sets of results in the three following chapters. I have selected six of my papers (two per subject) which I have reproduced in the second part of this memoir. When cited, the papers which have been reproduced are marked with an asterisk, as in [★BD15].

I have tried to write a self-contained document presenting in a consistent way several of the results I have obtained, and to do this I had to make a selection: not all the subjects I have been working on in the last years are presented in this memoir. For instance, I do not discuss my work on the accessibility percolation on the hypercube [BBS16; BBS14] (given the L -hypercube, assign random uniform numbers on $[0, 1]$ on the 2^L corners; is there a path along the edges of the hypercube, from the origin to the highest numbered corner, such that the numbers on the visited corners follow an increasing sequence? The $L \rightarrow \infty$ answer is yes if and only if the origin has a number smaller than $1 - \frac{1}{2} \operatorname{argsinh} 2$). I do not discuss either my contribution to estimate the speed of adaptation of a population under selection [RBW08; BRW08], nor my work [BBHHR15] on a BBM where the reproduction rate at position x is $|x|^p$ (the expected number of particles and the almost sure number of particles at time t both scale like $\exp[At^{(2+p)/(2-p)}]$, but the value of A is not the same for the “expected” and “almost-sure” cases.)

The literature on the FKPP and the BBM is huge, both in the physics and mathematics communities. I did not write a full review of this literature, and I simply cited the papers which, I felt, would give some useful context to my presentation.

I want to thank my collaborators for all the stimulating discussions we had while writing papers, and I give my love to my family.

Éric Brunet

Contents

Foreword	1
I Some aspects on the FKPP equation and the branching Brownian motion	5
1 Introduction	6
1.1 The BBM and the FKPP equation	6
1.1.1 The BBM	6
1.1.2 The FKPP equation	7
1.1.3 Duality	7
1.2 Basic properties: travelling waves and velocity selection	8
1.2.1 Bramson's result	8
1.2.2 Velocity selection	9
1.3 Universality	10
1.3.1 A branching random walk (BRW)	10
1.3.2 Other fronts in the FKPP class	11
1.4 Link with disordered systems	13
1.4.1 Spin glasses	13
1.4.2 Directed polymers in random medium	14
1.5 The necessity of adding a noise term to the FKPP equation	14
1.5.1 Reaction-diffusion	14
1.5.2 Directed polymers	15
1.5.3 Models of population dynamics	16
1.5.4 Universality of internal noise	17
1.5.5 The stochastic FKPP equation, the coalescing BBM and duality	17
1.5.6 The noisy FKPP equation in QCD	18
1.6 My contribution to the study of the BBM and the FKPP equation	18
2 Vanishing corrections for the position of the FKPP front	19
2.1 The Bramson term and the Ebert and van Saarloos term	19
2.1.1 The Bramson term	19
2.1.2 The Ebert and van Saarloos term	20
2.2 Understanding the Bramson term and the Ebert and van Saarloos term	21
2.2.1 The linear FKPP equation with an anchor	21
2.2.2 A model on the lattice	23
2.3 Summary, and a conjecture for the next order term	24
3 The limiting distribution of extremal points in the BBM	25
3.1 The limiting distribution seen from μ_t	25
3.1.1 Use and abuse of McKean's formula	25
3.1.2 The large time limit	26
3.2 Non-ergodicity and random shift	28
3.2.1 Lalley's and Sellke's result	28
3.2.2 Generalization to the extremal process	29
3.3 The structure of the limiting point process	30
3.3.1 Stability by juxtaposition and decorated Poisson point processes	30
3.3.2 Families and leaders	31

4	Position and genealogy in the noisy FKPP equation	33
4.1	The position of the front	34
4.1.1	The cut-off theory	34
4.1.2	Beyond the cut-off theory; a phenomenological description	36
4.1.3	Justifications for the guess (4.44)	40
4.2	Genealogies	45
4.2.1	Genealogies in models without selection	45
4.2.2	Genealogies in models with selection	46
4.2.3	Λ -coalescents	46
4.2.4	Phenomenological theory for the genealogical trees	47
4.3	Some other models around the noisy FKPP class	49
5	Conclusion	53
II	Facsimile of selected publications	54
[BDMM06a]:	Phenomenological theory giving the full statistics of the position of fluctuating pulled fronts	55
[BDMM07]:	Effect of selection on ancestry: an exactly soluble case and its phenomenological generalization	64
[BD09]:	Statistics at the tip of a branching random walk and the delay of traveling waves	84
[BD11]:	A branching random walk seen from the tip	89
[BD15]:	An Exactly Solvable Travelling Wave Equation in the Fisher-KPP Class	116
[BBHR15]:	Vanishing corrections for the position in a linear model of FKPP fronts	136
	Bibliography	166

Part I

Some aspects on the Fisher-KPP equation and the branching Brownian motion

Chapter 1

Introduction

1.1 The BBM and the FKPP equation

The main objects of this memoir are introduced: the BBM is a branching process, the FKPP a front equation. A wonderful relationship links one to the other.

1.1.1 The BBM

The branching Brownian motion, or BBM, was introduced in [INW68]. It describes a system where particles diffuse and branch independently from each other. Diffusion is assumed to be Brownian: the infinitesimal displacement dX of a particle during the infinitesimal time dt is Gaussian with $\langle dX \rangle = 0$ and $\langle dX^2 \rangle = 2 dt$. Branching is assumed to be Poisson: during each dt , a particle has a probability dt of being replaced by two particles at the same position which then move off independently, repeating the behaviour of the parent. One starts with a single particle at the origin to obtain something looking like Figure 1.1.

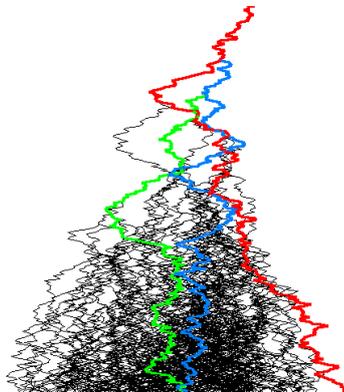


Figure 1.1: A BBM. Time increases downwards, and space is horizontal. At time $t = 0$ there is a single (red) particle, which branches after some time to create a blue particle. The blue branches to create a green, then the red branches again and then it becomes hard to keep track as more and more branching events occur.

The BBM can be seen as a model of unlimited growth and diffusion: a new species colonizes a rich environment with its population diffusing and growing exponentially fast:

$$A \rightarrow 2A. \tag{1.1}$$

If one calls $\rho(x, t)$ the expected density of particles (meaning that $\rho(x, t) dx$ is the expected number of particles in dx), then it is easy to see that ρ follows the heat equation with a linear growth term:

$$\partial_t \rho = \partial_x^2 \rho + \rho. \tag{1.2}$$

The initial condition being a delta at 0, one gets

$$\rho(x, t) = \frac{1}{\sqrt{4\pi t}} e^{t - \frac{x^2}{4t}}. \tag{1.3}$$

For large times, the expected density $\rho(x, t)$ is equal to 1 at positions $\pm x_t$ with $x_t = 2t - \frac{1}{2} \ln t + \mathcal{O}(1)$. This suggests that the frontiers of the BBM move asymptotically with a velocity 2, which turns out to be true. The logarithmic correction obtained from (1.3) is however incorrect, as explained in Section 1.2.

1.1.2 The FKPP equation

The Fisher [Fis37] and Kolmogorov, Petrovski, Piscounov [KPP37] equation (FKPP) is the heat equation with a non-linear growth term:

$$\partial_t h = \partial_x^2 h + h - h^2. \quad (1.4)$$

It can be seen as the deterministic limit of a reaction-diffusion process of type



Indeed, consider a medium homogeneously filled with particles that come into two types, A and B . Call $h(x, t) \in [0, 1]$ the local proportion of particles of type A around position x at time t . Then $1 - h(x, t)$ is the local proportion of particles of type B . The expected local increase in the number of A particles around x due to the $A + B \rightarrow 2A$ reaction is then proportional to $h(1 - h)$. Adding diffusion leads to the FKPP equation (1.4).

Both the BBM and the FKPP equation can be interpreted as models of population dynamics in a one-dimensional environment, but they are very different: in the BBM, there is no saturation, no limit to the number of A particles which reproduce indefinitely, while in the FKPP equation, the A particles simply replace the B particles and, therefore, their density remains bounded. Furthermore, the BBM is a stochastic model, while all fluctuations of the microscopic model (1.5) have been neglected to obtain the FKPP equation.

1.1.3 Duality

Even though the BBM and the FKPP equation look very different, there is a very remarkable relationship between the two models [McK75]. If one calls R_t the position of the rightmost particle in a BBM started from a single particle at the origin, and if one considers the solution $h(x, t)$ to the FKPP equation with an initial condition $h_0(x) = \mathbb{1}_{\{x < 0\}}$, then one has

$$h(x, t) = \mathbb{P}(R_t > x) \quad \text{if } h_0(x) = \mathbb{1}_{\{x < 0\}}. \quad (1.6)$$

In particular, the median position of the rightmost particle in the BBM at any given time t is equal to the position x such that $h(x, t) = \frac{1}{2}$.

McKean [McK75] wrote a more general relationship; assume that one starts the BBM with a single particle at the origin, let \mathcal{N}_t be the set of particles alive at time t and, for each $u \in \mathcal{N}_t$, call X_u the position of particle u . Assume also that one solves the FKPP equation for a given initial condition $h_0(x)$. Then

$$\left\langle \prod_{u \in \mathcal{N}_t} (1 - h_0(x - X_u)) \right\rangle = 1 - h(x, t). \quad (1.7)$$

In particular, for $h_0(x) = \mathbb{1}_{\{x < 0\}}$, the random variable being averaged on the left hand side is 1 if no particle is on the right of x and 0 otherwise. The left hand side is therefore $\mathbb{P}(R_t \leq x)$ and one recovers (1.6).

The incredibly useful relation (1.7) can be seen as a generalization to branching processes of the Feynman-Kac theorem. It can be understood in a couple of lines: let $g(x, t)$ be the left hand side of (1.7); one computes $g(x, t + dt)$ by looking at what happens during the first dt of the history of the system. During this first dt , there is a probability $1 - dt$ that the initial particle does not branch but simply moves by some random amount dX . The state of the BBM at time $t + dt$ is then the same as the state of another BBM at time t started from one single particle at dX . On the other hand, the first particle does branch during the initial dt with a probability dt . (The particles also move a little bit, but that is a negligible correction.) The state of the BBM at time $t + dt$ is then the same as the juxtaposition of two independent BBM at time t . Ignoring all terms smaller than dt , one concludes that

$$\begin{aligned} g(x, t + dt) &= \overbrace{(1 - dt) \langle g(x - dX, t) \rangle}^{\text{no branching during initial } dt} + \overbrace{dt \langle g(x, t) \rangle^2}^{\text{branching during initial } dt}, \\ &= g(x, t) + dt \left[\partial_x^2 g(x, t) + g(x, t)^2 - g(x, t) \right], \end{aligned} \quad (1.8)$$

where we expanded $g(x - dX, t)$ and used $\langle dX \rangle = 0$ and $\langle dX^2 \rangle = 2 dt$. One can see, then, that $h(x, t) = 1 - g(x, t)$ follows the FKPP equation.

1.2 Basic properties: travelling waves and velocity selection

For a steep initial condition, the solution to the FKPP converges to a front moving at velocity 2 with a shape ω , where $\omega(x) \sim Axe^{-x}$ for large x . When the initial condition decays like $e^{-\gamma x}$ with $\gamma < 1$, the solution to the FKPP converges to another front moving at velocity $v = \gamma + \gamma^{-1} > 2$.

1.2.1 Bramson's result

Recall that from the point of view of population dynamics, the solution $h(x, t)$ to the FKPP equation represents the local proportion of A in a $A + B \rightarrow 2A$ reaction-diffusion process. $h(x, t) = 0$ (only B particles) is a solution, but an unstable one with respect to the introduction of A particles. On the other hand, $h(x, t) = 1$ (only A particles) is a stable solution. Generally, we only consider initial conditions such that $0 \leq h_0(x) \leq 1$ because values outside that range do not make sense from the point of view of population dynamics. Then, at any time one still has $0 \leq h(x, t) \leq 1$.

Situations of interest include starting with a sea of B particles with a small number of A particles around the origin (a new and fitter species is introduced). It is clear that the number of A particles quickly increases and saturates around the origin, and then the A invade the medium leftwards and rightwards. In a growing region around the origin, $h(x, t)$ is nearly equal to 1 (only A particles) whereas it is nearly 0 very far away from the origin. At some time-dependent positions, there are two fronts (one for positive x moving to the right, and one for negative x moving to the left) where $h(x, t)$ is neither close to 0 nor to 1; see Figure 1.2.

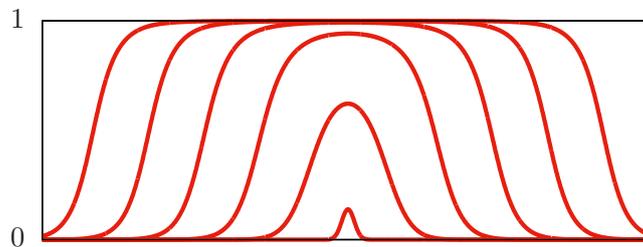


Figure 1.2: A small number of particles A invade a sea of B particles. The shape of the FKPP front is sketched at several times.

A simpler setting, which we adopt from now, is to consider only one invasion front by assuming that $h(x, t) \approx 1$ for large negative x and $h(x, t) \approx 0$ for large positive x : there are mostly A particles on the left, and mostly B particles on the right. It is clear that the A particles on the left invade the medium on the right, and one expects the appearance of an invasion front around some time-dependent position m_t , such that $h(x, t)$ is nearly 1 if x is sufficiently smaller than m_t , and $h(x, t)$ is nearly 0 if x is sufficiently larger than m_t .

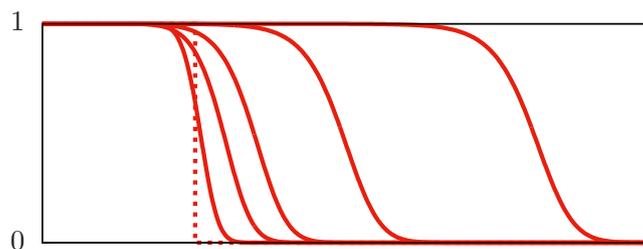


Figure 1.3: The A particles on the left invade the B particles on the right: starting from a step initial condition (drawn with dashed lines), the solution to the FKPP equation is drawn at times 1, 3, 5, 10 and 20.

In particular, if one starts the FKPP equation with a step initial condition $h_0(x) = \mathbb{1}_{\{x < 0\}}$ (only A on the left, and only B on the right), one obtains the shapes $h(x, t)$ drawn in Figure 1.3. The solution $h(x, t)$ develops into a moving front, meaning that there exists a non-trivial final shape $\omega(x)$ and a centring term m_t such that

$$h(m_t + z, t) \rightarrow \omega(z) \quad \text{as } t \rightarrow \infty. \quad (1.9)$$

The final shape is uniquely defined up to some translation by (1.9). On the other hand, the centring term m_t is not, and several choices are possible; a usual choice for m_t is to impose that $h(m_t, t) = \frac{1}{2}$. But whatever the choice, for any m_t verifying (1.9), it is known since Bramson [Bra78; Bra83] (see also [Rob13]) that

$$m_t = 2t - \frac{3}{2} \ln t + C + o(1) \quad \text{as } t \rightarrow \infty, \quad (1.10)$$

and that

$$\omega(z) \sim Aze^{-z} \quad \text{for large } z, \quad (1.11)$$

where the constants C and A are not known exactly and depend on the (translation) choice of $\omega(z)$. The front $h(x, t)$ moves at velocity 2.

Recall that we called R_t the position of the rightmost particle in the BBM. Using (1.6), the meaning of (1.9) and (1.10) is that $R_t - (2t - \frac{3}{2} \ln t)$ converges in distribution: the rightmost particle is around $2t - \frac{3}{2} \ln t$, with typical fluctuations of order 1, the amplitude of the typical fluctuations being the width of the travelling wave $\omega(z)$. (Note that, however, rare fluctuations will infinitely often send the rightmost particle close to position $2t - \frac{1}{2} \ln t$, at a distance $\ln t$ ahead of the typical position [HS09; Rob13].)

The result (1.9), (1.10) still holds (but with a different constant) for any non-zero initial condition $h_0(x) \in [0, 1]$ that converges “fast enough” to zero for large x . Here, fast enough means [Bra83] that

$$\int dx h_0(x) x e^x < \infty. \quad (1.12)$$

If $h_0(x)$ does not decay fast enough for (1.12) to hold, the front moves faster than in (1.10) (intuitively, there are already A particles far to the right, helping the invasion):

- If $h_0(x)$ decays like e^{-x} with some polynomial prefactor, the velocity of the front is still 2 but the logarithm term might be modified. (Some examples are given at the beginning of Chapter 2.)
- If $h_0(x)$ decays like $e^{-\gamma x}$ with some $\gamma < 1$, the velocity is $v(\gamma) > 2$.
- If $h_0(x)$ decays more slowly than any exponential, there is no velocity and the front keeps accelerating.

As long as $0 \leq h_0(x) \leq 1$, one cannot reach a velocity smaller than 2.

1.2.2 Velocity selection

One can understand heuristically this velocity selection for the FKPP front. Recall from (1.9) that $\omega(z)$ is the travelling wave moving at velocity 2. It is clear from (1.9) that $\omega(x - 2t)$ must be a solution to the FKPP equation, which implies that $\omega'' + 2\omega' + \omega - \omega^2 = 0$. More generally, if one calls ω_v the shape of a travelling wave at velocity v (which means that $\omega_v(x - vt)$ is solution to the FKPP equation), then one must have

$$\omega_v'' + v\omega_v' + \omega_v - \omega_v^2 = 0, \quad \omega_v(-\infty) = 1, \quad \omega_v(+\infty) = 0. \quad (1.13)$$

(Of course $\omega_2 = \omega$.) Far on the right, where ω_v is small, (1.13) can be linearised by neglecting the ω_v^2 term. One then obtains that, for large z ,

$$\omega_v(z) \sim \begin{cases} A_v e^{-\gamma z} & \text{if } v > 2, \text{ where } \gamma + \frac{1}{\gamma} = v, \\ A_v z e^{-z} \text{ or } A_v e^{-z} & \text{if } v = 2. \end{cases} \quad (1.14)$$

There is no solution that remains positive for $v < 2$. Notice that for any $v \geq 2$, the linear analysis gives two possible asymptotic behaviours for ω_v . There is however one unique solution up to translation to (1.13), and the (difficult!) full non-linear analysis leads to

$$\omega_v(z) \sim \begin{cases} A_v e^{-\gamma z} & \text{if } v > 2, \text{ where } \gamma < 1 \text{ is the smallest solution to } \gamma + \frac{1}{\gamma} = v, \\ A_v z e^{-z} & \text{if } v = 2, \end{cases} \quad (1.15)$$

with $A_v > 0$.

Consider now an initial condition $h_0(x)$ which decays for large x as $e^{-\gamma x}$ with $\gamma \leq 1$. Then $h_0(x)$ looks roughly like the travelling wave $\omega_v(x)$ with $v = \gamma + \gamma^{-1}$, and $h(x, t)$ eventually evolves into the front ω_v moving at velocity v : there exists a m_t such that

$$h(m_t + z, t) \rightarrow \omega_v(z) \quad \text{where} \quad \frac{m_t}{t} \rightarrow v. \quad (1.16)$$

The behaviour of $m_t - vt$ for large t then depends on the behaviour of $h_0(x)e^{\gamma x}$ for large x . For $v > 2$ (*i.e.* $\gamma < 1$), it can be understood to leading order simply by looking at the linearised equation. It is more difficult in the case $v = 2$.

If the initial condition decays for large x like $e^{-\gamma x}$ for some $\gamma > 1$, or if it decays faster than any exponential, then the initial condition is asymptotically very different from any travelling wave ω_v . What happens then is that the front evolves into the fastest decaying available travelling wave, which is the one for $v = 2$. The result is then that $h(m_t + z, t) \rightarrow \omega(z)$ with m_t as in (1.10).

1.3 Universality

The BBM can be generalized in many ways, and for each variant the McKean duality (1.7) allows to define a front equation which has properties similar to the FKPP. There are also many equations not obtained from the duality (1.7), but which describe the invasion front of a stable phase into an unstable phase and behave like the FKPP.

What makes the BBM so interesting is that its properties are shared by many branching processes. What makes FKPP so interesting is that many of the equations that describe a front $h(x, t)$ where a stable state invades an unstable state have the same properties as the FKPP equation.

The properties shared by all the fronts in the class of the FKPP equation include the following:

- There is an unstable constant solution and a stable constant solution; we always write the equation so that $h(x, t) = 0$ is unstable and $h(x, t) = 1$ is stable.
- There exist positive travelling wave solutions $h(x, t) = \omega_v(x - vt)$ if and only if $v \geq v_c$ for some critical velocity v_c . In the FKPP equation, $v_c = 2$.
- One has, for large z ,

$$\omega_{v_c}(z) \sim A_{v_c} z e^{-\gamma_c z}, \quad \omega_v(z) \sim A_v e^{-\gamma(v)z} \quad \text{for } v > v_c, \quad (1.17)$$

where $A_v > 0$ and where, for each $v \geq v_c$, the decay rate $\gamma = \gamma(v)$ is the smallest positive number such that $e^{-\gamma z}$ is solution to the equation followed by ω_v linearised around 0 (the unstable state). Naturally, $\gamma_c = \gamma(v_c)$.

- For a step initial condition, there exists a centring term m_t such that $h(m_t + z, t)$ converges as in (1.9) to the travelling wave $\omega_{v_c}(z)$. Furthermore,

$$m_t = v_c t - \frac{3}{2\gamma_c} \ln t + C + o(1) \text{ for large } t. \quad (1.18)$$

We now give several examples.

1.3.1 A branching random walk (BRW)

We consider a model in discrete time where, at each time-step, each particle is removed and replaced by two new particles with shifted positions:

$$\text{particle at position } X_u \text{ at time } t \text{ is replaced at time } t + 1 \text{ by two particles at } X_u + \epsilon_{u,1} \text{ and } X_u + \epsilon_{u,2}, \quad (1.19)$$

where the $\epsilon_{u,i}$ are identical independent random numbers drawn from some probability density function $\rho(\epsilon)$.

This can be seen as a model describing the growth of a population where, at discrete intervals, each individual has two children who wander a little and then settle down. Alternatively, the ‘‘position’’ of the particle could be in fact a measure of a given trait that changes from generation to generation due to random mutation.

Consider such a BRW started at time $t = 0$ by a single particle at the origin, choose a function $h_0(x)$ and introduce the function $h(x, t)$ as in (1.7) but only for integer t :

$$\left\langle \prod_{u \in \mathcal{N}_t} (1 - h_0(x - X_u)) \right\rangle = 1 - h(x, t), \quad (1.20)$$

where \mathcal{N}_t is the set of the 2^t particles present at time t and X_u is the position of particle u . Then

$$h(x, t + 1) = 1 - \left(1 - \int d\epsilon \rho(\epsilon) h(x - \epsilon, t) \right)^2, \quad h(x, 0) = h_0(x). \quad (1.21)$$

Indeed, let $u^{(1)}$ and $u^{(2)}$ be the particles at time 1. Given their positions, the product in (1.20) at time $t + 1$ can be decomposed into two independent products, one for the descendants of $u^{(1)}$ and one for the descendants of $u^{(2)}$. Each of these two products is the same as the product for a BRW at time t started from respectively $X_{u^{(1)}}$ or $X_{u^{(2)}}$, and one gets $1 - h(x, t + 1) = \langle 1 - h(x - X_{u^{(1)}}, t) \rangle \times \langle 1 - h(x - X_{u^{(2)}}, t) \rangle$, which is the same as (1.21).

For a well-behaved choice of the distribution $\rho(\epsilon)$ (for instance if it is uniform on some interval, or if it is Gaussian), then the equation (1.21) is in the class of the FKPP equation. In particular, the travelling waves ω_v are solution to

$$\omega_v(z - v) = 1 - \left(1 - \int d\epsilon \rho(\epsilon) \omega_v(z - \epsilon) \right)^2, \quad \omega_v(-\infty) = 1, \quad \omega_v(+\infty) = 0. \quad (1.22)$$

(compare to (1.13) for the FKPP.) For large z , when ω_v is small, one can linearise the equation

$$\omega_v(z - v) = 2 \int d\epsilon \rho(\epsilon) \omega_v(z - \epsilon) + \mathcal{O}(\omega_v^2). \quad (1.23)$$

Then, $e^{-\gamma z}$ is a solution to the linearised equation if

$$v = v(\gamma) := \frac{1}{\gamma} \ln \left[2 \int d\epsilon \rho(\epsilon) e^{\gamma \epsilon} \right]. \quad (1.24)$$

The minimum of this function $v(\gamma)$ is $v_c = v(\gamma_c)$.

To take an example, if one chooses for the random displacements $\epsilon_{u,i}$ a uniform distribution in $[0, 1]$, so that $\rho(\epsilon) = \mathbb{1}_{\{\epsilon \in [0,1]\}}$, one obtains the function $v(\gamma)$ represented in Figure 1.4 and one finds $v_c \approx 0.815172$ and $\gamma_c \approx 5.26208$. With these values, the rightmost particle in the BRW we have just defined is, for large t , typically at a distance of order 1 from m_t given in (1.18).

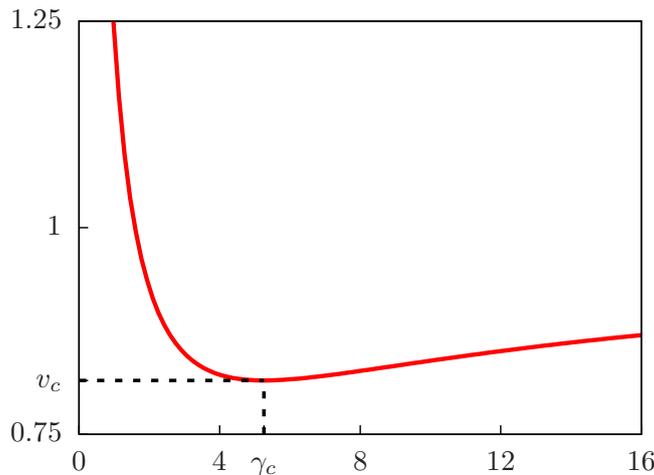


Figure 1.4: The function $v(\gamma)$ given by (1.24) with $\rho(\epsilon) = \mathbb{1}_{\{\epsilon \in [0,1]\}}$. One finds $v_c \approx 0.815172$ and $\gamma_c \approx 5.26208$.

1.3.2 Other fronts in the FKPP class

It is not difficult to build variants of the branching processes. Time can be continuous with branching events occurring at random as in the BBM, or time can be discrete with fixed generations as in the BRW. The particles might diffuse between two branching events (as in the BBM) or motion might only occur at branching events (as in the BRW). The number of children does not need to be 2; it might be any number, or even be random. Different children of a single branching event might have different laws for their displacements, and their displacements might be correlated. Under some rather not so restrictive conditions on the branching law, all of these variants lead for the quantity $h(x, t)$ defined as in (1.7) to a front equation in the FKPP class [Aïd13; BDZ14].

Of course, front equations have an existence outside of branching processes. A usual generalization of the FKPP equation is

$$\partial_t h = \partial_x^2 h + f(h), \quad (1.25)$$

for some function f (assumed to be differentiable) which meets certain necessary conditions:

- $f(0) = f(1) = 0$ so that both $h(x, t) = 0$ and $h(x, t) = 1$ are solutions,
 - $f(h) > 0$ for $h \in (0, 1)$ so that the population is growing,
 - $f'(0) > 0$ so that the solution $h(x, t) = 0$ is unstable,
 - $f'(1) < 0$ so that the solution $h(x, t) = 1$ is stable.
- (1.26)

With (1.25), one can look for travelling wave solutions $h(x, t) = \omega_v(x - vt)$ going at velocity v . They must verify

$$\omega_v'' + v\omega_v' + f(\omega_v) = 0, \quad \omega_v(-\infty) = 1, \quad \omega_v(+\infty) = 0. \quad (1.27)$$

The linearised equation for large z , in the unstable state, is

$$\omega_v'' + v\omega_v' + f'(0)\omega_v = o(\omega_v) \quad (1.28)$$

By looking for solutions of the form $\omega_v \propto e^{-\gamma z}$, one obtains the relationship between v and γ , the minimal velocity v_c and the value γ_c such that $v_c = v(\gamma_c)$:

$$v(\gamma) = \gamma + \frac{f'(0)}{\gamma}, \quad v_c = 2\sqrt{f'(0)}, \quad \gamma_c = \sqrt{f'(0)}. \quad (1.29)$$

An important question, of course, is to determine whether (1.25) is a front equation in the FKPP universality class (in which case the front moves at velocity v_c if the initial condition decays fast enough) or not (in which case the velocity is usually different).

In either case, as linear perturbations around the unstable state must move at velocity v_c (because they obey the linearised equation), the velocity of the front (1.25) cannot be smaller than v_c . The question is therefore whether the velocity of the front (for initial conditions that decay fast enough) is equal or larger than v_c .

There exists a rigorous proof [Bra83] that a sufficient condition for (1.25) to be in the FKPP class is to have (1.26) and:

$$f'(h) \leq f'(0) \text{ for all } h \in [0, 1], \quad f'(h) = f'(0) + \mathcal{O}(h^p) \text{ for some } p > 0 \text{ when } h \rightarrow 0. \quad (1.30)$$

This result is difficult to establish but, from (1.30), it is already easy to see that one has $f(h) \leq f'(0)h$ for all h , and therefore the front $h(x, t)$ must lie below the solution to the linearised equation $\partial_t h = \partial_x^2 h + f'(0)h$ with the same initial condition. For a step initial condition, the solution to this linearised equation moves at velocity v_c , so the front h must move at a velocity not larger than v_c . As the velocity cannot be smaller than v_c , it must be equal to v_c . Note that the linearised equation predicts the correct velocity, but fails at predicting the logarithmic correction $-3/(2\gamma_c) \ln t$ in (1.18).

Because the velocity is entirely determined by the linearised equation which is valid far on the right of the front, it is often said that the front in the FKPP class is “pulled” (by what happens on the right side).

When condition (1.30) is not realized, one can have situations where the reaction term $f(h)$ is at some place larger than the linearised reaction term $f'(0)h$. This means that some extra mass is created around the position of the front and, in certain cases, this extra mass can lead to a velocity larger than v_c . The front is sometimes said to be “pushed” (by what happens in the middle). What happens in the “pushed” scenario is that the front for a step initial condition moves at some velocity $v_* > v_c$ because none of the travelling waves ω_v for $v < v_*$ are positive; therefore, they cannot be reached by a positive initial condition $h_0(x)$. In fact, in the “pushed” scenario, the coefficient A_v in (1.17) is negative for $v < v_*$, positive for $v > v_*$ and zero at $v = v_*$ which means that the front $\omega_{v_*}(z)$ decays much faster as $z \rightarrow \infty$ as any other front:

$$\text{[“pushed” case; not in FKPP class]} \quad \begin{cases} \omega_{v_c}(z) \sim A_{v_c} z e^{-\gamma_c z} & \text{with } A_{v_c} < 0, \\ \omega_v(z) \sim A_v e^{-\gamma z} & \text{with } A_v < 0 \text{ for } v \in (v_c, v_*), \\ \omega_{v_*}(z) \sim B e^{-\gamma' z} & \text{with } B > 0, \\ \omega_v(z) \sim A_v e^{-\gamma z} & \text{with } A_v > 0 \text{ for } v > v_*, \end{cases} \quad (1.31)$$

where $\gamma \leq \gamma_c$ is as usual, for each v , the smallest solution of the relation (1.29) between v and γ and where, only for velocity v_* , the decay rate $\gamma' > \gamma_c$ is the other solution to (1.29).

When the sufficient condition (1.30) is not met, it can be difficult to determine whether a given function $f(h)$ which satisfies the basic conditions (1.26) leads or not for equation (1.25) to a front which is “pulled” (in the universality class of the FKPP equation) or “pushed” (not in the same universality class).

Let us end this paragraph with a final example of a front in the FKPP class. Consider a BBM where, at each branching event, the number n of children is random and is given by some distribution p_n . One finds easily that (1.7) leads to an equation of the type (1.25) with a reaction term given by

$$f(h) = 1 - h - \sum_n p_n (1 - h)^n. \quad (1.32)$$

Then $f'(0) = \sum_n n p_n - 1$ is the expected number of children at each branching event, minus one. If the second moment $\sum_n n^2 p_n$ is finite, the condition (1.30) is met and the resulting front equation (1.25) is in the FKPP class. This means that the front and the rightmost particle in the BBM move at the velocity $v_c = 2\sqrt{f'(0)}$ with the logarithmic correction (1.18).

1.4 Link with disordered systems

The BRW can be seen as a mean-field description of spin-glasses and of directed polymers in random medium. Then, the McKean relation and knowledge of FKPP universal properties allow to compute the free energies for these two mean-field models.

As an illustration, and to give some examples of models not related to population dynamics, we present an application of branching processes and front equations in the FKPP universality class to the theory of disordered systems. The model basically consists of interpreting the positions of the particles in the BRW as minus the energies of a disordered system. It can either be seen as a kind of mean-field spin glass or as a directed polymer in random medium on a tree.

1.4.1 Spin glasses

Consider a model of spin glasses; when the system has N spins, there are 2^N configurations. A configuration \mathcal{C} has an energy $E_{\mathcal{C}}$ which is the sum over all pairs of neighbouring spins of random couplings which depend only on the state of the pair of spins. Assume that the size of the system is increased by one spin without changing the random couplings between the N first spins. Each configuration \mathcal{C} of the N spin system is “split” into two configurations of the $(N+1)$ spin system: the configurations $\mathcal{C} \oplus \{+\}$ and $\mathcal{C} \oplus \{-\}$. To compute their energies, one needs to take into account all the couplings between the spins in \mathcal{C} (this gives the energy $E_{\mathcal{C}}$) and all the couplings between the new spin and the spins in \mathcal{C} . One can write

$$-E_{\mathcal{C} \oplus \{\pm\}} = -E_{\mathcal{C}} + \epsilon_{\mathcal{C}, \pm}, \quad (1.33)$$

where $\epsilon_{\mathcal{C}, \pm}$ represents the contribution to the energy of the new couplings. It is some random number depending on \mathcal{C} and on the orientation of the new spin.

The difficulty of course is that all the $\epsilon_{\mathcal{C}, \pm}$ are correlated numbers. As a kind of mean-field model, one could try to ignore these correlations and assume that all the $\epsilon_{\mathcal{C}, \pm}$ are identical independent variables. Then, (1.33) would define a BRW, the same as in (1.19): each particle in the BRW represents a spin configuration, the position of a particle is minus the energy of the corresponding configuration, and the time in the BRW is the number of spins. (We use minus the energies rather than the energies themselves because we focus on the rightmost particles of the BRW and we are interested in the states of lowest energies of the spin glass.)

In the following, we use the notations of the BRW rather than of the spin glasses. The partition function of the system at time t (*i.e.* when there are t spins) is given by

$$Z_t(\beta) = \sum_{u \in \mathcal{N}_t} e^{\beta X_u}, \quad (1.34)$$

where $\beta > 0$. Introduce the generating function $h(x, t)$:

$$1 - h(x, t) = \left\langle e^{-e^{-\beta x} Z_t(\beta)} \right\rangle = \left\langle \prod_{u \in \mathcal{N}_t} e^{-e^{-\beta(x - X_u)}} \right\rangle. \quad (1.35)$$

Using McKean’s relation (1.20), one gets that $h(x, t)$ is the solution to the front equation (1.21) with initial condition

$$h_0(x) = 1 - e^{-e^{-\beta x}}. \quad (1.36)$$

For large x , the initial condition (1.36) decays like $e^{-\beta x}$. From the usual results for fronts in the FKPP class, see Section 1.2.2, one knows that for each β there exists a velocity $v \geq v_c$ and a centring term $m_t = vt + o(t)$ such that $h(m_t + z, t)$ converges to the travelling wave $\omega_v(z)$ moving at velocity v . The value of the velocity is v_c if $\beta \geq \gamma_c$ and $v(\beta)$ if $\beta \leq \gamma_c$, where $v(\gamma)$ is the function defined in (1.24) and $v_c = v(\gamma_c)$ is its minimal value.

From the convergence of $h(m_t + z, t)$ to $\omega_v(z)$ and from (1.35), one already sees that $Z_t(\beta)e^{-\beta m_t}$ must converge in distribution to a non-trivial limit. One finally obtains [DS88; HS09] that

$$\lim_{t \rightarrow \infty} \frac{-\beta^{-1} \ln Z_t(\beta)}{t} = \begin{cases} -v(\beta) & \text{if } \beta \leq \gamma_c, \\ -v_c & \text{if } \beta \geq \gamma_c, \end{cases} \quad (1.37)$$

where the right hand side is the large time limit of $-m_t/t$. The system exhibits a phase transition.

Remark: The quantities $Z_t(\beta)e^{-\beta v(\beta)t}$ are, for each β , the so-called additive martingales for the BRW. As positive martingales, they must converge to some random constant for large times. From the fact that $Z_t(\beta)e^{-\beta m_t}$ converges in distribution (see discussion above) and the knowledge from Bramson of m_t for large t , one can obtain that the martingale $Z_t(\beta)e^{-\beta v(\beta)t}$ converges to a positive constant if $\beta < \gamma_c$, goes to 0 as $1/\sqrt{t}$ if $\beta = \gamma_c$, and goes to 0 exponentially fast if $\beta > \gamma_c$. Notice also that $Z_t(\beta)$ follows the recursion relation

$$Z_{t+1}(\beta) = e^{\beta \epsilon_1} Z_t^{(1)}(\beta) + e^{\beta \epsilon_2} Z_t^{(2)}(\beta), \quad (1.38)$$

where $Z_t^{(1)}$ and $Z_t^{(2)}$ are two independent copies of random variable Z_t and where ϵ_1 and ϵ_2 are two independent random shifts. This can be seen by considering the two children of the initial particle and by splitting the sum (1.34) into two sums, each over all the descendants of these two children. By using (1.38) in the definition (1.35) of $h(x, t)$, one can check directly that $h(x, t)$ is indeed solution to (1.21).

1.4.2 Directed polymers in random medium

Consider a directed polymer in $1+1$ dimensions on a lattice. A site $(x, t+1)$ is linked to (x, t) and to $(x-1, t)$, and independent identically distributed random energies are assigned to each link. The directed polymer, which starts somewhere from the level $t=0$, is a line made of several segments such that each segment follows a link, always in the direction where the t coordinate increases. The energy of the polymer is the sum of the energies of all the visited links. See the left part of Figure 1.5 for an illustration.

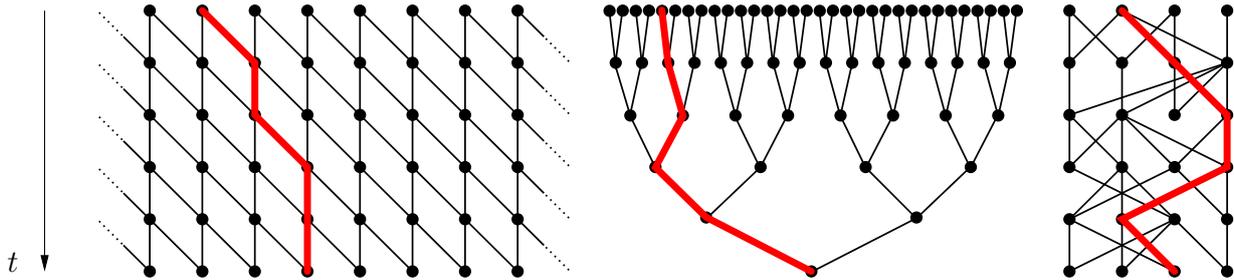


Figure 1.5: Illustration of a directed polymer. A site at level $t+1$ is linked to two sites at level t . (t increases when going downwards). The directed polymer is the bold red line going through the links, always in the direction such that t increases. On the left, the sites and the links are on a regular lattice. In the middle, the sites and the links form a tree structure. On the right, the sites are linked to randomly chosen sites at the previous level.

Let $Z_{x,t}(\beta)$ the partition function for a polymer of t segments ending at (x, t) . It satisfies the recursion

$$Z_{x,t+1}(\beta) = e^{\beta\epsilon_1(x,t+1)} Z_{x,t}(\beta) + e^{\beta\epsilon_2(x,t+1)} Z_{x-1,t}(\beta), \quad (1.39)$$

where $-\epsilon_1(x, t+1)$ and $-\epsilon_2(x, t+1)$ are the energies of the two possible links (respectively the vertical and diagonal link) leading to $(x, t+1)$.

For a fixed t , the different values of $Z_{x,t}(\beta)$ for different x are identically distributed random variables but they are correlated, which makes the problem difficult. In a mean-field variant of this model, one ignores these correlations (which means that, actually, one considers that the polymer lives on a Caley tree; see the middle part of Figure 1.5). Then [DS88]

$$Z_{t+1}(\beta) = e^{\beta\epsilon_1} Z_t^{(1)}(\beta) + e^{\beta\epsilon_2} Z_t^{(2)}(\beta), \quad (1.40)$$

which is the same as (1.38). The directed polymers on the Caley tree exhibit the same phase transition (1.37) as in the previous model of spin glasses without correlations.

1.5 The necessity of adding a noise term to the FKPP equation

The FKPP equation often appears as a $N \rightarrow \infty$ limit of some microscopic stochastic model involving some large parameter N (typically, a population size). When N is large but finite, an extra stochastic term needs to be added to the FKPP equation. In all the examples given, the noise term has an amplitude $\sqrt{h/N}$ when h is small; it is therefore relevant where h is of order $1/N$.

1.5.1 Reaction-diffusion

We introduced the FKPP equation as the deterministic limit of a reaction-diffusion process of type $A+B \rightarrow 2A$: if h is the local proportion of A particles, then the expected variation of h is due to diffusion and to a reaction term $h(1-h)$. Ignoring all fluctuations (which is the same as considering that there are infinitely many particles per unit length), this leads to the deterministic FKPP equation. Any actual microscopic model of reaction-diffusion is however bound to be stochastic, the total concentration is not infinity, and the propagation of the front must actually be noisy. It is natural to try to evaluate the effect of this noise.

Let us take a very simple example and consider a reaction process of the type



We assume for now that the reaction is well mixed: there is no relevant spatial dimension. Fix the total number of particles to be equal to N , and call $h(t)$ the proportion of particles of type A (so that there are $Nh(t)$ particles of type A and $N(1-h(t))$ particles of type B). The probability that a reaction occurs during dt must be proportional to $h(1-h)dt$ and, when it does occur, h increases or decreases by $1/N$. Let us write

$$dh = \begin{cases} \frac{1}{N} & \text{with probability } (1+\lambda)Nh(1-h)dt, \\ -\frac{1}{N} & \text{with probability } \lambda Nh(1-h)dt, \\ 0 & \text{otherwise.} \end{cases} \quad (1.42)$$

In front of $h(1-h)dt$, the terms $1+\lambda$ and λ give the relative rates of the two possible reactions, and the term N was chosen to give a nice large N limit. With h given, one obtains easily

$$\langle dh \rangle = h(1-h)dt, \quad \langle dh^2 \rangle = \frac{1+2\lambda}{N}h(1-h)dt. \quad (1.43)$$

We rewrite this as

$$dh = h(1-h)dt + \sqrt{\epsilon h(1-h)}dX, \quad \langle dX \rangle = 0, \quad \langle dX^2 \rangle = dt, \quad (1.44)$$

with $\epsilon = (1+2\lambda)/N$. One should be careful that dX is not a Gaussian variable, but some weird random variable such that dh is most of the time equal to 0 and very rarely equal to $\pm 1/N$.

When h is neither close to 0 or 1, the noise term is small and negligible compared to the $h(1-h)dt$ term. However, when h or $1-h$ is a $\mathcal{O}(\epsilon) = \mathcal{O}(N^{-1})$, both terms are important. We reach the perhaps unsurprising conclusion that noise only matters when the number of A particles or the number of B particles is not large.

Let us now put back a spatial dimension, on a lattice of spacing a . One allows any pair of two particles at two adjacent sites to exchange position with rate $dt/(Na^2)$. The same discussion as above leads to writing the following equation for the fraction $h(x,t)$ of A particles at site x :

$$\partial_t h(x,t) = \frac{h(x+a,t) + h(x-a,t) - 2h(x,t)}{a^2} + h(1-h) + \text{noise}, \quad (1.45)$$

where the expectation of the noise is zero and its amplitude is of order $\sqrt{h(1-h)/N}$ with N being now the total number of particles per site. The noise is small compared to the reaction term except where h or $1-h$ is of order $1/N$.

1.5.2 Directed polymers

In Section 1.4, we introduced directed polymers in $1+1$ dimension, which is a hard problem because for a given t there are important correlations between the partition functions $Z_{x,t}$ ending at different points x . We also introduced directed polymers on a tree, for which all the correlations are removed. It is naturally interesting to consider a model with a tiny amount of correlation between the partition functions.

A way to do this is to imagine that at each level t there is a large number N of possible passage points for the directed polymer. Each of these points at level t is linked to two other points chosen at random at level $t-1$, see the right part of Figure 1.5. As usual, a random energy is assigned independently to each link.

Let $Z_{i,t}(\beta)$ the partition function of a directed polymer of length t ending at point $i \in \{1, 2, \dots, N\}$ at level t . There are, as before, 2^t possible paths leading to that point from the line $t=0$, and the partition functions satisfy the recursion

$$Z_{i,t+1}(\beta) = e^{\beta\epsilon_1} Z_{j_1,t}(\beta) + e^{\beta\epsilon_2} Z_{j_2,t}(\beta), \quad (1.46)$$

where j_1 and j_2 are the random indices of the two points at level t linked to $(i, t+1)$ and where $-\epsilon_1$ and $-\epsilon_2$ are the corresponding energies of the links. Implicitly, ϵ_1 , ϵ_2 , j_1 and j_2 depend on i and t .

Two different points i and i' at the same level t have correlated partition functions, but the correlation becomes weak as N becomes large. Indeed, the probability that the two points i and i' share a common parent (which would mean that they are both linked to the same site at previous level), or a common grand-parent, is only of order $\mathcal{O}(N^{-1})$. In fact, one needs typically to go back $\mathcal{O}(\ln N)$ levels to find a link shared by the paths leading to i and i' . As $N \rightarrow \infty$, one never finds a shared link and the model becomes equivalent to the directed polymer on a tree.

To simplify a bit this hard problem, let us focus on the zero temperature limit $\beta \rightarrow \infty$. Calling $-X_i(t)$ the minimal energy of the 2^t paths leading to (i, t) , one has from (1.46)

$$X_i(t+1) = \max [X_{j_1}(t) + \epsilon_1, X_{j_2}(t) + \epsilon_2], \quad (1.47)$$

with the same notations as in (1.46).

Recall that t is here the length of the directed polymer, but (1.47) can also be seen as a dynamical system where t is a time: at a given time t , one has a cloud of N values $\{X_1(t), \dots, X_N(t)\}$. Given the state of the system at time t , the N new values at time $t + 1$ are obtained by repeating independently N times the same procedure: pick uniformly at random two possible parents from generation t , add some random numbers to their values and keep the largest. It is a kind of over-optimistic population model where $X_i(t)$ represents some desirable trait and where each individual “chooses” two parents at random and inherits from the best parent with some mutation.

Let $h(x, t)$ be the fraction of particles on the right of x at time t :

$$h(x, t) = \frac{1}{N} \sum_{i=1}^N \mathbb{1}_{\{X_i(t) > x\}}. \quad (1.48)$$

With the function $h(\cdot, t)$ given, the probability (with j and ϵ random) that $X_j(t) + \epsilon$ is on the right of x is $\int d\epsilon \rho(\epsilon) h(x - \epsilon, t)$. It is then easy to compute the expectation of $h(x, t + 1)$ given $h(\cdot, t)$.

$$\langle h(x, t + 1) \mid h(\cdot, t) \rangle = \mathbb{P}(X_i(t + 1) > x \mid h(\cdot, t)) = 1 - \left(1 - \int d\epsilon \rho(\epsilon) h(x - \epsilon, t)\right)^2. \quad (1.49)$$

With the function $h(\cdot, t)$ given, the $X_i(t + 1)$ are independent variables. Writing an expression for the variance of $h(x, t + 1)$ given the function $h(\cdot, t)$ is simple:

$$\text{var} [h(x, t + 1) \mid h(\cdot, t)] = \frac{1}{N} \text{var} [\mathbb{1}_{\{X_i(t) > x\}} \mid h(\cdot, t)] = \frac{1}{N} \left[\langle h(x, t + 1) \mid h(\cdot, t) \rangle - \langle h(x, t + 1) \mid h(\cdot, t) \rangle^2 \right].$$

A simple way to summarize these last two equations is to write

$$h(x, t + 1) = 1 - \left(1 - \int d\epsilon \rho(\epsilon) h(x - \epsilon, t)\right)^2 + \text{noise}. \quad (1.50)$$

It is a noisy version of the front equation (1.21) which was obtained for the directed polymer on the tree. The “noise” term in the equation above is complicated: it is discrete, because $h(x, t + 1)$ on the left hand side is a multiple of N^{-1} . It is correlated for different values of x , because the left hand side is a non-increasing function of x . The expectation of the noise is zero, of course, and its standard deviation is of order $\sqrt{h/N}$ where h is small. This means again that the noise is comparable to the deterministic term where h is of order $1/N$. As $N \rightarrow \infty$, the noise disappears and one recovers, as already argued, the directed polymer on a tree.

1.5.3 Models of population dynamics

The model of directed polymers led in the zero temperature limit to (1.47) which looks like a model of population dynamics: there is a constant population size N and we follow the value $X_i(t)$ of some trait of individual i at generation t . When building a new generation, each individual inherits the trait of one of its parent plus some mutation.

In the specific case of (1.47), the number of children of a given individual is random and position-dependent: all individuals have the same chance of being picked as a prospective parent, but because of the max in (1.47), the rightmost individuals (*i.e.* with the largest values for X) are more likely to be finally chosen. This is a selection mechanism: having a large value of the trait X relates to having more children.

One can devise models which are simpler and more natural than (1.47). For instance: each individual has at first exactly two children, but after the reproduction phase, there is a pruning phase where half the population is removed to keep its size constant at N individuals [BG10; DR11]. Without the pruning phase, the population size would double at each generation and one would recover exactly a BRW.

The pruning phase can be done in several ways. A first possibility is to choose uniformly the N survivors of the pruning phase. In this case there is no selection, the model we just defined is very close to the Wright-Fisher model [Wri31; Fis30].

The other extreme, which we consider here, is to have a flawless selection: for the pruning phase, one keeps the N rightmost individuals and remove the N leftmost:

$$\left\{ \begin{array}{l} \text{Reproduction phase:} \quad \text{particle } i \text{ at time } t \text{ is replaced by two particles at } X_i(t) + \epsilon_1 \text{ and } X_i(t) + \epsilon_2, \\ \text{Pruning phase:} \quad \text{only the } N \text{ rightmost are kept.} \end{array} \right. \quad (1.51)$$

Let us write an evolution equation for (1.51); define $h(x, t)$ as $1/N$ times the number of particles on the right of x at generation t , as in (1.48). With the function $h(\cdot, t)$ given, we want to compute $h(x, t + 1)$. We first compute

$h^*(x, t+1)$ defined as $1/N$ times the number of particles on the right of x after the reproduction phase leading to generation $t+1$, but before the pruning phase. One obtains easily that

$$\langle h^*(x, t+1) | h(\cdot, t) \rangle = 2 \int d\epsilon \rho(\epsilon) h(x - \epsilon, t). \quad (1.52)$$

(Nota: as there are $2N$ particles before the pruning phase, $h^*(-\infty, t) = 2$.) The variance of $h^*(x, t+1)$ given $h(\cdot, t)$ has an expression which is a bit more complicated, but one checks that it is of the order of h/N when h is small [BDMM07]. As before, we simply write that $h^*(x, t+1)$ is equal to its expectation given $h(\cdot, t)$ plus some noise. The pruning phase is then very easy to write: $h(x, t+1) = \min[1, h^*(x, t+1)]$. One finally obtains

$$h(x, t+1) = \min \left[1, 2 \int d\epsilon \rho(\epsilon) h(x - \epsilon, t) + \text{noise} \right], \quad (1.53)$$

where the noise, which has an amplitude of order $\sqrt{h/N}$, is again negligible compared to the deterministic term except where h is of order $1/N$.

1.5.4 Universality of internal noise

We have introduced three models (reaction-diffusion, directed polymer, population dynamics) where we take into account finite size effects: there is a large parameter N which is either the total number of particles per lattice site, the number of passage points per level for the directed polymer, or the total population of the system. Heuristically, when N is infinite, one recovers respectively the deterministic FKPP on a lattice, the directed polymer on a tree and the BRW. When N is finite, we obtain in each case (see (1.45), (1.50) and (1.53)) an equation in the FKPP universality class but with an extra noise term. A natural question is of course to determine how the picture described in Section 1.3 for the position of the FKPP front is modified by this noise term.

Notice that the noise term is horribly complicated in the three cases we have considered, but it is always of order $\sqrt{h/N}$ for small h . The region where h is small is of course the most important because we know from the general study of the (deterministic) FKPP universality class that the fronts are “pulled” by what happens in the linear region, *i.e.* far on the right, where h is small. To summarize, all the models we presented have noise terms with similar behaviours at the tip of the front, which is the only place where things matter for the dynamics of the front. One can therefore expect similar behaviours for all these models.

A common feature of the noisy models we have presented is that the noise comes from the stochastic nature of the motion of the particles defining the system. It is sometimes said to be an “internal noise”. One can also consider equations with an “external noise” which is due to randomness in the environment [RES00; BN12; Nad15]. To take an example, $\partial_t h = \partial_x^2 h + [1 + \eta(x, t)](h - h^2)$ would describe the evolution of a population where the reaction rate has a random contribution $\eta(x, t)$ because, maybe, the birth rate depends on temperature and temperature fluctuates.

In this memoir, we only consider noisy FKPP equations with an internal noise.

1.5.5 The stochastic FKPP equation, the coalescing BBM and duality

A reference noisy FKPP equation with an internal noise is the stochastic FKPP equation:

$$\partial_t h = \partial_x^2 h + h - h^2 + \sqrt{\frac{h - h^2}{N}} \eta(x, t), \quad (1.54)$$

where N is a large parameter and $\eta(x, t)$ is a delta-correlated Gaussian noise:

$$\langle \eta(x, t) \rangle = 0, \quad \langle \eta(x, t) \eta(x', t') \rangle = \delta(x - x') \delta(t - t'). \quad (1.55)$$

The stochastic FKPP (1.54) behaves in the same way as the other models we have presented in this section because its noise term has again the amplitude $\sqrt{h/N}$ when h is small.

It might be tempting to see the stochastic FKPP as an hydrodynamic limit of a reaction-diffusion process such as (1.41), but it is not directly the case [DMS03] because the randomness in the diffusion process makes things complicated. The stochastic FKPP has however been shown to be the limit of some long range voter model [MT95]. Furthermore, without taking any limit, it is also the dual process of a variant of the BBM called the coalescing BBM, in the same way as the (deterministic) FKPP is the dual of the BBM through McKean’s relation (1.7). In the coalescing BBM, particles diffuse and branch, as they do in the BBM, but they can also coalesce with a small rate ($\epsilon \ll 1$):



The meaning of (1.56) is clear when space is discrete: in a given box with n particles at time t , the probability to create a new particle during dt is $n dt$ and the probability to remove one particle is $\epsilon \frac{n(n-1)}{2} dt$ because there are $n(n-1)/2$ pairs of A particles that might interact. At long times, the number of particles in the box should oscillate around $2/\epsilon$. When space is continuous, it is more difficult but still possible to give a meaning to (1.56) using local times; the simplest way is to start from a variant of the model with discrete space and then take the limit where the lattice spacing goes to zero. Then, the model that one obtains is such that, at long times, the density of particles at any given position oscillates around $2/\epsilon \gg 1$.

For the coalescing BBM, call again \mathcal{N}_t the set of particles at time t and, for each $u \in \mathcal{N}_t$, call X_u the position of particle u . Then, if $h(x, t)$ is the solution to the stochastic FKPP equation (1.54) with initial condition $h_0(x)$, and if $\epsilon = 1/N$, one has [DMS03; SU86]

$$\left\langle \prod_{u \in \mathcal{N}_t} (1 - h_0(x - X_u)) \right\rangle = \left\langle \prod_{u \in \mathcal{N}_0} (1 - h(x - X_u, t)) \right\rangle. \quad (1.57)$$

On the left hand side, the product is over all the particles present at time t while on the right hand side it is a product on all the particles present in the initial condition. If the coalescing BBM is started from only one particle at the origin, then the right hand side reduces to $1 - \langle h(x, t) \rangle$ and (1.57) looks very similar to (1.7).

1.5.6 The noisy FKPP equation in QCD

Remarkably, branching processes and FKPP fronts have found some application in particle physics.

In 1988, a first link was established in [BP88] where it was shown that the “random cascading process” leading to the production of many particles during a scattering event can be analysed as a BRW using the methods developed the same year in [DS88]. It was later recognized in [MP03; MP04a; MP04b] that the Balitsky-Kovchegov equation (BK), which is used to describe high-energy scattering in quantum chromodynamics, is in fact in the universality class of the FKPP equation. (The front h would represent the scattering amplitude, x and t are not space and time, but parameters of the scattering. See also [MT02].) Finally, it has been suggested that high-energy scattering in QCD might in fact be similar to models in the class of the stochastic FKPP equation [IMM05], with the amplitude of the noise term related to the coupling constant.

The relationship between QCD and the FKPP equation is reviewed in [Mun09].

1.6 My contribution to the study of the BBM and the FKPP equation

In the following chapters, I present some of the results that my collaborators and I have obtained on the FKPP equation and on branching processes during the last years.

I explain in Chapter 2 how we studied in great details the position of a FKPP front (say, the position where the front is equal to $\frac{1}{2}$). The asymptotic expansion for large times of this position must of course start like Bramson’s result (1.10). The first vanishing term in this expansion was shown by Ebert and van Saarloos [ES00] to be of order $1/\sqrt{t}$ for initial conditions that decay “fast enough”. By introducing two independent models in the FKPP class that we were able to solve in two different ways, we could recover the result of Ebert and van Saarloos, determine with precision what “fast enough” meant, make a new prediction for the first vanishing term when the initial condition does not decay “fast enough” and find the second vanishing correction to the position of the front.

I present in Chapter 3 my work on the limiting distribution of extremal points in the BBM: for any integer n , the joint probability distribution function of the distances between the n rightmost particles in the BBM converges as $t \rightarrow \infty$ to some limiting distribution. This limiting distribution has a nice description as a “decorated exponential Poisson point process”, and any feature of this distribution (for instance, the expected distance between the two rightmost particles, or the distribution of this distance) can be measured by integrating numerically the FKPP equation with some well-chosen initial condition.

Chapter 4 concerns the study of noisy FKPP equations, with an internal noise as described in Section 1.5. These noisy fronts move more slowly than their deterministic counterparts; I explain how one can compute, to leading order, the correction to the velocity, the diffusion constant and, in fact, all the cumulants of the position of these fronts. When the noisy front comes from a model of population dynamics with N individuals, as in Section 1.5.3, we also studied the statistical properties of the genealogical tree of the particles. We showed that, properly rescaled and in the $N \rightarrow \infty$ limit, the genealogical tree can be described by the Bolthausen-Sznitman coalescent, the same tree as occurs in the replica method for studying spin glasses.

Chapter 2

Vanishing corrections for the position of the FKPP front

In this chapter, I discuss the position of a front described by the FKPP equation as a function of its initial condition h_0 . Throughout this chapter, it is assumed that

$$h_0 \in [0, 1], \quad h_0(-\infty) = 1, \quad h_0(+\infty) = 0. \quad (2.1)$$

The large time asymptotic position of the front depends on how fast h_0 converges to 0 at infinity.

The asymptotic position of the front is known up to a $o(1)$ since Bramson [Bra83]. Ebert and van Saarloos [ES00] predicted the first vanishing term. I will present two different models which allowed us to better understand this asymptotic position, and in particular to precise and refine the prediction from Ebert and van Saarloos.

2.1 The Bramson term and the Ebert and van Saarloos term

We recall Bramson's result and describe in detail the prediction by Ebert and van Saarloos.

2.1.1 The Bramson term

For a non-lattice equation in the FKPP universality class, if the initial condition h_0 decays fast enough so that

$$\int dx h_0(x) x e^{\gamma_c x} < \infty, \quad (2.2)$$

(and assuming (2.1), of course), then

$$\lim_{t \rightarrow \infty} h(m_t + z, t) \text{ exists and is not trivial (neither 0 nor 1)} \quad (2.3)$$

if and only if

$$m_t = v_c t - \frac{3}{2\gamma_c} \ln t + C + o(1), \quad (2.4)$$

where $v_c = v(\gamma_c)$ is the critical (minimal) velocity and C is any constant. The $-3/(2\gamma_c) \ln t$ in (2.4) is the famous “Bramson term”.

The value of the limit (2.3) as a function of z is the critical travelling wave, which is only defined up to translation because changing C in (2.4) shifts the limit in (2.3). To fix the invariance by translation, we insist now that the limit (2.3) is $\frac{1}{2}$ for $z = 0$. This determines uniquely the constant C in (2.4) for each initial condition h_0 . Calling $\omega(z)$ the (now unique) critical travelling wave, the convergence (2.3) is uniform in z :

$$\lim_{t \rightarrow \infty} h(m_t + z, t) = \omega(z) \quad \text{uniformly in } z, \quad (2.5)$$

with of course $\omega(0) = \frac{1}{2}$. It is furthermore known that there exists a positive constant A such that

$$\omega(z) \sim A z e^{-\gamma_c z} \quad \text{for large } z. \quad (2.6)$$

For the FKPP equation itself, $\partial_t h = \partial_x^2 h + h - h^2$, one has $v(\gamma) = \gamma + \gamma^{-1}$, see (1.29), so that $v_c = 2$ and $\gamma_c = 1$. From (1.13) with $v = 2$, $\omega(z)$ is solution to

$$\omega'' + 2\omega' + \omega - \omega^2 = 0, \quad \omega(-\infty) = 1, \quad \omega(+\infty) = 0, \quad \omega(0) = \frac{1}{2}. \quad (2.7)$$

The results above have been proved by Bramson [Bra78; Bra83] for the FKPP equation and the FKPP equation with the reaction term $h - h^2$ replaced by any $f(h)$ which satisfies some properties, see Section 1.3.2. They have also mostly been proved by Aidékon in the case of a step initial condition $h_0(x) = \mathbb{1}_{\{x < 0\}}$ for any front equation which is the dual through McKean (see Section 1.3) of a non-lattice BRW with some mild hypotheses on the reproduction law [Aid13]; see also [HS09; BDZ14]. Notice that (2.3) only makes sense if the equation is non-lattice; the lattice case needs to be addressed separately [BDZ14].

The position of the front is not only known when (2.2) holds, but also for many other initial conditions; the velocity and the Bramson term need however to be modified. To give some examples, the limit (2.3) exists and is non-trivial if and only if m_t is of the following given form, depending on the initial condition:

$$\text{if } h_0(x) \sim ax^\nu e^{-\gamma_c x}, \quad \text{then } m_t = \begin{cases} v_c t + \frac{\nu-1}{2\gamma_c} \ln t + C + o(1) & \text{if } \nu > -2, \\ v_c t - \frac{3}{2\gamma_c} \ln t + \frac{1}{\gamma_c} \ln \ln t + C + o(1) & \text{if } \nu = -2, \\ v_c t - \frac{3}{2\gamma_c} \ln t + C + o(1) & \text{if } \nu < -2, \end{cases} \quad (2.8)$$

$$\text{if } h_0(x) \sim ax^\nu e^{-\gamma x} \text{ with } \gamma \in (0, \gamma_c), \quad \text{then } m_t = v(\gamma)t + \frac{\nu}{\gamma} \ln t + C + o(1),$$

with, for each line, a different constant term C .

2.1.2 The Ebert and van Saarloos term

As the front h correctly centred looks more and more like the critical wave ω , one can show that for any $\alpha \in (0, 1)$ and any large enough time t , there exists a unique position where the front is equal to α . Let $\mu_t^{(\alpha)}$ be that position:

$$h(\mu_t^{(\alpha)}, t) = \alpha. \quad (2.9)$$

It is clear that $\mu_t^{(\alpha)}$ must be, up to some translation, a valid m_t in the sense of (2.5) and that, therefore,

$$\mu_t^{(\alpha)} = v_c t - \frac{3}{2\gamma_c} \ln t + C^{(\alpha)} + o(1), \quad h(\mu_t^{(\alpha)} + z, t) \rightarrow \omega(\omega^{-1}(\alpha) + z), \quad (2.10)$$

for some α -dependent constant $C^{(\alpha)}$. It makes sense to look for the next terms in the asymptotic expansion of $\mu_t^{(\alpha)}$. In a paper from 2000, Ute Ebert and Wim van Saarloos argued [ES00] that for any front in the FKPP class, any $\alpha \in (0, 1)$, and any initial condition that decays fast enough, one has

$$\mu_t^{(\alpha)} = v_c t - \frac{3}{2\gamma_c} \ln t + C^{(\alpha)} - 3\sqrt{\frac{2\pi}{\gamma_c^5 v''(\gamma_c) t}} + \dots, \quad (2.11)$$

where, remarkably, the prefactor of the new $1/\sqrt{t}$ term does not depend on α ; see also [MM14]. Ebert and van Saarloos claimed furthermore that the next term was a non-universal $\mathcal{O}(t^{-1})$ term, but I now think there is first a universal $\mathcal{O}[(\ln t)/t]$ term, see Section 2.3. The method used in [ES00] consisted in studying $e^{\gamma_c z \sqrt{t}} h(m_t + z\sqrt{t}, t)/\sqrt{t}$ as a function of z for large times. Derrida and I argued [BD97] in 1997 that this quantity was converging to $Az \exp(-cz^2)$ for some A and c , and this observation allowed us to recover the Bramson term. Looking at the way this limit is reached allowed Ebert and van Saarloos to obtain their $1/\sqrt{t}$ term. What is still missing is a rigorous derivation; in fact, one of the starting points of their argument is that one can obtain an asymptotic expansion of the instantaneous velocity by taking the derivative of (2.10): $\dot{\mu}_t^{(\alpha)} = v_c - 3/(2\gamma_c t) + \dots$, but this has never been proved. [ES00] does not state precisely for which initial conditions (2.11) is expected to hold. At the very least, one needs (2.2) because otherwise the Bramson term is not even there, but this is not sufficient. In the following section, I explain why (2.11) holds if and only if

$$\int dx h_0(x) x^2 e^{\gamma_c x} < \infty. \quad (2.12)$$

When (2.2) holds but not (2.12), then the Bramson term is present but the Ebert and van Saarloos term needs to be modified.

2.2 Understanding the Bramson term and the Ebert and van Saarloos term

The presence of a saturation term is important, but its nature is not. Taking advantage of that fact, we study two different models in the FKPP class with a simplified saturation term. The models can be solved.

The dynamics of equations in the FKPP class are essentially controlled by what happens where h is small. However, the linearised FKPP equation alone is not even sufficient to understand Bramson's $-3/(2\gamma_c) \ln t$ term. Indeed, consider the solution to the linearised equation $\partial_t h = \partial_x^2 h + h$ with a step initial condition $h_0(x) = \mathbb{1}_{\{x \leq 0\}}$. One obtains easily that

$$h(x, t) = \frac{e^t}{\sqrt{4\pi t}} \int_x^\infty dy e^{-\frac{y^2}{4t}} \quad [\text{linearised equation}], \quad (2.13)$$

and that

$$h(m_t + z, t) \rightarrow e^{-z} \quad \text{with } m_t = 2t - \frac{1}{2} \ln t - \frac{1}{2} \ln(4\pi) + o(1) \quad [\text{linearised equation}]. \quad (2.14)$$

The prefactor of the logarithmic correction is wrong (1/2 instead of 3/2) and the limit of $h(m_t + z, t)$ for large z is missing a z prefactor in front of the exponential, as in (2.6). The conclusion is that even though the FKPP front is essentially controlled in the linear region (where h is small), it is necessary to introduce a non-linear saturation term so that the value of the front does not grow indefinitely in the region where h is not small. The nature of this saturation term is not important (except for the value of C , the asymptotic expansion (2.4) of m_t would remain unchanged if the reaction rate $h - h^2$ in the FKPP equation was replaced by $h - h^3$), but it is important to have a saturation mechanism.

If one believes in the universality of this problem and in the fact that the nature of the saturation mechanism is not important, one is led to look for simple and, hopefully, solvable models in the FKPP class. I have studied two such models.

2.2.1 The linear FKPP equation with an anchor

Since the position of a FKPP front is determined by what happens in the linear part, we look for a solvable model which is essentially linear. However, as we have just seen, it cannot be linear all the way and one needs a saturation mechanism. Consider the real FKPP equation and call $\mu_t = \mu_t^{(\alpha)}$ the position where the front is equal to α . On the right of μ_t , one can linearise the equation without changing the behaviour of the position. One does not need to look on the left of μ_t ; one expects it should not matter as long as there is some saturation. It therefore seems reasonable to look at the following linear equation with boundary:

$$\begin{cases} \partial_t h = \partial_x^2 h + h & \text{if } x > \mu_t, \\ h(\mu_t, t) = \alpha & \text{for any } t > 0. \end{cases} \quad (2.15)$$

Of course, (2.15) is simply an initial-boundary problem with a Dirichlet boundary condition at μ_t , and one knows that a solution to (2.15) exists for any (reasonable) choice of μ_t . As the goal is to determine the asymptotic expansion of μ_t , (2.15) is clearly not sufficient, and an extra constraint needs to be added. A natural choice is to fix the derivative of h at position μ_t :

$$\begin{cases} \partial_t h = \partial_x^2 h + h & \text{if } x > \mu_t, \\ h(\mu_t, t) = \alpha & \text{for any } t > 0, \\ \partial_x h(\mu_t, t) = \beta & \text{for any } t > 0, \end{cases} \quad (2.16)$$

with α and β two given numbers. In (2.16), both h and μ_t are the quantities to be solved for. In other words, μ_t is determined as being the (hopefully unique) function such that the solution $h(x, t)$ to (2.15) has its derivative at position μ_t equal to β at all times.

For any front, a travelling wave solution is such that $\mu_t = vt$ and $h(\mu_t + z, t) = \omega_v(z)$ independent of t . In the case of (2.16), one checks easily that one must have $\omega_v'' + v\omega_v' + \omega_v = 0$ with $\omega_v(0) = \alpha$ and $\omega_v'(0) = \beta$. The travelling waves ω_v are non-oscillatory only for $v \geq v_c = 2$. For the critical velocity $v_c = 2$, one obtains that

$$\omega_2(z) = [\alpha + (\alpha + \beta)z]e^{-z}, \quad (2.17)$$

and for $v > 2$ one gets $\omega_v(z) = Ae^{-\gamma z} + Be^{-\gamma^{-1}z}$ where $\gamma + \gamma^{-1} = v$ and where A and B are some coefficients that depend on α , β and γ .

Because all the ingredients (diffusion, growth, saturation) are in place, one expects (2.16) to be in the universality class of FKPP for any choice of α and β , as long as $\alpha \geq 0$ and $\alpha + \beta > 0$, because the critical

travelling wave should be positive and behave like ze^{-z} at infinity. The case $\alpha + \beta < 0$ would correspond to a “pushed” front, as in (1.31).

In [★BBHR15], we studied (2.16) with the parameters $\alpha = 0$ and $\beta = 1$. (The same model was also studied in [Hen14] using partial differential equations methods in the special case where the initial condition is compactly supported.) Taking $\alpha = 0$ might not seem very natural, but this choice should lead to a front in the FKPP class and it simplifies a lot the analysis. Indeed, for $\alpha = 0$ and only for $\alpha = 0$, the solution $h(x, t)$ to (2.15) for a given boundary μ_t and a given initial condition h_0 can be written by linearity as

$$h(x, t) = e^t \int_0^\infty dy h_0(y) q(x, t; y), \quad (2.18)$$

(we assume $\mu_0 = 0$) where q is solution to

$$\partial_t q = \partial_x^2 q \quad \text{for } x > \mu_t, \quad q(x, t = 0; y) = \delta(x - y), \quad q(\mu_t, t; y) = 0. \quad (2.19)$$

Then, q is simply the density of probability that a Brownian path $s \mapsto B_s$ started from y arrives at time t at position x without ever hitting the boundary:

$$q(x, t; y) = \mathbb{E}^y [\delta(B_t - x) \mathbb{1}_{\{B_s > \mu_s \ \forall s \in [0, t]\}}]. \quad (2.20)$$

(The Brownian path is normalized in such a way that $\langle dB_s^2 \rangle = 2 ds$.) Write (2.20) with x replaced by $\mu_t + x$ as a path integral, and make the change of variable $B_s = \mu_s + \tilde{B}_s$, where the tilde is dropped after two lines:

$$q(\mu_t + x, t; y) = \int_{B_0=y}^{B_t=\mu_t+x} \mathcal{D}B_s e^{-\int_0^t \frac{(dB_s)^2}{4ds}} \mathbb{1}_{\{B_s > \mu_s \ \forall s \in [0, t]\}}, \quad (2.21)$$

$$= \int_{\tilde{B}_0=y}^{\tilde{B}_t=x} \mathcal{D}\tilde{B}_s e^{-\int_0^t \frac{(\dot{\mu}_s ds + d\tilde{B}_s)^2}{4ds}} \mathbb{1}_{\{\tilde{B}_s > 0 \ \forall s \in [0, t]\}}, \quad (2.22)$$

$$= e^{-\frac{1}{4} \int_0^t ds (\dot{\mu}_s)^2} \int_{B_0=y}^{B_t=x} \mathcal{D}B_s e^{-\frac{1}{2} \int_0^t \dot{\mu}_s dB_s - \int_0^t \frac{(dB_s)^2}{4ds}} \mathbb{1}_{\{B_s > 0 \ \forall s \in [0, t]\}}, \quad (2.23)$$

$$= e^{-\frac{1}{4} \int_0^t ds (\dot{\mu}_s)^2} \mathbb{E}^y [\delta(B_t - x) e^{-\frac{1}{2} \int_0^t \dot{\mu}_s dB_s} \mathbb{1}_{\{B_s > 0 \ \forall s \in [0, t]\}}], \quad (2.24)$$

where in the last expression $s \mapsto B_s$ is again a Brownian motion. The passage from (2.20) to (2.24) is called a Girsanov transform. The boundary is now a straight line at position 0, but there is a complicated extra term inside the expectation. Write (2.24) as a conditional expectation

$$q(\mu_t + x, t; y) = e^{-\frac{1}{4} \int_0^t ds (\dot{\mu}_s)^2} \mathbb{E}^y [\delta(B_t - x) \mathbb{1}_{\{B_s > 0 \ \forall s \in [0, t]\}}] e^{-\frac{x-y}{2t} \mu_t} \psi(x, t; y), \quad (2.25)$$

with

$$\psi(x, t; y) = e^{\frac{x-y}{2t} \mu_t} \mathbb{E}^y \left[e^{-\frac{1}{2} \int_0^t \dot{\mu}_s dB_s} \Big|_{B_t = x \text{ and } B_s > 0 \ \forall s \in [0, t]} \right] = \mathbb{E}^{y \rightarrow x} \left[e^{-\frac{1}{2} \int_0^t \dot{\mu}_s (d\xi_s - \frac{x-y}{t} ds)} \right], \quad (2.26)$$

where the last expectation is over a Bessel bridge ξ_s going from y to x in a time t , which is of course the same as a Brownian started from y , conditioned to end at x and to remain positive at all times. The expectation which remains in (2.25) is easy to compute: it is the probability that a Brownian started from y ends at x without ever touching the origin, and one obtains its expression through the method of mirrors:

$$\mathbb{E}^y [\delta(B_t - x) \mathbb{1}_{\{B_s > 0 \ \forall s \in [0, t]\}}] = \frac{1}{\sqrt{\pi t}} \sinh \frac{xy}{2t} e^{-\frac{x^2+y^2}{4t}}. \quad (2.27)$$

Therefore, it “only” remains to compute $\psi(x, t; y)$.

In [★BBHR15], we have shown under some mild conditions on μ_t that the function $\psi(x, t; y)$ has a large time limit which is independent of x and can be written as an expectation over a Bessel process (not a bridge). Then, with all the other terms in (2.25) and (2.18) known, it is easy to see that there exists a unique possible asymptotic expansion for μ_t up to $o(1)$ such that $\partial_x h(\mu_t, t) \rightarrow 1$ in the large time limit. This asymptotic expansion for μ_t is given by (2.8) and the limit of $h(\mu_t + x, t)$ is of course the travelling wave with the corresponding velocity.

It is interesting to see that, after all, we did not have to solve (2.16); rather, we looked at (2.15) with the extra condition that in the large time limit (only!) the solution h must satisfy $\partial_x h(\mu_t, t) \rightarrow 1$. This turned out to be sufficient to determine, for any initial condition, the asymptotic expansion of the position μ_t of the front up to $o(1)$, and to show that the results are exactly the same as Bramson’s results for the real FKPP equation.

The second and more difficult part of [★BBHR15] was to estimate the speed at which ψ in (2.26) converges to its limit. Used in (2.25) and (2.18), this allowed to determine the speed at which $h(\mu_t + x, t)$ would converge

to its limit, the travelling wave. For instance, we showed that when solving (2.15) for a compactly supported initial condition and for $\mu_t = 2t - (3/2) \ln t + C - r/\sqrt{t}$ when $t > 1$, then $h(\mu_t + x, t)$ is equal to its limit plus a correction of order $1/\sqrt{t}$, except when $r = 3\sqrt{\pi}$, in which case $h(\mu_t + x, t)$ is equal to its limit plus a correction of order at most $(\ln t)/t$. This suggests strongly that the solution to (2.16) satisfies

$$\mu_t = 2t - \frac{3}{2} \ln t + C - \frac{3\sqrt{\pi}}{\sqrt{t}} + \dots \quad (2.28)$$

which is the Ebert and van Saarloos prediction (2.11). We also showed that for an initial condition of the form $h_0(x) \sim Ax^\nu e^{-x}$, then (2.28) only holds if $\nu < -3$, which is nearly the same as condition (2.12). We also computed the first vanishing term for $\nu \in [-3, -2)$, which is the domain where Bramson's term is present but not the $1/\sqrt{t}$ correction from Ebert and van Saarloos. (Our findings are summarized in Section 2.3.)

Again, it is interesting to notice that, in order to obtain (2.28) or its generalisation, we did not after all solve (2.16), but rather we determined the function μ_t which, in (2.15), maximises the speed of the convergence of $h(\mu_t + x, t)$ to its limit.

2.2.2 A model on the lattice

Bernard Derrida and I have also studied the following model [★BD15]: space x is discrete and time t is continuous. The values $h(x, t)$ of the front, with $x \in \mathbb{Z}$, evolve according to

$$\partial_t h(x, t) = \begin{cases} h(x, t) + h(x-1, t) & \text{if } h(x, t) < 1, \\ 0 & \text{if } h(x, t) = 1. \end{cases} \quad (2.29)$$

In this model, the front undergoes a linear growth for all values of h smaller than 1 (not only in the limit $h \ll 1$). There is some mixing (which plays the rôle of diffusion) because the growth rate of h on a given lattice site depends on the lattice site on its left. Finally, the front saturates at $h = 1$, the stable phase.

Looking for travelling wave solutions of the form $e^{-\gamma(x-vt)}$ in the region where $h < 1$, one obtains the relation between v and γ :

$$v(\gamma) = \frac{1}{\gamma} [1 + e^\gamma]. \quad (2.30)$$

Looking for the minimum of this function gives $v_c = 3.59112\dots$ and $\gamma_c = 1.27846\dots$

We always assume that the initial condition h_0 satisfies $h_0(x) = 1$ for $x \leq 0$ and $h_0(x) \in [0, 1)$ for $x \geq 1$, and we introduce, for $x \geq 1$, the time t_x at which the front reaches the value 1 on site x :

$$h(x, t) < 1 \quad \text{if } t < t_x, \quad h(x, t) = 1 \quad \text{if } t \geq t_x. \quad (2.31)$$

We also assume that $h_0(x)$ is a non-increasing function of x ; this implies that $h(x, t)$ remains at all time a non-increasing function of x and that $x \mapsto t_x$ is an increasing function of x .

The model (2.29) is sufficiently linear so that one can solve it [★BD15] in the following sense: for a given initial condition $h_0(x)$, the times t_x satisfy for λ small enough

$$\sum_{x=1}^{\infty} h_0(x) \lambda^x = -\frac{\lambda}{1+\lambda} + \frac{2}{1+\lambda} \sum_{x=1}^{\infty} e^{-(1+\lambda)t_x} \lambda^x. \quad (2.32)$$

This remarkable equality allows to explore the question of the position of the front as a function of time.

One shortcoming, however, of (2.29), is that the model is defined on the lattice. This means that one cannot consider the quantity $h(m_t + x, t)$ as in (2.3) because m_t does not take integer values. Similarly, one cannot define $\mu_t^{(\alpha)}$ as the position where the front is α because, most of the time, such a position does not exist. One way to overcome this problem is to take advantage of the fact that time is continuous and, in some sense, exchange the rôles of space and time. For instance, instead of considering the limit (2.3), one could investigate the limit of $h(x, t_x - \tau)$ for fixed τ as x goes to infinity. Furthermore, the quantity t_x (time at which the front reaches 1 on a given site x) plays a rôle similar to $\mu_t^{(\alpha)}$ (position at which the front is equal to α at a given time t).

What we did in [★BD15] is to compute from (2.32) the large x asymptotic of t_x for various initial conditions. We argued that, if and only if

$$\sum_{x \geq 1} h_0(x) x^2 e^{\gamma_c x} < \infty, \quad (2.33)$$

(this is the discrete version of (2.12)), then

$$t_x = \frac{x}{v_c} + \frac{1}{\gamma_c v_c} \left[\frac{3}{2} \ln x + C + 3 \sqrt{\frac{2\pi v_c}{\gamma_c^3 v''(\gamma_c)}} x^{-1/2} + \dots \right], \quad (2.34)$$

where we were not able to compute the constant C . In (2.34), the position x is an integer, but if one forgets this point and formally inverts the asymptotic expansion to obtain x as a function of t , one obtains the same asymptotic expansion as (2.11).

The method we used to extract the asymptotic (2.34) from (2.32) consists in matching singularities in the left and right hand sides of (2.32) when $\lambda = e^{\gamma_c - \epsilon}$ with $\epsilon \rightarrow 0^+$. Indeed, if and only if (2.33) holds, the left hand side of (2.32) with $\lambda = e^{\gamma_c - \epsilon}$ must have two finite derivatives at $\epsilon = 0^+$:

$$\sum_{x \geq 1} h_0(x) e^{(\gamma_c - \epsilon)x} = a_0 + a_1 \epsilon + a_2 \epsilon^2 + o(\epsilon^2) \quad \text{with } a_i = \frac{1}{i!} \sum_{x \geq 1} h_0(x) (-x)^i e^{\gamma_c x}, \quad (2.35)$$

where the three numbers a_0 , a_1 and a_2 are finite by hypothesis. But we could show that without the $x^{-1/2}$ term of (2.34), there would be a $\epsilon^2 \ln \epsilon$ term appearing in the $\epsilon \rightarrow 0^+$ expansion of the right hand side of (2.32) with $\lambda = e^{\gamma_c - \epsilon}$, in contradiction with (2.35).

By generalizing this method, the vanishing corrections to the asymptotic expansion of t_x for large x can be obtained for any initial condition, not just when (2.33) is satisfied. In particular, one can recover a recent result [MM14] giving the first vanishing term for an initial condition satisfying $h_0(x) \sim e^{-\gamma_c x}$.

2.3 Summary, and a conjecture for the next order term

When computing the vanishing terms in the asymptotic expansion of $\mu_t^{(\alpha)}$, we obtained the same results for the two models introduced in the previous section. We now summarize these results. A conjecture for a $(\ln t)/t$ correction is also introduced.

From the results of [★BD15; ★BBHR15], the following picture emerges in FKPP fronts for the position $\mu_t^{(\alpha)}$ where the front is equal to α :

$$\begin{cases} \text{if } \int dx h_0(x) x^2 e^{\gamma_c x} < \infty & \mu_t^{(\alpha)} = v_c t - \frac{3}{2\gamma_c} \ln t + C^{(\alpha)} - 3 \sqrt{\frac{2\pi}{\gamma_c^5 v''(\gamma_c) t}} + \dots, \\ \text{if } h_0(x) \sim A x^\nu e^{-x} \text{ with } -3 \leq \nu < -2 & \mu_t^{(\alpha)} = v_c t - \frac{3}{2\gamma_c} \ln t + C^{(\alpha)} - D(\nu) t^{1+\frac{\nu}{2}} + \dots, \end{cases} \quad (2.36)$$

where, in the second line, the coefficient $D(\nu)$ depends on ν but not on α or on A . It is remarkable that we obtained the same result in both models, and it suggests strongly that these results are universal for FKPP fronts.

Notice that in the second line of (2.36), when $\nu = -3$, the first vanishing correction $D(-3)/\sqrt{t}$ is of the same order as the correction from Ebert and van Saarloos on the first line. The prefactor is however different.

In a paper with Julien Berestycki [BB16], we considered again the model (2.29) on the lattice and pushed the expansion of (2.32) with $\lambda = e^{\gamma_c - \epsilon}$ up to terms of order ϵ^3 . We noticed that an $\epsilon^3 \ln \epsilon$ term appears in the expansion unless one adds a $(\ln x)/x$ term to (2.34) with a well-chosen coefficient. By inverting (2.34) with this new term, we obtained a conjecture for the next-order term in the expansion of $\mu_t^{(\alpha)}$ for equations in the FKPP class: if and only if

$$\int dx h_0(x) x^3 e^{\gamma_c x} < \infty, \quad (2.37)$$

then

$$\mu_t^{(\alpha)} = v_c t - \frac{3}{2\gamma_c} \ln t + C^{(\alpha)} - 3 \sqrt{\frac{2\pi}{\gamma_c^5 v''(\gamma_c) t}} + \frac{54 - 54 \ln 2 + 3 \frac{v'''(\gamma_c)}{v''(\gamma_c)}}{4\gamma_c^3 v''(\gamma_c)} \frac{\ln t}{t} + \mathcal{O}\left(\frac{1}{t}\right). \quad (2.38)$$

(The condition (2.37) is the continuous equivalent of the condition needed for the $o(\epsilon^2)$ in (2.35) to be in fact a $a_3 \epsilon^3 + o(\epsilon^3)$.) For the FKPP equation, $\gamma_c = 1$, $v_c = 2$, $v''(\gamma_c) = 2$, $v'''(\gamma_c) = -6$, the conjecture (2.38) gives

$$\mu_t^{(\alpha)} = 2t - \frac{3}{2} \ln t + C^{(\alpha)} - \frac{3\sqrt{\pi}}{\sqrt{t}} + \frac{9}{8}(5 - 6 \ln 2) \frac{\ln t}{t} + \mathcal{O}\left(\frac{1}{t}\right). \quad (2.39)$$

The following term, of order $\mathcal{O}(1/t)$ cannot be universal: it is easy to see that it must depend on α , on the initial condition, etc.

Numerical simulations [BB16] are difficult to perform but give results compatible with (2.39).

Chapter 3

The limiting distribution of extremal points in the BBM

The positions of particles in the BBM can be seen as the actual positions of individuals in an ongoing invasion by a population reproducing without limits, or as the value of some traits of these individuals, or, in a completely different setting, as the energy spectrum of some disordered system similar to a spin glass. In any case, the extremal values are important to study and understand.

The previous chapter focused in great details on the position $\mu_t^{(1/2)}$ of a FKPP front, which is the same from McKean (1.6) as the median position of the rightmost particle. But what about the second rightmost particle, and the next ones? I have been working [★BD09; ★BD11; ABBS12] on the distribution of extremal points in a BBM as seen from the rightmost tip of the system. The results we obtained can be extended to any BRW [Mad15] and were the subject of a Bourbaki seminar [Gou13].

3.1 The limiting distribution seen from μ_t

We first consider the distribution of extremal points seen from the median position μ_t of the rightmost particle. McKean's formula is the keystone of this work.

3.1.1 Use and abuse of McKean's formula

Recall McKean's formula (1.7): if $h(x, t)$ is the solution of the FKPP equation with initial condition $h_0(x)$, then

$$\left\langle \prod_{u \in \mathcal{N}_t} (1 - h_0(x - X_u)) \right\rangle = 1 - h(x, t), \quad (3.1)$$

where the product on the left is over all the particles present at time t in a BBM and where X_u is the position of particle u . In particular, see (1.6),

$$h(x, t) = \mathbb{P}(R_t > x) \quad \text{if } h_0(x) = \mathbb{1}_{\{x < 0\}}, \quad (3.2)$$

because, for that choice of h_0 , the product in (3.1) is 1 if no X_u is larger than x and 0 otherwise.

In two papers [★BD09; ★BD11], Bernard Derrida and I have shown how to use (3.1) to obtain a description of the statistics of extremal points; the same method was later used by Arguin, Bovier and Kistler to establish a proven rigorous result [ABK11; ABK12; ABK13b]. To illustrate the method, consider the following example; pick $\lambda \in [0, 1)$ and assume that h_0 is given by

$$h_0(x) = \begin{cases} 1 - \lambda & \text{if } x < 0, \\ 0 & \text{if } x \geq 0. \end{cases} \quad (3.3)$$

The product in (3.1) is simply $\lambda^{N(x,t)}$, where we introduce

$$N(x, t) = [\text{number of particles on the right of } x \text{ at time } t \text{ in the BBM}]. \quad (3.4)$$

Then, of course,

$$h(x, t) = 1 - \langle \lambda^{N(x,t)} \rangle. \quad (3.5)$$

A small λ expansion gives

$$h(x, t) = 1 - \mathbb{P}(N(x, t) = 0) - \lambda \mathbb{P}(N(x, t) = 1) + \mathcal{O}(\lambda^2) = \mathbb{P}(R_t > x) - \lambda \mathbb{P}(S_t \leq x < R_t) + \mathcal{O}(\lambda^2), \quad (3.6)$$

where S_t is the position of the second rightmost particle. At time t , there is a probability e^{-t} that there is still only one particle in the BBM, and then the position S_t of the second rightmost particle is not defined. The expression above is still valid with the convention that $S_t = -\infty$ if the number of particles is still 1.

Differentiate with respect to x

$$\partial_x h(x, t) dx = -\mathbb{P}(R_t \in dx) + \lambda [\mathbb{P}(R_t \in dx) - \mathbb{P}(\mathcal{N}_t > 1 \text{ and } S_t \in dx)] + \mathcal{O}(\lambda^2), \quad (3.7)$$

where \mathcal{N}_t is also the number of particles in the system at time t . Multiply by x and integrate:

$$\int dx x \partial_x h(x, t) = -\langle R_t \rangle + \lambda [\langle R_t \rangle - \langle \mathbb{1}_{\{\mathcal{N}_t > 1\}} S_t \rangle] + \mathcal{O}(\lambda^2). \quad (3.8)$$

The coefficient of λ is the expected distance between the rightmost and second rightmost particles in the BBM at time t (with a small and vanishing subtlety: there is a small probability e^{-t} that the second rightmost does not exist yet). Similarly, one checks that the term of order λ^2 gives the expected distance between the second and third rightmost particles, etc.

Note that the small λ expansion can be performed *a priori*: using $h(x, t) = A(x, t) - \lambda B(x, t) + \mathcal{O}(\lambda^2)$ in the FKPP equation, one gets with (3.3):

$$\partial_t A = \partial_x^2 A + A - A^2, \quad A_0(x) = \mathbb{1}_{\{x < 0\}}, \quad \partial_t B = \partial_x^2 B + B - 2AB, \quad B_0(x) = \mathbb{1}_{\{x < 0\}}, \quad (3.9)$$

and the expected distance $\langle R_t \rangle - \langle \mathbb{1}_{\{\mathcal{N}_t > 1\}} S_t \rangle$ is simply $-\int dx x \partial_x B$. It is easy to integrate numerically the coupled equations (3.9) to compute the expected distance at any time.

By choosing other initial conditions h_0 , one can compute other statistical properties of the rightmost particles in the BBM at any time t . For instance, by choosing

$$h_0(x) = \begin{cases} 1 & \text{if } x < -a, \\ 1 - \lambda & \text{if } -a \leq x < 0, \\ 0 & \text{if } x \geq 0, \end{cases} \quad (3.10)$$

one obtains from (3.1)

$$h(x, t) = 1 - \langle \lambda^{N(x, t)} \mathbb{1}_{\{R_t \leq x + a\}} \rangle. \quad (3.11)$$

From there, it is quite easy to get the joint probability distribution of the positions of the rightmost and the n -th rightmost particles at any time t ; in particular, one can obtain the probability distribution function of the distance between the rightmost and the n -th rightmost particles at time t .

By choosing more and more complicated initial conditions, one can obtain in such a way the joint probability distribution of the positions of the n rightmost particles at time t for any value of n . It is sufficient, actually, to consider initial conditions h_0 such that $h_0(x)$ is constant by pieces and $h_0(x) = 0$ for $x \geq 0$, similarly to (3.10).

3.1.2 The large time limit

Bramson's result allows to conclude that the distribution of points at the tip in the BBM does reach a large time limit. Indeed, let

$$\mu_t = [\text{median position of the rightmost particle in the BBM at time } t] = 2t - \frac{3}{2} \ln t + C + o(1). \quad (3.12)$$

(When considering FKPP with a step initial condition, then μ_t is where the front is equal to $\frac{1}{2}$. In the previous section, we called this quantity $\mu_t^{(1/2)}$.) Recall, from McKean's relation (3.2) and from Bramson's convergence theorem (2.5), that one has

$$\mathbb{P}(R_t > \mu_t + z) = h(\mu_t + z, t) \xrightarrow{t \rightarrow \infty} \omega(z), \quad \text{for a front } h \text{ started from } h_0(x) = \mathbb{1}_{\{x < 0\}}. \quad (3.13)$$

The meaning of (3.13) is that the distribution of R_t relative to the median position μ_t converges in law.

When considering other statistical properties of the positions of the particles at the tip of the BBM, for instance the joint probability distribution of the positions of the n rightmost particles, one needs to consider initial conditions such as (3.10) which are equal to zero for $x \geq 0$. These initial conditions certainly decay fast enough to apply Bramson's result and one knows, see Section 1.2, that $h(m_t + z, t) \rightarrow \omega(z)$ if and only if $m_t = \mu_t - C[h_0] + o(1)$, where $C[h_0]$ is a constant which depends on the initial condition h_0 . This leads to

$$h(\mu_t + z, t) \xrightarrow{t \rightarrow \infty} \omega(z + C[h_0]), \quad \text{for a front } h \text{ started from any } h_0 \text{ that decays fast enough.} \quad (3.14)$$

In words, pick a h_0 such as (3.10); both the front started from h_0 and the front started from the step initial condition converge to the same shape. However, as h_0 is smaller than the step initial condition, the front started from h_0 remains behind the front started from the step initial condition. In the large time limit, both fronts have the shape $\omega(z)$, but shifted from each other by the quantity $C[h_0] \geq 0$ which represents the large time delay, or lag, taken by the front started from h_0 as compared to the front started from the step initial condition.

One can combine (3.14) with the conclusions of McKean's relation; for instance, with h_0 given by (3.10), one gets from (3.11)

$$\left\langle \lambda^{N(\mu_t+z,t)} \mathbb{1}_{\{R_t \leq \mu_t+z+a\}} \right\rangle \xrightarrow[t \rightarrow \infty]{} 1 - \omega(z + C[h_0]). \quad (3.15)$$

The delay $C[h_0]$ is an unknown function of λ and of a . From the left hand side, one can extract the joint probability density that, at time t , the rightmost particle is at position $\mu_t + X$ and the n -th rightmost at position $\mu_t + Y$. The right hand side means that this joint probability density has a large time limit which could be deduced from the knowledge of $C[h_0]$.

Similarly, by picking more complicated h_0 , one can show that the joint probability distribution of the positions of the n rightmost particles relative to μ_t reaches a large time limit. Furthermore, the description of this limiting distribution is entirely contained in the knowledge of the lags $C[h_0]$ for a well chosen family of initial conditions h_0 .

Using this method, we studied in [★BD09] the front started from (3.3) and showed that, for this choice of h_0 , one had $C[h_0] = -\ln(1-\lambda) - \ln[-\ln(1-\lambda)] + \mathcal{O}(1)$ as $\lambda \rightarrow 1$. This allowed us to conclude that the expected distance in the limiting distribution between the n -th and $(n+1)$ -th rightmost particles in the BBM was, for n large, of order $1/n - 1/[n \ln n]$. By numerical integration of the FKPP equation, we also computed in [★BD11] the limiting probability distribution function of the distance $d_{1,2}$ between the two rightmost particles. We found a distribution which is extremely close to a simple exponential distribution: $\mathbb{P}(d_{1,2} \in da) \approx 2e^{-2a} da$, see Figure 3.1. However, the distribution of $d_{1,2}$ is not so simple [★BD11] and we found from a large deviation analysis (confirmed by the numerical simulations) that the tail of the distribution is of order $e^{-(1+\sqrt{2})a}$ for large a . A precise measure of the expectation of $d_{1,2}$ gives

$$\langle d_{1,2} \rangle = 0.497\dots \quad (3.16)$$

We still do not understand why the distribution of $d_{1,2}$ is so close to a simple exponential for its typical values.

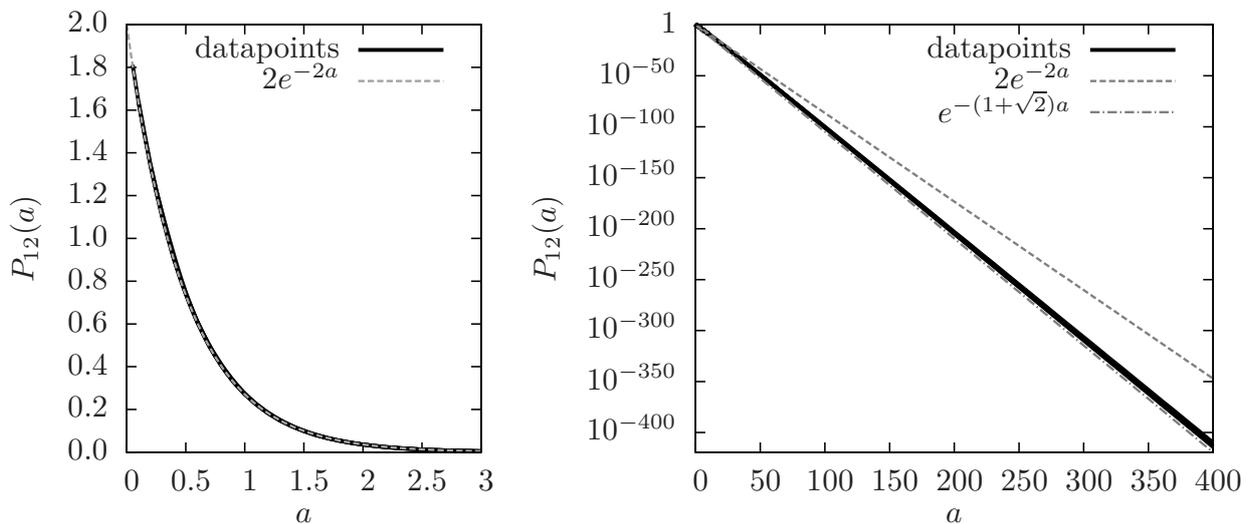


Figure 3.1: The limiting probability density that the distance $d_{1,2}$ between the two rightmost particles in the BBM is equal to a , compared to $2e^{-2a}$ and to $e^{-(1+\sqrt{2})a}$. The left graph focuses on short values of a while the right graph focuses on large values.

The curves in Figure 3.1 were obtained with the method outlined in this section by numerical integration of the (deterministic) FKPP front equation with various initial conditions h_0 in order to compute the delays $C[h_0]$. A direct Monte Carlo simulation of the (stochastic) BBM would never have allowed to reach such an accuracy.

3.2 Non-ergodicity and random shift

Lalley and Sellke have shown that the position of the rightmost particle around its median was not ergodic, but was in fact given by a random shift and a fluctuating Gumble. We extend this result to the limiting distribution of extremal points.

3.2.1 Lalley's and Sellke's result

The position of the rightmost particle relative to its median position converges in law to the distribution defined by the critical travelling wave $\omega(x)$, see (3.13). As noticed by Lalley and Sellke in 1987 [LS87], the distribution of the rightmost particle is not ergodic:

$$\lim_{t \rightarrow \infty} \mathbb{P}(R_t > \mu_t + z) = \omega(z) \neq [\text{fraction of the time where } R_t > \mu_t + z \text{ for one given realization}]. \quad (3.17)$$

The reason for this is that even at large times, the BBM “remembers” what happens at small times. For instance, consider a realization where by chance all the particles drift to the left while there are few of them. In such a realization, the BBM is at all times shifted to the left, and the rightmost particle spends most of its time on the left of the median position μ_t ; this is illustrated in Figure 3.2

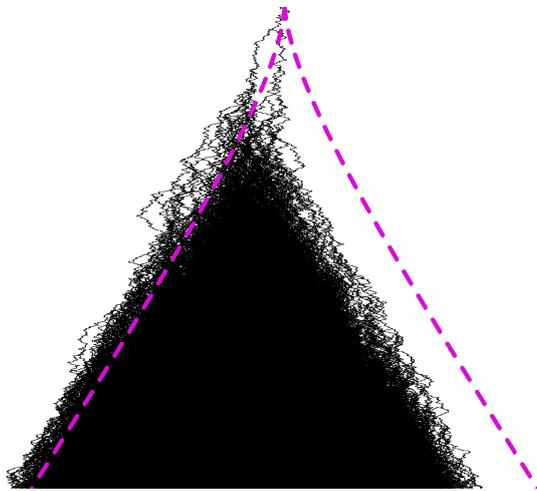


Figure 3.2: A realization of the BBM up to time 15. Time goes downwards and space is horizontal. The dashed lines represent the median positions of the leftmost and rightmost particles. For this realization, both the leftmost and the rightmost particles are nearly always on the left of their median positions.

Lalley and Sellke [LS87] gave a nice description of the position R_t of the rightmost particle, which amounts to the following: for each realization of the BBM, there exists a random variable Z (which is a function of the BBM) such that

$$R_t = \mu_t + \ln(AZ) + g_t, \quad (3.18)$$

where μ_t is the median of R_t , A is the constant appearing in $\omega(z) \sim Aze^{-z}$ for large z , see (2.6), and where, for large time, g_t is a Gumble variable:

$$\lim_{t \rightarrow \infty} \mathbb{P}(g_t \leq z) = e^{-e^{-z}}. \quad (3.19)$$

For instance, the value of $\ln(AZ)$ appearing in (3.18) is negative for the realization of Figure 3.2. Lalley and Sellke conjectured that g_t was ergodic (meaning that, for a given realization, the fraction of the time where g_t is smaller than z is also given by $\exp(-e^{-z})$), but this was proved only quite recently [ABK13a].

The variable Z is the large time limit of the so called derivative Martingale [Nev88]

$$Z_t = \sum_{u \in \mathcal{N}_t} (2t - X_u) e^{X_u - 2t}, \quad Z = \lim_{t \rightarrow \infty} Z_t. \quad (3.20)$$

(Recall that \mathcal{N}_t is the set of all particles present at time t and X_u for $u \in \mathcal{N}_t$ is the position of particle u .) One shows that Z_t has a finite and positive large time limit Z with probability 1. This limit depends (of course) on the realization of the BBM and, in fact, mostly depends on what happens in the early stages of the BBM, when the number of particles is small. Combining (3.13), (3.18) and (3.19), one can relate the distribution of Z to the travelling wave [LS87]:

$$\omega(z) = 1 - \langle e^{-AZe^{-z}} \rangle. \quad (3.21)$$

3.2.2 Generalization to the extremal process

As explained above, in the BBM, the position of the rightmost particle for large times fluctuates as an ergodic Gumble random variable around position $\mu_t + \ln(AZ)$, where μ_t is the (deterministic) median position of the rightmost particle and $\ln(AZ)$ is a random shift which is decided at the early stages of the BBM. Actually, this random shift also applies to the whole stationary measure of the positions of the rightmost particles. Writing

$$[\text{position of the } n\text{-th rightmost particle at time } t] = \mu_t + \ln(AZ) + Y^{(n)}(t), \quad (3.22)$$

we have argued [★BD11] that the joint probability distribution of $\{Y^{(1)}(t), \dots, Y^{(n)}(t)\}$ for any n reaches a large time limit which is independent of Z . We conjectured that the distribution of the $Y^{(n)}$ was ergodic, but this was proved only in [ABK15]. The positions of the rightmost points have therefore a deterministic contribution μ_t , a random shift $\ln(AZ)$ which is decided at early times, and random fluctuations $Y^{(n)}(t)$ which depend on the recent history of the process.

To obtain this result, we combined Lalley's and Sellke's proof [LS87] with the methods exposed in Section 3.1. Pick an initial condition $h_0(x)$ which looks like (3.3) or (3.10), and start a BBM with a single particle at the origin. Write an expression for the expectation of $\prod_u [1 - h_0(x - X_u)]$ at time t as in McKean's relation (3.1), but this time condition about the state \mathcal{F}_s of the BBM at a given time $s < t$:

$$\left\langle \prod_{u \in \mathcal{N}_t} (1 - h_0(x - X_u)) \middle| \mathcal{F}_s \right\rangle = \prod_{w \in \mathcal{N}_s} \left\langle \prod_{\substack{u \in \mathcal{N}_t \\ u \text{ desc. of } w}} (1 - h_0(x - X_u)) \middle| X_w \right\rangle, \quad (3.23)$$

$$= \prod_{w \in \mathcal{N}_s} (1 - h(x - X_w, t - s)), \quad (3.24)$$

where h is the solution of the FKPP with an initial condition h_0 . Indeed, in the left hand side, the product can be split into products over the descendants of each of the \mathcal{N}_s particles present at time s as in the right hand side of (3.23). Each of these particles spawns, independently of the others, a simple BBM started from their position X_w , and one applies McKean's formula (3.1) for each of these BBM.

Use $x = \mu_t + z$ in (3.24) and take the large time limit:

$$\lim_{t \rightarrow \infty} \left\langle \prod_{u \in \mathcal{N}_t} (1 - h_0(\mu_t + z - X_u)) \middle| \mathcal{F}_s \right\rangle = \prod_{w \in \mathcal{N}_s} \lim_{t \rightarrow \infty} (1 - h(\mu_{t-s} + (\mu_t - \mu_{t-s}) + z - X_w, t - s)), \quad (3.25)$$

$$= \prod_{w \in \mathcal{N}_s} (1 - \omega(2s + z - X_w + C[h_0])), \quad (3.26)$$

where we used Bramson's convergence with a lag (3.14), and also that $\mu_t - \mu_{t-s} \rightarrow 2s$ as $t \rightarrow \infty$, see (3.12). We now take s large. As the rightmost particle at time s is around position $2s - \frac{3}{2} \ln s + C$, one knows that $2s - X_w$ must be large for all $w \in \mathcal{N}_s$. Writing that $\omega(z) \sim (Az + B)e^{-z}$ for large z , one gets from (3.26) for large s :

$$\lim_{t \rightarrow \infty} \left\langle \prod_{u \in \mathcal{N}_t} (1 - h_0(\mu_t + z - X_u)) \middle| \mathcal{F}_s \right\rangle \sim \prod_{w \in \mathcal{N}_s} \left(1 - [A(2s + z - X_w + C[h_0]) + B] e^{-2s - z + X_w - C[h_0]} \right), \quad (3.27)$$

$$\sim \exp \left(- \sum_{w \in \mathcal{N}_s} [A(2s + z - X_w + C[h_0]) + B] e^{-2s - z + X_w - C[h_0]} \right),$$

$$\sim \exp \left(- [AZ_s + (B + z + C[h_0])Y_s] e^{-z - C[h_0]} \right), \quad (3.28)$$

with Z_s the derivative Martingale as in (3.20) and where $Y_s = \sum_{w \in \mathcal{N}_s} e^{X_w - 2s}$ is the critical additive Martingale. With probability one, for large s , one has [LS87] $Y_s \rightarrow 0$ and $Z_s \rightarrow Z > 0$, and hence

$$\lim_{s \rightarrow \infty} \lim_{t \rightarrow \infty} \left\langle \prod_{u \in \mathcal{N}_t} (1 - h_0(\mu_t + z - X_u)) \middle| \mathcal{F}_s \right\rangle = \exp(-AZe^{-z - C[h_0]}) = \exp(-e^{-z - C[h_0] + \ln(AZ)}). \quad (3.29)$$

By choosing different values of h_0 , the left hand side gives some information of the limiting distribution as $t \rightarrow \infty$ of the rightmost particles conditioned by the state at some large time s of the BBM. For instance, with h_0 as in (3.10), one gets

$$\lim_{s \rightarrow \infty} \lim_{t \rightarrow \infty} \left\langle \lambda^{N(\mu_t + z, t)} \mathbb{1}_{\{R_t \leq \mu_t + z + a\}} \middle| \mathcal{F}_s \right\rangle = \exp(-e^{-z - C[h_0] + \ln(AZ)}), \quad (3.30)$$

which allows to determine the limiting joint probability distribution of the rightmost and n -th rightmost particles. This distribution features a random shift $\ln(AZ)$, and is otherwise entirely determined by the knowledge of the delays $C[h_0]$.

3.3 The structure of the limiting point process

The stationary point process of the rightmost particles seen from $\mu_t + \ln(AZ)$ can be described as a juxtaposition of independent families

3.3.1 Stability by juxtaposition and decorated Poisson point processes

At large times, the rightmost particles in the BBM can be described by a random limiting point process ν which is shifted by a deterministic time-dependent value μ_t and by a random (realization-dependent) constant value $\ln(AZ)$, see (3.22). The rightmost point in this limiting point process is Gumble distributed, see (3.18) and (3.19).

The limiting point process of the BBM must feature a remarkable property: consider the BBM just after its first branching event: there are two particles (say, a red and a blue particles) which evolve independently and spawn a red and a blue BBM. When time is large, the rightmost red particles are distributed according to the randomly shifted limiting point process; let ν_{red} be the realization of the process and $\mu_t + a_{\text{red}}$ be the random shift. The rightmost blue particles are also distributed according to the randomly shifted limiting point process, and we call ν_{blue} the realization of the process and $\mu_t + a_{\text{blue}}$ the random shift.

Of course, taken together, the rightmost particles (no matter whether they are blue or red) form a third realization ν_{all} of the limiting point process, shifted by $\mu_t + a_{\text{all}}$. Then

$$\{\nu_{\text{all}} \text{ shifted by } a_{\text{all}}\} = \{\nu_{\text{red}} \text{ shifted by } a_{\text{red}}\} \cup \{\nu_{\text{blue}} \text{ shifted by } a_{\text{blue}}\}. \quad (3.31)$$

This is illustrated in Figure 3.3.

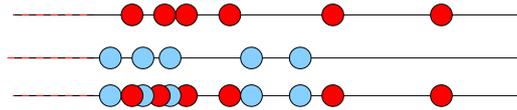


Figure 3.3: The superposability property: the shifted red realization (first line) and the shifted blue realization (second line) of the limiting point process taken together give a third shifted realization (third line) of the same point process.

But $a_{\text{all}} = \ln(AZ_{\text{all}})$. By decomposing Z over the red and blue particles and writing $Z_{\text{all}} = Z_{\text{red}} + Z_{\text{blue}}$, one gets easily that

$$a_{\text{all}} = \ln(e^{a_{\text{red}}} + e^{a_{\text{blue}}}). \quad (3.32)$$

Because the realizations ν_{red} and ν_{blue} of the process and the shifts a_{red} and a_{blue} are independent, (3.31) means that the superposition of two independent realizations of the point process ν shifted by two arbitrary amounts a_{red} and a_{blue} must be equal in law to the same point process shifted by another amount given by (3.32). We called this the “superposability” property [★BD11]; a better name seems to be “exp-1-stability” [Mai13].

There exists a very simple exp-1-stable point process: it is the Poisson point process of density $\rho(x) = e^{-x}$, which we call the “exponential Poisson point process”. Indeed, $\rho(x-a) + \rho(x-b) = \rho(x-c)$ with $c = \ln(e^a + e^b)$, which yields the exp-1-stability property (3.31) with (3.32).

Starting from the exponential Poisson point process, it is not difficult to build other exp-1-stable point processes. For instance, the exponential Poisson point process where each point is replaced by two points at the same position is obviously exp-1-stable. More generally, choose an arbitrary point process σ such that there is always one point at the origin and no point for positive positions; consider then the “ σ -decorated exponential Poisson point process” defined in the following way (see also Figure 3.4):

1. Pick a realization $\{X_1 > X_2 > X_3 > \dots\}$ of the exponential Poisson point process,
 2. For each point X_i , pick an independent realization σ_i of the decorating point process σ ,
 3. The realization of the process is given by $\bigcup_i \{\sigma_i \text{ shifted by } X_i\}$.
- (3.33)

Obviously, any σ -decorated exponential Poisson point process is exp-1-stable. Conversely, it turns out that any exp-1-stable point processes can be built as a σ -decorated exponential Poisson point process [Mai13; DMZ08]; we conjectured that result in [★BD11]. Hence, there must exist a point process σ , the decoration, such that the limiting point process of the BBM can be described as a σ -decorated exponential Poisson point process (3.33).

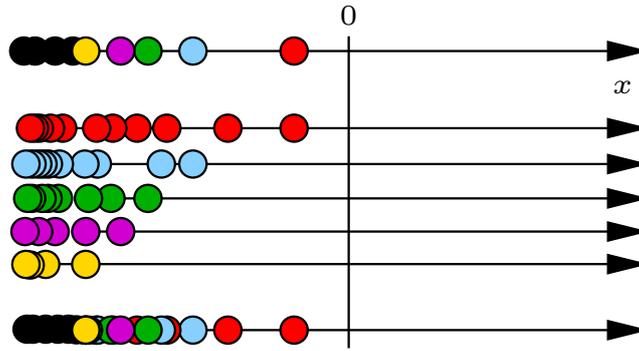


Figure 3.4: A realization of the σ -decorated exponential Poisson point process (3.33). On the first line, a realization of the exponential Poisson point process is picked. On the five intermediate lines, for each of the five rightmost points of the exponential Poisson point process, a realization of the decoration point process σ is picked and then appropriately shifted. On the last line, all the points are put together.

3.3.2 Families and leaders

The most precise description known up to now of the limiting point process of the rightmost points in the BBM is then the following: there exists a decoration σ such that the distribution of the rightmost points converge to a σ -decorated exponential Poisson point process shifted by $\mu_t + \ln(AZ)$, where we recall that μ_t is the median position of the rightmost particle and $\ln(AZ)$ is some random variable.

The rightmost particles in the BBM are then naturally decomposed into “clusters”, or “families”: two particles are in the same family if they are both parts of the same decoration. Each family has a “leader”, which is the rightmost particle of each family. All the leaders taken together are the points in the exponential Poisson point process.

Derrida and Spohn had already shown [DS88] in 1988 that two particles in a BBM picked around the right tip have their first common ancestor either a “short” time ago (within a time of order 1 from the current time) or a “long” time ago (within a time of order 1 from the origin of times). This defines equivalence classes of particles at the tip having branched recently. Arguin Bovier and Kistler [ABK11; ABK12] have proved that in fact, the equivalence classes of particles at the tip having branched recently and the families of the σ -decorated exponential Poisson point process are the same thing: two particles at the tip have their first common ancestor a short time ago if and only if they are in the same family. Otherwise, their first common ancestor is close to the beginning of the process. This leads to two natural implicit descriptions of the point measure σ :

- One can look at the path followed by one leader, in reverse time [ABBS12]. There is some branching on this line which generate sub-trees (in forward time), and the set of all the particles issued from all the sub-trees form the decoration, because only one family was constructed, see Figure 3.5. However, the law of the reversed path followed by the leader, the branching law and the law for generating the sub-trees are complicated because they must ensure that, in the end, the leader is indeed their rightmost particle.

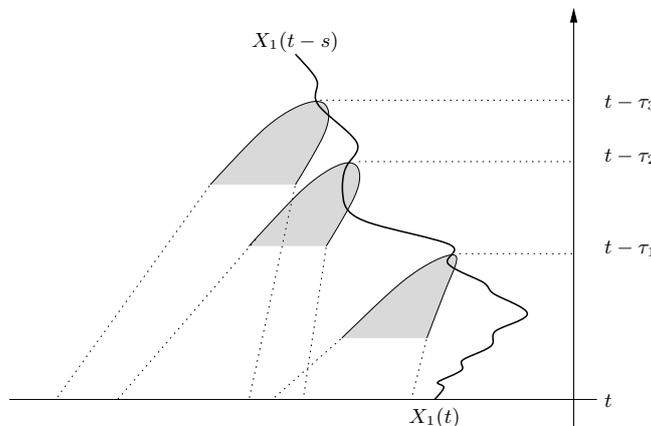


Figure 3.5: From the position $X_1(t)$ of the leader of one family at time t , the trajectory of the leader can be reconstructed backward in time. Some branching occur on the line and the descendants of these branching events are on the left of $X_1(t)$ at time t .

- As shown by Arguin Bovier and Kistler [ABK13b], one can consider a BBM conditioned to have its

rightmost particle at time t on the right of $2t$. This BBM goes unusually fast (the rightmost should be around $2t - \frac{3}{2} \ln t$), so that, with overwhelming probability, only one family is present (it would be too expensive to have two independent lines going unusually fast). However, because this BBM does not go too fast (the velocity is not modified), the law of distances between the rightmost particles and the other particles in its family is not modified. Hence, the set of rightmost points in this fast-but-not-too-fast BBM has the same law as the decoration σ .

Unfortunately, these two descriptions of the decoration σ are not very explicit, and it is difficult to extract useful information from them. The methods discussed in Section 3.1 can still be used to recover some information on σ ; for instance, recall that $d_{1,2}$ is the distance for large times between the two rightmost particles in the BBM, and call $d_{1,2}^\sigma$ the distance between the leader and the next particle in the decoration σ . Recall that the probability that the two rightmost points in an exponential Poisson point process have a distance larger than a is e^{-a} ; then:

$$\mathbb{P}(d_{1,2} > a) = \mathbb{P}(d_{1,2}^\sigma > a) \times e^{-a}. \quad (3.34)$$

Indeed, $d_{1,2} > a$ if and only if the distance between the rightmost particle and its first runner-up in the same family is larger than a (first term in the right hand side) and if, at the same time, the distance between the two rightmost leaders is larger than a (second term). These two events are independent, hence (3.34). Then, from Figure 3.1 and the discussion above it, one sees that the density of probability that $d_{1,2}^\sigma$ is around a is nearly given by e^{-a} when a is not too large, but is of order $\exp(-\sqrt{2}a)$ for large values of a .

One can draw another unexpected relation from (3.34); the probability that the two rightmost particles are of two different families is clearly

$$\begin{aligned} \mathbb{P}(\text{the 2 rightmost are in different families}) &= \int_0^\infty \mathbb{P}(\text{the 2 rightmost are in different families and } d_{1,2} \in da), \\ &= \int_0^\infty \mathbb{P}(d_{1,2}^\sigma > a) \times e^{-a} da, \end{aligned} \quad (3.35)$$

$$= \int_0^\infty \mathbb{P}(d_{1,2} > a) da, \quad (3.36)$$

$$= \langle d_{1,2} \rangle = 0.497\dots, \quad (3.37)$$

see (3.16) for the last equality. Similarly, with a bit more work, one obtains

$$\mathbb{P}(\text{the } n \text{ rightmost particles in the BBM are in } n \text{ different families}) = \langle d_{1,2} \rangle^{n-1}. \quad (3.38)$$

There are probably more relations such as (3.34) or (3.38) that one could derive in order to gain some partial knowledge on the decoration σ , but the truth is that our understanding of σ is very incomplete.

Chapter 4

Position and genealogy in the noisy FKPP equation

Consider the microscopic models described in Sections 1.5.2 (directed polymers) and 1.5.3 (population dynamics): there is a cloud of N points or sites or particles following a Markovian evolution. At each time-step, each particle is replaced by a number of children, and the position of each child is the position of the parent plus some random number. The total number N of particles remains constant. The two models differ in the way the number of children of each particle is chosen; in both cases, the particles on the right tend to have more children than the particles on the left. In either model, one introduces $h(x, t)$ the fraction of particles on the right of x at time t , and one finds that $h(x, t)$ follows a stochastic equation:

- For the model of Section 1.5.2 (directed polymers), to find the position of particle i at time $t + 1$, one picks randomly $k > 1$ prospective parents $j_{1;i,t}, \dots, j_{k;i,t}$ at generation t (we used $k = 2$ in the introduction), and the position $X_i(t + 1)$ of the new particle is the largest of the k positions of the parents shifted by random amounts $\epsilon_{1;i,t}, \dots, \epsilon_{k;i,t}$. One finds, see (1.50),

$$\begin{aligned} X_i(t + 1) &= \max [X_{j_{1;i,t}}(t) + \epsilon_{1;i,t}, \dots, X_{j_{k;i,t}}(t) + \epsilon_{k;i,t}], \\ h(x, t + 1) &= 1 - \left(1 - \int d\epsilon \rho(\epsilon) h(x - \epsilon, t)\right)^k + \text{noise}. \end{aligned} \quad (4.1)$$

- For the model of Section 1.5.3 (population dynamics), each particle first has $k > 1$ children (we used $k = 2$ in the introduction), whose positions are the position of the parent plus some random amounts. The population size is then brought back from kN to N by keeping only the N rightmost individuals. One finds, see (1.53),

$$\begin{aligned} [\text{population at time } t + 1] &= [\text{the } N \text{ rightmost of } \{X_i(t) + \epsilon_{i,j;t}, i \in \{1, \dots, N\}, j \in \{1, \dots, k\}\}], \\ h(x, t + 1) &= \min \left[1, k \int d\epsilon \rho(\epsilon) h(x - \epsilon, t) + \text{noise}\right]. \end{aligned} \quad (4.2)$$

In both (4.1) or (4.2), the noise term at a position x where h is small is of order $\sqrt{h(x, t)/N}$. It is non-Gaussian and correlated in space, because it ensures that $x \mapsto h(x, t)$ is a non-increasing function which only takes values that are multiples of $1/N$. As $N \rightarrow \infty$, the noise term goes to zero and one is left with a deterministic equation which behaves like the FKPP for a large class of $\rho(\epsilon)$. (See [Aid13]. In particular, any non-lattice distribution with exponential tails works. We always assume $\rho(\epsilon)$ to be in that class.) These equations are, respectively,

$$h(x, t + 1) = 1 - \left(1 - \int d\epsilon \rho(\epsilon) h(x - \epsilon, t)\right)^k, \quad \text{and} \quad h(x, t + 1) = \min \left[1, k \int d\epsilon \rho(\epsilon) h(x - \epsilon, t)\right]. \quad (4.3)$$

In the large time limit, for a step initial condition, a front described by one of the equations in (4.3) moves at the critical velocity v_c given by

$$v_c = \min_{\gamma > 0} v(\gamma) = v(\gamma_c), \quad \text{where } v(\gamma) = \frac{1}{\gamma} \ln \left[k \int d\epsilon \rho(\epsilon) e^{\gamma \epsilon} \right], \quad (4.4)$$

see Section 1.3. It is the same function $v(\gamma)$ in both models because the linearised equations are the same. The limiting shape $\omega(z)$ of the front depends however on which equation is chosen, but in any case one has

$$h(\mu_t + z, t) \xrightarrow[t \rightarrow \infty]{} \omega(z), \quad \text{with } \omega(z) \sim Aze^{-\gamma_c z} \text{ for large } z \quad [\text{for (4.3), without the noise term}], \quad (4.5)$$

with μ_t the position of the front.

The question, naturally, is what happens for noisy equations such as (4.1) or (4.2) when N is large but finite.

4.1 The position of the front

Because of the noise, the velocity of the front is significantly smaller than v_c , even for large N . Furthermore, the front diffuses. We introduce a phenomenological description to compute the velocity correction and the diffusion constant. In order to justify some step, several models which are nearly (but not quite!) in the FKPP case must be investigated.

We focus on the models (4.1) and (4.2) which describe a cloud of N particles. The first thing to remark is that the N particles stay together: the distance between the rightmost and the leftmost particles reaches for large times a (N dependent) stationary distribution. Indeed, if the N particles were separated into several well-distant clouds, the particles in the rightmost cloud would have more descendants than the particles in the leftmost clouds and, as the population is conserved, the leftmost clouds would eventually disappear.

Define the “position of the cloud” μ_t at time t as the empirical average of the positions of the N particles:

$$\mu_t = \frac{1}{N} \sum_{i=1}^N X_i(t). \quad (4.6)$$

One expects the cloud of particles to have a velocity v_N which is close to v_c :

$$\lim_{t \rightarrow \infty} \frac{\mu_t}{t} = \lim_{t \rightarrow \infty} \frac{\langle \mu_t \rangle}{t} = v_N, \quad \lim_{N \rightarrow \infty} v_N = v_c. \quad (4.7)$$

Furthermore, the position of the cloud diffuses with some small diffusion constant D_N :

$$\lim_{t \rightarrow \infty} \frac{\langle \mu_t^2 \rangle - \langle \mu_t \rangle^2}{t} = D_N > 0, \quad \lim_{N \rightarrow \infty} D_N = 0. \quad (4.8)$$

It is clear indeed that the system is diffusive: as the width of the front remains finite, the distribution of the N points around μ_t must converge to some stationary regime. Then, for T large enough, the position increments $\mu_{nT+T} - \mu_{nT}$ for $n \in \mathbb{N}$ are nearly independent and identically distributed random variables. μ_t can then (nearly) be seen as the sum of t/T identical independent variables and, hence, the variance of μ_t must scale like t . This argument also allows to conclude that all the cumulants of μ_t must asymptotically increase like t .

Notice also that one could have defined μ_t as the position of the rightmost particle, for instance, or of the leftmost particle: as the width of the cloud remains finite, all these possible definitions of the position differ by some quantity that does not grow in time, so that the values of the velocity v_N and of the diffusion constant D_N in (4.7) and (4.8) are not affected.

4.1.1 The cut-off theory

During my PhD, my supervisor Bernard Derrida and I proposed [BD97] the cut-off theory as a description of the main mechanism to understand the asymptotic value of $v_c - v_N$.

The idea was to notice that one of the effects of the noise in (4.1) or (4.2) is to ensure that $h(x, t)$ is for all x and t an integer multiple of $1/N$, because $h(x, t)$ is defined as the fraction of the number of particles on the right of x . In particular, if $h(x, t)$ is non-zero, it cannot be smaller than $1/N$. It is not hard to see that this last fact alone must have a huge effect on the mechanism of velocity selection because, as we have seen in Section 1.2.2, it is the asymptotic behaviour of the front for large x that determines the velocity.

For this reason, we proposed to replace the noise in noisy FKPP equations by a cut-off at $1/N$; for instance, (4.1) with $k = 2$ becomes

$$h(x, t+1) = \begin{cases} 1 - \left(1 - \int d\epsilon \rho(\epsilon) h(x - \epsilon, t)\right)^2 & \text{if this is larger than } \frac{1}{N}, \\ 0 & \text{otherwise.} \end{cases} \quad (4.9)$$

This equation is deterministic: the noise has been completely removed and replaced by the cut-off. Let $v_N^{(\text{cut-off})}$ be the velocity of (4.9); we conjectured (and checked numerically) that for large N

$$v_c - v_N \sim v_c - v_N^{(\text{cut-off})}. \quad (4.10)$$

Let us compute $\Delta = v_c - v_N^{(\text{cut-off})}$, which we expect to be a small positive number for large N . We look for a travelling wave $\tilde{\omega}(z)$ such that $h(x, t) = \tilde{\omega}(x - (v_c - \Delta)t)$ is solution to (4.9). Let L be the position of the cut-off ($\tilde{\omega}$ and L depend implicitly on N); one has

$$\begin{cases} \tilde{\omega}(z) = 1 - \left(1 - \int d\epsilon \rho(\epsilon) \tilde{\omega}(z + v_c - \Delta - \epsilon)\right)^2 & \text{if } z < L, \\ \tilde{\omega}(z) = 0 & \text{if } z > L, \\ \tilde{\omega}(L) = \frac{1}{N}. \end{cases} \quad (4.11)$$

Compare this to the equation satisfied by the travelling wave ω_v at velocity v for the equation without cut-off:

$$\omega_v(z) = 1 - \left(1 - \int d\epsilon \rho(\epsilon) \omega_v(z + v - \epsilon)\right)^2. \quad (4.12)$$

We recall from Section 1.2.2 that a linear analysis of ω_v leads for large z to:

$$\begin{cases} \omega_v(z) = [A_v + o(1)]e^{-\gamma z} & \text{if } v > v_c, \text{ with } \gamma \text{ smallest solution of } v(\gamma) = v, \\ \omega_{v_c}(z) = [Az + o(1)]e^{-\gamma_c z} & \text{if } v = v_c, \end{cases} \quad (4.13)$$

with $A > 0$, $A_v > 0$ and with $v(\gamma)$ defined in (4.4). (Because of the invariance by translation of (4.12), there are many travelling wave solutions obtained by shifting the origin of z . In the $v = v_c$ case, we should have written in all generality $\omega_{v_c}(z) = [A'z + B + o(1)]e^{-\gamma_c z}$, but we fixed the translation invariance by choosing the solution for which $B = 0$.)

We recall also that there exists no travelling wave ω_v with values in $[0, 1]$ for $v < v_c$. However, it turns out that if we allow ω_v to become negative, there exists also travelling waves for $v < v_c$. As in (4.13), a linear analysis leads to $\omega_v \propto e^{-\gamma z}$ with $v(\gamma) = v$ but, with $v < v_c$, the solutions γ to $v(\gamma) = v$ are complex numbers. Then, writing $\gamma_R > 0$ and $\gamma_I > 0$ for the real and imaginary parts of γ , one can complete (4.13); for large z :

$$\omega_v(z) = [A_v \sin(\gamma_I z + \Phi_v) + o(1)]e^{-\gamma_R z} \quad \text{if } v < v_c, \quad \text{with } v(\gamma_R \pm i\gamma_I) = v, \quad (4.14)$$

for some A_v and Φ_v . We emphasize again that the solutions (4.14) for $v < v_c$ are negative in some regions and cannot be reached by the FKPP equation from a non-negative initial condition.

Comparing (4.11) and (4.12), it is clear that for z small enough compared to L , one should have $\tilde{\omega}(z) \approx \omega_{v_c - \Delta}(z)$ but, as $\Delta > 0$, we need to consider the travelling waves (4.14) rather than (4.13). For $v = v_c - \Delta$ with Δ small, and recalling that $v_c = v(\gamma_c)$ and $v'(\gamma_c) = 0$, one finds easily that

$$\gamma_R = \gamma_c + \mathcal{O}(\Delta), \quad \gamma_I \sim \sqrt{\frac{2\Delta}{v''(\gamma_c)}}. \quad (4.15)$$

To leading order in (4.14), this gives

$$\omega_{v_c - \Delta}(z) \approx A \sqrt{\frac{v''(\gamma_c)}{2\Delta}} \sin\left(\sqrt{\frac{2\Delta}{v''(\gamma_c)}} z\right) e^{-\gamma_c z} \quad \text{for large } z. \quad (4.16)$$

Notice that we fixed the invariance by translation by choosing the solution such that $\Phi_v = 0$. Furthermore, we determined the leading behaviour of $A_{v_c - \Delta}$ by noticing that one must have $\omega_{v_c - \Delta} \rightarrow \omega_{v_c}$ as $\Delta \rightarrow 0$.

The travelling wave $\omega_{v_c - \Delta}$ cancels for the first time when the argument of the sinus is π . This travelling wave must be very close to $\tilde{\omega}$ (the travelling wave of the problem with cut-off) which cancels at position L . Hence, one expects

$$\sqrt{\frac{2\Delta}{v''(\gamma_c)}} L \sim \pi, \quad (4.17)$$

or

$$\Delta = v_c - v_N^{(\text{cut-off})} \sim \frac{\pi^2 v''(\gamma_c)}{2L^2}. \quad (4.18)$$

It remains to determine the value of L . The solution (4.16) to (4.11) for large z decays as $e^{-\gamma_c z}$ (with a sinus prefactor which varies slowly). The cut-off is reached when $\tilde{\omega}$ becomes close to $1/N$ and hence

$$e^{-\gamma_c L} \sim \frac{1}{N} \quad \text{or } L \sim \frac{\ln N}{\gamma_c}. \quad (4.19)$$

Finally,

$$v_c - v_N^{(\text{cut-off})} \sim \frac{\pi^2 \gamma_c^2 v''(\gamma_c)}{2 \ln^2 N}. \quad (4.20)$$

Thus, a small cut-off of order $1/N$ at the tip of the front translates into a relatively huge correction to the velocity of order $1/\ln^2 N$. The precise way the cut-off is introduced is not so important: a cut-off at $2/N$ rather than at $1/N$ gives the same leading correction to the velocity.

Since [BD97], some rigorous proofs of the cut-off effect have been written. For instance, both [BDL08; DPK07] consider the following FKPP equation with cut-off

$$\partial_t h = \partial_x^2 h + h(1 - h) \mathbb{1}_{\{h > 1/N\}}, \quad (4.21)$$

and prove in two different ways that its velocity $v_N^{(\text{cut-off})}$ is given by

$$v_N^{(\text{cut-off})} = 2 - \frac{\pi^2}{\ln^2 N} + \mathcal{O}\left(\frac{1}{\ln^3 N}\right), \quad (4.22)$$

as expected from (4.20).

Going back to noisy equations (4.1) and (4.2), the conjecture of [BD97], which was amply checked by numerical simulations, is that the cut-off theory describes the main effect of introducing noise into FKPP-like equations, as in (4.1) or (4.2); when N is large:

- the distance between the centre of the front and the rightmost particle scales like $(\ln N)/\gamma_c$ as in (4.19),
- the empirical density of particles relative to the centre of the cloud is typically given by the sin times exponential shape (4.16),
- the leading term to the velocity correction $v_c - v_N$ to the velocity is the same as in the cut-off theory:

$$v_c - v_N \sim v_c - v_N^{(\text{cut-off})} \sim \frac{\pi^2 \gamma_c^2 v''(\gamma_c)}{2 \ln^2 N}. \quad (4.23)$$

The correction (4.23) to the velocity has rigorously been proved [BG10] by Bérard and Gouéré in 2010 for the noisy FKPP model (4.2) under some assumptions on $\rho(\epsilon)$. Other proofs exist for other models in the same class, see Section 4.3.

Notice however that the cut-off equation (4.9) is deterministic. Hence, there is no diffusion and the cut-off theory provides no help for computing the diffusion constant D_N .

4.1.2 Beyond the cut-off theory; a phenomenological description

In 2006, with Bernard Derrida, Stéphane Munier and Alfred H. Mueller, we proposed [\star BDMM06a] a phenomenological description of the effect of the noise on the position of the front. This description leads to a better understanding of the velocity v_N , and predicts the leading term for the diffusion constant D_N and for all the higher cumulants:

$$\begin{aligned} v_N &\approx v_c - \frac{\pi^2 \gamma_c^2 v''(\gamma_c)}{2 [\ln N + 3 \ln \ln N + \mathcal{O}(1)]^2} \approx v_N^{(\text{cut-off})} + 3 \pi^2 \gamma_c^2 v''(\gamma_c) \frac{\ln \ln N + \mathcal{O}(1)}{\ln^3 N}, \\ D_N &\approx \pi^2 \gamma_c v''(\gamma_c) \frac{\pi^2/3}{\ln^3 N}, \quad \lim_{t \rightarrow \infty} \frac{[n\text{-th cumulant of } \mu_t]}{t} \approx \pi^2 \gamma_c^{3-n} v''(\gamma_c) \frac{n! \zeta(n)}{\ln^3 N} \text{ for } n \geq 2, \end{aligned} \quad (4.24)$$

where $\zeta(n)$ is the Riemann zeta function.

The main ingredients of the theory leading to (4.24) are the following:

- The cut-off theory is essentially correct: the noise is negligible everywhere except at the tip of the front, and the main effect of the noise at the tip is to bring back the front to 0 when it becomes of order $1/N$. The shape of the front is most of the time given by the sinus-exponential shape (4.16):

$$h(\mu_t + z, t) \approx \frac{AL}{\pi} \sin\left(\frac{\pi z}{L}\right) e^{-\gamma_c z} \mathbb{1}_{\{z \leq L\}} \quad \text{if } z \text{ is large,} \quad (4.25)$$

where μ_t is the position of the front and $L = (\ln N)/\gamma_c$.

- Fluctuations around the cut-off shape (4.25) are, for most of the time, negligible.
- At rare times, a large fluctuation occurs that changes significantly the shape of the front. The relaxation of the shape of the front towards the sinus-exponential shape gives rise to a significant shift in the position of the front.
- Taken together, all these shifts due to large fluctuations lead to (4.24).

Figure 4.1 shows one of these rare and large fluctuations that do have an effect on the position of the front.

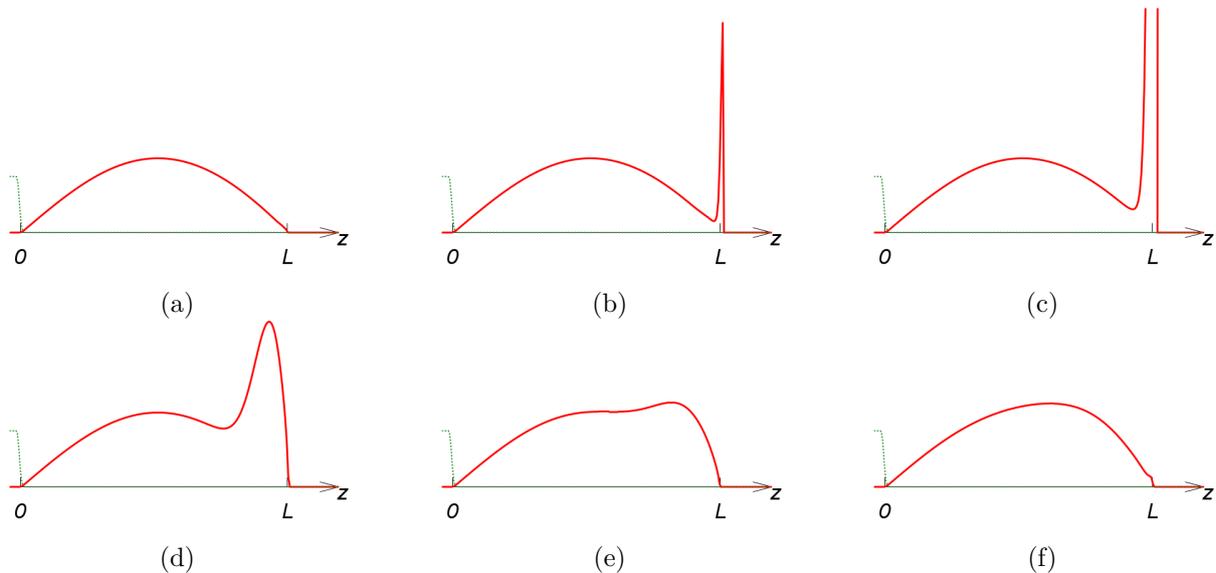


Figure 4.1: Snapshots of the simulation of a noisy front for some large value of N . The green dotted lines are $h(\mu_t + z, t)$ and the red solid lines are $h(\mu_t + z, t)e^{\gamma_c z}$ as a function of z , where we recall that μ_t is the position of the front and $L = (\ln N)/\gamma_c$. Each plot corresponds to a different time t . In plot (a), one recognizes the sinus shape (4.25) predicted by the cut-off theory. In fact, most of the time, the front looks like plot (a). Plots (b) and (c) represent a rare, large, fluctuation rising quickly, and (d), (e) and (f) are steps in the relaxation of that fluctuation. In plot (f), the front is nearly back to its sinus shape. Notice that the fluctuation is invisible on the green dotted plot representing $h(\mu_t + z, t)$; it is only on the red solid plot representing $h(\mu_t + z, t)e^{\gamma_c z}$ that there is something to see.

Understanding the relaxation of a fluctuation

Consider a deterministic front equation with cut-off; if the initial shape of the front is the sinus-exponential travelling wave, then the position of the front at all times is $\mu_t = v_N^{(\text{cut-off})}t$. If the initial shape is different, the function $t \mapsto \mu_t$ is no longer linear at all times. However, at large times, one still expects to have $\mu_t = v_N^{(\text{cut-off})}t + \text{Cste} + o(1)$, where the constant depends on the initial condition. We focus here on computing this constant. The idea is to identify this constant when the initial condition is the shape of the front just after a large fluctuation (as in plot (c) of Figure 4.1) with the effect on the position of the noisy front that such a fluctuation conveys.

So far, in the cut-off theory, we only focused on the travelling wave (4.25), but we need now to understand the time evolution of the front. We did write a dynamical equation (4.9), but it is rather complicated. We now write a simpler dynamical equation which also represents a front with cut-off:

$$\partial_t h = \partial_x^2 h + h - h^2 \quad \text{if } x < \mu_t + L, \quad h(\mu_t + L, t) = 0, \quad (4.26)$$

where μ_t is the position of the front defined in such a way that $h(\mu_t + z, t)$ converges to the travelling wave (4.25) and where $L \approx (\ln N)/\gamma_c$. The cut-off, which was originally formulated as “the front cancels when it reaches $1/N$ ” has been replaced by a boundary condition at $\mu_t + L$. (4.26) can be simplified even more by getting rid of the non-linearity. Indeed, one important effect of the non-linear saturation term $-h^2$ in the FKPP is that the travelling wave behaves for $z \gg 1$ as Aze^{-z} , with a z prefactor in front of the exponential. This z prefactor is present for any non-linear saturation term, and absent if one removes them. So, rather than putting the non-linear term $-h^2$ in (4.26), we enforce the presence of the Az prefactor in a more simple way by writing:

$$\partial_t h = \partial_x^2 h + h \quad \text{for } x \in (\mu_t, \mu_t + L), \quad h(\mu_t, t) = 0, \quad \partial_x h(\mu_t, t) = A, \quad h(\mu_t + L, t) = 0, \quad (4.27)$$

where the non-linearity in (4.26) has been replaced by the two boundary conditions at μ_t , as in Section 2.2.1.

The equation (4.27) is linear. Writing

$$h(\mu_t + z, t) = As(z, t)e^{-z}, \quad (4.28)$$

one obtains

$$\partial_t s = \partial_z^2 s - (2 - \mu_t)\partial_z s + (2 - \mu_t)s, \quad s(0, t) = s(L, t) = 0, \quad \partial_z s(0, t) = 1. \quad (4.29)$$

The shape $s(z, t)$ varies over the large space scale L . One therefore expects $\partial_z s$ to be typically L times smaller than s and $\partial_z^2 s$ to be typically L^2 times smaller than s . Then, necessarily, $\partial_t s$ must be L^2 times smaller than

s and $2 - \dot{\mu}_t$ must be of order $1/L^2$. This implies that the $\partial_z s$ term must be negligible for large L . Rewriting the equation without it, one gets

$$\partial_t s = \partial_z^2 s + (2 - \dot{\mu}_t)s, \quad s(0, t) = s(L, t) = 0, \quad \partial_z s(0, t) = 1. \quad (4.30)$$

Ignore for the moment the condition $\partial_z s(0, t) = 1$. (4.30) can be solved by decomposing over eigenmodes:

$$s(z, t) = \sum_{n \geq 1} a_n \frac{L}{\pi} \sin\left(\frac{n\pi z}{L}\right) e^{2t - \mu_t - \frac{n^2 \pi^2}{L^2} t}, \quad (4.31)$$

where the coefficients a_n can be obtained by decomposing the initial condition $s_0(z)$ over the eigenmodes; assuming $\mu_0 = 0$:

$$a_n = \frac{2\pi}{L^2} \int_0^L dz \sin\left(\frac{n\pi z}{L}\right) s_0(z). \quad (4.32)$$

The position μ_t of the front is now entirely determined by imposing the last boundary condition $\partial_z s(0, t) = 1$ at all times.

We are mostly interested in the large t limit. For large times, only the first eigenmode $n = 1$ in (4.31) contributes:

$$s(z, t) \approx a_1 \frac{L}{\pi} \sin\left(\frac{\pi z}{L}\right) e^{2t - \mu_t - \frac{\pi^2}{L^2} t} \quad \text{for } t \gg L^2. \quad (4.33)$$

With the boundary condition $\partial_z s(0, t) = 1$, one concludes that

$$\mu_t \approx \left(2 - \frac{\pi^2}{L^2}\right)t + \ln a_1 \quad \text{for } t \gg L^2. \quad (4.34)$$

Recalling that, for the equation considered, one has $v_c = 2$, $\gamma_c = 1$ and $v''(\gamma_c) = 2$, one recovers the cut-off velocity (4.18), or (4.20) when using $L \approx \ln N$. In fact, the derivation above could be made on a more general equation and one would find

$$\mu_t \approx v_N^{(\text{cut-off})} t + \frac{1}{\gamma_c} \ln a_1 \quad \text{for } t \gg L^2. \quad (4.35)$$

One recognizes also in (4.33) the same sinus shape as in (4.25); this shape is really a signature of the diffusion equation in a strip of finite width L . The new information is the fact that one can now predict μ_t for all times (depending on the initial condition) from (4.31) and, in particular, one finds in (4.35) the shift in the position of the front for an arbitrary initial condition.

Understanding the fluctuations

In the previous paragraph, we only considered the deterministic FKPP equation with cut-off to determine in (4.35) how the front relaxes from a fluctuation. To determine how fluctuations occur, the cut-off theory is obviously insufficient and one needs to consider a model with an actual noise, such as model (4.2) describing the evolution of a cloud of N points.

It is important to understand that the only place where noise is important is at the tip of the front, close to the rightmost particle, where h is of order $1/N$ because the noise term is negligible in the bulk of the front. In fact, we checked numerically [BD01], with models on the lattice, that if the noise is cancelled everywhere except at the rightmost occupied site, the velocity v_N and diffusion constant D_N and even the noise correction $v_N - v_N^{(\text{cut-off})}$ are asymptotically not modified.

Recall that μ_t is the position of the front chosen in such a way that $h(\mu_t + z, t)$ looks (most of the time) like the travelling wave (4.25), without any phase in the sinus. When considering FKPP with cut-off, we used to say that the cut-off was at position $\mu_t + L$ with $L \approx (\ln N)/\gamma_c$; now, with a microscopic noisy model, we say that the typical position of the rightmost particle is around position $\mu_t + L$. Let us call

$$[\text{position of the rightmost particle}] = \mu_t + L + \delta_t, \quad (4.36)$$

where δ_t is some random number which is typically of order 1.

The model we consider describes a cloud of particles which diffuse and reproduce. The number of particles remains equal to N by removing some of the leftmost particles. For the rightmost particles, the effect of maintaining the population at N individuals has little incidence on time scales which are small compared to L^2 : locally, for times not too large, the right boundary of the front does not “feel” the saturation around the left boundary. For this reason, it is reasonable to assume that the distribution of δ_t becomes independent of N as $N \rightarrow \infty$ and is actually given by the same law as the underlying branching process, without saturation. As we have already discussed in Section 3.2.1, this distribution is Gumble, see (3.19). Introducing $p(\delta) d\delta$ as the probability per unit time that a fluctuation of size δ develops, we write

$$p(\delta) \approx C_1 e^{-\gamma_c \delta} \quad \text{for } \delta \text{ large}, \quad (4.37)$$

where C_1 is some constant. (We only care for the large values of δ , for which the Gumble is like an exponential.)

Suppose now that at time t_0 , a large fluctuation occurs with some large value of δ in the position (4.36) of the rightmost particle. The density of particles around $\mu_{t_0} + L$ and up to $\mu_{t_0} + L + \delta$ has increased by a small amount of order $1/N$. For large z we modify (4.25) into

$$h(\mu_{t_0} + z, t_0) \approx \frac{AL}{\pi} \sin\left(\frac{\pi z}{L}\right) e^{-\gamma c z} \mathbb{1}_{\{z \leq L\}} + f(z), \quad \text{where } f(z) = \begin{cases} \mathcal{O}\left(\frac{1}{N}\right) & \text{for } z \approx L \text{ and } z < L + \delta, \\ 0 & \text{for } z \ll L \text{ or } z > L + \delta. \end{cases} \quad (4.38)$$

Multiply everything by $e^{\gamma c z}/A$ to recover the sinus shape. Around where f is non zero, one has $e^{\gamma c z} = N\mathcal{O}(e^{\gamma c \delta})$:

$$\frac{1}{A} h(\mu_{t_0} + z, t_0) e^{\gamma c z} \approx \frac{L}{\pi} \sin\left(\frac{\pi z}{L}\right) \mathbb{1}_{\{z \leq L\}} + \tilde{f}(z), \quad \text{where } \tilde{f}(z) = \begin{cases} \mathcal{O}(e^{\gamma c \delta}) & \text{for } z \approx L \text{ and } z < L + \delta, \\ 0 & \text{for } z \ll L \text{ or } z > L + \delta. \end{cases} \quad (4.39)$$

Even though the fluctuation is small on the level of h , it can easily become huge on the level of the sinus shape, see Figure 4.1, plots (b) and (c). The sinus shape has a height L and the fluctuation a height $e^{\gamma c \delta}$; one can easily have $e^{\gamma c \delta} \gg L$ even though $\delta \ll L$.

One can now estimate a_1 appearing in (4.35) by using (4.39) as the initial shape $s_0(z)$ in (4.32), see (4.28). A small difficulty is that (4.32) was established for a model with a hard cut-off at $\mu_t + L$: the value of $s_0(z)$ for $z > \mu_t + L$ is irrelevant. Here, the fluctuation extends up to $\mu_t + L + \delta$ with $0 \leq \delta \ll L$, and obviously it has an effect on the position of the front, even though it is on the right of $\mu_t + L$. The point is that we have never been very precise on the definition of the cut-off and, anyway, the microscopic model does not have a hard cut-off as what was assumed to establish (4.32). What we do to fix this formal difficulty is to assume, just for applying (4.32), that the big fluctuation of height $e^{\gamma c \delta}$ is a little bit on the left of the cut-off, rather than a little bit on the right. Then,

$$a_1 = 1 + \frac{2\pi}{L^2} \int_0^L dz \sin\left(\frac{\pi z}{L}\right) \tilde{f}(z) \approx 1 + C_2 \frac{e^{\gamma c \delta}}{L^3}, \quad (4.40)$$

where C_2 is some constant. The 1 in the right hand side is the contribution of the sinus in (4.39), while the other term is the contribution of \tilde{f} . The denominator is L^3 because only values of z very close to L contribute to the integral and one can expand the sinus into $\pi(L - z)/L$ with $L - z$ of order 1.

Finally, using (4.35), one can get the long time effect $R(\delta)$ on the position of the front caused by the relaxation of a fluctuation of size δ :

$$R(\delta) \approx \frac{1}{\gamma c} \ln\left(1 + C_2 \frac{e^{\gamma c \delta}}{L^3}\right). \quad (4.41)$$

Phenomenological picture of the fluctuating front

From (4.41), small fluctuations have little effect on the position of the front: one needs to take $\delta \approx \frac{1}{\gamma c} 3 \ln L$ to have an effect $R(\delta)$ which is not small. The probability to observe at a random instant such a fluctuation is of order $e^{-\gamma c \delta} \propto L^{-3}$; in other words, a relevant fluctuation occurs typically every $\mathcal{O}(L^3)$ units of time.

The fluctuations, which occur locally at the tip of the front, rise quickly in a time scale of order 1 but, because of the diffusive nature of the front, take a time of order L^2 to relax back into the typical sinus-exponential shape. The picture is therefore the following:

every $\mathcal{O}(L^3)$ units of time, a relevant fluctuation rises quickly and then relaxes in a time $\mathcal{O}(L^2)$.

This separation of time scales implies that there is asymptotically no need to consider overlapping fluctuations; it is sufficient to assume that they occur and relax completely before the next one occurs. Then, it is easy to obtain that [*BDMM06a]

$$v_N - v_N^{(\text{cut-off})} \approx \int d\delta p(\delta) R(\delta), \quad D_N \approx \int d\delta p(\delta) R(\delta)^2, \quad \lim_{t \rightarrow \infty} \frac{[n\text{-th cumulant of } \mu_t]}{t} \approx \int d\delta p(\delta) R(\delta)^n. \quad (4.42)$$

Using (4.37) for $p(\delta)$ and (4.41) for $R(\delta)$, one finds

$$v_N - v_N^{(\text{cut-off})} \approx C_3 \frac{3 \ln L}{\gamma c L^3}, \quad D_N \approx C_3 \frac{\pi^2/3}{\gamma c^2 L^3}, \quad \lim_{t \rightarrow \infty} \frac{[n\text{-th cumulant of } \mu_t]}{t} \approx C_3 \frac{n! \zeta(n)}{\gamma c^n L^3}, \quad (4.43)$$

where $C_3 = C_1 C_2 / \gamma c$. Note that in computing (4.43), all the values of δ up to $(3 \ln L) / \gamma c + \mathcal{O}(1)$ contribute for computing the velocity. For the diffusion constant and the higher cumulants, only the values of δ equal to $(3 \ln L) / \gamma c + \mathcal{O}(1)$ contribute.

The correction to the cut-off velocity and all the cumulants are therefore obtained up to some constant C_3 . The last step is of course to determine that constant and, unfortunately, we had no analytical argument for determining C_3 . But we could guess that

$$C_3 = \frac{C_1 C_2}{\gamma_c} = \pi^2 v''(\gamma_c). \quad (4.44)$$

With this value of C_3 , one recovers the prediction (4.24).

4.1.3 Justifications for the guess (4.44)

Modified cut-off theory

The guess (4.44) is such that the velocity of the noisy front is

$$v_N = v_c - \frac{\pi^2 v''(\gamma_c)}{2L^2} \quad \text{with } L = \frac{1}{\gamma_c} [\ln N + 3 \ln \ln N + \mathcal{O}(1)]. \quad (4.45)$$

This is the cut-off velocity, except that the length L of the front (the distance between the middle of the front and the rightmost particle) is $(\ln N + 3 \ln \ln N)/\gamma_c$ instead of $(\ln N)/\gamma_c$. This length makes sense because it corresponds to $\delta = (3 \ln L)/\gamma_c$, which is the maximal displacement of the rightmost particle that contributes to (4.43). When applying the cut-off theory, we first used for L (defined as the distance to the cut-off) the typical distance $(\ln N)/\gamma_c$ of the rightmost particle to the centre of the front. It turns out that one should use instead the maximum distance $(\ln N + 3 \ln N)/\gamma_c$ which is likely enough to have an effect.

BBM with an absorbing wall

What makes the BBM a tractable model is that all the particles, once born, behave independently from each other. What makes the models in the noisy FKPP class such as (4.2) so difficult is that this independence is lost: a particle disappears if there are N other particles on its right.

In an appendix of [★BDMM06a] we considered the BBM with an absorbing wall, which is a model first introduced [Kes78] by Kesten in 1978. The rules are the following:

- one starts at $t = 0$ with an initial finite set of particles on the positive axis,
- the particles diffuse and branch independently as in the BBM,
- there is an absorbing wall started from the origin and moving at velocity v ; any particle touching the wall is removed.

The idea behind this model is that it shares the “simplicity” of the BBM (particles once born behave independently), but the particles on the left are removed in a way similar to what happens in models in the noisy FKPP class. This model is illustrated in Figure 4.2.

The population is not kept at a constant size and, in the large time limit, one out of two things may happen:

- either the system goes extinct in finite time and all the particles are absorbed,
- either the number of particles diverges in spite of the wall.

When the velocity v of the wall is larger or equal to 2 (the velocity of a BBM), the particles do not stand a chance and the probability of extinction is 1. When the velocity of the wall is smaller than 2, either outcome has non-zero probability. Call $p_v(x)$ the probability of long time survival when the initial state consists of a single particle located at $x > 0$ and when the wall goes at velocity v . A simple analysis similar to (1.8) leads to

$$p_v'' - v p_v' + p_v - p_v^2 = 0, \quad p_v(0) = 0, \quad \text{with } v < 2. \quad (4.46)$$

This equation looks like the equation (1.13) for the travelling waves ω_v of the FKPP at velocity v , except for the boundary condition $p_v(0) = 0$ and the fact that left and right are reversed (the sign of v is changed, and, $p_v(\infty) = 1$ whereas all the fronts we usually consider are equal to 1 far on the left and to 0 far on the right). Remember that ω_v for $v < 2$ has an oscillatory behaviour at large x , see (4.16). Then, one sees easily that

$$p_v(x) = \omega_v(X_v - x), \quad \text{with } X_v = [\text{smallest } x \text{ such that } \omega_v(x) = 0]. \quad (4.47)$$

Assume now that the velocity of the wall is given by

$$v = 2 - \frac{\pi^2}{L^2} \quad \text{with } L \text{ large.} \quad (4.48)$$

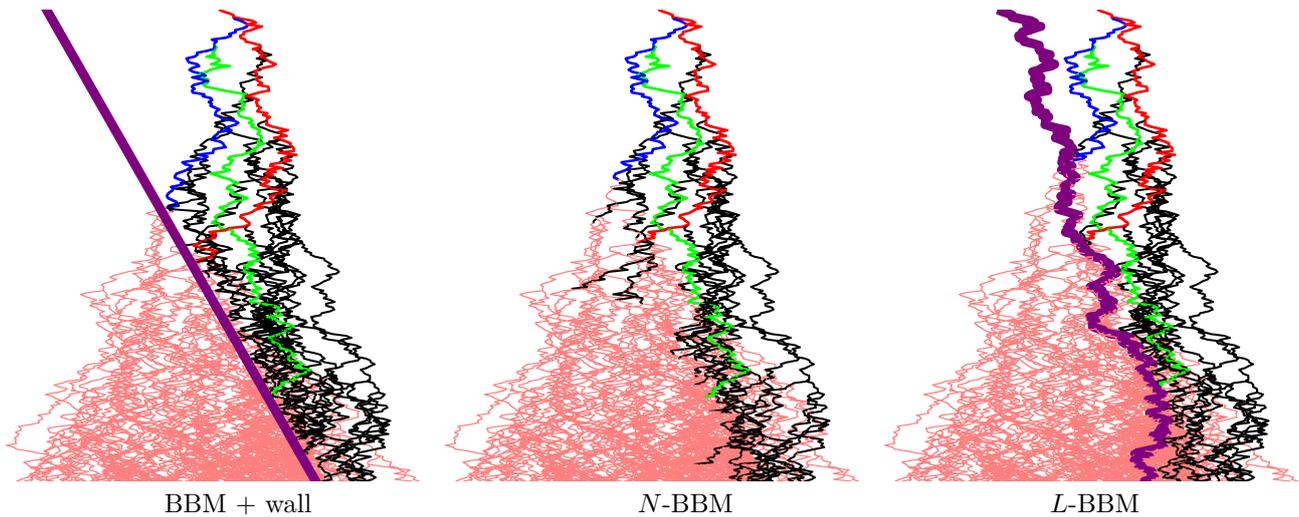


Figure 4.2: Three variations on the BBM where some particles on the left are removed (the removed particles and the descendants they would have had are still drawn in a light colour.) On the left, a BBM with a wall: the wall, a thick purple line, moves at constant speed and kills the particles it encounters. In the middle, a N -BBM: at most N particles may live, and when this number is reached and a branching occurs, the leftmost particle is removed. On the right, a L -BBM: all the particles at a distance larger than L from the rightmost are removed. The N -BBM and L -BBM are discussed in Section 4.3.

(This looks like the velocity with a cut-off, of course, but here L is a free parameter: we have not introduced anything called N yet in this model of BBM with wall.) The function $\omega_v(z)$ for large z is given by the usual sinus-exponential shape (see (4.16) with $\Delta = \pi^2/L^2$, $\gamma_c = 1$, $v_c = 2$ and $v''(\gamma_c) = 2$), one obtains $X_v \approx L$ and

$$p_v(x) \approx \begin{cases} \frac{AL}{\pi} \sin \frac{\pi x}{L} e^{x-L} & \text{for } x \geq 0 \text{ and } L - x \gg 1, \\ 1 & \text{for } x - L \gg 1. \end{cases} \quad (4.49)$$

(See also [DS07], and [BBS11] for a rigorous statement and proof, and [HH07] for the survival probability up to time t when $v > 2$, and [GHS11; BG11] for the BRW case.) We are now ready to use our understanding of the BBM with absorbing wall to get some insight on the FKPP equation with noise. The idea is to choose, for the BBM with wall, the velocity of the wall and the initial condition in a way as similar as possible as the typical state of the FKPP with noise:

- for the velocity of the wall, choose $v = 2 - \frac{\pi^2}{L^2}$,
- for the initial condition, start with N particles with independent positions chosen according to a distribution that looks like the distribution of particles in the front with cut-off: $\rho(x) \approx B \ln(N) \sin \frac{\pi x}{\ln N} e^{-x} \mathbb{1}_{\{x < \ln N\}}$ for x large.

Up to that point, L and N are two independent parameters. For times not too large, one expects the particles to move as in the FKPP equation with noise at velocity v_N . But the wall moving at velocity v eats particles from the left. If $v > v_N$, the system has a high probability of going extinct. On the other hand, if $v < v_N$, the particles escape from the wall and survive indefinitely with high probability. If $v = v_N$, finally, the probability of survival should be neither close to 0 nor to 1.

Fortunately, the probability of survival of the cloud of particles is easy to compute from (4.49) because all the particles are independent. With the initial condition we have just described, one finds, for large N and L ,

$$[\text{probability that the system survives}] \approx \begin{cases} 1 & \text{If } L \ll \ln N, \\ 0 & \text{If } L \gg \ln N, \\ 1 - e^{-CN(\ln N)^3 e^{-L}} & \text{If } L \sim \ln N, \end{cases} \quad (4.50)$$

for some constant C . The probability of survival transitions from 1 to 0 when

$$L = \ln N + 3 \ln \ln N + \mathcal{O}(1), \quad (4.51)$$

so that the velocity of the front in the noisy FKPP equation, equal to the velocity of the absorbing wall at the survival transition, should indeed be given by (4.45).

In a more precise study of the BBM with an absorbing wall, Damien Simon and Bernard Derrida [SD08] have considered a BBM with an absorbing wall started with one particle at distance 1 from the wall and conditioned

to have exactly one particle at a very large time T . The conditioning prevents the system from going extinct and also from having an exploding population size. At times t large enough to forget the initial condition, and such that $T - t$ is large enough to not feel too much the conditioning, the system reaches a quasi-stationary state regime. As the velocity of the wall gets close to 2, they found that the expected number of particles in the quasi-stationary state was given by

$$\langle N \rangle \sim (2 - v)^{3/2} \exp \left[\frac{\pi}{\sqrt{2 - v}} \right]. \quad (4.52)$$

Inverting this relation to give v as a function of $\langle N \rangle$ gives back the velocity (4.45).

Julien Berestycki, Nathanaël Berestycki and Jason Schweinsberg [BBS13] have also studied this model of BBM with absorbing wall (but with no conditioning on the state of the system at large times). They showed that, starting from N particles not too much on the right, the number of particles at time $u \ln^3 N$ was of order N in the limit $N \rightarrow \infty$ with $u > 0$ fixed, if and only if the velocity is given by (4.45).

The exponential model

In the noisy FKPP model (4.2), there is a cloud of N particles. At each time-step, each particle is replaced by k children, but only the N global rightmost individuals are kept. Here, k is a fixed number (typically $k = 2$) and N is sent to infinity while k is kept constant.

We consider now a model where, at each time-step, each particle has infinitely many children (with however a finite number of descendants on the right), but only the N global rightmost individuals are kept. Occasionally, these N rightmost might be the children of a single particle.

We assume that the descendants of a single particle are located according to a Poisson point process of density $\rho(\epsilon)$ shifted by the position of the parent. The density ρ must satisfy $\int d\epsilon \rho(\epsilon) = \infty$ (because having infinitely many particles is the only way to make sure there are at least N of them) but $\int_0^\infty d\epsilon \rho(\epsilon) < \infty$ (because there must only be finitely many particles on the right of any point so that “keeping the N rightmost” makes sense).

Calling, as usual, $h(x, t)$ the fraction of particles on the right of x , one finds easily that

$$h(x, t + 1) = \min \left[1, \int d\epsilon \rho(\epsilon) h(x - \epsilon, t) + \text{noise} \right]. \quad (4.53)$$

One can compare (4.53) to (4.2), but one needs to be careful: the $\rho(\epsilon)$ in (4.2) is a density of probability (normalized to 1) while the $\rho(\epsilon)$ in (4.53) is the density of the Poisson point process and its integral is infinity.

As N goes to infinity, the noise term goes to 0. Applying the usual techniques, one looks for travelling waves that decay as $h(x, t) \sim e^{-\gamma(x-vt)}$, which leads to

$$e^{\gamma v} = \int d\epsilon \rho(\epsilon) e^{\gamma \epsilon}. \quad (4.54)$$

Then, one looks at the minimal velocity $v_c = v(\gamma_c)$. To take an example, if $\rho(\epsilon) = \mathbb{1}_{\{\epsilon < 1\}}$, one finds $v(\gamma) = 1 - (\ln \gamma)/\gamma$ and one finds that the minimum is reached for $\gamma_c = e$ and that $v_c = 1 - 1/e$.

There exist however choices of $\rho(\epsilon)$ which are problematic. For instance, pick

$$\rho(\epsilon) = e^{-\epsilon}. \quad (4.55)$$

For this choice, the integral in (4.54) diverges for all values of γ . There is no function $v(\gamma)$, no minimal velocity v_c and one concludes that (4.53) with (4.55) is not an equation in the FKPP class. It has however the significant advantage of being a solvable model for any value of N [BDMM06b; *BDMM07]. To see this, let us go back to the definition of the model. Call X_i , $i = \{1, 2, \dots, N\}$ the positions of the N particles at time t . Before selection, the children of particle i are given by a Poisson point process of density $\rho(x - X_i)$. Then, the children of all the particles before selection are given by a Poisson point process of density $\sum_i \rho(x - X_i)$. Finally comes the selection process, and:

$$\left\{ \text{positions of the } N \text{ particles at time } t + 1 \right\} = \left\{ \text{the } N \text{ rightmost of P.P.P. of density } \sum_i \rho(x - X_i) \right\}. \quad (4.56)$$

With the choice (4.55), the density of the Poisson point process can simply be written as

$$\left\{ \text{positions of the } N \text{ particles at time } t + 1 \right\} = \left\{ \text{the } N \text{ rightmost of P.P.P. of density } e^{-(x-\mu_t)} \right\}, \quad (4.57)$$

with

$$\mu_t = \ln \left(\sum_i e^{X_i} \right). \quad (4.58)$$

The quantity μ_t can be interpreted as the position of the cloud of N particles. It is the only quantity needed to compute the positions of the N particles at time $t + 1$. In particular, it is easy to see that

$$\mu_{t+1} - \mu_t = \ln \left(\sum_{i=1}^N e^{z_i} \right), \quad \text{where } \{z_1, \dots, z_N\} \text{ are the } N \text{ rightmost of a P.P.P. of density } e^{-z}. \quad (4.59)$$

Furthermore, the increments $\mu_{t+1} - \mu_t$ for different times t are independent. It is then possible to compute exactly their cumulants, and one finally finds, for large N ,

$$\begin{aligned} v_N = \langle \mu_{t+1} - \mu_t \rangle &= \ln \left[\ln N + \ln \ln N + \mathcal{O}(1) \right] = \ln \ln N + \frac{\ln \ln N + \mathcal{O}(1)}{\ln N}, \\ D_N = \text{var}(\mu_{t+1} - \mu_t) &\sim \frac{\pi^2/3}{\ln N}, \quad \frac{[n\text{-th cumulant of } \mu_t]}{t} \sim \frac{n! \zeta(n)}{\ln N} \text{ for } n \geq 2, \end{aligned} \quad (4.60)$$

where $\zeta(n)$ is the Riemann zeta function.

This is of course very similar to the result (4.41) in the FKPP case. The most interesting point, however, is that the phenomenological theory used to explain the FKPP case can also explain (4.60). Indeed, consider the sum in (4.59): typically, the smallest z_i is at position $-\ln N$ (because $\int_{-\ln N}^{\infty} dz e^{-z} = N$) and the largest z_i is around the origin (because $\int_0^{\infty} dz e^{-z} = 1$). Then, the sum is typically equal to $\int_{-\ln N}^0 dz e^{-z} \times e^z = \ln N$, and one could have expected *a priori* a velocity asymptotically equal to $\ln \ln N$. Now, imagine a rare fluctuation where the largest z_i has a large value δ . Singling out this value and approximating the rest of the sum by $L = \ln N$, one gets

$$\mu_{t+1} - \mu_t \approx \ln \left(L + e^\delta \right) = \ln L + R(\delta) \quad \text{with } R(\delta) = \ln \left(1 + \frac{e^\delta}{L} \right) \quad \text{and } L = \ln N. \quad (4.61)$$

Of course, δ , as the rightmost point of a Poisson point process with exponential density, is Gumble distributed. In particular, with $p(\delta)$ the law of δ ,

$$p(\delta) \sim e^{-\delta} \quad \text{for large } \delta. \quad (4.62)$$

Compare to (4.37) and (4.41). One can then recover (4.60) by writing $v_N \approx \ln L + \int d\delta p(\delta) R(\delta)$, $D_n \approx \int d\delta p(\delta) R(\delta)^2$, etc. The similarity with the usual FKPP case is striking: there is a base velocity computed from the typical shape ($v_N^{\text{(cut-off)}}$ in the FKPP case, $\ln L$ here). Corrections to that velocity are due to rare fluctuations where the rightmost particle goes some large distance δ to the right of its typical position. The distribution $p(\delta)$ is the same in both cases. The effect of that fluctuation is a shift $R(\delta)$. Comparing (4.41) for the FKPP case to (4.61) above, the functions $R(\delta)$ are identical except that the denominator is L^3 in the former case and L in the latter. The most important contribution is therefore due to δ of order $3 \ln L$ in the FKPP case but only of order $\ln L$ here in the exponential model, and the effect of noise on the velocity is that L should be replaced by $L + 3 \ln L$ in the FKPP case and by $L + \ln L$ here. A relevant fluctuation (for which $R(\delta) \approx 1$) occurs every L^3 steps in the FKPP case, and every L steps here, so that the diffusion constant and all the higher cumulants per unit time scale respectively like $1/L^3$ and $1/L$. The relaxation time to recover from a fluctuation is L^2 in the FKPP case and is 1 unit time here. In both cases, this relaxation time is L times smaller than the time between two fluctuations, so that a relevant fluctuation has the time to relax before another one occurs.

The Figure 4.3 summarizes the comparison between the noisy FKPP and the exponential model.

The Gumble model

In the noisy FKPP model (4.1), there are N particles at each generation. Each particle of a new generation picks k prospective parents, looks at their positions, adds random amounts and chooses the highest value for its own position. This procedure is repeated N times at each time-step to generate the N new particles.

Usually, one picks for k a fixed number (typically 2) and sends N to ∞ . With Bernard Derrida, we considered in [BD04] the case $k = N$ where every particle in the previous generation is a prospective parent of each new particle:

$$X_i(t+1) = \max [X_1(t) + \epsilon_1, \dots, X_N(t) + \epsilon_N]. \quad (4.63)$$

Given the state of the system at generation t , the positions of the new particles at generation $t+1$ are independent and distributed according to:

$$\mathbb{P} [X_i(t+1) < x] = \prod_{j=1}^N \mathbb{P} [\epsilon_j < x - X_j(t)]. \quad (4.64)$$

In all generality, this is a difficult problem. However, if one chooses a Gumble distribution for the ϵ_j :

$$\mathbb{P}[\epsilon_j < \epsilon] = e^{-e^{-\epsilon}}, \quad (4.65)$$

	noisy FKPP	exponential model
v_N	$v_c - \frac{K}{(\ln N + 3 \ln \ln N)^2} + \dots$	$\ln(\ln N + \ln \ln N) + \dots$
D_N	$\frac{2K}{\gamma_c} \frac{\pi^2}{3(\ln N)^3} + \dots$	$\frac{\pi^2}{3 \ln N} + \dots$
$p(\delta)$	$C_1 \gamma_c e^{-\gamma_c \delta}$	$e^{-\delta}$
$R(\delta)$	$\frac{1}{\gamma_c} \ln \left(1 + C_2 \frac{e^{\gamma_c \delta}}{(\ln N)^3} \right)$	$\ln \left(1 + \frac{e^\delta}{\ln N} \right)$
Relaxation time	$(\ln N)^2$	1
Most relevant fluctuation size	$\frac{1}{\gamma_c} 3 \ln \ln N$	$\ln \ln N$

Figure 4.3: A comparison between the phenomenological theory of fluctuations in the FKPP and in the exponential model. The constant K is given by $K = \pi^2 \gamma_c^2 v''(\gamma_c)/2$.

then (4.64) becomes

$$\mathbb{P} [X_i(t+1) < x] = e^{-e^{-(x-\mu_t)}} \quad \text{with } \mu_t = \ln \left(\sum_{j=1}^N e^{X_j(t)} \right). \quad (4.66)$$

μ_t can be interpreted as the position of the front at time t . It has the same expression (4.58) as in the exponential model. One concludes that, for the choice (4.65), the positions of the particles at generation $t+1$ are N independent Gumble variables relative to the position μ_t of the front at time t . From here, it is clear that the velocity, diffusion constant and higher cumulants can all be computed. One finds the same results (4.60) as for the exponential model, except that the velocity is increased by $\ln N$:

$$v_N = \ln N + \ln [\ln N + \ln \ln N + \mathcal{O}(1)] = \ln N + \ln \ln N + \frac{\ln \ln N + \mathcal{O}(1)}{\ln N}, \quad (4.67)$$

$$D_N \sim \frac{\pi^2/3}{\ln N}, \quad \frac{[n\text{-th cumulant of } \mu_t]}{t} \sim \frac{n! \zeta(n)}{\ln N} \text{ for } n \geq 2,$$

where $\zeta(n)$ is the Riemann zeta function.

Again, one checks easily that these results can also be explained, to leading order, by the same phenomenological theory as in the exponential model and the noisy FKPP case, with the same scaling as in the exponential model.

Fronts with global constraints

In [Hal11], Oskar Hallatschek considers noisy fronts following the stochastic equation

$$\partial_t \tilde{h} = \partial_x^2 \tilde{h} + v \partial_x \tilde{h} + f(x) \tilde{h} + \sqrt{\frac{2\tilde{h}}{N}} \eta + [\text{saturation term}], \quad (4.68)$$

where η is a space and time dependent Gaussian white noise and $f(x)$ a space dependent reaction rate. If one chooses $f(x) = 1$ for $x > 0$, this is very similar to the stochastic FKPP equation (1.54) seen in a frame moving at velocity v . ([Hal11] is mainly concerned with the case $f(x) = x$, but he also considers briefly the FKPP case.)

The originality of [Hal11] lies in the saturation term: it is a non-local term depending on the noise η tailored in such a way that the global constraint $\int dx \tilde{h}(x, t) u(x) = 1$ is satisfied at all times for some given function u . For a specific choice of u , a miracle occurs and $\langle \tilde{h}(x, t) \rangle$ obeys a closed equation. Looking at the stationary regime $\tilde{\omega}(x) = \lim_{t \rightarrow \infty} \langle \tilde{h}(x, t) \rangle$ for that specific choice of u , one finds

$$0 = \tilde{\omega}''(x) + v \tilde{\omega}'(x) + f(x) \tilde{\omega}(x) - \frac{2\tilde{\omega}^2(x) e^{vx}}{N \int dy \tilde{\omega}^2(y) e^{vy}}. \quad (4.69)$$

From the cut-off theory, if one assumes that $\tilde{\omega}(x) \approx L\phi(x/L)e^{-x}$ and $v \approx 2$ (in fact, ϕ is the sinus shape), then the integral in (4.69) is of order L^3 . If one assumes also that the growth term and the non-linear term are both equal to e^{-L} at the cut-off position $x = L$, one gets (remember that $f(x) = 1$)

$$NL^3e^{-L} \sim \text{Cste}, \quad (4.70)$$

which is the same relation as in (4.45). Notice however that $\tilde{h}(x, t)$ in (4.68) is supposed to mimic $h(vt + x, t)$ for a noisy FKPP front h , and that $\langle h(vt + x, t) \rangle$ for any choice of v does not reach a non-trivial stationary state, so that the applicability of (4.68) to noisy FKPP fronts is debatable. It would be interesting anyway to analyse in more detail (4.69).

4.2 Genealogies

Some models in the noisy FKPP class describe N particles diffusing and branching. These particles have a genealogical tree which is, once rescaled properly, described by the Bolthausen-Sznitman coalescent.

In this chapter, we have considered several models in the noisy FKPP class describing the evolution of a cloud of N particles. We focused up to now on the statistics of the position of the cloud but, in the models we discussed, any given particle at generation t has one parent at generation $t - 1$ and zero, one or several children at generation $t + 1$. It makes sense, then, to study the statistics of the genealogical tree of the population.

4.2.1 Genealogies in models without selection

The study of genealogical trees in simplified models of population dynamics has of course a long history. One of the most important model was the Wright-Fisher model [Wri31; Fis30] which describes the evolution of a population of constant size N . The rule is simple: at each generation, each new individual chooses a parent at random (with replacement) in the previous generation.

Consider two arbitrary individuals in the Wright-Fisher model, and call T_2 the number of generations one needs to go back in time to find the first common ancestor of these two individuals. There is clearly a probability $\mathbb{P}(T_2 = 1) = 1/N$ that the two individuals have the same parent. If they do not, there is a probability $1/N$ that their parents have the same parent. It is then easy to see that $\mathbb{P}(T_2 = g) = (1 - 1/N)^{g-1}/N$, and that the expected number of generations needed to find a common ancestor is

$$\langle T_2 \rangle = N \quad [\text{Wright-Fisher}]. \quad (4.71)$$

More generally, calling

$$T_n = \left(\begin{array}{l} \text{the number of generations one needs to go back in time} \\ \text{to find the first common ancestor of } n \text{ given individuals} \end{array} \right), \quad (4.72)$$

one finds easily that, as $N \rightarrow \infty$,

$$\frac{\langle T_3 \rangle}{\langle T_2 \rangle} \rightarrow \frac{4}{3}, \quad \frac{\langle T_4 \rangle}{\langle T_2 \rangle} \rightarrow \frac{3}{2}, \quad \frac{\langle T_n \rangle}{\langle T_2 \rangle} \rightarrow 2 - \frac{2}{n} \quad [\text{Wright-Fisher}]. \quad (4.73)$$

The results (4.73) are very robust within models of population dynamics without selection. For instance, in the Moran model [Mor58], one still has (4.73) while (4.71) is replaced by $\langle T_2 \rangle = N/2$. (We recall that the Moran model is a time-continuous version of the Wright-Fisher model: there is a population of N individuals, and during dt , each individual has a probability dt of branching. To keep the population constant, a random individual is removed at each branching event.)

In 1982, Kingman [Kin82] introduced what is now called ‘‘Kingman’s coalescent’’ as the most simple and quintessential model of population dynamics without selection. Kingman’s coalescent is essentially the $N \rightarrow \infty$ limit of the Wright-Fisher model with time rescaled by a factor N and running backwards. In Kingman’s coalescent, two particles have during dt a probability dt of coalescing, so that their coalescence time τ is exponentially distributed. (In Wright-Fisher language: two individuals have their most recent common ancestor $N\tau$ generations ago with τ exponentially distributed.) If one considers n particles, the probability that during dt one coalescence occurs where two particles merge into one is $\binom{n}{2}dt$ because there are $\binom{n}{2}$ pairs of particles. Three or more particles may not coalesce at the same instant. Several pairs of particles may not coalesce at the same instant (these events are possible in Wright-Fisher, but with a very small probability when N is large). In Kingman’s coalescent, one has by construction $\langle T_2 \rangle = 1$ and the ratios (4.73) are exact.

In general, a model of population dynamics with a fixed population size N and without any selection mechanism looks, as $N \rightarrow \infty$, more and more like Kingman’s coalescent after rescaling of time, and one expects, for instance, $\langle T_3 \rangle / \langle T_2 \rangle$ to converge to $4/3$ when N becomes large.

4.2.2 Genealogies in models with selection

One of the defining features of models in the noisy FKPP class is that there is a strong selection mechanism: particles on the right have typically more children than particles on the left.

To pick an example, consider again model (4.2): each particle first has k children, whose positions are the position of the parent plus some random amount, then comes the strong selection phase: out of the kN children, only the N rightmost are kept. One could imagine replacing that selection phase by a simple neutral pruning: out of the kN children, N chosen at random are kept. The model thus obtained would be very close to the Wright-Fisher model. One would find $\langle T_2 \rangle = (kN - 1)/(k - 1)$, the time ratios would be given by (4.73) and the genealogical trees properly rescaled would be those of Kingman's coalescent. Also, the model would not be in the noisy FKPP class.

In the original model (4.2) with its strong selection phase (only the N rightmost are kept) and which is in the noisy FKPP class, the genealogical trees are different. In 2006, with Bernard Derrida, Stéphane Munier and Alfred H. Mueller [BDMM06b], we measured numerically the different coalescence times $\langle T_n \rangle$; see also [BD13]. We found that they scaled like a power of $\ln N$ rather than like N as in (4.71), and that the ratios of the times were very different from the values in (4.73). In fact, the numerical results were consistent with

$$\langle T_p \rangle \propto (\ln N)^3 \quad [\text{noisy FKPP}], \quad (4.74)$$

and

$$\frac{\langle T_3 \rangle}{\langle T_2 \rangle} \rightarrow \frac{5}{4}, \quad \frac{\langle T_4 \rangle}{\langle T_2 \rangle} \rightarrow \frac{25}{18} \quad [\text{noisy FKPP}]. \quad (4.75)$$

The ratios (4.75) are those obtained in the Bolthausen-Sznitman coalescent which was introduced [BS98] in the context of the mean-field theory of spin-glasses.

We also simulated the exponential model of Section 4.1.3, page 42, and found results compatible with (4.75) but not (4.74): we found, instead of (4.74),

$$\langle T_p \rangle \propto \ln N \quad [\text{exponential model}]. \quad (4.76)$$

Recall that the exponential model is not in the noisy FKPP class, but that it shares some important features with models in that class, see Figure 4.3 on page 44. In particular, the diffusion constant D_N scales like $1/(\ln N)^3$ in the noisy FKPP case and like $1/\ln N$ in the exponential case, and one remarks that the coalescence times $\langle T_p \rangle$ scale as the inverse of the diffusion constant in both the noisy FKPP case and the exponential model.

One can solve exactly the exponential model [BDMM06b; \star BDMM07] and write explicit expressions for the $\langle T_p \rangle$ as functions of N . Then, one can show that (4.75) and (4.76) hold for the exponential model. (They also hold for the Gumble model of page 43 [Cor16].)

Even more interestingly, the phenomenological theory of Section 4.1.2, which allowed us to describe the position of the fronts in both the exponential model and in the noisy FKPP case, can be extended to explain how the genealogical trees in both models look, as $N \rightarrow \infty$, more and more like a rescaled Bolthausen-Sznitman coalescent [\star BDMM07]. This is explained in Section 4.2.4 after a brief recall on the Λ -coalescents.

4.2.3 Λ -coalescents

Kingman's coalescent and the Bolthausen-Sznitman coalescent are two particular cases of the large family of Λ -coalescents [Pit99; Sag99]. See [Ber09] for an introduction and a review. Informally, in a Λ -coalescent, when considering b particles, any given group of $k \in \{2, 3, \dots, b\}$ particles coalesce into one single particle with a rate $\lambda_{b,k}$. As explained in Figure 4.4, the $\lambda_{b,k}$ must satisfy the relation

$$\lambda_{b,k} = \lambda_{b+1,k} + \lambda_{b+1,k+1}, \quad (4.77)$$

which can be shown to imply that there exists a unique finite measure Λ on $[0, 1]$ such that

$$\lambda_{b,k} = \int_0^1 f^k (1-f)^{b-k} \frac{\Lambda(df)}{f^2}. \quad (4.78)$$

In Kingman's coalescent, the measure Λ is a delta at 0, only coalescences of pairs ($k = 2$) occur and

$$\lambda_{b,k} = \mathbb{1}_{\{k=2\}}, \quad \langle T_2 \rangle = 1, \quad \langle T_3 \rangle = \frac{4}{3}, \quad \langle T_4 \rangle = \frac{3}{2}, \quad \langle T_n \rangle = 2 - \frac{2}{n} \quad [\text{Kingman}]. \quad (4.79)$$

In the Bolthausen-Sznitman coalescent, one has $\Lambda(df) = df$, any number of particles have a chance of coalescing together during one single event and

$$\lambda_{b,k} = \frac{(k-2)!(b-k)!}{(b-1)!}, \quad \langle T_2 \rangle = 1, \quad \langle T_3 \rangle = \frac{5}{4}, \quad \langle T_4 \rangle = \frac{25}{18}, \quad \langle T_n \rangle \sim \ln \ln n \quad (n \text{ large}) \quad [\text{Bolthausen-Sznitman}].$$

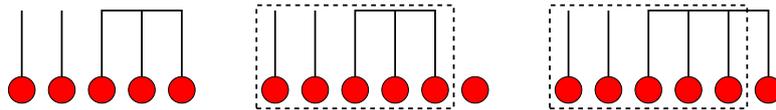


Figure 4.4: On the left drawing, the three rightmost particles coalesce together but without the two others with rate $\lambda_{5,3}$. One adds a sixth particle. The same three particles as before may coalesce together in two different ways: without the sixth particle (see middle drawing) with rate $\lambda_{6,3}$ or with the sixth particle (see right drawing) with rate $\lambda_{6,4}$. To ensure consistency, one must have $\lambda_{5,3} = \lambda_{6,3} + \lambda_{6,4}$.

The value for $\langle T_2 \rangle$ is clear: when there are only two particles, the probability that they coalesce during dt is $\lambda_{2,2} dt = dt$ and the coalescence time of the system is a simple exponential variable. When there are three particles (for instance), there is a probability $\lambda_{3,3} dt$ that they coalesce at once during dt . There is also a probability $3\lambda_{3,2} dt$ (because there are three pairs) that exactly two particles coalesce. Then, by considering what happens during the initial dt , one gets

$$T_3 = dt + \begin{cases} 0 & \text{with probability } \lambda_{3,3} dt, \\ T_2 & \text{with probability } 3\lambda_{3,2} dt, \\ T_3 & \text{with probability } 1 - \lambda_{3,3} dt - 3\lambda_{3,2} dt, \end{cases} \quad (4.80)$$

and, after averaging and using $\langle T_2 \rangle = 1$, one obtains

$$\langle T_3 \rangle = \frac{1 + 3\lambda_{3,2}}{3\lambda_{3,2} + \lambda_{3,3}}. \quad (4.81)$$

This is valid for any Λ -coalescent. In Kingman's case, $\lambda_{3,2} = 1$ and $\lambda_{3,3} = 0$ so that $\langle T_2 \rangle = \frac{4}{3}$. In the Bolthausen-Sznitman case, $\lambda_{3,2} = \lambda_{3,3} = \frac{1}{2}$ and $\langle T_2 \rangle = \frac{5}{4}$. Figure 4.5 shows side by side one realization of Kingman's coalescent and one realization of the Bolthausen-Sznitman coalescent when starting with 50 particles.

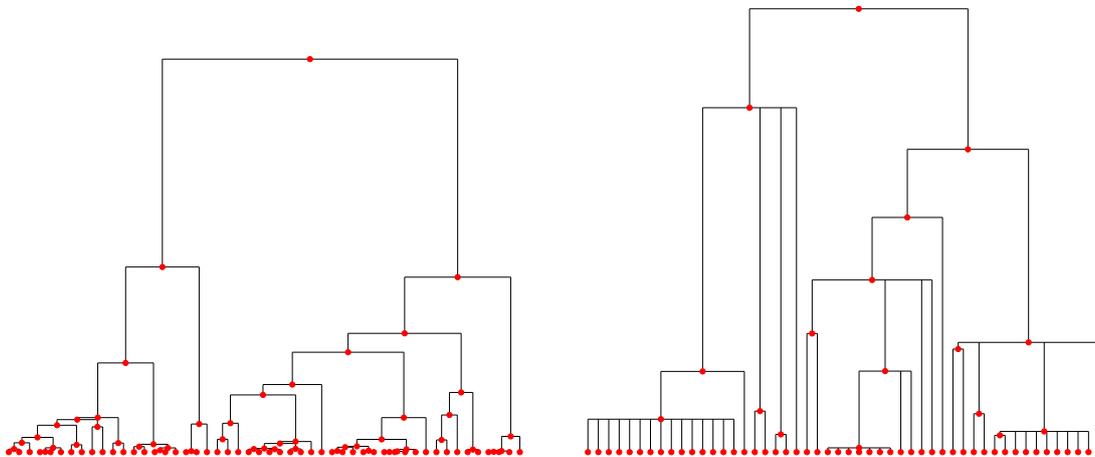


Figure 4.5: On the left, one realization of Kingman's coalescent. On the right, one realization of the Bolthausen-Sznitman coalescent. The total coalescence times are respectively 1.823 and 2.056.

4.2.4 Phenomenological theory for the genealogical trees

Recall the phenomenological theory for the position of the front: the front moves typically at the velocity of the cut-off theory, but every $\mathcal{O}(\ln^\alpha N)$ time-steps a large fluctuation occurs where a particle moves at a distance δ ahead of its typical position. The distribution of δ is $p(\delta) \propto e^{-\gamma_c \delta}$, see (4.37), and the effect of such a fluctuation is a shift $R(\delta) \approx \frac{1}{\gamma_c} \ln(1 + C_2 e^{\gamma_c \delta} / L^\alpha)$ in the position of the front, see (4.41) and (4.61). Here, $\alpha = 3$ in the noisy FKPP case, $\alpha = 1$ in the exponential model, and we recall that $L = (\ln N) / \gamma_c$.

The idea behind adapting this phenomenological theory to genealogical trees is the following: when a large fluctuation occurs, there is a shift in the position of the front due to the numerous descendants of the particle that jumped ahead of the front. In fact, after relaxation, the descendants of that particle represent a finite fraction f of the population. Indeed, for a front $h(x, t)$, the number of particles in a small Δx is $N[h(x, t) - h(x + \Delta x, t)] \approx -N \partial_x h(x, t) \Delta x$. Imagine now a front \tilde{h} which followed the same history as h except

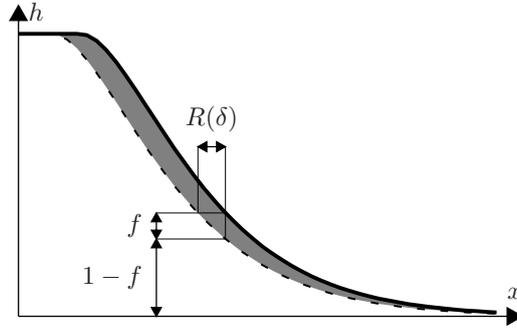


Figure 4.6: Relation between $R(\delta)$ and f . The dashed line (on the left of the grey area) is the front $h(x, t)$. The solid line (on the right of the grey area) is the front $\tilde{h}(x, t) = h(x - R(\delta), t)$ which followed the same evolution as h , except that a large fluctuation of size δ just occurred and relaxed. The grey area represents the particles that originated in the fluctuation, and $f \approx 1 - e^{-\gamma_c R(\delta)}$ is the fraction of the particles that originated in that fluctuation.

that a fluctuation of size δ occurred and relaxed recently. One has $\tilde{h}(x, t) = h(x - R(\delta), t)$ and the number of particles in Δx for \tilde{h} is approximately $-N\partial_x h(x - R(\delta), t)\Delta x$. The difference in the number of particles in Δx between h and \tilde{h} is due to the descendants of the particle that started the large fluctuation, see Figure 4.6. The fraction f of particles in Δx issued from the fluctuation is therefore

$$f \approx \frac{\partial_x h(x - R(\delta), t) - \partial_x h(x, t)}{\partial_x h(x - R(\delta), t)}. \quad (4.82)$$

But h (and $\partial_x h$) are essentially proportional to $e^{-\gamma_c x}$ in the interesting region (the x prefactor does not matter much). Hence, f is largely independent of x and is given by

$$f \approx 1 - e^{-\gamma_c R(\delta)} \approx 1 - \left(1 + C_2 \frac{e^{\gamma_c \delta}}{L^\alpha}\right)^{-1} = \left(1 + \frac{L^\alpha e^{-\gamma_c \delta}}{C_2}\right)^{-1}. \quad (4.83)$$

Recall that $p(\delta) d\delta$ is the probability per unit time that a fluctuation of size δ develops and that $p(\delta) \approx C_1 e^{-\gamma_c \delta}$ for $\delta \gg 1$, see (4.37). One then gets from (4.83) $df \approx \frac{L^\alpha \gamma_c}{C_2} e^{-\gamma_c \delta} f^{-2} d\delta$ and the probability $\tilde{p}(f) df$ per unit time that a large fluctuation occurs where a fraction f of the population is replaced by the descendants of one single runaway particle:

$$\tilde{p}(f) \approx \frac{C_1 C_2}{\gamma_c L^\alpha} \times \frac{1}{f^2} \quad \text{for } f \gg L^{-\alpha}. \quad (4.84)$$

For the exponential model, one has $C_1 = C_2 = \gamma_c = \alpha = 1$, so that

$$\tilde{p}(f) \approx \frac{1}{L} \times \frac{1}{f^2} \quad \text{for } f \gg 1/L \quad [\text{exponential model}]. \quad (4.85)$$

In the noisy FKPP case, $C_1 C_2 / \gamma_c = \pi^2 v''(\gamma_c)$ according to (4.44) and $\alpha = 3$:

$$\tilde{p}(f) \approx \frac{\pi^2 v''(\gamma_c)}{L^3} \times \frac{1}{f^2} \quad \text{for } f \gg L^{-3} \quad [\text{noisy FKPP}]. \quad (4.86)$$

Assume that a large fluctuation just occurred which replaced a fraction f of the population, and consider $b \geq 2$ particle. The probability that, out of these b particles, the $k \in \{2, 3, \dots, b\}$ first ones coalesce (without the $b - k$ other) within the fluctuation is $f^k (1 - f)^{b-k}$. As a fluctuation of size in df occurs per unit time with a probability $\tilde{p}(f) df$, one concludes that the rate at which k particles out of b coalesce is given by

$$\tilde{\lambda}_{b,k} = \int f^k (1 - f)^{b-k} \tilde{p}(f) df. \quad (4.87)$$

With $\tilde{p}(f)$ given by either (4.85) or (4.86), one recognizes the rates $\lambda_{b,k}$ of the Bolthausen-Sznitman coalescent scaled by a factor $1/L$ or $\pi^2 v''(\gamma_c)/L^3$, see (4.78) with $\Lambda(df) = df$. The prediction (validated by numerical simulation) is then that

$$\langle T_2 \rangle \sim L \quad [\text{exponential model}] \quad \langle T_2 \rangle \sim \frac{L^3}{\pi^2 v''(\gamma_c)} \quad [\text{noisy FKPP}], \quad (4.88)$$

with the ratios (4.75). Combined to (4.24) and (4.60), and recalling that $L = (\ln N)/\gamma_c$, one gets, for either the exponential model or the noisy FKPP,

$$D_N \times \langle T_2 \rangle = \frac{\pi^2}{3\gamma_c^2}, \quad (4.89)$$

with $\gamma_c = 1$ in the exponential model.

One needs to be a little bit careful with the meaning of (4.87). In the standard coalescent setting, there is for each dt a probability $\lambda_{b,k} dt$ that k out of b particles coalesce at once. Here, (4.87) means that for each dt there is a probability $\tilde{\lambda}_{b,k} dt$ that a large fluctuation develops in which k out of b particles find a common ancestor. But, certainly, the k particles cannot coalesce at once for models in the noisy FKPP class (there are models where each particle has at most two children at each time-step). What happens is that large fluctuations occur every L^3 unit of times and relax within a time L^2 . When following the ancestry of the k particles out of b , there is no coalescence during the time of order L^3 needed to go back to the large fluctuation, and then the particles coalesce in several steps during the relatively short time L^2 that the fluctuation takes to relax. This time L^2 becomes instantaneous after rescaling time by L^3 , and only with this rescaling can one recover the Bolthausen-Sznitman coalescent with its multiple coalescences.

4.3 Some other models around the noisy FKPP class

Some of the results presented so far have been rigorously proved in several models. We make a small review of these results.

The models (4.1) and (4.2) on which we have focused so far describe the stochastic evolution of a cloud of N particles in discrete time. The random front $h(x, t)$ is then defined as the fraction of particles on the right of x . However, in all our heuristic explanations, we did not need to go into the specifics of the model. Therefore, one might expect our results to extend to other models in the noisy FKPP class.

In this section, we briefly describe several models which, except for the last, fall in the universality class of the noisy FKPP equation and, in some cases, review rigorous results that have been obtained.

Reaction-diffusion

Consider the reaction-diffusion model on the lattice described in Section 1.5.1. There are two types of particles (A and B), a total of N particles per site, and the number of A particles at site x at time t is by definition $Nh(x, t)$. Particles in adjacent sites may exchange positions and in any given site a A particle may contaminate a B particle and change it into another A :



(One can also allow $A+B \rightarrow 2B$ with a smaller rate.) The quantity $h(x, t)$ follows the noisy front equation (1.45):

$$\partial_t h(x, t) = \frac{h(x+a, t) + h(x-a, t) - 2h(x, t)}{a^2} + h - h^2 + \text{noise}, \quad (4.91)$$

where a is the lattice spacing and where the noise is of order $\sqrt{h(1-h)/N}$. As in (4.1) and (4.2), the front values $h(x, t)$ are multiples of $1/N$, the noise term is non-Gaussian, correlated over different lattice sites and of order $\sqrt{h/N}$ for h small. An important difference with (4.1) or (4.2) is that $x \mapsto h(x, t)$ is no longer a non-increasing function, because h is no longer the fraction of particles on the right of x .

The noise in (4.91) is negligible everywhere except where h or $1-h$ is of order $1/N$. One does not expect what happens for h close to 1 to matter, because the velocity selection mechanism of a front equation is controlled by the region where h is small. The cut-off theory applies directly: there are no allowed positive values of h that are smaller than $1/N$. All the arguments of Section 4.1.1 apply and one expects the same $1/\ln^2 N$ correction to the velocity as in (4.23) with the sinus shape for the front as in (4.16). The phenomenological theory of Section 4.1.2 also applies with no modification, and one expects (4.24) to hold, because on the right of the front, where the population per site of A particles is small compared to N , the stochastic evolution of the A particles looks like a BRW on the lattice.

The reaction (4.90) can be interpreted as a parent A giving birth to two A children, and one can clearly draw genealogical trees. The results of Section 4.2 are however difficult to apply, because the total number of A increases indefinitely instead of staying equal to N . The coalescence time between two A particles clearly depends on the (arbitrarily large) distance between the two particles.

The stochastic FKPP equation

It is tempting to send a to 0 in (4.91) and to replace the complicated noise by a nice delta-correlated Gaussian noise term. One then gets the stochastic FKPP equation (1.54) introduced in Section 1.5.5:

$$\partial_t h = \partial_x^2 h + h - h^2 + \sqrt{\frac{h-h^2}{N}} \eta(x, t), \quad (4.92)$$

where $\eta(x, t)$ is a delta-correlated Gaussian noise:

$$\langle \eta(x, t) \rangle = 0, \quad \langle \eta(x, t) \eta(x', t') \rangle = \delta(x - x') \delta(t - t'). \quad (4.93)$$

This equation does not occur in a very natural way and, in particular, it is not the hydrodynamic limit of (4.91) [DMS03]. N in (4.92) is no longer a total population size, as in the models (4.1) and (4.2), nor a population size per site, as in (4.91), but simply some large parameter. The values $h(x, t)$ are no longer constrained to be multiples of $1/N$, as in (4.1), (4.2) and (4.91), but can take any value in $[0, 1]$.

Because h can take all the values in $[0, 1]$, applying the cut-off argument is less straightforward. It can however still be done because one effect of the noise is to bring back very quickly to zero any value which is small compared to $1/N$. One way to understand this is to remove the spatial dimension and to consider the following equation for $h(t)$:

$$dh = (h - h^2) dt + \sqrt{\frac{h - h^2}{N}} dW. \quad (4.94)$$

The process $h(t)$ eventually gets stuck at either $h = 0$ or $h = 1$. It is an easy computation to check that

$$\mathbb{P}(h \text{ gets eventually stuck at } 0) = \frac{e^{-2Nh_0} - e^{-2N}}{1 - e^{-2N}}, \quad (4.95)$$

where $h_0 \in [0, 1]$ is the starting value of the process. (Simply call $f(h_0)$ the probability (4.95); then by looking at what happens during the first dt one gets $f(h_0) = \langle f(h_0 + (h_0 - h_0^2) dt + \sqrt{(h_0 - h_0^2)/N} dW) \rangle$, which leads to $2Nf'(h_0) + f''(h_0) = 0$. With the boundary conditions $f(0) = 1$ and $f(1) = 0$ one recovers (4.95). See also [PL99; DMS03].) When N is large, what (4.95) means is that the process takes off and reaches 1 with a large probability if h_0 is large compared to $1/N$ and, on the other hand, the process goes back to 0 with large probability if h_0 is small compared to $1/N$.

Going back to (4.92), with the spatial dimension, the same mechanism occurs: the noise term quickly brings h to zero in any region where h is small compared to $1/N$. It was actually shown [MS95; MMQ11] that for an initial condition that decays fast enough, then at all positive times there exists a random position $r(t)$ such that $h(x, t) = 0$ for $x > r(t)$. The intuition is that, because of the noise term, $h(x, t)$ cancels at a small distance on the right of the point where it is of order $1/N$. We are then back to a situation where one can apply the cut-off theory: the front h cancels very close to the point where it reaches $1/N$, hence it must look like the travelling wave (4.16) which moves at the velocity (4.23), with the $1/\ln^2 N$ correction.

For the stochastic FKPP equation (4.92), Mueller, Mytnik and Quastel have shown [MMQ08; MMQ11] that, for N large enough,

$$2 - \frac{\pi^2}{\ln^2 N} - 28\pi^2 \frac{\ln \ln N}{\ln^3 N} \leq v_N \leq 2 - \frac{\pi^2}{\ln^2 N} + 9\pi^2 \frac{\ln \ln N}{\ln^3 N}, \quad (4.96)$$

see also [CD05]. This result is in agreement with the prediction from the cut-off (4.23) and is compatible with the prediction (4.24) for the next order term from the phenomenological theory of Section 4.1.2.

Note that it is not easy to see directly on (4.92) why the phenomenological theory should apply to the stochastic FKPP equation, as there are no particles in that model. For the same reason, there does not seem to be a straightforward way to apply the results on genealogical trees of Section 4.2 on that equation.

The coalescing BBM

Recall the coalescing BBM also introduced in Section 1.5.5: it is a standard BBM with the added rule that two particles meeting may coalesce with a small rate $\epsilon = 2/N$. At the tip of the front, where particles are few, the coalescences do not occur and everything behaves as in an usual BBM. Far on the left, however, the number of particles per unit length fluctuates around N , which is the value for which branching and coalescence occur at the same rate. As in the reaction-diffusion model on the lattice (4.90), the cut-off theory and the phenomenological description are expected to apply and lead to the result (4.24) for the velocity and diffusion constant of the rightmost particle. However, again as in (4.90), it is not clear how to apply the results on genealogical trees.

Because of the duality (1.57) between the stochastic FKPP and the coalescing BBM, the velocity, diffusion constant and other cumulants of the position per unit time must be the same in both models.

The N -BBM

In the N -BBM, particles diffuse and branch as in a BBM but, each time the population size is larger than N , the leftmost particle is removed; see the illustration in Figure 4.2, page 41. The N -BBM is really a time-continuous version of model (4.2); it describes the motion of a cloud of N particles, and the cut-off theory, the phenomenological theory and the results on genealogical trees are expected to hold. The N -BBM was rigorously

studied by Maillard in 2016 [Mai16], who showed that if one starts with N particles distributed along a well chosen sinus-exponential shape similar to (4.16) then, for fixed $u > 0$ and fixed $\alpha \in (0, 1)$,

$$\left\langle \text{position of the } (\alpha N)\text{-th rightmost particle at time } u \ln^3 N \right\rangle = x_\alpha + \tilde{v}_N \times u \ln^3 N + o(1) \quad \text{as } N \rightarrow \infty, \quad (4.97)$$

for some explicit function x_α and with \tilde{v}_N given by

$$\tilde{v}_N = 2 - \frac{\pi^2}{\ln^2 N} + 6\pi^2 \frac{\ln \ln N}{\ln^3 N} + \frac{c}{\ln^3 N}, \quad (4.98)$$

where c has a complicated expression. \tilde{v}_N is the same as the velocity v_N predicted (4.24) by the phenomenological theory with an extra $c/\ln^3 N$ term. (Whether or not that extra term is universal is an open question.) Maillard also shows that, to leading order and on the same time scale $\ln^3 N$, the diffusion constant and all the other cumulants per unit time have the values predicted by (4.24).

This is a beautiful result, which validates the phenomenological theory. Unfortunately, it does not work for an arbitrary initial condition and it does not prove that the front goes at the velocity (4.98) on different time scales: there is no proof that \tilde{v}_N defined in (4.97) is the same quantity (up to negligible terms) as the velocity v_N defined, for a fixed N , as the $t \rightarrow \infty$ limit of $1/t$ times the position of the front at time t .

One expects the rescaled genealogical trees in the N -BBM to converge to the Bolthausen-Sznitman coalescent, but no proof exists.

The L -BBM

In the L -BBM, particles diffuse and branch as in the BBM, but all particles that reach a distance L from the rightmost particle are removed; see the illustration in Figure 4.2, page 41. The number of particles in the L -BBM fluctuates. Recall however that, in the cut-off theory of a model with N particles, the distance between the centre of the cloud and the rightmost particle is $(\ln N)/\gamma_c$. As $\gamma_c = 1$, one can then expect the L -BBM to look most of the time very much like the N -BBM if

$$L \sim \ln N. \quad (4.99)$$

The results (4.24) from the cut-off theory and phenomenological description should apply with the simple substitution (4.99), and the genealogical trees with time rescaled by L^3 are expected to converge to the Bolthausen-Sznitman coalescent.

In [Pai16], Michel Pain proved that the velocity of the L -BBM is given by

$$v_L = 2 - \frac{\pi^2}{L^2} + o\left(\frac{1}{L^2}\right), \quad (4.100)$$

which is the same as the prediction from the cut-off theory (4.23) with the mapping (4.99) from L to N .

BBM with an absorbing wall

The BBM with an absorbing wall has already been discussed a lot in Section 4.1.3 in order to obtain the velocity v_N of the noisy FKPP front. Genealogies in this model have also been studied in [BBS13], where the authors found that, for the BBM with an absorbing wall at velocity v_N and with, initially, around N particles located not too far on the right, the genealogy of the particles seen on a time scale of order $(\ln N)^3$ converges to the Bolthausen-Sznitman coalescent.

Models of evolution with fitness

One of the most studied classes of models of evolution is the one in which individuals carry a fitness value. They reproduce (or branch) with a rate which increases linearly with the difference between their fitness and the population's average fitness, and die with a constant rate which is adjusted in such a way that the population size remains constant. Furthermore, the fitness values evolve randomly because of mutations. (There is of course much variability in the precise definition of models in that class: is time discrete or continuous? Is the fitness space discrete or continuous? Are deleterious mutations allowed, or only beneficial mutations? Do mutations only occur at branching events or at all times? etc.)

Representing the fitness values as positions on a spatial dimension, these models of evolution are branching processes which bear some similarities with models in the noisy FKPP class: in either case, particles on the right are at an advantage over particles on the left. The difference comes from the way this advantage plays out: in the N -BBM (to take an example in the FKPP class), all the particles have the same birth rate but only particles on the left die. In models with fitnesses, all the particles have the same death rate but particles on the right have a higher birth rate.

The models of evolution with fitness do not enter in the universality class of FKPP. In particular, as N increases, the velocity v_N (or rate at which beneficial mutations accumulate) diverges. The way v_N diverges depends on how the mutation rate and the selective advantage of a mutation both scale with N . It has been estimated in several scaling regimes [DF07; RBW08; BRW08; PSK10; YEC10; Hal11; HG15; GD13; Sch15a]. (The literature is huge, and this list of references is incomplete. To find more, [PSK10; YEC10; Sch15a] give many pointers.)

Nevertheless, when time is properly rescaled (not by $(\ln N)^3$ but by some more complicated expression), it turns out that the genealogical trees are still described by the Bolthausen-Sznitman coalescent [DWF13; NH13; Sch15b]. The mechanism by which the Bolthausen-Sznitman coalescent appears is roughly the same as in the phenomenological theory of the FKPP case: from time to time, a lucky individual gets a higher fitness than what is typical and it replaces over some short time a fraction of the population.

Chapter 5

Conclusion

I have reviewed several aspects of branching processes, and of fronts described by an equation in the FKPP class or in the noisy FKPP class. An equation in the FKPP class must feature some kind of diffusion, a growth term, and a saturation mechanism. A remarkable feature is the apparent robustness of the results that can be obtained; for instance, with a front in the FKPP class, one only has to linearise the equation to find the function $v(\gamma)$. Then, the velocity of the front, its shape and its asymptotic position can be determined with great accuracy from the knowledge of this $v(\gamma)$ only. If the front has an internal noise of order $\sqrt{h/N}$ where h is small, the correction to the velocity and, in fact, all the cumulants of the position can be determined from $v(\gamma)$ and N only.

The main reason for this universality seems to be that all that matters in a FKPP front takes place in the linear region, where the equation reduces to diffusion plus growth. The presence of a saturation term is however essential: the fully linear equation has a different behaviour.

As a physicist, I took advantage many times of this universality by first choosing, for each problem, the most practical model available in order to obtain a result, and then by generalizing the result to other fronts in the FKPP class. The problem of course is that this universality is not a proven fact and, rigorously, one cannot extend results derived within one model to another.

An interesting open problem, among many others, is to determine the scope of this universality: which property is universal, and which is not? To pick an example, consider the asymptotic expansion of the position $\mu_t^{(1/2)}$ of the front; even though a rigorous proof of the $1/\sqrt{t}$ Ebert and van Saarloos term exists only for one toy model (see Section 2.2.1), it is widely believed to be true for any equation in the FKPP class. But what about the next term, which we argued to be of order $(\ln t)/t$, and the second next, of order $1/t$? My intuition is that the former is universal, but not the latter. This is however only an unconvincing guess which needs to be substantiated. One way to proceed could be to manage to derive the same result for several other models in the FKPP class, which would make the guess more convincing. A more ambitious goal would be to understand better the universality class itself and to determine which properties are universal or not.

Part II

Facsimile of selected publications

Phenomenological theory giving the full statistics of the position of fluctuating pulled fronts

E. Brunet,¹ B. Derrida,¹ A. H. Mueller,² and S. Munier^{3,4}

¹Laboratoire de Physique Statistique, École Normale Supérieure, 24 rue Lhomond, 75231 Paris cedex 05, France

²Department of Physics, Columbia University, New York, New York 10027, USA

³Centre de Physique Théorique, Unité mixte de recherche du CNRS (UMR 7644), École Polytechnique, 91128 Palaiseau, France

⁴Dipartimento di Fisica, Università di Firenze, via Sansone 1, 50019 Sesto F., Florence, Italy

(Received 1 December 2005; published 26 May 2006)

We propose a phenomenological description for the effect of a weak noise on the position of a front described by the Fisher-Kolmogorov-Petrovsky-Piscounov equation or any other traveling-wave equation in the same class. Our scenario is based on four hypotheses on the relevant mechanism for the diffusion of the front. Our parameter-free analytical predictions for the velocity of the front, its diffusion constant and higher cumulants of its position agree with numerical simulations.

DOI: 10.1103/PhysRevE.73.056126

PACS number(s): 02.50.-r, 05.40.-a

I. INTRODUCTION

The Fisher Kolmogorov-Petrovsky-Piscounov (FKPP) equation [1]

$$\partial_t h = \partial_x^2 h + h - h^2 \quad (1)$$

describes how a stable phase [$h(x,t)=1$ for $x \rightarrow -\infty$] invades an unstable phase [$h(x,t)=0$ for $x \rightarrow +\infty$] and how the front between these two phases builds up and travels [2]. This equation was first introduced in a problem of genetics, but equations similar to Eq. (1) appear in much broader contexts like reaction-diffusion problems [3,4], optimization [5], disordered systems [6,7], and even particle physics [8–10]. A remarkable example is the problem of the high-energy scattering of a projectile consisting of a small color dipole on a target in the framework of quantum chromodynamics (QCD): in Ref. [8] it was recognized that the Balitsky-Kovchegov (BK) equation [9], a mean-field equation for high-energy scattering in QCD, is in the same class as the FKPP equation with h being the scattering amplitude, t the rapidity of the scattering, and x the logarithm of the inverse projectile size.

It is well known [2,11] that equations like Eq. (1) have a family of traveling-wave solutions of the form $h(x,t)=h(z)$ with $z=x-vt$. There is a relation between the exponential decay of each solution [$h(z) \sim \exp(-\gamma z)$ for large z] and its velocity: $v=v(\gamma)$. For example, $v(\gamma)=\gamma+1/\gamma$ for the FKPP equation (1). Other front equations would give different expressions of $v(\gamma)$. See, for example, Sec. IV or Refs. [12,13].

If one starts with a steep enough initial condition, the front converges to the traveling wave with the minimal velocity. Therefore,

$$v_{\text{deterministic}} = \min_{\gamma} v(\gamma) = v(\gamma_0) \quad \text{where } v'(\gamma_0) = 0, \quad (2)$$

$$h_{\text{deterministic}}(z) \approx A z e^{-\gamma_0 z}.$$

(The multiplicative factor z in $h_{\text{deterministic}}$ is present only for this slowest moving solution.)

There is a large class (the FKPP class) of equations describing the propagation of a front into an unstable state which select the minimal velocity, as described by (2). [There exist also equations of fronts propagating into an unstable state, called “pushed” or “type II,” for which the velocity selected by the front is not the slowest one and equations of fronts propagating into a stable state. The properties of these fronts are quite different [2,14,15] from the properties of Eq. (1), and we will not consider them in the present paper.]

Deterministic front equations such as Eq. (1) usually occur as the limit of a stochastic reaction-diffusion model [16] when the number of particles (or bacterias or reactants) involved becomes infinite. In a physical situation, all numbers remain finite and a small noise term should be added to Eq. (1) to represent the fluctuations at the microscopic scale. One might write, for instance [17],

$$\partial_t h = \partial_x^2 h + h - h^2 + \sqrt{h(1-h)/N} \eta(x,t), \quad (3)$$

where $\eta(x,t)$ is a normalized Gaussian white noise and N is the number of particles involved.

The effect of such a noise is to make the shape of the traveling-wave fluctuate in time [4]. It affects also its velocity and makes the front diffuse [2,16,18].

For a chemical problem, N might be of the order of the Avogadro number and one could think that such a small noise term should give small corrections, of order $1/\sqrt{N}$, to the shape and position of the front. However, because the front motion is extremely sensitive to small fluctuations in the region where $h \approx 1/N$, this is not the case. In the presence of noise as in Eq. (3), the front has an exponential decay if $h(x,t) \gg 1/N$, but it vanishes much faster than this exponential in the region where $h(x,t)$ is of order $1/N$ [4]. (This is obvious in a particle model, as there cannot be less than one particle at a given place.) As an approximation to understand the effect of the microscopic stochastic details of the system, it has been suggested to replace the noise term by a deterministic cutoff which makes the front vanish very quickly

when $h \approx 1/N$ [12]. For instance, for the FKPP equation (1), one way of introducing the cutoff is

$$\partial_t h = \partial_x^2 h + (h - h^2)a(Nh),$$

$$\text{with } a(r) = 1 \text{ for } r > 1 \text{ and } a(r) \rightarrow 0 \text{ for } r \rightarrow 0. \quad (4)$$

In the presence of such a cutoff, the velocity and shape (2) become, for any equation in the FKPP class,

$$h_{\text{cutoff}}(z) \approx A \frac{L}{\pi} \sin\left(\frac{\pi z}{L}\right) e^{-\gamma_0 z} \quad \text{where } L = \frac{1}{\gamma_0} \ln N, \quad (5a)$$

$$v_{\text{cutoff}} \approx v(\gamma_0) - \frac{\pi^2 v''(\gamma_0)}{2L^2}. \quad (5b)$$

[The shape (5a) is valid only in the linear region, where h is small enough for the nonlinear term h^2 to be negligible but still larger than $1/N$. Note that for $z \ll L$, the shape coincides with (2). A way to interpret the sine is to say that the front moves slower than the minimal velocity $v_{\text{deterministic}} = v(\gamma_0)$ and that the decay rate becomes complex: $\gamma = \gamma_0 \pm i\pi/L$. Then, the expression of v_{cutoff} results from an expansion of $v(\gamma)$ for large L .]

The prediction (5) does not depend on the details of the microscopic model. It only depends on the deterministic equation and on the existence of a microscopic scale. This cutoff picture is also present in the mean field QCD context in [19], where it was introduced to avoid unitarity violating effects in the BK equation at intermediate stages of rapidity evolution. In this context, N is $1/\alpha_{\text{QCD}}^2$ where α_{QCD} is the strong-coupling constant.

Extensive numerical simulations of noisy fronts have been performed over the years [3,18], and the large correction (5b) to the velocity found in the cutoff picture seems to give the correct leading correction to the velocity of noisy fronts. (See [20] for rigorous bounds.) Being a deterministic approximation, the cutoff theory gives, however, no prediction for the diffusion constant of the front.

In the present paper, we develop a phenomenological description which leads to a prediction for this diffusion constant. This description tries to capture the rare relevant events which give the dominant contribution to the fluctuations in the position of the front. The prediction is that the full statistics of the front position in the noisy model depends only on the amplitude $1/N$ of the noise at the microscopic scale and on $v(\gamma)$, a property of the deterministic equation. For large N , all the other details of the underlying microscopic model do not contribute to the leading order. Our description leads to the following prediction for the velocity and for the diffusion constant of the front for large N :

$$v - v_{\text{cutoff}} = \pi^2 \gamma_0^3 v''(\gamma_0) \frac{3 \ln \ln N}{\gamma_0^2 \ln^3 N} + \dots, \quad (6a)$$

$$D = \pi^2 \gamma_0^3 v''(\gamma_0) \frac{\pi^2/3}{\gamma_0^2 \ln^3 N} + \dots. \quad (6b)$$

Actually, our phenomenological approach also gives a prediction to the leading order for all the cumulants of the position of the front. For $n \geq 2$,

$$\frac{[n\text{th cumulant}]}{t} = \pi^2 \gamma_0^3 v''(\gamma_0) \frac{n! \zeta(n)}{\gamma_0^2 \ln^3 N} + \dots, \quad (6c)$$

where $\zeta(n) = \sum_{k \geq 1} k^{-n}$.

The $1/\ln^3 N$ dependence of the diffusion constant was already observed in numerical simulations [18]. In the QCD context, it was proposed in [21] to identify the full QCD problem with a stochastic evolution, such as Eq. (3), and the dependence of the diffusion constant was used to suggest a new scaling law for QCD hard scattering at, perhaps, ultra-high energies.

We do not have, at present, a mathematical proof of the results (6). Rather, we believe that we have identified the main effects contributing to the diffusion of the front. We present our scenario in Sec. II where we state a set of four hypotheses from which the results (6) follow. We give arguments to support these hypotheses in Secs. III A–III D. Finally, to check our claims, we present numerical simulations in Sec. IV for the five first cumulants of the position of the front. These simulations match very well the predictions (6).

II. PICTURE AND ITS QUANTITATIVE CONSEQUENCES

To simplify the discussion, we consider, in this section, more specifically a microscopic particle model rather than a continuous stochastic model such as Eq. (3). This is merely a convenience to make our point clearer, but the discussion below could be rephrased for other models in the stochastic FKPP class.

We consider models where particles diffuse on the line and, occasionally, duplicate. If one takes, for $h(x, t)$, the density of particles or, alternatively, the number of particles on the right of x , it is clear that it is not yet described by a front equation, because it grows exponentially fast with time; one needs to introduce a saturation rule. For instance, one can (i) keep the number of particles fixed by removing the leftmost particles if necessary, or (ii) remove all the particles which are at a distance larger than L behind the rightmost particle, or (iii) limit the density by allowing, with a small probability, that two particles meeting recombine into one single particle [4].

A. Scenario for the propagation of the front

The main picture of our phenomenological description is the following. The evolution of the front is essentially deterministic, and its typical shape and velocity are given by Eq. (5). But from time to time, a fluctuation sends a small number of particles at some distance δ ahead of the front. At first, the position of the front, determined by where most of the particles are, is only modified by a negligible amount of order $1/N$ by this fluctuation. However, as the system relaxes, the number of wandering particles grows exponentially and they start contributing to the position of the front. Meanwhile, the bulk catches up and absorbs the wandering

particles and their many offsprings; finally, the front relaxes back to its typical shape (5a). The net effect of a fluctuation is therefore to shift the position of the front by some amount $R(\delta)$ which depends, obviously, on the size δ of the fluctuation. A useful quantity to characterize the fluctuations is the width of the front. It can easily be defined as the distance between the leading particle (where $h \approx 1/N$) and some position in the bulk of the front—for instance, where $h=0.5$. (Changing this reference point would change the width by a finite amount, independent of N .) This width is typically of order L , where L is given by the cutoff theory (5a). During a fluctuation that sends particles at a distance δ ahead of the front, the width of the front increases quickly to $L+\delta$ and then relaxes slowly back to L .

We emphasize that, in this scenario, the effect of noise is so weak that, most of time, it can be ignored and the cutoff theory describes accurately the evolution of the front. It is only occasionally, when a rare sequence of random microscopic events sends some particles well ahead of the front that the cutoff theory is no longer valid. The way this fluctuation relaxes is, however, well described by the deterministic cutoff theory.

We shall encode this scenario in the following quantitative assumptions.

(i) If we write the instantaneous fluctuating width of the front as $L+\delta$, then the probability distribution function for δ is given by

$$p(\delta)d\delta = C_1 e^{-\gamma_0 \delta} d\delta, \quad (7)$$

where C_1 is some constant. Note that we assume this form only over some relevant range of values: δ large enough (compared to 1) but much smaller than L (typically of order $\ln L$). Fluctuations where δ is “too small” are frequent but do not contribute much to the front position. Fluctuations where δ is “too large” are so rare that we do not need to take them into account. Only for “moderate” values of δ do we assume the above exponential probability distribution function.

(ii) The long-term effect of a fluctuation of size δ (assuming that there are no other fluctuations in between) is a shift of the front position by the quantity

$$R(\delta) = \frac{1}{\gamma_0} \ln \left(1 + C_2 \frac{e^{\gamma_0 \delta}}{L^3} \right), \quad (8)$$

where C_2 is another constant.

(iii) The fluctuations of the position of the front are dominated by large and rare fluctuations of the shape of the front. We assume that they are rare enough that a given relevant fluctuation has enough time to relax before another one occurs.

From these three hypotheses alone, one can derive our results (6) up to a single multiplicative constant. This constant can be determined with the help of a fourth hypothesis.

(iv) For the aim of computing the first correction to the front velocity obtained in the cutoff theory (5), one can simply use the expression (5b) with L replaced by L_{eff} where

$$L_{\text{eff}} = \frac{1}{\gamma_0} \ln N + \frac{3}{\gamma_0} \ln \ln N + \dots \quad (9)$$

It is important to appreciate that the typical width of the front is still L and not L_{eff} . The latter quantity is just what should be used in (5b) to give the correct velocity.

A similar scenario was used by Kloster [22] to obtain the $\ln \ln(N)/\ln^3 N$ correction to the cutoff velocity of Eq. (6a). However, his prefactor differs from ours: in our notations, Ref. [22] would give a “2” in the numerator of the right-hand side of Eq. (6a) instead of our “3.”

B. How Eqs. (6) follow from these hypotheses

We are now going to see how the results (6) follow from these four hypotheses.

First, we argue that the probability to observe a fluctuation of size δ during a time interval Δt can be written as $p(\delta)d\delta\Delta t/\tau$, where $p(\delta)$ is the distribution (7) of the increase of the width of the front and where τ is some typical time characterizing the rate at which these fluctuations occur. Indeed, during a fluctuation of a given size, the width of the front increases to that size and then relaxes back. For a large δ , observing a front of size $L+\delta$ is very rare, but when it happens, the most probable is that one is observing the maximum expansion of a fluctuation with a size close to δ ; the contribution from fluctuations of sizes significantly larger than δ is negligible as they are much less likely.

Second, as a fluctuation builds up at the very tip of the front where the saturation rule (see beginning of Sec. II) can be neglected, we argue that the typical time τ introduced in the previous paragraph and the time it takes to build a fluctuation of a given size do not depend on N . (However, the relaxation time of a fluctuation depends on N as the bulk of the front is involved in the relaxation.)

Let X_t be the position of the front, δ_0 the minimal size of a fluctuation giving a relevant contribution to the position of the front, and Δt a time much smaller than the time between two relevant fluctuations, but much larger than the time it takes to build up such a fluctuation and have it relax. (This is authorized by the third hypothesis.) We have

$$X_{t+\Delta t} = \begin{cases} X_t + v_{\text{cutoff}}\Delta t + R(\delta) & \text{prob. } \frac{\Delta t}{\tau} p(\delta)d\delta \text{ for } \delta > \delta_0, \\ X_t + v_{\text{cutoff}}\Delta t & \text{prob. } 1 - \frac{\Delta t}{\tau} \int_{\delta_0}^{\infty} p(\delta)d\delta. \end{cases}$$

[Note that $\frac{\Delta t}{\tau} \int_{\delta_0}^{\infty} p(\delta)d\delta$ is the probability of observing a relevant fluctuation during the time Δt . By definition of Δt , this is much smaller than 1.]

One can then compute the average, denoted $\langle \cdot \rangle$, of $\exp(\lambda X_{t+\Delta t})$. One gets, for λ small enough,

$$\partial_t \ln \langle e^{\lambda X_t} \rangle = \lambda v_{\text{cutoff}} + \frac{1}{\tau} \int p(\delta) [e^{\lambda R(\delta)} - 1] d\delta. \quad (10)$$

Expanding in powers of λ , one recognizes on the left-hand-side the cumulants of X_t . Therefore, one gets

$$v - v_{\text{cutoff}} = \frac{1}{\tau} \int p(\delta) R(\delta) d\delta,$$

$$\frac{[n\text{th cumulant}]}{t} = \frac{1}{\tau} \int p(\delta) R^n(\delta) d\delta \text{ for } n \geq 2. \quad (11)$$

At this point, one can notice from the expressions of $p(\delta)$ and $R(\delta)$ that the values of δ such that $e^{\gamma_0 \delta} \gg L^3$ have a negligible contribution to the integrals giving the velocity and the cumulants. Thus appears naturally a $\delta_{\text{max}} = (3/\gamma_0) \ln L$ which is exactly the effective correction to the width of the front appearing in Eq. (9).

The integrals in Eq. (11) can be evaluated, and one gets

$$\int p(\delta) R^n(\delta) d\delta = \frac{C_1 C_2}{\gamma_0^{n+1} L^3} \int_0^{L^3/C_2} \ln^n \left(1 + \frac{1}{x} \right) dx, \quad (12)$$

with $x = (L^3/C_2) \exp(-\gamma_0 \delta)$. For $n=1$, this integral gives $\ln(L^3/C_2)$. For $n \geq 2$, one can integrate from 0 to ∞ (the correction is at most of order $1/L^6$) and one recognizes $n! \zeta(n)$. Finally,

$$v - v_{\text{cutoff}} = \frac{C_1 C_2}{\tau \gamma_0} \frac{3 \ln L}{\gamma_0 L^3},$$

$$\frac{[n\text{th cumulant}]}{t} = \frac{C_1 C_2}{\tau \gamma_0} \frac{n! \zeta(n)}{\gamma_0^n L^3}. \quad (13)$$

Everything is determined up to *one* numerical constant $C_1 C_2 / \tau$. As the fourth hypothesis gives the velocity, one can easily determine that constant and recover Eqs. (6).

All the cumulants (except the first one) are of the same order of magnitude, as the fluctuations are due to rare big events.

III. ARGUMENTS TO SUPPORT THE HYPOTHESES

A. First hypothesis

This first hypothesis is not very surprising if one considers that $\exp(-\gamma_0 \delta)$ is the natural decay rate of the deterministic equation. A more quantitative way to understand Eq. (7) is that building up a fluctuation is an effect which is very localized at the tip of the front, where saturation effects can be neglected. We present in Appendix A a calculation using this property.

Moreover, numerical simulations [23] of that probability distribution function give evidence that for large enough N , the decay is exponential with the rate γ_0 as in Eq. (7).

B. Second hypothesis

To obtain Eq. (8), we need to compute the response of the deterministic model with a cutoff (4) to a fluctuation at the tip of the front. This is a purely deterministic problem: starting with a fluctuation (i.e., a configuration slightly different from the stationary shape), we let the system evolve with a cutoff and relax back to its stationary shape (5a), and we would like to compute the shift in position due to this fluctuation.

Although the evolution is purely deterministic, the problem remains a difficult one. For simplicity, we discuss here the case of the FKPP equation (4). The extension to other traveling wave equations in the FKPP class is straightforward.

There are two nonlinearities in Eq. (4): one is the $-h^2$ term, which is important when h is of order 1, and the other one is the cutoff term $a(Nh)$, which is important when h is of order $1/N$. Between these two points, there is a large length of order $L = \ln N$ where one can neglect both nonlinearities. This means that, for all practical purpose, one can simply use the linearized version of the FKPP equation for the whole front except for two small regions with a size of order 1 at both ends of the front.

Let X_t be the position of the front and L_t its length. There are many equivalent ways of defining precisely these quantities; for instance, we can take X_t such that $h(X_t, t) = 10^{-5}$ and L_t such that $h(X_t + L_t, t) = \frac{1}{N}$. We expect that $X_t - v_{\text{cutoff}} t$ and $L_t - L$, which are quantities of order 1, have a relaxation time of order L^2 , as for the shape of the front [14,24].

For $X_t < x < X_t + L_t$, the problem is linear:

$$\partial_t h = \partial_x^2 h + h. \quad (14)$$

Using the ansatz

$$h(x, t) = L_t G \left(\frac{x - X_t}{\frac{L_t}{\gamma}}, \frac{t}{\frac{L_t^2}{\tau}} \right) e^{-(x - v_{\text{cutoff}} t)}, \quad (15)$$

with $v_{\text{cutoff}} = 2 - \frac{\pi^2}{L^2}$ [see (5b) for $v(\gamma) = \gamma + 1/\gamma$] and keeping only the dominant terms in L , the function $G(y, \tau)$ evolves according to

$$\partial_\tau G = \partial_y^2 G + \pi^2 G, \quad (16)$$

with the boundary conditions

$$G(0, \tau) \approx 0, \quad G(1, \tau) \approx 0. \quad (17)$$

[More precisely, $G(0, \tau)$ and $G(1, \tau)$ would be nonzero only at the next order in a $1/L$ expansion.]

The problem reduces to a diffusion problem with absorbing boundary conditions. The stationary configuration is the sine shape (5a), as expected.

If at time $t=0$ the shape is different from this stationary configuration, it will relax back to it in the long-time limit up to a multiplicative constant:

$$G(y, \infty) = \frac{B}{\pi} \sin(\pi y). \quad (18)$$

As the stationary shape for $h(x, t)$ must be of the form given by (5a), we obtain, using Eq. (15) that the final shift in position is given by

$$R(\delta) = \lim_{t \rightarrow \infty} (X_t - v_{\text{cutoff}} t) = \ln \frac{B}{A}. \quad (19)$$

To compute the value of B , one simply needs to project the initial condition on the sine shape:

$$B = Ae^{R(\delta)} = 2\pi \int_0^1 dy \sin(\pi y) G(y, 0). \quad (20)$$

We now proceed to use this expression for the perturbations we are interested in: perturbations localized near the cutoff.

We do not have a full information on the initial condition $h(x, 0)$ or, equivalently, $G(y, 0)$. However, as we expect a perturbation to grow at the very tip of the front, we expect that $h(x, 0)$ is identical to its stationary shape, except in a region of size of order $\Delta x \approx 1$ near the tip. On the scale we consider, this means that $G(y, 0)$ is perturbed over a region of size $\Delta y \approx 1/L$ —in other words,

$$G(y, 0) = A \left[\frac{1}{\pi} \sin(\pi y) + p(1-y) \right], \quad (21)$$

where the perturbation $p(y')$ is nonzero only for $y' = 1-y$ of order $1/L$. Therefore, from Eq. (20),

$$e^{R(\delta)} = 1 + 2\pi \int_0^{b/L} dy' \pi y' p(y'), \quad (22)$$

where b is a number of order 1 representing the extent over which a perturbation initially affects the shape of the front. [$p(y') \approx 0$ if $y' > b/L$.]

The precise shape of $p(y')$ is not known, but its amplitude can be easily understood in a stochastic particle model: if some particles are sent at a distance $\delta \ll L$ ahead of the front, $h(x, t)$ increases by $1/N$ at position $x = X_t + L + \delta$. Because of the exponential factor in Eq. (15), this translates to an increase of order $p(y') \approx \exp(\delta)/L$ for the reduced shape $G(y, \tau)$. Combining everything, one finally gets

$$e^{R(\delta)} = 1 + C_2 \frac{e^\delta}{L^3}, \quad (23)$$

where C_2 is some number of order 1 which depends on the precise shape $p(y')$. Expression (23) is just our second hypothesis, up to factors γ_0 , which can be put back by dimensional analysis.

One consequence of the argument above is that C_2 is of order 1 compared to L . However, it gives no information about the dependence of C_2 on δ or on the shape of the fluctuation. We think that if C_2 depends on δ , it is a weak dependence that we can ignore. A simple situation where this can be checked is when δ is large: if a particle jumps sufficiently far ahead, it will start a front of its own that will completely replace the original front. For such a front, it is well known [12,25] that the position for large t is given at first (while the cutoff is not relevant) by $\delta + 2t - \frac{3}{2} \ln t$. When the velocity $2 - \frac{3}{2t}$ matches v_{cutoff} , that is, at a time $t_0 \approx L^2$, a crossover occurs and the position becomes $R(\delta) + v_{\text{cutoff}} t$. Matching the two expressions for the position at time $t = t_0$, one obtains $R(\delta) \approx \delta - \ln L^3$, as predicted by Eq. (8). This indicates that, at least for large δ , the number C_2 has no δ dependence.

C. Third hypothesis

From Sec. II and Eq. (8), the size δ of the fluctuations that contribute significantly to the diffusion of the front is such that $\exp(\gamma_0 \delta) \sim L^3$. From Eq. (7), the typical time between two such fluctuations is therefore L^3 . On the other hand, from Sec. III B, the relaxation time of a fluctuation is of order L^2 . It is therefore safe to assume that a relevant fluctuation has enough time to relax before another one occurs.

D. Fourth hypothesis

The fourth hypothesis states that, to compute the shift in velocity, one should use a front width L_{eff} that is larger than what is predicted by the cutoff theory by an amount $\frac{3}{\gamma_0} \ln \ln N$. The hypothesis is plausible as this length is precisely the distance δ at which the relevant fluctuations occur: the main effect of the fluctuations would then be to increase the effective width of the front that enters the cutoff theory (5). We present in Appendix B a simplified model to support this claim.

Remarkably, the front width L_{eff} emerges naturally in the QCD context [19].

IV. NUMERICAL SIMULATIONS

We consider here a reaction-diffusion model with saturation which was introduced in [13] as a toy model for high-energy scattering in QCD. Particles are evolving in discrete time on a one-dimensional lattice. At each time step, a particle may jump to the nearest position on the left or on the right with respective probabilities p_l and p_r and may divide into two particles with probability λ . We also impose that each of the $n(x, t)$ particles piled up at x at time t may die with probability $\lambda n(x, t)/N$.

Between times t and $t+1$, $n_l(x, t)$ particles out of $n(x, t)$ move to the left and $n_r(x, t)$ move to the right. Furthermore, $n_+(x, t)$ particles are replaced by their two offsprings at x and $n_-(x, t)$ particles disappear. Hence the total variation in the number of particles on site x reads

$$\begin{aligned} n(x, t+1) - n(x, t) = & -n_l(x, t) - n_r(x, t) - n_-(x, t) + n_+(x, t) \\ & + n_l(x+1, t) + n_r(x-1, t). \end{aligned} \quad (24a)$$

The numbers describing a time step at position x have a multinomial distribution:

$$\begin{aligned} P(\{n_l, n_r, n_+, n_-\}) = & \frac{n!}{n_l! n_r! n_+! n_-! \Delta n!} p_l^{n_l} p_r^{n_r} \lambda^{n_+} (\lambda n/N)^{n_-} \\ & \times (1 - p_l - p_r - \lambda - \lambda n/N)^{\Delta n}, \end{aligned} \quad (24b)$$

where $\Delta n = n - n_l - n_r - n_+ - n_-$ and all quantities in the previous equation are understood at site x and time t . The mean evolution of $u \equiv n/N$ in one step of time reads

$$\begin{aligned} \langle u(x, t+1) | \{u(x, t)\} \rangle = & u(x, t) + p_l [u(x+1, t) - u(x, t)] \\ & + p_r [u(x-1, t) - u(x, t)] + \lambda u(x, t) [1 - u(x, t)]. \end{aligned} \quad (25)$$

When N is infinitely large, one can replace the u 's in Eq. (25)

BRUNET *et al.*PHYSICAL REVIEW E **73**, 056126 (2006)

by their averages. One obtains then a deterministic front equation in the FKPP class with

$$v(\gamma) = \frac{1}{\gamma} \ln[1 + \lambda + p_l(e^{-\gamma} - 1) + p_r(e^{\gamma} - 1)], \quad (26)$$

and γ_0 is defined by $v'(\gamma_0) = 0$; see (2).

For the purpose of our numerical study, we set

$$p_l = p_r = 0.1 \quad \text{and} \quad \lambda = 0.2. \quad (27)$$

From Eq. (26), this choice leads to

$$\begin{aligned} \gamma_0 &= 1.3521 \dots, \quad v(\gamma_0) = 0.25538 \dots, \\ v''(\gamma_0) &= 0.16773 \dots \end{aligned} \quad (28)$$

Predictions for all cumulants of the position of the front are obtained by replacing the values of these parameters in Eqs. (6).

Technically, in order to be able to go to very large values of N , we replace the full stochastic model by its deterministic mean-field approximation $u \rightarrow \langle u \rangle$, where $\langle u \rangle$ is given by Eq. (25), in all bins in which the number of particles is larger than 10^3 (that is, in the bulk of the front). Whenever the number of particles is smaller, we use the full stochastic evolution (24). We add an appropriate boundary condition on the interface between the bins described by the deterministic equation and the bins described by the stochastic equation so that the flux of particles is conserved [26]. This setup will be called “model I.” Eventually, we shall use the mean-field approximation everywhere except in the rightmost bin (model II): at each time step, a new bin is filled immediately on the right of the rightmost nonempty site with a number of particles given by a Poisson law of expectation $\theta = N \langle u(x, t+1) | \{u(x, t)\} \rangle$. In the context of a slightly different model in the same universality class [18], this last approximation was shown numerically to give indistinguishable results from those obtained with the full stochastic version of the model, as far as the front velocity and its diffusion constant were concerned. We shall confirm this observation here.

We define the position of the front at time t by

$$X_t = \sum_{x=0}^{\infty} u(x, t). \quad (29)$$

We start at time $t=0$ from the initial condition $u(x, 0) = 1$ for $x \leq 0$ and $u(x, 0) = 0$ for $x > 0$. We evolve it up to time $t = \ln^2 N$ to get rid of subasymptotic effects related to the building of the asymptotic shape of the front, and we measure the mean velocity between times $\ln^2 N$ and $16 \times \ln^2 N$. For model I (many stochastic bins), we average the results over 10^4 such realizations. For model II (only one stochastic bin), we generate 10^5 such realizations for $N \leq 10^{50}$ and 10^4 realizations for $N > 10^{50}$. In all our simulations, models I and II give numerically indistinguishable results for the values of N where both models were simulated, as can be seen on the figures (results for model I are represented by a circle and for model II by a cross).

First, we check that the effective width of the front is L_{eff} given by Eq. (9). We extract the latter from the measured mean velocity v using the formula

$$L_{\text{eff}} = \pi \sqrt{\frac{v''(\gamma_0)}{2(v(\gamma_0) - v)}}. \quad (30)$$

We subtract from L_{eff} the width of the front obtained in the cutoff theory $L = (\ln N) / \gamma_0$ and compare the numerical result with the analytical formula

$$L_{\text{eff}} - L = \frac{3 \ln(\ln N)}{\gamma_0} + c + d \frac{\ln(\ln N)}{\ln N}. \quad (31)$$

The first term on the right-hand side is suggested by our fourth assumption [see Eq. (9)]. We have added two subleading terms which go beyond our theory: a constant term and a term that vanishes at large N . The latter are naturally expected to be the next terms in the asymptotic expansion for large N . We include them in this numerical analysis because in the range of N in which we are able to perform our numerical simulations, they may still bring a significant contribution.

We fit Eq. (31) to the numerical data obtained in the framework of model II, restricting ourselves to values of N larger than 10^{30} . In the fit, each data point is weighted by the

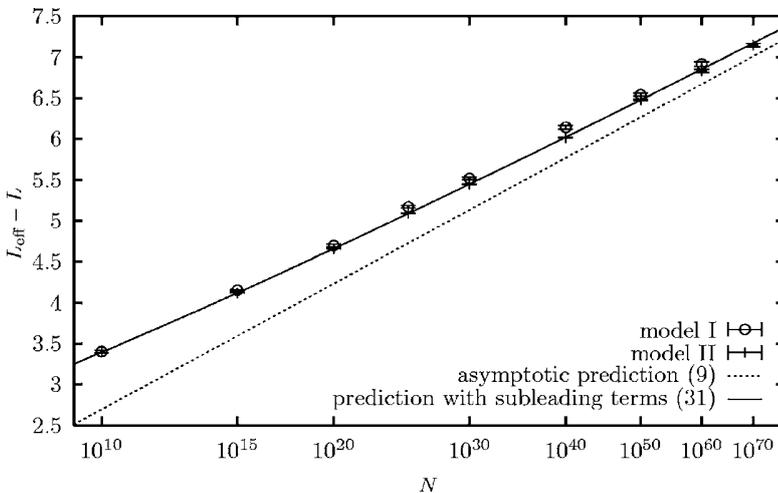


FIG. 1. Measured L_{eff} , defined by Eq. (30), from which we have subtracted the width L in the cutoff theory, as a function of N . The dotted line represents the leading terms $3 \ln \ln N / \gamma_0$; see Eq. (9). The subleading terms (31) of the solid line have been determined by a fit.

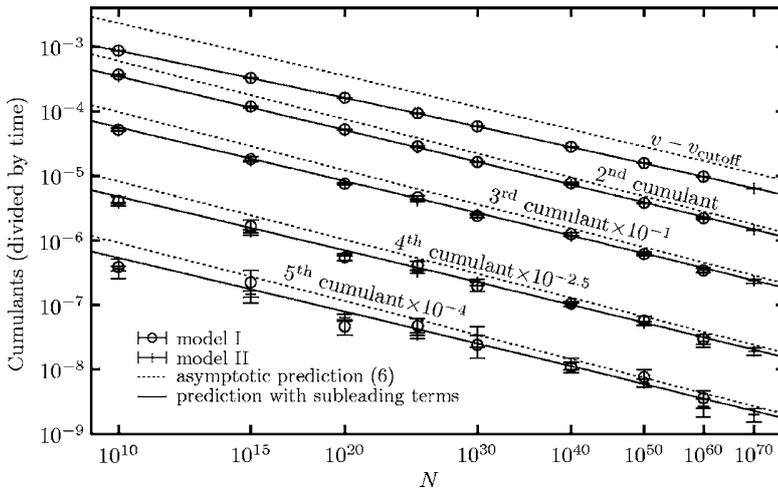


FIG. 2. From top to bottom, the correction to the velocity given by the cutoff theory and the cumulants of orders 2–5 of the position of the front in the stochastic model. The numerical data are compared to our parameter-free analytical predictions (6), represented by the dotted line. The subleading terms of the solid lines are numerically the same as in Fig. 1; no further fit has been performed for the present figure.

statistical dispersion of its value in our sample of data. We obtain a determination of the values of the free parameters $c = -4.26 \pm 0.01$ and $d = 5.12 \pm 0.27$, with a good quality of the fit ($\chi^2/N_{DF} \sim 1.15$). The numerical data together with the theoretical predictions are shown in Fig. 1. We see a clear convergence of the data to the predicted asymptotics at large N (dotted line in the figure), but subleading corrections that we have accounted for phenomenologically here are sizable over the whole range of N .

We now turn to the higher-order cumulants. Our numerical data are shown in Fig. 2 together with the analytical predictions obtained from Eqs. (6) (dotted lines in the figure). We see that the numerical simulations get very close to the analytical predictions at large N . However, like in the case of L_{eff} , higher-order corrections are presumably still important for the lowest values of N displayed on the plot.

We try to account for these corrections by replacing the factor $(\ln N)/\gamma_0 = L$ in the denominator of the expression for the cumulants in Eqs. (6) by the ansatz for L_{eff} given in Eq. (31), *without retuning the parameters*. The results are shown in Fig. 2 (solid lines) and are in excellent agreement with the numerical data over the whole range of N . We could also have refitted the parameters c and d for each cumulant separately, as, *a priori*, they are not predicted by our theory. We observe that this is not required by our data.

This last observation suggests that all the cumulants can be computed, with a good accuracy, with the effective width L_{eff} as the only parameter. We check this in Fig. 3, which represents the ratio of the n th cumulant (divided by time) by the correction to the velocity $v_{\text{deterministic}} - v$ to the power $3/2$. If one supposes that the correction to the velocity varies like $1/L_{\text{eff}}^2$ and the cumulants like $1/L_{\text{eff}}^3$ for some effective width L_{eff} , this width disappears from the ratio plotted and one can compare the numerical results to our analytical prediction with no free parameter or unknown subleading terms. Within statistical error, the data seem to agree for N large enough with our prediction, suggesting that, indeed, all the cumulants can be described with a good accuracy with only the effective width L_{eff} .

Simulations, not shown here, for the model introduced in [18] support also our predictions (6).

V. CONCLUSION

The main idea that we have put forward in the present work is that all the fluctuations of the front position, and in particular the diffusion constant, are dominated by large but rare fluctuations at the tip of the front.

Under some more precise assumptions (hypotheses of Sec. II) on these fluctuations, we were able to obtain explicit

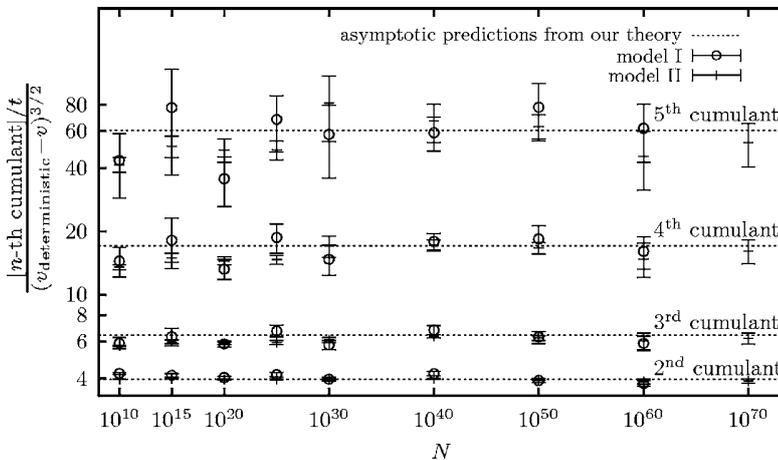


FIG. 3. The ratio of cumulants 2–5 divided by the correction to the velocity to the power $3/2$. The dotted lines are the analytical prediction assuming only the cutoff theory (5b) for the velocity and the predictions (6) for the cumulants.

expressions (6) of the cumulants of the position of the front. We checked these predictions in our numerical simulations of Sec. IV. In Sec. III, we gave some arguments in support of the four hypotheses of Sec. II. None of these arguments can be regarded as a mathematical derivation, and we can imagine that some details, such as the precise shape of the distribution of fluctuations (7) or the explicit expression (8), could be slightly modified by a more precise analysis. We believe, however, given the good agreement of the predictions (6) with the numerical simulations, that our picture is very close, if not identical, to the actual behavior of the front for large values of N .

An alternative picture to explain the $1/\ln^3 N$ scaling of the diffusion constant was proposed in [16]. So far, we have not been able to relate the two approaches. One of the claims of the present work is that all the cumulants have the same $1/\ln^3 N$ dependence; it would be interesting to know if this is also predicted by Panja's theory [16].

To conclude, we would like to point out the remarkable similarity between the predictions (6) and the exact results obtained recently [27] in the context of directed polymers. Basically, the results of [27] are the same, *mutatis mutandis*, as our present results (6), for all the cumulants. The only significant change is that the $3 \ln \ln N$ for the velocity and the $1/\ln^3 N$ dependence for all the cumulants in Eqs. (6) corresponds, in [27], to a $\ln \ln N$ for the velocity and a $1/\ln N$ for all the cumulants (see Eq. (23) of [27] with $L = \ln N$ and where the term $L + \ln L$ in the velocity corresponds to v_{cutoff} , as seen from Eq. (28) of [27]). What is interesting is that our scenario of Sec. II for FKPP fronts applies also for the system studied in [27]: Indeed, the fluctuations of the position are mainly due to the rare big events taking place at the tip of the "front" (see the last paragraph before the conclusion of [27]), the position of the rightmost particle is given by Eq. (7) (see Eq. (32) of [27] with $\delta = -\ln q$ and $X_t = \ln B_t$), the effect of a large fluctuation can be written as Eq. (8) with the L^3 term replaced by L (the logarithm of Eq. (34) of [27] can be written as in $X_{t+1} - X_t = L + \ln L + R(\delta)$), relevant fluctuations (of size $\ln L$ instead of $3 \ln L$) appear every L time steps (instead of every L^3 time steps) and the relaxation time is 1 instead of L^2 . This similarity may add a further piece of evidence for our results.

ACKNOWLEDGMENT

This work was partially supported by the U.S. Department of Energy.

APPENDIX A: LIMIT $N \rightarrow \infty$

In this appendix, we try to provide an argument for the exponential decay (7) of the distribution for the width of the front. To this aim, we consider a very simple model of reaction-diffusion: particles diffuse on the line, and during each time interval dt , each particle duplicates with a probability dt . The motions of all the particles are uncorrelated.

If one added a saturation rule as described at the beginning of Sec. II, the density of particles (or the number of particles on the right of x , depending on the precise satura-

tion rule) would be described by a stochastic FKPP equation. However, the saturation affects only the motion of particles in the bulk of the front, where the density is high. As the fluctuations develop in the low-density region, it is reasonable to assume that the distribution of the size of the fluctuations are well described by the model *without any saturation*.

For this model without saturation, let $P_t(x)$ the probability that, at time t , no particles are present on the right of x given that, at $t=0$, there is a single particle at the origin: $P_0(x) = \theta(x)$. During the first "time step" dt , the only particle in the system moves by a quantity $\eta\sqrt{dt}$ where η is a Gaussian number of variance 2 and duplicates with a probability dt . If it duplicates, the probability $P_{t+dt}(x)$ is the probability that the offsprings of both particles are on the left of x . As the particles have uncorrelated motion, this is the product of the probabilities for each offspring. Finally, one gets [28]

$$P_{t+dt}(x) = \langle P_t(x - \eta\sqrt{dt})(1 - dt) + P_t^2(x - \eta\sqrt{dt})dt \rangle,$$

where the average is on η . After simplification,

$$\partial_t P = \partial_x^2 P - P + P^2. \quad (\text{A1})$$

One notices that $1 - P_t(x)$ is solution of the deterministic FKPP equation (1). Therefore, for large t and x [2,12,25],

$$1 - P_t(x) \sim z e^{-z - z^2/4t} \quad \text{for } z = x - 2t + \frac{3}{2} \ln t.$$

Let $Q_t(x)$ be the probability that there are no particles on the right of x when the initial condition is a given density of particles $\rho_0(x)$. Using the fact that all the particles are independent, one gets easily

$$Q_t(x) = \exp \left[- \int dy \rho_0(y) [1 - P_t(x - y)] \right]. \quad (\text{A2})$$

$\rho_0(y)$ needs to reproduce the shape of the front seen from the tip. Starting from (5a), we write $\rho(y) = N h_{\text{cutoff}}(L + y)$ and take the large- N limit. One gets $\rho_0(y) = -y \exp(-y)$ for $y < 0$ and $\rho_0(y) = 0$ for $y > 0$. Evaluating the integral in Eq. (A2), one gets, for large t and $x - 2t \ll \sqrt{t}$,

$$Q_t(x) \approx \exp[-C e^{-(x-2t)}]. \quad (\text{A3})$$

[Notice that the $(3/2)\ln t$ factor canceled out.]

The probability distribution function of the rightmost particle is clearly $\partial_x Q_t(x)$. We see that in this stochastic model, the front moves at a deterministic velocity equal to 2 and that the position of the rightmost particle around the position of the front is given by a Gumbel distribution.

From (A3), the distribution $\partial_x Q_t(x)$ gives our first hypothesis (7) for large fluctuations ($\delta = x - 2t \gg 1$). Our attempts to check numerically (A3) by simulating fronts with a large but finite number of particles confirmed this exponential decay for large δ , but showed some discrepancy for $\delta < 0$, which we do not understand. This, however, does not affect the hypothesis (7).

APPENDIX B: MOVING WALL

We consider again the reaction-diffusion model introduced in Appendix A. As we said, one needs to add a satura-

ration effect to obtain a propagating front equation for the density, but doing so introduces correlations in the motions of the particles that make the model hard to solve. In this appendix, we introduce an approximate way of adding a saturation effect which does not introduce any such correlation.

In a real front, the tip is subject to huge fluctuations happening on short time scales. On the other hand, the bulk of the front moves smoothly and adjusts very slowly to the fluctuations happening at the tip. Therefore, we believe that, for times not too large, it is a reasonable approximation to assume that the bulk of the front moves at a constant velocity.

To implement this idea, our model is the following: a wall starting at the origin is moving to the right at a constant velocity v . Particles are present on the right of the wall. The particles are evolving as in Appendix A, except that whenever a particle crosses the wall, it is removed.

We first consider a single particle starting at a distance z of the wall. After a time t , either all the offsprings of this particle have been caught by the wall or some have survived. We want to compute the probability $E_t(z)$ that all the particles have been caught at time t . The original particle, after a time dt , is at a distance $z-vdt + \eta\sqrt{dt}$ from the wall, and it might have duplicated with probability dt . Using the same method as in Appendix A, one gets

$$\partial_t E_t = \partial_z^2 E_t - v \partial_z E_t - E_t + E_t^2, \quad (\text{B1})$$

with the conditions

$$E_0(z) = 0 \quad \text{for } z > 0 \quad \text{and } E_t(z) = 1 \quad \text{for } z < 0. \quad (\text{B2})$$

In the long-time limit, $E_t(z)$ converges to the stationary solution $E_\infty(z)$ and one recognizes that $h(z) = 1 - E_\infty(-z)$ is the stationary solution of the FKPP equation (1) if $z = x - vt$. In other words, $1 - E_\infty(-z)$ is the shape of a traveling front. As this shape reaches 0 for $z=0$, it must be a front with a sine arch and a velocity v smaller than 2, as in (5). So if $v < 2$, the probability $1 - E_\infty(-z)$ is the shape of the front with a cutoff:

$$1 - E(z) \sim L e^{-L} \sin\left(\frac{\pi z}{L}\right) e^z \quad \text{where } v = 2 - \frac{\pi^2}{L^2}. \quad (\text{B3})$$

[The extra factor e^{-L} comes from the fact that $z=0$ is the tip of the front in (B3) while it is the bulk of the front in (5a). If $v > 2$, all the particles eventually die and $E_t(z)$ converges to 1.]

If one starts with a density $\rho(z)$ of particles at time $t=0$, the probability E_t^* that everybody dies is given, similarly to Eq. (A2), by

$$E_t^* = \exp\left[-\int_0^{+\infty} dz \rho(z)[1 - E_t(z)]\right]. \quad (\text{B4})$$

We consider, as an initial condition, the situation in the real front with $\rho(z) = Nh(z) \sim NL \sin(\pi z/L) \exp(-z)$ for $z < L$ as in (5a). One gets, for long times,

$$E_t^* \rightarrow \exp[-CNL^3 e^{-L}]. \quad (\text{B5})$$

We see that the system survives if

$$NL^3 e^{-L} \gtrsim 1 \quad \text{or } L \gtrsim \ln N + 3 \ln \ln N, \quad (\text{B6})$$

which, given $\gamma_0 = 1$, is exactly Eq. (9).

-
- [1] R. A. Fisher, *Ann. Eugenics* **7**, 355 (1937); A. Kolmogorov, I. Petrovsky, and N. Piscounov, *Bull. Univ. État Moscou A* **1**, 1 (1937).
- [2] W. van Saarloos, *Phys. Rep.* **386**, 29 (2003).
- [3] L. Pechenik and H. Levine, *Phys. Rev. E* **59**, 3893 (1999).
- [4] C. R. Doering, C. Mueller, and P. Smereka, *Physica A* **325**, 243 (2003).
- [5] S. N. Majumdar and P. L. Krapivsky, *Physica A* **318**, 161 (2003).
- [6] B. Derrida and H. Spohn, *J. Stat. Phys.* **51**, 817 (1988).
- [7] D. Carpentier and P. Le Doussal, *Nucl. Phys. B* **588**, 531 (2000).
- [8] S. Munier and R. Peschanski, *Phys. Rev. Lett.* **91**, 232001 (2003).
- [9] I. Balitsky, *Nucl. Phys. B* **463**, 99 (1996); Y. V. Kovchegov, *Phys. Rev. D* **60**, 034008 (1999); **61**, 074018 (2000).
- [10] C. Marquet, R. Peschanski, and G. Soyez, *Nucl. Phys. A* **756**, 399 (2005).
- [11] D. G. Aronson and H. F. Weinberger, *Lect. Notes Math.* **446**, 5 (1975).
- [12] É. Brunet and B. Derrida, *Phys. Rev. E* **56**, 2597 (1997).
- [13] R. Enberg, K. Golec-Biernat, and S. Munier, *Phys. Rev. D* **72**, 074021 (2005).
- [14] D. A. Kessler, Z. Ner, and L. M. Sander, *Phys. Rev. E* **58**, 107 (1998).
- [15] J. Armero, J. Casademunt, L. Ramírez-Piscina and J. M. Sancho, *Phys. Rev. E* **58**, 5494 (1998); A. Rocco, J. Casademunt, U. Ebert and W. van Saarloos, *Phys. Rev. E* **65**, 012102 (2002).
- [16] D. Panja, *Phys. Rev. E* **68**, 065202(R) (2003); *Phys. Rep.* **393**, 87 (2004).
- [17] C. Mueller and R. B. Sowers, *J. Funct. Anal.* **128**, 439 (1995).
- [18] É. Brunet and B. Derrida, *Comput. Phys. Commun.* **121-122**, 376 (1999); *J. Stat. Phys.* **103**, 269 (2001).
- [19] A. H. Mueller and A. I. Shoshi, *Nucl. Phys. B* **692**, 175 (2004).
- [20] J. G. Colon and C. R. Doering, *J. Stat. Phys.* **120**, 421 (2005).
- [21] E. Iancu, A. H. Mueller, and S. Munier, *Phys. Lett. B* **606**, 342 (2005).
- [22] M. Kloster, *Phys. Rev. Lett.* **95**, 168701 (2005).
- [23] E. Moro (private communication).
- [24] D. Panja and W. van Saarloos, *Phys. Rev. E* **65**, 057202 (2002).
- [25] M. D. Bramson, *Mem. Am. Math. Soc.* **44** (1983).
- [26] E. Moro, *Phys. Rev. E* **69**, 060101(R) (2004); **70**, 045102(R) (2004).
- [27] É. Brunet and B. Derrida, *Phys. Rev. E* **70**, 016106 (2004).
- [28] H. P. McKean, *Commun. Pure Appl. Math.* **28**, 323 (1975).

Effect of selection on ancestry: An exactly soluble case and its phenomenological generalization

É. Brunet,¹ B. Derrida,¹ A. H. Mueller,² and S. Munier³

¹Laboratoire de Physique Statistique, École Normale Supérieure, 24 rue Lhomond, 75231 Paris cedex 05, France

²Department of Physics, Columbia University, New York, New York, 10027, USA

³Centre de Physique Théorique, École Polytechnique, CNRS, 91128 Palaiseau, France

(Received 25 April 2007; published 3 October 2007)

We consider a family of models describing the evolution under selection of a population whose dynamics can be related to the propagation of noisy traveling waves. For one particular model that we shall call the exponential model, the properties of the traveling wave front can be calculated exactly, as well as the statistics of the genealogy of the population. One striking result is that, for this particular model, the genealogical trees have the same statistics as the trees of replicas in the Parisi mean-field theory of spin glasses. We also find that in the exponential model, the coalescence times along these trees grow like the logarithm of the population size. A phenomenological picture of the propagation of wave fronts that we introduced in a previous work, as well as our numerical data, suggest that these statistics remain valid for a larger class of models, while the coalescence times grow like the cube of the logarithm of the population size.

DOI: 10.1103/PhysRevE.76.041104

PACS number(s): 02.50.-r, 05.40.-a, 89.75.Hc

I. INTRODUCTION

It has been recognized for a long time that there is a strong analogy between neo-Darwinian evolution and statistical mechanics [1]. For an evolving population, there is an ongoing competition between the mutations which make individuals explore larger and larger regions of genome space and selection which tends to concentrate them at the optimal fitness genomes. This is very similar to the competition between the energy and the entropy in statistical mechanics.

In the simplest models of evolution, one associates to each individual [2,3] (or to each species [4]) a single number which represents how fit this individual is to its environment. This fitness is transmitted to the offspring, up to small variations due to mutations. A higher fitness usually means a larger number of offspring [2,3,5–9]. If the size of the population is limited by the available resources, survivors are chosen at random among all the offspring. This leads in the long term to a selection effect: the descendants of individuals with low fitness are eliminated whereas the offspring of the individuals with high fitness tend to overrun the whole population.

Our focus in this paper is a class of such models [5–9] describing the evolution of a population of fixed size N under asexual reproduction. The i th individual is characterized by a single real number, $x_i(g)$, which represents its adequacy to the environment. [This $x_i(g)$ plays a role similar to fitness in the sense that offspring with higher $x_i(g)$ will be selected; in the following, we shall simply call it the position of the individual.] At a generation g , the population is thus represented by a set of N real numbers $x_i(g)$ for $1 \leq i \leq N$. At each new generation, all individuals disappear and are replaced by some of their offspring: the j th descendant of individual i has position $x_i(g) + \epsilon_{i,j}(g)$ where $\epsilon_{i,j}(g)$ represents the effect of mutations from generation g to generation $g+1$. Then comes the selection step: at generation $g+1$, one keeps only the N rightmost offspring among the descendants of all individuals at generation g . One may consider two particular variants of this model.

Model A. Each individual has a fixed number k of offspring and all the $\epsilon_{i,j}(g)$ are independently distributed according to a given distribution $\rho(\epsilon)$. For example, $\rho(\epsilon)$ may be the uniform distribution between 0 and 1. A realization of such an evolution is shown in Fig. 1.

Model B. Each individual has infinitely many offspring: the $\epsilon_{i,j}(g)$ are distributed according to a Poisson process of density $\psi(\epsilon)$ [this means that, with probability $\psi(\epsilon)d\epsilon$, there is one offspring of individual i with position between $x_i(g) + \epsilon$ and $x_i(g) + \epsilon + d\epsilon$]. The density $\psi(\epsilon)$ is *a priori* arbitrary. The only constraints we impose are that $\psi(\epsilon)$ decays fast enough, when ϵ increases, for the position not to diverge after one generation, and that $\int_{-\infty}^{\infty} \psi(\epsilon)d\epsilon = \infty$, for the survival probability to be 1. (This latter constraint implies in fact that each individual i has infinitely many offspring before the selection step. After selection, however, each individual has a finite number of surviving offspring in the next generation, and the model would remain the same if each individual had

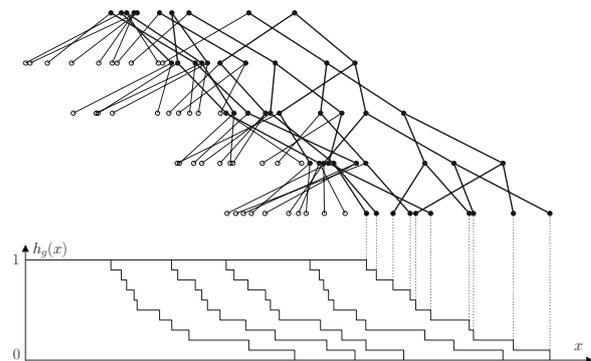


FIG. 1. Numerical simulation of the evolution of model A, with $k=2$ and $\rho(\epsilon)$ uniform between $-\frac{1}{2}$ and $\frac{1}{2}$ for $N=10$. Upper plot: The filiation between each individual and its two offspring is shown. At each generation, the N rightmost survive. Lower plot: The noisy traveling wave front $h_g(x)$, constructed as in (1), is shown for the five generations of the upper plot.

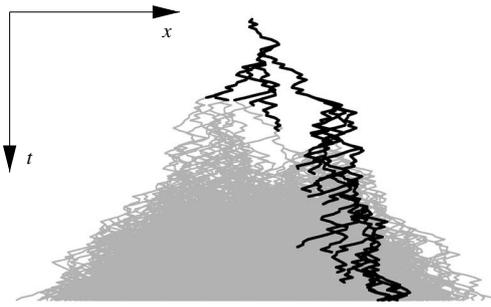


FIG. 2. A branching process for which the size N of the population is limited to five. Each time the number of walks reaches six, the leftmost walk is eliminated. Time goes downwards and the horizontal direction represents space. The actual population is represented in black, while the grey lines represent what the population would be for infinite N (i.e., in the absence of selection).

N offspring located at the N rightmost positions of the Poisson process.)

Another example would be N branching random walks where the size of population is kept constant by eliminating the leftmost walk each time a branching event occurs. A visual representation of this latter example is shown in Fig. 2.

As discussed in Sec. II, these models are related to noisy traveling wave equations, of the Fisher–Kolmogorov–Petrovsky–Piscounov (Fisher–KPP) type [10–12], which appear in many contexts: disordered systems [13,14], reaction-diffusion [15–18], fragmentation [19] or QCD [20–22]. A number of recent works [8,18,23–30] focused on the fluctuations of the position of these fronts, and this will allow us to predict how the fitness of the population evolves with the number of generations.

Another interesting aspect of these models with stochastic evolution is their genealogy [9]: one can associate to any group of individuals, at a given generation, its genealogical tree. One can then study how this tree fluctuates, and in particular what is the number of generations needed to reach their most recent common ancestor. The relationship between noisy traveling waves and genealogies is the main purpose of the present paper.

Note that in the present work, we will limit the discussion to models A and B where selection is strict in the sense that the N rightmost offspring are selected at each time step. In the numerical simulations presented in [9], we showed that the behavior is robust, as it remains unchanged when selection is less strict, for example, when the N survivors are chosen at random among the $3N/2$ rightmost offspring.

While the models we consider here are difficult to solve for arbitrary $\rho(\epsilon)$ and $\psi(\epsilon)$, one particular case of model B, with $\psi(\epsilon) = e^{-\epsilon}$, turns out to be analytically solvable both for the statistics of the position of the population and for the properties of the genealogical trees. We shall call this case the “exponential model” and present its solution in Sec. III.

As explained at the end of Sec. II, the exponential model is, however, nongeneric in the sense that it does not behave like a Fisher–KPP front. The generic case (which behaves like a noisy Fisher–KPP equation that we are not able to

solve) and the exponential model can, however, be both described by a similar phenomenological theory [8] that we develop in Sec. IV. As a consequence, we argue that both the generic case and the exponential model have the same cumulants for the position of the front (up to a change of scale), and that the genealogical trees have the same statistics in both models (up to a change of time scale). Numerical results, presented in Sec. V, support these claims.

II. THE LINK WITH NOISY FISHER-KPP FRONTS

Our models are nothing but stochastic models for the evolution of the positions of N individuals along the real axis. These positions form a cloud that does not spread: if an individual happens to fall far behind the cloud, it will have no surviving offspring, whereas the descendants of an individual far ahead of the cloud grow until they replace the whole population. With this picture in mind, it makes sense to describe the population by a front. Let $Nh_g(x)$ be the number of individuals with a position larger than x ,

$$h_g(x) = \frac{1}{N} \int_x^\infty dz \sum_{i=1}^N \delta[z - x_i(g)]. \quad (1)$$

Clearly, $h_g(x)$ is a decreasing function with $h_g(-\infty) = 1$ and $h_g(+\infty) = 0$. In this section, we write the noisy equation that governs the evolution of this front.

Let $Nh_{g+1}^*(x)$ be the number of offspring on the right-hand side of x at generation $g+1$ before the selection step. [So, for instance, $h_{g+1}^*(-\infty)$ is k in model A and ∞ in model B]. Once $h_{g+1}^*(x)$ is known, the selection step to get $h_{g+1}(x)$ is simply

$$h_{g+1}(x) = \min[1, h_{g+1}^*(x)]. \quad (2)$$

Let us write the average and variance of $h_{g+1}^*(x)$ for both models.

A. Statistics of $h_{g+1}^*(x)$ for model A

In model A, one can write

$$Nh_{g+1}^*(x) = \sum_{i=1}^N n_{g+1}^{(i)}(x), \quad (3)$$

where $n_{g+1}^{(i)}(x)$ is the total number of offspring before selection of the i th individual of generation g that fall on the right-hand side of x . The probability that an offspring of i falls on the right-hand side of x is $\int_x^\infty d\epsilon \rho(\epsilon - x_i)$ and, as the k offspring of $x_i(g)$ are independent, $n_{g+1}^{(i)}(x)$ has a binomial distribution. The average and variance are therefore given by

$$\overline{n_{g+1}^{(i)}(x)} = k \int_x^\infty d\epsilon \rho[\epsilon - x_i(g)], \quad (4)$$

$$\begin{aligned} \text{Var}[n_{g+1}^{(i)}(x)] &= k \int_x^\infty d\epsilon \rho[\epsilon - x_i(g)] \\ &\quad \times \left(1 - \int_x^\infty d\epsilon \rho[\epsilon - x_i(g)] \right). \end{aligned}$$

As the variables $n_{g+1}^{(i)}(x)$ are uncorrelated, the average and

variance of $Nh_{g+1}^*(x)$ are simply from (3) the sums over i of the averages and variances of the $n_{g+1}^{(i)}(x)$. For the average, one has

$$\begin{aligned} \overline{Nh_{g+1}^*(x)} &= k \int_x^\infty d\epsilon \sum_i \rho[\epsilon - x_i(g)] \\ &= -k \int_x^\infty d\epsilon \int dz \rho(\epsilon - z) N h'_g(z), \end{aligned} \quad (5)$$

where we used, from (1),

$$\sum_{i=1}^N \delta[x - x_i(g)] = -N h'_g(x). \quad (6)$$

Simplifying, and doing the same transformation for the variance, one finally gets for model A

$$\overline{h_{g+1}^*(x)} = k \int d\epsilon h_g(x - \epsilon) \rho(\epsilon), \quad (7a)$$

$$\begin{aligned} \text{Var}[h_{g+1}^*(x)] &= \frac{k}{N} \int d\epsilon h_g(x - \epsilon) \rho(\epsilon) \\ &\quad \times \left(1 - 2 \int_\epsilon^\infty dz \rho(z) \right). \end{aligned} \quad (7b)$$

[Note that these average and variance are obtained for a given $h_g(x)$: they are not computed for the whole history.]

B. Statistics of $h_{g+1}^*(x)$ for model B

In model B, before the selection step, an individual at position $x_i(g)$ has infinitely many offspring given by a Poisson process of density $\psi[x - x_i(g)]$. As Poisson processes are additive, the whole population (before selection) at generation $g+1$ is also given by a Poisson process of density $\Psi(x)$ with

$$\Psi(x) = \psi[x - x_1(g)] + \cdots + \psi[x - x_N(g)]. \quad (8)$$

The number of individuals on the right-hand side of x is therefore a Poisson random number of average $\int_x^\infty d\epsilon \Psi(\epsilon)$, thus

$$\overline{Nh_{g+1}^*(x)} = \text{Var}[Nh_{g+1}^*(x)] = \int_x^\infty d\epsilon \Psi(\epsilon). \quad (9)$$

One can rewrite $\Psi(\epsilon)$ using the same trick as in (6) and (5). One finally gets for model B

$$\overline{h_{g+1}^*(x)} = \int d\epsilon h_g(x - \epsilon) \psi(\epsilon) \quad (10a)$$

and

$$\text{Var}[h_{g+1}^*(x)] = \frac{1}{N} \int d\epsilon h_g(x - \epsilon) \psi(\epsilon). \quad (10b)$$

Front equations for both models and comparison to Fisher-KPP fronts

Comparing (10) and (7), one sees that one can write, for both models

$$h_{g+1}^*(x) = \overline{h_{g+1}^*(x)} + \eta_g(x) \sqrt{\text{Var}[h_{g+1}^*(x)]}, \quad (11)$$

where $\eta_g(x)$ is a noise with $\overline{\eta_g(x)}=0$ and $\text{Var}[\eta_g(x)]=1$. Using (2) one finally gets for model A,

$$\begin{aligned} h_{g+1}(x) &= \min \left[1, k \int d\epsilon h_g(x - \epsilon) \rho(\epsilon) \right. \\ &\quad \left. + \frac{\eta_g(x)}{\sqrt{N}} \sqrt{k \int d\epsilon h_g(x - \epsilon) \rho(\epsilon) \left(1 - 2 \int_\epsilon^\infty dz \rho(z) \right)} \right] \end{aligned} \quad (12a)$$

and, for, model B,

$$\begin{aligned} h_{g+1}(x) &= \min \left[1, \int d\epsilon h_g(x - \epsilon) \psi(\epsilon) \right. \\ &\quad \left. + \frac{\eta_g(x)}{\sqrt{N}} \sqrt{\int d\epsilon h_g(x - \epsilon) \psi(\epsilon)} \right]. \end{aligned} \quad (12b)$$

The precise distribution of $\eta_g(x)$ depends on N and on the choice of the model. Far from both tips of the front, this distribution is Gaussian. At the tip, however, where $h_g(x)$ is of order $1/N$, both $h_g(x)$ and its variance are comparable and the noise cannot be approximated by a Gaussian. [This is because the number of individuals is small and the discrete character of $h_g(x)$ cannot be forgotten anymore.] Furthermore, the noise is correlated in space but uncorrelated for different g .

Thus, the precise expression of the noise $\eta_g(x)$ is rather complicated, but its variance is 1, so that the amplitude of the whole noise term in (12) decays as $1/\sqrt{N}$ as N becomes large.

Equations (12) are very similar to the noisy Fisher-KPP equation

$$\frac{\partial h_g(x)}{\partial g} = \frac{\partial^2 h_g(x)}{\partial x^2} + h_g(x) - h_g(x)^2 + \frac{\eta_g(x)}{\sqrt{N}} \sqrt{h_g(x) - h_g(x)^2}, \quad (13)$$

where $\eta_g(x)$ is a Gaussian noise with $\overline{\eta_g(x)}=0$ and $\overline{\eta_g(x) \eta_{g'}(x')} = \delta(g - g') \delta(x - x')$. The noisy Fisher-KPP equation appears as a dual equation for the branching process $A \rightarrow 2A$ (rate 1) and $2A \rightarrow A$ (rate $1/N$) or, more simply, is an approximate equation, valid for large N , describing the fraction of A in the chemical reaction $A + B \rightarrow 2A$ when the concentration of reactants is of order N [16,31,32].

Comparing (12) and (13), the convolution of $h_g(x)$ by $k\rho(\epsilon)$ or $\psi(\epsilon)$ in (12) spreads the front in the same way as the diffusion term in (13). The same convolution induces the growth, similarly to the linear $h_g(x)$ term in (13), as $k\phi(\epsilon)$ and $\psi(\epsilon)$ both have an integral larger than 1. Thus, the fixed point $h_g(x)=0$ is unstable. To balance the indefinite growth of $h_g(x)$, both (12) and (13) have a saturation mechanism

[respectively, the $\min(1, \dots)$ and the $-h_g(x)^2$ term] which makes $h_g(x)=1$ a stable fixed point. So, ignoring the noise terms ($N \rightarrow \infty$), both (12) and (13) describe a front that propagates from a stable phase $h_g(x)=1$ into an unstable phase $h_g(x)=0$. Finally, the noise terms in (12) and (13) have a similar amplitude of the order of $\sqrt{h_g(x)}/N$ in the unstable region $h_g(x) \ll 1$.

It is clear from the definitions of our models that the average velocity of the front is an increasing function of N . We first consider the limiting case $N \rightarrow \infty$, which is equivalent to removing the noise term ($\eta_g=0$) from (12) and (13). To determine [12] the velocity of such traveling wave equations, it is usually sufficient to consider the linearized equation in the unstable region $h_g(x) \ll 1$ (where the saturation mechanism can be neglected). Looking for solutions of the form $h_g(x) \approx \exp[-\gamma(x-vg)]$, one gets a relation between the decay rate γ and the velocity $v=v(\gamma)$ that reads

$$v(\gamma) = \frac{1}{\gamma} \ln \left(k \int d\epsilon \rho(\epsilon) e^{\gamma\epsilon} \right) \quad \text{for model A,} \quad (14a)$$

$$v(\gamma) = \frac{1}{\gamma} \ln \left(\int d\epsilon \psi(\epsilon) e^{\gamma\epsilon} \right) \quad \text{for model B.} \quad (14b)$$

[For Fisher-KPP (13), one has $v(\gamma) = \gamma^{-1} + \gamma$.]

In many cases, when $v(\gamma)$ is finite over some range of γ and reaches a minimal value $v(\gamma_0)$ for some finite positive decay rate γ_0 , the selected velocity of the front for a steep enough initial condition [12] is this minimal velocity $v(\gamma_0)$. For instance, for (13), one has $\gamma_0=1$ and the selected velocity is $v(\gamma_0)=2$. Whenever this minimal velocity exists, we shall say that the model is in the universality class of the Fisher-KPP equation (13). For finite N , i.e., in the presence of noise, there is a correction to this velocity and the front diffuses. We shall recall [8] in Sec. IV that for the generic Fisher-KPP case, the correction to the velocity is of order $1/\ln^2 N$ and that the diffusion constant is of order $1/\ln^3 N$.

There are, however, some choices of $\rho(\epsilon)$ or $\psi(\epsilon)$ for which $v(\gamma)$ is everywhere infinite or has no minimum. An example that we study in some detail in Sec. III is model B with $\psi(\epsilon) = e^{-\epsilon}$, for which $v(\gamma) = \infty$ for all γ . We shall see that, in presence of noise, the velocity of that front diverges as $\ln \ln N$ for large N instead of converging to a finite value. Another case would be model A with $\rho(\epsilon) = p\delta(\epsilon-1) + (1-p)\delta(\epsilon)$ for which $v(\gamma)$ has no minimum when $p \geq 1/k$. [Note, however, that for $p < 1/k$, the function $v(\gamma)$ has a minimum and the model belongs to the Fisher-KPP class.]

It has been known for a long time that traveling wave equations are related to branching random walks [33,34]. This can be seen by considering a single individual at the origin at generation 0 and by looking at the evolution of the probability $Q_g(x)$ that all of its descendants at generation g are on the left-hand side of x . In the case of model B with $N = \infty$, one has

$$\begin{aligned} Q_{g+1}(x) &= \prod_y [1 - \psi(y)dy + \psi(y)dy Q_g(x-y)] \\ &= \exp \left(\int dy \psi(y) [Q_g(x-y) - 1] \right). \end{aligned} \quad (15)$$

This equation describes the propagation of a front of the Fisher-KPP type, but where the unstable fixed point is at $Q_g=1$ instead of 0. For Q_g close to 1, one gets exponentially decaying traveling wave solutions of the form $1 - Q_g(x) \propto \exp[-\gamma(x-vg)]$, with $v=v(\gamma)$ given by (14b). [A similar calculation for model A leads to $v(\gamma)$ given by (14a).]

III. EXACT RESULTS FOR THE EXPONENTIAL MODEL

In this section, we derive exact expressions (for large N) of the velocity, diffusion constant, and coalescence times for model B with $\psi(\epsilon) = e^{-\epsilon}$. We first write some expressions valid for model B with an arbitrary density function $\psi(\epsilon)$, which we shall later apply to the exponential model.

Before selection, the positions of the individuals at generation $g+1$ are distributed according to a Poisson process of density $\Psi(x)$ defined in (8). We now wish to know the distribution of the N rightmost individuals of this Poisson process (i.e., of the offspring who survive the selection step). We first consider the probability that there are no offspring on the right-hand side of x . Clearly, it is given by

$$\prod_{x < z < \infty} [1 - \Psi(z)dz] = \exp \left(- \int_x^\infty \Psi(z)dz \right). \quad (16)$$

Then, the probability that the rightmost offspring at generation $g+1$ is in the interval $[x_1, x_1+dx_1]$, and the second rightmost is in $[x_2, x_2+dx_2]$, up to the $(N+1)$ st rightmost particle is, for $x_{N+1} < x_N < \dots < x_1$,

$$\Psi(x_{N+1})dx_{N+1} \Psi(x_N)dx_N \cdots \Psi(x_1)dx_1 \exp \left(- \int_{x_{N+1}}^\infty \Psi(z)dz \right). \quad (17)$$

It will be more convenient not to specify the ordering of the N rightmost particles. Then, the probability that the $(N+1)$ st rightmost particle is in the interval $[x_{N+1}, x_{N+1}+dx_{N+1}]$ (as before) and that the N rightmost particles are in the intervals $[x_k, x_k+dx_k]$ for $1 \leq k \leq N$, with no constraint on the order of x_1, \dots, x_N , becomes, for $k=1, \dots, N$,

$$\begin{aligned} &\frac{1}{N!} \Psi(x_{N+1})dx_{N+1} \Psi(x_N)dx_N \cdots \Psi(x_1)dx_1 \\ &\times \exp \left(- \int_{x_{N+1}}^\infty \Psi(z)dz \right) \quad \text{when } x_{N+1} < x_k. \end{aligned} \quad (18)$$

One obtains the probability that the $(N+1)$ st rightmost particle is in the interval $[x_{N+1}, x_{N+1}+dx_{N+1}]$ by integrating (18) over x_1, \dots, x_N ,

$$\frac{1}{N!} \Psi(x_{N+1}) dx_{N+1} \left(\int_{x_{N+1}}^{\infty} \Psi(x) dx \right)^N \exp \left(- \int_{x_{N+1}}^{\infty} \Psi(z) dz \right). \quad (19)$$

[As we imposed $\int_{-\infty}^{+\infty} \psi(\epsilon) d\epsilon = \infty$ in the definition of the model, this distribution is normalized; see (8).] Finally, the probability of x_1, \dots, x_N given x_{N+1} is the ratio of (18) by (19). One can see that, given the value of x_{N+1} , the distributions of $x_1(g+1), \dots, x_N(g+1)$ are independent and one gets that, given x_{N+1} , each of the N rightmost particles is in $[x, x+dx]$ with probability

$$\frac{\Psi(x) dx}{\int_{x_{N+1}}^{\infty} \Psi(x) dx} \quad \text{for } x_{N+1} < x. \quad (20)$$

Therefore, to generate the whole population after selection at generation $g+1$, one needs to calculate the density $\Psi(x)$ according to (8), then to choose the position of the $(N+1)$ st rightmost particle according to (19) and, finally, to generate *independently* the N rightmost particles $x_1(g+1), \dots, x_N(g+1)$ with the distribution (20). Note that the $(N+1)$ st particle is not selected and is therefore eliminated after the N rightmost particles have been generated. This procedure is valid for any $\psi(\epsilon)$, but is in general complicated because (8) is not easy to handle analytically.

A. Statistics of the position of the front in the exponential model

In the exponential model $\psi(\epsilon) = e^{-\epsilon}$, however, everything becomes simpler: the Poisson process (8) becomes

$$\Psi_{\text{exp}}(x) = e^{-(x-X_g)} \quad \text{with } X_g = \ln(e^{x_1(g)} + e^{x_2(g)} + \dots + e^{x_N(g)}), \quad (21)$$

which means that the offspring of the whole population are distributed as if they were the offspring of a single effective individual located at position X_g . The distribution of the $(N+1)$ st rightmost particle (19) becomes

$$x_{N+1} = X_g + z \quad \text{with } \text{Prob}(z) = \frac{1}{N!} \exp[-(N+1)z - e^{-z}], \quad (22)$$

and, once x_{N+1} has been chosen, the distribution (20) of the $x_k(g+1)$ for $k=1, \dots, N$ becomes

$$x_k(g+1) = x_{N+1} + y_k \quad \text{with } \text{Prob}(y_k) = e^{-y_k} \quad \text{for } y_k > 0. \quad (23)$$

We now recall the calculation of the statistics of the position of the front [9] which was done for a similar model in [14], because we shall use later the same approach to calculate the statistics of the genealogical trees.

There are many ways of defining the position of the front at a given generation g . One could consider the position of its center of mass, or the position of the rightmost or leftmost individual, or actually, any function of the positions $x_k(g)$

such that a global shift of all the $x_k(g)$ leads to the same shift in the position of the front. Because the front does not spread, the difference between two such definitions of the position does not grow with time so that, in the limit $g \rightarrow \infty$, all these definitions lead to the same velocity, diffusion constant, and higher cumulants.

For the exponential model, it is convenient to use X_g , defined in (21), as the position of the front. Indeed, one can write

$$\Delta X_g = X_{g+1} - X_g = z + \ln(e^{y_1} + e^{y_2} + \dots + e^{y_N}), \quad (24)$$

where the definitions and probability distributions of z and y_k are given in (22) and (23). From (24), the shifts ΔX_g are uncorrelated random variables, and the average velocity v_N and diffusion constant D_N of the front are given by

$$v_N = \langle \Delta X_g \rangle, \quad D_N = \langle \Delta X_g^2 \rangle - \langle \Delta X_g \rangle^2. \quad (25)$$

More generally, all cumulants of the front position at a long time g are simply g times the cumulants of ΔX_g . To compute these cumulants, we evaluate the generating function $G(\beta)$ defined as

$$e^{G(\beta)} = \langle e^{-\beta \Delta X_g} \rangle = \int_{-\infty}^{+\infty} dz \text{Prob}(z) e^{-\beta z} \int_0^{+\infty} dy_1 \text{Prob}(y_1) \dots \times \int_0^{+\infty} dy_N \text{Prob}(y_N) (e^{y_1} + \dots + e^{y_N})^{-\beta}, \quad (26)$$

and one obtains the cumulants by doing a small β expansion,

$$G(\beta) = \sum_{n \geq 1} \frac{(-\beta)^n}{n!} \langle \Delta X_g^n \rangle_c. \quad (27)$$

Using (22), the integral over z is easy,

$$\int_{-\infty}^{+\infty} dz \text{Prob}(z) e^{-\beta z} = \frac{1}{N!} \int_{-\infty}^{+\infty} dz \exp[-(\beta + N + 1)z - e^{-z}] = \frac{\Gamma(N + 1 + \beta)}{\Gamma(N + 1)}. \quad (28)$$

To calculate the integrals over y_i in (26), one can use the representation (valid for $\beta > 0$)

$$Z^{-\beta} = \frac{1}{\Gamma(\beta)} \int_0^{+\infty} d\lambda \lambda^{\beta-1} e^{-\lambda Z} \quad (29)$$

with $Z = e^{y_1} + \dots + e^{y_N}$. This leads to the factorization of the integrals over y_1, \dots, y_N . Replacing $\text{Prob}(y_k)$ by its explicit expression from (23), one gets for $\beta > 0$ (a similar calculation can be made for $\beta > -1$),

$$e^{G(\beta)} = \frac{\Gamma(N + 1 + \beta)}{\Gamma(N + 1)\Gamma(\beta)} \int_0^{+\infty} d\lambda \lambda^{\beta-1} I_0(\lambda)^N, \quad (30)$$

where

$$I_0(\lambda) = \int_0^{+\infty} dy e^{-y - \lambda e^y}. \quad (31)$$

One can rewrite $I_0(\lambda)$ in several ways,

BRUNET *et al.*PHYSICAL REVIEW E **76**, 041104 (2007)

$$\begin{aligned}
I_0(\lambda) &= \lambda \int_{\lambda}^{+\infty} \frac{du}{u^2} e^{-u} \\
&= e^{-\lambda} + \lambda(\ln \lambda + \gamma_E) - \lambda \int_0^{\lambda} du \frac{1 - e^{-u}}{u} \\
&= 1 + \lambda(\ln \lambda + \gamma_E - 1) - \sum_{k=0}^{+\infty} \frac{(-1)^k}{(k+1)(k+2)!} \lambda^{k+2},
\end{aligned} \tag{32}$$

where $\gamma_E = -\Gamma'(1)$ is the Euler constant. (It is easy to check that the derivatives of these expressions divided by λ coincide and the integration constant can be checked in the large λ limit.) The expansion (32) can also be found in the literature [35] as I_0 is an exponential integral.

As $I_0(\lambda)$ is a monotonous decreasing function, the integral (30) is dominated by λ close to 0. In fact, using (30), one can check that the range of values of λ which dominate (30) is of the order of $1/[N \ln N]$. Indeed, if one makes the change of variables

$$\mu = \lambda N \ln N, \tag{33}$$

one gets $I_0(\lambda)^N$ for values of μ of order 1,

$$\begin{aligned}
[I_0(\lambda)]^N &\simeq \exp[N\lambda(\ln \lambda + \gamma_E - 1)] \\
&\simeq \exp\left(\frac{\mu}{\ln N}(\ln \mu - \ln N - \ln \ln N + \gamma_E - 1)\right) \\
&\simeq e^{-\mu} \left[1 + \mu \frac{\ln \mu - \ln \ln N + \gamma_E - 1}{\ln N} \right. \\
&\quad \left. + \frac{1}{2} \left(\mu \frac{\ln \mu - \ln \ln N + \gamma_E - 1}{\ln N} \right)^2 + \dots \right],
\end{aligned} \tag{34}$$

where terms of order $1/N$ have been dropped. Replacing this expression into (30) and using

$$\int_0^{\infty} d\mu \mu^{x-1} e^{-\mu} (\ln \mu)^k = \frac{d^k}{dx^k} \Gamma(x), \tag{35}$$

one gets

$$\begin{aligned}
e^{G(\beta)} &\simeq \frac{\Gamma(N+1+\beta)}{\Gamma(N+1)\Gamma(\beta)} \frac{1}{(N \ln N)^\beta} \left(\Gamma(\beta) \right. \\
&\quad \left. + \frac{\Gamma'(\beta+1) + \Gamma(\beta+1)(-\ln \ln N + \gamma_E - 1)}{\ln N} + \dots \right) \\
&\simeq \frac{\Gamma(N+1+\beta)}{\Gamma(N+1)} \frac{1}{(N \ln N)^\beta} \left[1 + \frac{\beta}{\ln N} \left(\frac{\Gamma'(\beta+1)}{\Gamma(\beta+1)} \right. \right. \\
&\quad \left. \left. - \ln \ln N + \gamma_E - 1 \right) + \dots \right].
\end{aligned} \tag{36}$$

(The next order is obtained in Appendix A.) The Stirling formula allows to simplify the expression

$$\frac{\Gamma(N+1+\beta)}{\Gamma(N+1)} \frac{1}{N^\beta} = 1 + O\left(\frac{1}{N}\right). \tag{37}$$

Then, one gets from (36) the following expression for the generating function:

$$\begin{aligned}
G(\beta) &= -\beta \ln \ln N - \frac{\beta}{\ln N} \left(\ln \ln N + 1 - \gamma_E - \frac{\Gamma'(1+\beta)}{\Gamma(1+\beta)} \right) \\
&\quad + o\left(\frac{1}{\ln N}\right).
\end{aligned} \tag{38}$$

[This expression was obtained assuming $\beta > 0$, but one can show that it remains valid for $\beta > -1$ by using, instead of (29), a different representation of $Z^{-\beta}$.] Now one simply reads off the expressions of the cumulants of the position of the front by comparing the expansion of (38) in powers of β and (27),

$$\begin{aligned}
v_N &= \frac{\langle X_g \rangle}{g} = \langle \Delta X_g \rangle = \ln \ln N + \frac{1}{\ln N} (\ln \ln N + 1) + \dots, \\
D_N &= \frac{\langle X_g^2 \rangle_c}{g} = \langle \Delta X_g^2 \rangle_c = \frac{\pi^2}{3 \ln N} + \dots \\
\frac{\langle X_g^n \rangle_c}{g} &= \langle \Delta X_g^n \rangle_c = \frac{n! \zeta(n)}{\ln N} = \frac{n!}{\ln N} \sum_{i \geq 1} \frac{1}{i^n} + \dots,
\end{aligned} \tag{39}$$

up to terms of order $\ln \ln N / \ln^2 N$ that are computed in Appendix A. The velocity v_N diverges for large N , in contrast with models of the Fisher-KPP class for which v_N has a finite large N limit. Note that velocities which become infinite in the large N limit occur in other models of evolution with selection [2].

B. Trees in the exponential model

Let us now consider the ancestors of a group of $p \geq 2$ individuals chosen at random in the population (of size N). Looking at their genealogy, one observes a tree which fluctuates with the choice of the p individuals and which is characterized by its shape and coalescence times.

For model B with an arbitrary density $\psi(\epsilon)$, the probability of finding, at generation $g+1$ before selection, an offspring in $[x, x+dx]$ is $\Psi(x)dx$ with Ψ given by (8). On the other hand, the probability of finding in $[x, x+dx]$ an offspring of $x_i(g)$ is, by definition, $\psi[x-x_i(g)]dx$. Therefore, given an offspring at generation $g+1$ and position x , the probability that its parent was the i th individual [at position $x_i(g)$] is

$$W_i(x) = \frac{\psi[x-x_i(g)]}{\Psi(x)}. \tag{40}$$

For general $\psi(\epsilon)$, these probabilities $W_i(x)$ depend on x , making the calculation of these coalescence times difficult. In the exponential model, however, (40) becomes

$$W_i = e^{x_i(g) - X_g} = \frac{e^{x_i(g)}}{e^{x_1(g)} + \dots + e^{x_N(g)}} = \frac{e^{y_i}}{e^{y_1} + \dots + e^{y_N}}, \quad (41)$$

where the $y_k = x_k(g) - x_{N+1}(g)$ are the exponential variables of (23). Therefore the W_i do not depend on x . It follows that the probability q_p that p individuals at generation $g+1$ have the same ancestor at generation g is simply

$$q_p = \left\langle \sum_{i=1}^N W_i^p \right\rangle, \quad (42)$$

where the average is over the y_i of (41). After performing this average, all the terms in the sum over i become equal since the y_i are identically distributed. Therefore

$$\begin{aligned} q_p &= N \langle W_1^p \rangle \\ &= N \int_0^{+\infty} dy_1 e^{-y_1} \dots \int_0^{+\infty} dy_N e^{-y_N} e^{p y_1} (e^{y_1} + \dots + e^{y_N})^{-p}. \end{aligned} \quad (43)$$

Using the representation (29), one obtains

$$q_p = \frac{N}{(p-1)!} \int_0^{+\infty} d\lambda \lambda^{p-1} I_p(\lambda) I_0(\lambda)^{N-1} \quad (44)$$

in terms of the function $I_0(\lambda)$ introduced in (31) and of its derivatives

$$\begin{aligned} I_p(\lambda) &= \int_0^{+\infty} dy e^{(p-1)y - \lambda e^y} = (-)^p \frac{d^p}{d\lambda^p} I_0(\lambda) \\ &= \lambda^{1-p} \int_{\lambda}^{+\infty} du u^{p-2} e^{-u}. \end{aligned} \quad (45)$$

For small λ one gets, by taking derivatives of Eq. (32), to leading order,

$$I_0(\lambda) \simeq 1 + \lambda(\ln \lambda + \gamma_E - 1), \quad I_1(\lambda) \simeq -(\ln \lambda + \gamma_E),$$

$$I_p(\lambda) \simeq \frac{(p-2)!}{\lambda^{p-1}} \quad \text{for } p \geq 2. \quad (46)$$

So far, (44) is an exact expression and valid for arbitrary N . From now on, we will work at leading order in $\ln N$, leaving the extension to subleading orders to Appendix A.

As for the obtention of (38) from (30), the integral over λ is dominated by the region where λ is of order $1/[N \ln N]$. Doing the same change of variable $\mu = \lambda N \ln N$, one gets $I_0(\lambda)^N \simeq e^{-\mu}$ and, using (46), $\lambda^{p-1} I_p(\lambda) \simeq (p-2)!$. Therefore, we obtain for $p \geq 2$,

$$q_p \simeq \frac{1}{\ln N} \frac{1}{p-1}. \quad (47)$$

We see that for large N the probability that p branches merge is of the same order for all p , in contrast to the neutral model ([36,37] and Appendix C) for which q_p is of order $1/N^{p-1}$, so that $q_2 \gg q_3 \gg q_4 \gg \dots$.

To calculate the moments of the coalescence times, it is convenient to introduce the probability $r_p(k)$ that p randomly chosen individuals at generation $g+1$ have exactly k ancestors at generation g . In one generation, at leading order in N , only a single coalescence may occur among the p individuals, and (47) tells us that the coalescence probability goes like $1/\ln N$ (any additional coalescence at the same generation would in fact cost an additional power of $1/\ln N$; see Appendix A). Consequently, we just need that $p-k+1$ individuals coalesce to one ancestor, say individual number i (the probability is W_i^{p-k+1}), and that none of the other individuals have i as an ancestor [probability $(1-W_i)^{k-1}$]. Altogether, this reads¹

$$r_p(k) = \binom{p}{k-1} \left\langle \sum_{i=1}^N W_i^{p-k+1} (1-W_i)^{k-1} \right\rangle. \quad (48)$$

The factor $(1-W_i)^{k-1}$ may be expanded and the average may be expressed with the help of the q_p defined in (42),

$$r_p(k) = \binom{p}{k-1} \sum_{j=0}^{k-1} \binom{k-1}{j} (-1)^{k-1-j} q_{p-j}. \quad (49)$$

Replacing (47) in (49), one gets after some algebra

$$r_p(k) \simeq \frac{1}{\ln N} \frac{p}{(p-k)(p-k+1)}, \quad (50)$$

which holds for $k < p$. The probability $r_p(p)$ that there is no coalescence at all among the p individuals (that is to say, that all p have distinct ancestors) has a simple expression, which is obtained from a completeness relation,

$$r_p(p) = 1 - \sum_{k=1}^{p-1} r_p(k) \simeq 1 - \frac{p-1}{\ln N}. \quad (51)$$

The knowledge of the probabilities $r_p(k)$ in (50) and (51) allows one to determine (in the large N limit) all the statistical properties of the trees.

We introduce the probability $P_p(g)$ that p individuals have their first common ancestor a number of generations g in the past. For $p \geq 2$, one may write a recursion for $P_p(g)$ in the form

$$P_p(g+1) = \sum_{k=2}^p r_p(k) P_k(g) + r_p(1) \delta_g^0. \quad (52)$$

Using (50) and (51), this becomes

¹In the mathematical literature, one would rather use the transition rates $\lambda_{b,q}$ which give the probability that out of b individuals, the only event is the coalescence of the q first individuals [38,39]. Clearly, $r_p(k) = \binom{p}{k-1} \lambda_{p,p-k+1}$. All the $\lambda_{b,q}$ can be obtained through a measure Λ through $\lambda_{b,q} = \int_0^1 x^{q-2} (1-x)^{b-q} \Lambda(dx)$. The exponential model corresponds to a uniform measure Λ , studied in [40].

$$\begin{aligned}
P_p(g+1) - P_p(g) &= -\frac{p-1}{\ln N} P_p(g) \\
&+ \sum_{k=2}^{p-1} \frac{1}{\ln N} \frac{p}{(p-k)(p-k+1)} P_k(g) \\
&+ r_p(1) \delta_g^0. \tag{53}
\end{aligned}$$

In the large- N limit, the number of generations g over which the coalescence occurs is typically $\ln N \gg 1$ (since the coalescence probabilities scale like $1/\ln N$). It is then natural to introduce the rescaled variable $t = g/\ln N$ and the corresponding coalescence probability $R_p(t) dt = P_p(g) dg$. In this new variable, the recursion becomes for $t > 0$,

$$\frac{dR_p(t)}{dt} = -(p-1)R_p(t) + \sum_{k=2}^{p-1} \frac{p}{(p-k)(p-k+1)} R_k(t). \tag{54}$$

This equation may be solved by introducing the generating function

$$\Psi(\lambda, t) = \sum_{p \geq 2} \lambda^{p-1} R_p(t), \tag{55}$$

which turns the summation over k in (54) into

$$\frac{\partial \Psi}{\partial t} = [(1-\lambda) \ln(1-\lambda)] \frac{\partial \Psi}{\partial \lambda} - [\ln(1-\lambda)] \Psi. \tag{56}$$

The general solution (which can be obtained by the method of characteristics) reads

$$\Psi(\lambda, t) = \frac{1}{1-\lambda} \phi[e^{-t} \ln(1-\lambda)], \tag{57}$$

where ϕ is an arbitrary function. The initial condition for (54) is the probability that all p individuals coalesce between times 0 and dt [see (47)],

$$R_p(t=0) dt = q_p \frac{dg}{dt} dt = \frac{dt}{p-1}, \tag{58}$$

and thus, (55) becomes

$$\Psi(\lambda, t=0) = -\ln(1-\lambda). \tag{59}$$

This leads to

$$\Psi(\lambda, t) = \frac{d}{dt} (1-\lambda)^{e^{-t}-1}. \tag{60}$$

The expansion of (60) in powers of λ using

$$(1-\lambda)^{-a} = \frac{1}{\Gamma(a)} \sum_{p=0}^{+\infty} \frac{\Gamma(p+a)}{\Gamma(p+1)} \lambda^p, \tag{61}$$

leads through (55) to

$$\begin{aligned}
R_p(t) &= \frac{1}{(p-1)!} \frac{d}{dt} \frac{\Gamma(p-e^{-t})}{\Gamma(1-e^{-t})} \\
&= \frac{1}{(p-1)!} \frac{d}{dt} [(1-e^{-t})(2-e^{-t}) \cdots (p-1-e^{-t})], \tag{62}
\end{aligned}$$

which is just a polynomial of order $p-1$ in the variable e^{-t} . More explicitly, for the first values of p , one finds

$$R_2(t) = e^{-t}, \quad R_3(t) = \frac{3}{2}e^{-t} - e^{-2t}, \tag{63}$$

$$R_4(t) = \frac{11}{6}e^{-t} - 2e^{-2t} + \frac{1}{2}e^{-3t}, \quad \dots$$

The average coalescence times [using (62)] are

$$\begin{aligned}
\langle T_p \rangle &= \sum_{g=0}^{\infty} g P_p(g) = \ln N \int_0^{+\infty} dt t R_p(t) \\
&= \ln N \int_0^{\infty} dt \left[1 - (1-e^{-t}) \left(1 - \frac{e^{-t}}{2} \right) \cdots \left(1 - \frac{e^{-t}}{p-1} \right) \right] \tag{64}
\end{aligned}$$

and one gets

$$\langle T_2 \rangle = \ln N, \quad \langle T_3 \rangle = \frac{5}{4} \langle T_2 \rangle, \quad \langle T_4 \rangle = \frac{25}{18} \langle T_2 \rangle, \quad \dots \tag{65}$$

These expressions contrast with a neutral model of coalescence with no selection [37,41], where at each generation one would choose the N survivors at random among all the offspring at generation $g+1$ (see Appendix C),

$$\langle T_2^{\text{neutral}} \rangle = \mathcal{O}(N), \quad \langle T_3^{\text{neutral}} \rangle = \frac{4}{3} \langle T_2^{\text{neutral}} \rangle, \tag{66}$$

$$\langle T_4^{\text{neutral}} \rangle = \frac{3}{2} \langle T_2^{\text{neutral}} \rangle, \quad \dots$$

(Table I compares the frequencies of the trees in the cases with and without selection.)

As shown in Appendix B, the ratios (65) are on the other hand identical to those that would be computed if the genealogical trees had the same statistical properties as mean-field spin glasses [40,42].

We also see that $\langle T_p \rangle$ in (65) scales like $\ln N$ for any fixed value of p , which means that on average, a given number of individuals have their first common ancestor at order $\ln N$ generations in the past. It is, however, interesting to note that for large p ,

$$\langle T_p \rangle \simeq \ln N \times \ln \ln p \tag{67}$$

which is obtained by using, from (62), $R_p(t) \simeq \frac{d}{dt} p^{-\exp(-t)} \simeq \frac{d}{dt} e^{-\exp[-(t-\ln \ln p)]}$ for large p ; $R_p(t)$ becomes a Gumbel distribution of width of order 1 centered at $\ln \ln p$.

IV. PHENOMENOLOGICAL EXTENSION TO GENERIC MODELS

The exponential model had the advantage of being exactly solvable, but as already mentioned, it is nongeneric because

TABLE I. Probabilities of observing each of the possible genealogical trees for three and four individuals in the neutral case and in the exponential model.

	Neutral case	Exponential model
	$\frac{2}{3}$	$\frac{1}{3}$
	$\frac{1}{3}$	$\frac{1}{6}$
	0	$\frac{1}{6}$
	0	$\frac{2}{9}$
	0	$\frac{1}{9}$

	Neutral case	Exponential model
	1	$\frac{3}{4}$
	0	$\frac{1}{4}$

the velocity $v_{N \rightarrow \infty}$ as $N \rightarrow \infty$, in contrast to models of the Fisher-KPP type. We do not know how to calculate directly the velocity v_N , diffusion constant D_N , or the coalescence times of the generic Fisher-KPP case. One can, however, use a phenomenological picture of front propagation [8] and ancestry, which is consistent with exact calculations in the case of the exponential model, and agrees with numerical simulations in the generic case.

A. Picture of the propagation of fluctuating pulled fronts

Let us recall briefly the phenomenological picture of front propagation that emerged from [8,43]. In this picture, most of the time, the front evolves in a deterministic way well reproduced by an equation obtained from (12) by removing the noise term, and by adding a cutoff that takes into account the discreteness of the number of individuals: This ensures that $h_g(x)$ cannot take values less than $1/N$. The evolution equation in the case of model B reads [43]

$$h_{g+1}(x) = \begin{cases} \min\left(1, \int d\epsilon \psi(\epsilon) h_g(x - \epsilon)\right) & \text{if that number} \\ 0 & \text{is larger than } 1/N, \\ & \text{otherwise.} \end{cases} \quad (68)$$

{Note that in the exponential model [$\psi(\epsilon) = e^{-\epsilon}$], it is easy to see that the solution to (68) is

$$h_g(x) = \begin{cases} 1 & \text{for } x < Y_g, \\ e^{-(x-Y_g)} & \text{for } Y_g < x < Y_g + \ln N, \\ 0 & \text{for } x > Y_g + \ln N, \end{cases} \quad (69)$$

where the parameter Y_g can be used as the definition of the position of the front. Substituting (69) into (68), one obtains the velocity

$$v_{\text{cutoff}}^{\text{exp}} = Y_{g+1} - Y_g = \ln(\ln N + 1) \simeq \ln \ln N, \quad (70)$$

which does agree, to leading order, with the exact expression (39).}

For fronts in the Fisher-KPP class [including (68)], the cutoff theory can also be worked out [43]. One obtains

$$h_g(x) \propto L_0 \sin\left(\pi \frac{x - Y_g}{L_0}\right) e^{-\gamma_0(x - Y_g)} \quad (71a)$$

and

$$v_{\text{cutoff}}^{\text{F-KPP}} = Y_{g+1} - Y_g \simeq v(\gamma_0) - \frac{\pi^2 v''(\gamma_0)}{2L_0^2}, \quad (71b)$$

where $v(\gamma)$ is given by (14), γ_0 is the value of γ which minimizes $v(\gamma)$, and $L_0 = (\ln N) / \gamma_0$ is the length of the front, from the region where h_g is of order 1 to the region where it cancels. The expression of $h_g(x)$ in (71a) is only valid for $h_g(x) \ll 1$ and $x - Y_g < L_0$.

BRUNET *et al.*PHYSICAL REVIEW E **76**, 041104 (2007)

By convention, we shall define $\gamma_0=1$ in the exponential case. Then, both in (69) and in (71a), the front has essentially an exponential decay with rate γ_0 and its length is $L_0 = (\ln N)/\gamma_0$.

So far, (70) and (71) have been obtained from a purely deterministic calculation (68), where only the discreteness of $h_g(x)$ has been taken into account. Stochasticity may be put back in the picture for the generic (Fisher-KPP) case in the following phenomenological way, as developed in [8].

From time to time, a rare fluctuation sends a few individuals ahead of the front at a distance δ from its tip. This occurs during the time interval dt with a probability $p(\delta) d\delta dt$ where $p(\delta)$ was assumed [8] to be

$$p(\delta) = C_1 e^{-\gamma_0 \delta} \quad (72)$$

for δ large enough. C_1 is a given constant.

These individuals then multiply and build up their own front in an essentially deterministic way. After about L_0^2 generations, the descendants of these individuals have mixed up with the individuals that stem from the rest of the front. The effect of this rare fluctuation is therefore to pull ahead the front by a quantity $R(\delta)$ which, in the generic (Fisher-KPP) case, is given [8] by

$$R(\delta) = \frac{1}{\gamma_0} \ln \left(1 + C_2 \frac{e^{\gamma_0 \delta}}{L_0^\alpha} \right), \quad (73)$$

where C_2 is another constant and $\alpha=3$. Finally, in [8] it was argued that

$$C_1 C_2 = \pi^2 \gamma_0 v''(\gamma_0). \quad (74)$$

[Note that (72)–(74) have been obtained in [8] on heuristic arguments and we do not know how to properly derive them.]

As we shall show in the next section, the same picture applies to the exponential model with some slight modifications: in (73), one needs to take $\alpha=1$ instead of $\alpha=3$, everywhere γ_0 must be replaced by 1, one should replace (74) by $C_1=C_2=1$ and the relaxation time of a fluctuation by 1 instead of L_0^2 .

With these ingredients, it is not difficult to write the generating function of the position Y_g of the front,

$$\langle e^{-\beta Y_g} \rangle \sim e^{gG(\beta)} \quad \text{where} \quad (75)$$

$$G(\beta) = -\beta v_{\text{cutoff}} + \int d\delta p(\delta) (e^{-\beta R(\delta)} - 1).$$

The first term in $G(\beta)$ is due to the deterministic motion, while the integral represents the effect of the forward rare fluctuations. In the case of the exponential model, this expression leads to (39), up to terms of order $1/\ln N$ for the velocity and of order $\ln \ln N / \ln^2 N$ for the other cumulants. In the generic Fisher-KPP case, the average front velocity, diffusion constant, and higher order cumulants are found from (75) to be [8]

$$v_N = v(\gamma_0) - \frac{\pi^2 \gamma_0^2 v''(\gamma_0)}{2 \ln^2 N} + \gamma_0^2 v''(\gamma_0) \pi^2 \frac{3 \ln \ln N}{\ln^3 N} + \dots$$

$$= v(\gamma_0) - \frac{\pi^2 \gamma_0^2 v''(\gamma_0)}{2 (\ln N + 3 \ln \ln N)^2} + \dots,$$

$$D_N = \gamma_0 v''(\gamma_0) \frac{\pi^4}{3 \ln^3 N} + \dots,$$

$$\frac{\langle (Y_g - Y_0)^n \rangle_c}{g} = \gamma_0^{3-n} v''(\gamma_0) \frac{\pi^2 n! \zeta(n)}{\ln^3 N} + \dots \quad \text{for } n \geq 2. \quad (76)$$

One important aspect of (73) is that when δ is of order $(\alpha \ln L_0)/\gamma_0$, the front is shifted by one additional unit in position due to this fluctuation. This means that a large fraction of the population is replaced by the descendants of the individuals produced by this fluctuation. Thus, when one considers a given number of individuals at generation g , the most probable is that their most recent common ancestor belongs to one of these fluctuations that triggered shifts of order 1 in the position of the front in the past generations. According to (72), such events occur once every $\Delta g \sim L_0^\alpha$ generations. Δg is likely to give the order of magnitude of the average coalescence times. In Sec. IV C, we shall build on this observation to obtain the statistics of the genealogical trees and the coalescence times in the generic Fisher-KPP case. But first, we show that this phenomenological picture is consistent with the exact results (39) for the exponential model.

B. Exponential model

Since the exponential model can be solved exactly (Sec. III), we are now going to test in this case our phenomenological picture of Sec. IV A. Let us first show that (72) gives the correct distribution of fluctuations.

In the exponential model at any generation g , the front is built according to (23) by drawing N independent exponential random numbers y_k , which represent the positions of the particles relative to a common origin x_{N+1} . There is a probability $(1 - e^{-y})^N$ that none of the y_k are on the right of y ; therefore the distribution of the rightmost y_k is

$$\begin{aligned} \text{Prob}(y_{\text{rightmost}}) &= N(1 - e^{-y_{\text{rightmost}}})^{N-1} e^{-y_{\text{rightmost}}} \\ &\simeq \exp[-(y_{\text{rightmost}} - \ln N) - e^{-(y_{\text{rightmost}} - \ln N)}]. \end{aligned} \quad (77)$$

$y_{\text{rightmost}}$ is the distance between the rightmost particle and the $(N+1)$ st rightmost particle (before selection). We define the length l of the front as $l = y_{\text{rightmost}}$. (A more natural definition could have been the distance between the rightmost and the leftmost particles, which is obtained by replacing N by $N-1$ in the previous equation. For large N , the difference between these two definitions is negligible.) The average length of the front is therefore $\langle l \rangle \simeq \ln N + \gamma_E$ with fluctuations of order 1 given by a Gumbel distribution, and the probability to observe a large fluctuation where $l = \ln N + \delta$ with $\delta \gg 1$ is given by

$$p(\delta) \simeq \exp(-\delta - e^{-\delta}) \simeq \exp(-\delta), \quad (78)$$

which is the same as (72).

We now wish to know the effect of such a fluctuation on the position of the front. As the shape of the front is isocorrelated between two successive generations, the relaxation time of a fluctuation is 1 and it is sufficient to compute ΔX_g

given the value of δ at generation g . Given the value of $l = y_{\text{rightmost}}$, the distribution (23) of the $N-1$ other y_k becomes

$$\text{Prob}(y_k) = \frac{e^{-y_k}}{1 - e^{-l}} \quad \text{for } 0 < y_k < l. \quad (79)$$

As in (26), we introduce the generating function of the displacement ΔX_g , given the value of l ,

$$\begin{aligned} \langle e^{-\beta \Delta X_g | l} \rangle &= \int dz \text{Prob}(z) e^{-\beta z} \int dy_1 \text{Prob}(y_1) \cdots \int dy_{N-1} \text{Prob}(y_{N-1}) (e^{y_1} + \cdots + e^{y_{N-1}} + e^l)^{-\beta} \\ &= \frac{\Gamma(N+1+\beta)}{\Gamma(N+1)} \frac{1}{(1 - e^{-l})^{N-1}} \int_0^l dy_1 e^{-y_1} \cdots \int_0^l dy_{N-1} e^{-y_{N-1}} (e^{y_1} + \cdots + e^{y_{N-1}} + e^l)^{-\beta}, \end{aligned} \quad (80)$$

where (28) and (79) were used. By using the same representation (29) that led to (30), one gets

$$\begin{aligned} \langle e^{-\beta \Delta X_g | l} \rangle &= \frac{\Gamma(N+1+\beta)}{\Gamma(N+1)\Gamma(\beta)} \int_0^\infty d\lambda \lambda^{\beta-1} \\ &\times \left(\frac{1}{1 - e^{-l}} \int_0^l dy e^{-y-\lambda e^y} \right)^{N-1} e^{-\lambda e^l}, \end{aligned} \quad (81)$$

which, in terms of $I_0(\lambda)$ defined in (31), is the same as

$$\begin{aligned} \langle e^{-\beta \Delta X_g | l} \rangle &= \frac{\Gamma(N+1+\beta)}{\Gamma(N+1)\Gamma(\beta)} \int_0^\infty d\lambda \lambda^{\beta-1} \\ &\times \left(\frac{I_0(\lambda) - e^{-l} I_0(\lambda e^l)}{1 - e^{-l}} \right)^{N-1} e^{-\lambda e^l}, \end{aligned} \quad (82)$$

where, using (32),

$$\begin{aligned} \frac{I_0(\lambda) - e^{-l} I_0(\lambda e^l)}{1 - e^{-l}} &= 1 - \lambda \frac{l}{1 - e^{-l}} \\ &+ \sum_{k=0}^{+\infty} \frac{(-1)^k}{(k+1)(k+2)!} \lambda^{k+2} \frac{e^{l(k+1)} - 1}{1 - e^{-l}}. \end{aligned} \quad (83)$$

Expressions (82) and (83) are valid for any value of l . We now consider a large fluctuation $l = \ln N + \delta$ with $1 \ll \delta \leq \ln \ln N$. As for (30), the integral is dominated by values of λ of order $1/(N \ln N)$. Making as before the change of variable $\mu = \lambda N \ln N$, and dropping all the terms of order $1/N$, one gets

$$\begin{aligned} \left(\frac{I_0(\lambda) - e^{-l} I_0(\lambda e^l)}{1 - e^{-l}} \right)^{N-1} &\simeq \exp \left[-\mu \left(1 + \frac{\delta}{\ln N} \right) \right. \\ &+ \sum_{k=0}^{+\infty} \frac{(-1)^k}{(k+1)(k+2)!} \\ &\left. \times \left(\frac{\mu}{\ln N} \right)^{k+2} e^{\delta(k+1)} \right]. \end{aligned} \quad (84)$$

We are only interested in the leading order in $1/\ln N$. Dropping higher order terms, one gets, in (82),

$$\begin{aligned} \langle e^{-\beta \Delta X_g | \delta} \rangle &\simeq \frac{\Gamma(N+1+\beta)}{\Gamma(N+1)\Gamma(\beta)} \frac{1}{(N \ln N)^\beta} \int_0^\infty d\mu \mu^{\beta-1} \\ &\times \exp \left[-\mu \left(1 + \frac{\delta + e^\delta}{\ln N} \right) \right] \\ &\simeq \frac{1}{(\ln N)^\beta} \left(1 + \frac{e^\delta}{\ln N} \right)^{-\beta}, \end{aligned} \quad (85)$$

where (37) has been used and where δ was neglected compared to e^δ .

This means that up to the order $1/(\ln N)$ we are considering, ΔX_g given δ is deterministic with

$$\Delta X_g(\delta) \simeq \ln \ln N + \ln \left(1 + \frac{e^\delta}{\ln N} \right) \simeq v_{\text{cutoff}} + R(\delta), \quad (86)$$

where we used (70) and (73) with $C_2 = \alpha = \gamma_0 = 1$.

The phenomenological picture we developed for the generic case is therefore justified for the exponential case: each rare fluctuation of size δ in the length of the front leads to a shift $R(\delta)$, given by (73), for the position of the front.

C. Genealogical trees

With the above scenario, one can also build a simplified picture for the evolution of a population. We assume that, at each generation, there is with a small probability a fluctua-

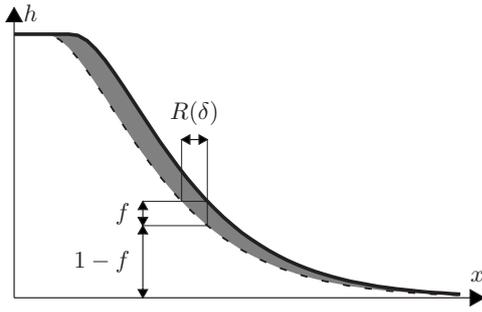


FIG. 3. Effect of a fluctuation of a front. The dashed line is the front (87) in the absence of a fluctuation. The plain line is the front (88) if a rare fluctuation occurred. The grey area represents the contribution to the front from the descendants of the fluctuation. After the front has relaxed, they represent a proportion f of the whole population.

tion of amplitude f produced by an individual ahead of the front. The long term effect of this fluctuation is that a fraction f of the population is replaced by the descendants of this individual.

One can now relate the probability distribution of f to the phenomenological picture of front propagation. Starting with a front at position Y_{g_0} at generation g_0 , we consider its position Y_g at a generation $g > g_0$. If no important fluctuation has occurred, the tail of the front is given by

$$h_{\text{no fluctuation}}(x, g) \propto e^{-\gamma_0(x - Y_g^{\text{no fluctuation}})} \quad (87a)$$

with

$$Y_g^{\text{no fluctuation}} = Y_{g_0} + v_{\text{cutoff}}(g - g_0). \quad (87b)$$

[See (71); for simplicity, we neglect the sine prefactor in the tail as it is a slowly varying factor which, to the leading order, does not change our final result.]

If instead a fluctuation has occurred, generated by an individual ahead of the front by a distance δ , then the shape is eventually described by

$$h_{\text{fluctuation}}(x, g) \propto e^{-\gamma_0(x - Y_g^{\text{fluctuation}})} \quad (88a)$$

with

$$Y_g^{\text{fluctuation}} = Y_{g_0} + v_{\text{cutoff}}(g - g_0) + R(\delta). \quad (88b)$$

that is, the front is pulled ahead by $R(\delta)$. If one assumes that the extra mass in the front with fluctuation (in gray in Fig. 3) is due to the fraction f of descendants originating from the fluctuation, then one gets $h_{\text{no fluctuation}} = (1-f)h_{\text{fluctuation}}$. The substitution of (87) and (88) yields

$$f = 1 - e^{-\gamma_0 R(\delta)}. \quad (89)$$

This equation defines the mapping between the f and the δ representations of the phenomenological model. The probability distribution of δ in (72) and the expression (73) of $R(\delta)$ implies the following distribution of f :

$$\text{Prob}(f) = \frac{C_1 C_2}{\gamma_0 L_0^\alpha} \frac{1}{f^2}. \quad (90)$$

[Note that this expression cannot be valid down to $f=0$ for the distribution to be normalized. One should therefore consider that (90) is valid above a certain small threshold f_{min} . This threshold has no effect on the correlations calculated below.]

Using (74) and $\alpha=3$ in the Fisher-KPP case, and $C_1=C_2=\gamma_0=\alpha=1$ in the exponential case (see Sec. IV B), one gets

$$\text{Prob}(f) = \begin{cases} \frac{1}{\ln N} \frac{1}{f^2} & \text{for the exponential model.} \\ \frac{\pi^2 \gamma_0^3 v''(\gamma_0)}{\ln^3 N} \frac{1}{f^2} & \text{for the generic Fisher-KPP case.} \end{cases} \quad (91)$$

In this model, p individuals may coalesce if they belong to the fraction f of individuals that are the descendants of a fluctuation. The probability of such an event thus reads

$$q_p = \int_0^1 df \text{Prob}(f) f^p = \frac{C_1 C_2}{\gamma_0 L_0^\alpha} \frac{1}{p-1} \quad (92)$$

which, for the exponential model, is identical to the exact asymptotic result in (47).

The coalescence probabilities in one generation $r_p(k)$ may be obtained in a straightforward way in this model. One first chooses the $k-1$ individuals among p that do not have a common ancestor in the previous generation. The latter must be part of the fraction $1-f$ of individuals, while the remaining $p-k+1$ individuals that have their common ancestor in the previous generation must belong to the fraction f . Thus

$$r_p(k) = \binom{p}{k-1} \int_0^1 df \text{Prob}(f) f^{p-k+1} (1-f)^{k-1} = \frac{C_1 C_2}{\gamma_0 L_0^\alpha} \frac{p}{(p-k)(p-k+1)}, \quad (93)$$

with the same result as in (50) for the exponential model.² At this point, the combinatorics to get the coalescence probabilities and average times are the same as in the exact calculation for the exponential model in Sec. III B. So, for the exponential model we recover the results of Sec. III B and for the generic Fisher-KPP case, we get instead

$$\langle T_2 \rangle \simeq \frac{\ln^3 N}{\pi^2 \gamma_0^3 v''(\gamma_0)}, \quad (94)$$

while the ratios $\langle T_i \rangle / \langle T_2 \rangle$ are the same (65) as for the exponential model, in agreement with the results of numerical simulations of [9] and of Sec. V below. Indeed, the $r_p(k)$'s

²In the language of the transition rates $\lambda_{b,q}$ defined in [38,39], one would write $\lambda_{b,q} = \int_0^1 df p(f) f^q (1-f)^{b-q} \propto \int_0^1 df f^{q-2} (1-f)^{b-q}$. It is the Λ -coalescent with the uniform measure, i.e., the Bolthausen-Sznitman coalescent.

given in (50) and (93) are identical except for an overall constant which cancels out in the ratios.

We note an interesting relation between the average coalescence time and the front diffusion constant, valid both in the exponential model and in the generic Fisher-KPP case,

$$D_N \times \langle T_2 \rangle \approx \frac{\pi^2}{3\gamma_0^2}. \quad (95)$$

We will test numerically this identity in Sec. V.

As a side remark, we note that if $\text{Prob}(f)$ of (91) is replaced by $C\text{ste} f^{-a}$ with $a \rightarrow 3$ (instead of $a=2$ in our selective evolution models), then the ratios of the coalescence times are identical to those obtained for evolution models without selection, see Appendix C.

V. NUMERICAL SIMULATIONS

A. Algorithms

In order to measure the velocity and diffusion constant of our models, it is sufficient to follow the evolution of the positions of the individuals. In the case of model A, at each generation, one first draws at random the k offspring of each individual and then one keeps the N rightmost offspring as the new population. This can be done in a computer time linear in N . For model B, one can start by drawing at random the two rightmost offspring of each individual. If Z is the position of the N th rightmost offspring out of this first set of $2N$, then one draws for each individual all its remaining offspring which are larger than Z . Then, taking the N rightmost individuals among those drawn gives the new population.

We measured the velocity v_N using $v_N = \langle X_{g_0+g} - X_{g_0} \rangle / g$ and the diffusion constants D_N as in [44], using $D_N = \langle (X_{g_0+g} - X_{g_0} - v_N g)^2 \rangle / g$ for a large g . (These expressions are in principle only valid in the $g \rightarrow \infty$ limit.) In practice, we must choose an appropriate value of g and average over many runs. For each value of N , we measured the diffusion constant twice, once with $g \approx 2 \ln^3 N$ and once with $g \approx 10 \ln^3 N$, and we have plotted both values with the same symbol. The fact that one cannot distinguish the two sets of data indicates that the values of g we took are large enough and that we accumulated enough statistics.

To measure the statistics of the genealogical trees in the population, one needs to memorize more information than simply the positions of the individuals in the current generation. The most naive method would be to record the whole history of the population, keeping for all individuals in all generations their positions and parents, and then to analyze at the end the whole genealogical tree. This is clearly too time and memory consuming. Instead, we used the three following algorithms.

The first algorithm consists in working with a matrix T_g , the element $T_g(i, i')$ being the age of the most recent common ancestor of the pair of individuals i and i' at generation g . This matrix is simple to update: if j and j' are the parents of i and i' , then $T_{g+1}(i, i') = 1 + T_g(j, j')$ for $i \neq i'$ and $T_{g+1}(i, i) = 0$. By sampling random elements of the matrix at different generations, one obtains the average value of the coalescence time between two individuals. The nice thing is

that, due to the ultrametric structure of the tree {for any i, j and k , $T_g(i, j) \leq \max[T_g(i, k), T_g(j, k)]$ }, no more information is needed to compute the coalescence times of three or more individuals: the age of the most recent ancestor of p individuals i_1, \dots, i_p is simply given by $\max[T_g(i_1, i_2), T_g(i_1, i_3), \dots, T_g(i_1, i_p)]$. This method is appropriate for values of N up to about 10^3 as it takes a long time of order N^2 to update the matrix at each generation.

In the second algorithm, instead of working with this matrix $T_g(i, j)$, we take advantage of the tree structure of the genealogy by recording only its "relevant" nodes: at generation g , we say that a node is "relevant" if it is an individual of the current generation g or if it is the first common ancestor of any pair of individuals of the current generation. Clearly, the "relevant" nodes have a tree structure (the first common ancestor of any two "relevant" nodes is a "relevant" node), which we record as well. The leaves of this tree are the current generation, and the root is the most recent common ancestor of the whole population. This tree is simple to update: if, after one timestep, a node has no child, it is removed and its parent is updated. If a node has only one child, it is removed as well and its child and parent get directly connected. If the root of the tree has only one child, it is removed and its child becomes the new root. As can be seen easily, the tree has at most $2N-1$ nodes and it can be updated in a time of order N . The extraction of the interesting information from the tree is also very fast: if a node has p children, and these children are the ancestors of $\alpha_1, \dots, \alpha_p$ individuals of the current generation, then this node is the most common ancestor of $\sum_{i \neq j} \alpha_i \alpha_j$ pairs of individuals. More generally, this node is the most common ancestor of $\binom{\sum_i \alpha_i}{q} - \sum_i \binom{\alpha_i}{q}$ groups of q individuals in the current generation. By computing this quantity on each node of the tree, one obtains the average (or even the distribution) of all the coalescence times within the current generation in a computer time of order N . This algorithm turns out to be very fast and we used it for N up to about 10^6 .

The third algorithm only works for a limited class of models, for which the positions $x_i(g)$ are integers: instead of recording the N positions, one only needs to record the number of individuals at a given site. The typical width of the front and, therefore, the number of variables to handle, are only of order $\ln N$. Let us, then, consider model B with $\psi(\epsilon)$ given as a sum of Dirac functions: $\psi(\epsilon) = \sum_q \phi_q \delta(\epsilon - q)$. This means that, before selection, an individual at position x has a number of offspring at position $x+q$ which has a Poisson distribution of average ϕ_q . Considering now the whole population, the number of offspring at time $g+1$ and site y is also a random Poisson number of average $\sum_x n(x, g) \phi_{y-x}$, where $n(x, g)$ is the number of individuals at site x and generation g [compare to (8)]. To simplify, we consider only cases where $\phi_q = 0$ for q larger than some q_0 , so that one can easily update the system from right to left by drawing Poisson numbers and stopping when the total number of individuals at time $g+1$ reaches N . So far, the method described allows us to update the positions of the particles, and therefore to extract the velocity and the diffusion constant, in a time proportional to $\ln N$ per generation. A similar method has already been used in [43,44] to simulate populations up to $N \approx 10^{100}$. To

BRUNET *et al.*

 PHYSICAL REVIEW E **76**, 041104 (2007)

extract the coalescence times, one needs to keep more information. The difficulty resides in the fact that the many individuals at a given position usually have different ancestors. To overcome this difficulty, one can consider the average coalescence times $\bar{T}_g(x, x')$ of two different individuals at respective positions x and x' . To update that matrix, one starts from the probability that an individual of generation $g+1$ and position y is the offspring of an individual who was at position x ,

$$\text{Prob}(y \text{ comes from } x) = \frac{n(x, g)\phi_{y-x}}{\sum_{x'} n(x', g)\phi_{y-x'}}. \quad (96)$$

[Compare to (40).] Then, one obtains that

$$\begin{aligned} \bar{T}_{g+1}(y, y') &= 1 + \sum_{x, x'} \text{Prob}(y \text{ comes from } x) \\ &\quad \times \text{Prob}(y' \text{ comes from } x') \bar{T}_g(x, x') \\ &\quad \times \left(1 - \frac{\delta_x^{y'}}{n(x, g)}\right). \end{aligned} \quad (97)$$

(The term in parentheses is the probability that individuals at positions y and y' come from two different parents given the parents' positions x and x' .) Then, the average coalescence time of two individuals in the population is simply given by

$$\frac{1}{N(N-1)} \sum_{x, x'} \bar{T}_g(x, x') n(x, g) n(x', g) \left(1 - \frac{\delta_x^{x'}}{n(x, g)}\right). \quad (98)$$

Therefore, by storing a matrix of size $\ln^2 N$, which can be updated in a time $\ln^4 N$, one can obtain the average coalescence time of two individuals. An interesting observation is that this algorithm simulates one possible realization of the positions of the particles; however, the quantity $\bar{T}_g(x, y)$ is

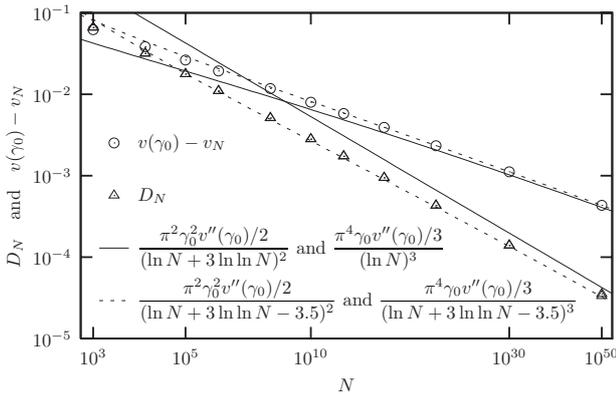


FIG. 4. Numerical simulations of model B with $\psi(\epsilon) = \frac{1}{4} \sum_{n=0}^{\epsilon} \delta(\epsilon-n)$. For this model, one has $\gamma_0 = \ln(2)$, $v(\gamma_0) = -1$, and $v''(\gamma_0) = 2/\ln(2)$. The circles are the correction to the velocity and the triangles the diffusion constant, as a function of N . The plain lines are the predictions (76). The dotted lines are the predictions (76) with, for both quantities, the same subleading terms added in the denominators. (The scale on the N axis is proportional to $\ln \ln N$.)

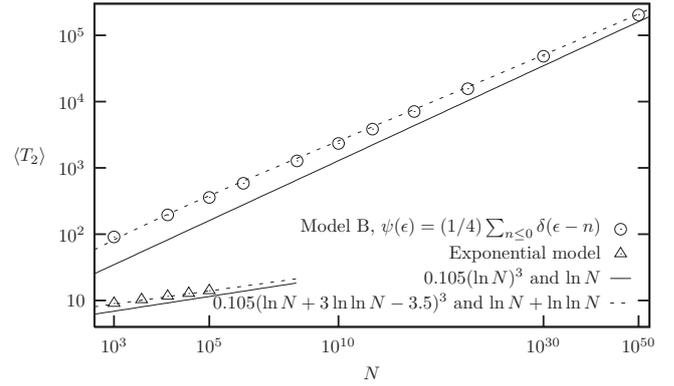


FIG. 5. Numerical simulations of $\langle T_2 \rangle$ for model B with $\psi(\epsilon) = \frac{1}{4} \sum_{n=0}^{\epsilon} \delta(\epsilon-n)$ (circles) and for the exponential model (triangles). The plain lines are the predictions (65) for $\langle T_2 \rangle$ and (94), while the dotted lines are the same predictions with some subleading term: for the generic case, we used subleading terms suggested by (95) and the fit of Fig. 4, and for the exponential model the exact results (A15).

actually an average over all the possible genealogical trees in the population given that realization of the positions over time of the particles. A complexity in time of order $\ln^4 N$ allows already to simulate rather large systems. However, a further optimization is possible in the special case where ϕ_q is constant for $q \leq q_0$. For that specific model, additional simplifications occur (one can write a recursion on the matrix elements) and the matrix $\bar{T}_g(x, x')$ can be updated in a time of only $\ln^2 N$. This allows one to study systems of size N up to about 10^{50} in a few weeks time on standard desktop computers. There is, unfortunately, not enough information in the matrix $T_g(x, x')$ to extract the average coalescence time of three (or more) individuals: to that purpose, one needs to simulate a tensor with three (or more) indices which can be updated with rules very similar to (97). Because of this extra complexity, we only measured the average coalescence time of three individuals for values of N up to 10^{20} .

B. Results

Using this last algorithm, we have simulated model B for $\psi(\epsilon) = \frac{1}{4} \sum_{n=0}^{\epsilon} \delta(\epsilon-n)$ up to $N=10^{50}$. The velocity and diffusion constants are shown in Fig. 4, compared to the predictions (76) in plain lines. There is still a small visible difference between numerics and theory, but this difference gets smaller as N increases. In order to obtain a better fit, we have included subleading corrections by changing the denominator $(\ln N + 3 \ln \ln N)^2$ for the velocity in (76) into $(\ln N + 3 \ln \ln N - 3.5)^2$. Similarly, we changed the denominator $(\ln N)^3$ for the diffusion constant in (76) into $(\ln N + 3 \ln \ln N - 3.5)^3$. With these subleading terms (in dotted lines on the figure), the fit is almost perfect over more than 40 orders of magnitude.

We have no theory to justify these extra subleading terms, but we simply notice that it is possible to fit both the correction to the velocity and the diffusion constant using the same

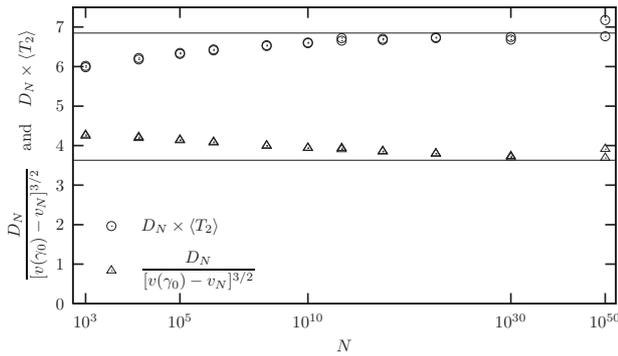


FIG. 6. Numerical simulations of model B with $\psi(\epsilon) = \frac{1}{4} \sum_{n=0} \delta(\epsilon-n)$. The circles represent the product $D_N \times \langle T_2 \rangle$ compared to the prediction (95). The triangles are the ratio of the diffusion constant and the correction to the velocity to the power 3/2, compared to $\pi \sqrt{8/v''(\gamma_0)}/(3\gamma_0^2)$, which is the prediction obtained from (76). The predictions (95) and (76) are represented by the horizontal lines.

subleading terms in the denominators of their respective expressions.

For the same model, $\langle T_2 \rangle$ is shown in Fig. 5 (using circles), compared to the prediction (94) in plain lines. As for the velocity and diffusion constant, there is still a small visible difference and we obtain a better fit if we include subleading terms (in dotted lines): guided by (95) and the fit used for the diffusion constant in Fig. 4, we changed the numerator of (94) from $(\ln N)^3$ into $(\ln N + 3 \ln \ln N - 3.5)^3$. On the same figure, $\langle T_2 \rangle$ for the exponential model is shown (using triangles), compared with the exact prediction (65) $\langle T_2 \rangle \approx \ln N$. Here again, the fit is improved by including the subleading corrections (A15) $\langle T_2 \rangle \approx \ln N + \ln \ln N$ obtained in Appendix A.

Figure 6 combines data from Figs. 4 and 5. The triangles are the ratio of the diffusion constant and of the correction to the velocity to the power 3/2. For large N , this should converge to a constant which we can compute from (76). The

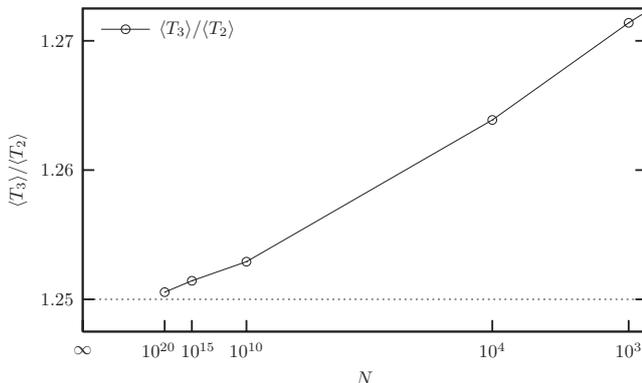


FIG. 7. Numerical simulations of model B with $\psi(\epsilon) = \frac{1}{4} \sum_{n=0} \delta(\epsilon-n)$. The circles represent the ratio $\langle T_3 \rangle / \langle T_2 \rangle$ as a function of N , compared to the result 5/4 suggested by the phenomenological theory of Sec. IV C. (The scale on the N axis is proportional to $1/\ln N$.)

circles are the product of the diffusion constant and of the coalescence time $\langle T_2 \rangle$, which we expect to converge to the value given in (95). The horizontal lines on the figure represent both predictions.

Finally, Fig. 7 shows the ratio $\langle T_3 \rangle / \langle T_2 \rangle$ as a function of N up to $N=10^{20}$. The ratio is very close to 1.25 for large N , which is the prediction of the phenomenological theory of Sec. IV C [see also (65)].

VI. CONCLUSION

In the present work, we have solved exactly a simple model of evolution with selection, the exponential model of Sec. III. For this model, we have calculated the velocity and the diffusion constant (39) of the parameter representing the adequacy of the population to its environment, as well as the coalescence times (64),(65) which characterize the genealogy. We have shown that the statistical properties of the genealogical trees are identical to those trees which appear in the Parisi mean-field theory of spin glasses [45,46]. They, therefore, follow the Bolthausen-Sznitman statistics [40,47], in contrast to the case of evolution without selection which obeys the statistics of the Kingman coalescent.

The reason why the exponential model is exactly soluble is that, going from one generation to the next, the only relevant information on the position of the individuals is contained in one single variable X_g defined in (21). The exponential model belongs to a larger class of models parametrized by a single function ρ (for model A) or ψ (for model B). We have not been able to solve the generic case and, unfortunately, the exponential model is special: while the generic case can be described by a Fisher-KPP front, with a velocity which converges when $N \rightarrow \infty$, the velocity of the front associated to the exponential model diverges when $N \rightarrow \infty$. We have however constructed a phenomenological picture (Sec. IV) of front propagation which can be used both for the exponential model and for the generic Fisher-KPP case, and which also provides predictions for the genealogy. Within this picture, the average coalescence times scale like $\ln^3 N$ with the size N of the population for the generic Fisher-KPP case (while it grows like $\ln N$ for the exponential model), and the structure of the trees is the same as in the Parisi mean-field theory of spin glasses.

Proving the validity of the phenomenological picture for generic models is an interesting open question for future research. Understanding more deeply why our models of selective evolution are related to spin glasses would also deserve some efforts. Last, it would be interesting to study genealogies in other models of selective evolution [2] to test the robustness of our results.

ACKNOWLEDGMENT

This work was partially supported by the Department of Energy.

APPENDIX A: EXACT RESULTS FOR THE EXPONENTIAL MODEL INCLUDING SUBLEADING ORDERS

In this appendix, we obtain higher orders in the large $\ln N$ expansion, for the statistics of the position of the front and

for the coalescence probabilities in the exponential model.

1. Front position statistics

The exact expression for the cumulants of the front velocity was given in (30) in terms of the function I_0 defined in (31). Discarding all the terms of order $1/N$ or smaller, one can use directly the expression (34) of $[I_0(\lambda)]^N$ as a function of the rescaled variable μ in (30). Keeping terms up to the order $1/\ln^2 N$, one gets, using also (37),

$$e^{G(\beta)} = \frac{1}{\ln^\beta N} \frac{1}{\Gamma(\beta)} \int_0^\infty d\mu \mu^{\beta-1} e^{-\mu} \times \left[1 + \mu \frac{\ln \mu - \ln \ln N + \gamma_E - 1}{\ln N} + \frac{1}{2} \left(\mu \frac{\ln \mu - \ln \ln N + \gamma_E - 1}{\ln N} \right)^2 + \dots \right]. \quad (\text{A1})$$

The integrals of each term can be computed using (35). One gets

$$e^{G(\beta)} = \frac{1}{\ln^\beta N} \left[1 + \frac{\beta}{\ln N} \left(\frac{\Gamma'(\beta+1)}{\Gamma(\beta+1)} - l \right) + \frac{\beta(\beta+1)}{2 \ln^2 N} \left(\frac{\Gamma''(\beta+2)}{\Gamma(\beta+2)} - 2 \frac{\Gamma'(\beta+2)}{\Gamma(\beta+2)} l + l^2 \right) + \dots \right] \quad (\text{A2})$$

with $l = \ln \ln N - \gamma_E + 1$. Taking the logarithm of (A2), one obtains $G(\beta)$. By expanding in powers of β and comparing with (27), one gets the cumulants of the position of the front. We give the velocity and diffusion constant,

$$v_N = \ln \ln N + \frac{\ln \ln N + 1}{\ln N} - \frac{(\ln \ln N)^2 - 1 + \frac{\pi^2}{6}}{2 \ln^2 N} + \dots, \quad D_N = \frac{\pi^2}{3} \frac{1}{\ln N} - \frac{1}{\ln^2 N} \left(\frac{\pi^2}{3} \ln \ln N - \frac{\pi^2}{6} + 2\zeta(3) \right) + \dots \quad (\text{A3})$$

Note that the first correction to the leading term can be in both cases obtained by replacing in the leading term $\ln N$ by $\ln N + \ln \ln N$: $v_N \simeq \ln(\ln N + \ln \ln N)$ and $D_N \simeq (\pi^2/3)/(\ln N + \ln \ln N)$. This is reminiscent of the observation in Fig. 4 that, in the generic case, the fit was better by replacing the $\ln N$ by $\ln N + 3 \ln \ln N$ in the theoretical prediction for the diffusion constant.

2. Tree statistics

To get subleading orders for the statistics of the tree in the exponential case, one needs to generalize the discussion in Sec. III B where we derived the leading term in the large $\ln N$ expansion. The central quantity is still the probability $r_p(k)$ that p individuals at generation $g+1$ have exactly k ancestors in the previous generation. But while at leading order it was enough to consider one coalescence at each step,

one needs to take into account up to n simultaneous coalescences when one wishes to keep terms of arbitrary order $1/\ln^n N$.

One must assign an ancestor at generation g to each individual at generation $g+1$. We start from the probability $W_i(x)$ given in (40) that the parent of an individual at position x and generation $g+1$ was the i th individual of generation g . In the exponential model, $W_i(x)$ does not depend on x [see (41)]. We consider p individuals of generation $g+1$ and we note p_i , the number of these individuals that are descendants of the i th individual of generation g . The probability distribution of the p_i is

$$\text{Prob}(p_1, \dots, p_N) = \frac{p!}{p_1! \dots p_N!} \delta_{p_1 + \dots + p_N}^p W_1^{p_1} \dots W_N^{p_N}. \quad (\text{A4})$$

One now averages over the positions of individuals at generation g , and $r_p(k)$ is simply the probability that there are exactly k nonzero p_i 's. After relabeling the individuals at generation g , one gets

$$r_p(k) = \binom{N}{k} \sum_{p_1 \geq 1, \dots, p_k \geq 1} \frac{p!}{p_1! \dots p_k!} \delta_{p_1 + \dots + p_k}^p \langle W_1^{p_1} \dots W_k^{p_k} \rangle. \quad (\text{A5})$$

It is actually convenient to call n the number of p_i that are strictly larger than 1 and to write $r_p(k)$ as a sum over n : after another relabeling,

$$r_p(k) = \binom{N}{k} \sum_{n \geq 0} \binom{k}{n} \sum_{p_1 \geq 2, \dots, p_n \geq 2} \frac{p!}{p_1! \dots p_n!} \delta_{p_1 + \dots + p_n}^{p-k+n} \langle W_1^{p_1} \dots W_n^{p_n} W_{n+1} \dots W_k \rangle. \quad (\text{A6})$$

Indeed, as we shall see, each term in the sum over n gives a contribution of order $1/\ln^n N$ in the final result. The averaged term can be expressed using the probability W_i given in (41),

$$J_{p,k,n}^{p_1, \dots, p_n} = \langle W_1^{p_1} \dots W_n^{p_n} W_{n+1} \dots W_k \rangle = \int_0^\infty dy_1 e^{-y_1} \dots \int_0^\infty dy_n e^{-y_n} \frac{e^{p_1 y_1 + \dots + p_n y_n + y_{n+1} + \dots + y_k}}{(e^{y_1} + \dots + e^{y_n})^p}. \quad (\text{A7})$$

The technique to evaluate the integrals involved here is essentially the same as in Sec. III. We first use the standard representation (29) for the denominator in the integrand. Then the integral over y_i may be expressed with the help of the functions $I_p(\lambda)$ defined in (45):

$$J_{p,k,n}^{p_1, \dots, p_n} = \frac{1}{(p-1)!} \int_0^{+\infty} d\lambda \lambda^{p-1} I_{p_1}(\lambda) \dots \times I_{p_n}(\lambda) I_1(\lambda)^{k-n} I_0(\lambda)^{N-k}. \quad (\text{A8})$$

As before, for large N , the term $I_0(\lambda)^N$ makes the integral (A8) dominated by values of λ of order $1/(N \ln N)$. It is sufficient to use the leading order (46) for the $I_p(\lambda)$ as next

orders in λ would generate terms of order $1/N$, which we discard throughout. Making the change of variables μ

$=\lambda N \ln N$ [see (33)], and using the fact that $p_1 + \dots + p_n = p - k + n$, one gets for the integrand of (A8),

$$\lambda^{p-1} I_{p_1} \dots I_{p_n} I_1^{k-n} I_0^{N-k} \simeq \frac{(p_1-2)! \dots (p_n-2)!}{N^{k-1} \ln^{n-1} N} \mu^{k-1} e^{-\mu} \left(1 + \frac{(k-n-\mu)(\ln \ln N - \ln \mu - \gamma_E) - \mu}{\ln N} + \dots \right), \quad (\text{A9})$$

(A8) can then be evaluated using (35). One gets

$$J_{p,k,n}^{p_1, \dots, p_n} = \frac{(p_1-2)! \dots (p_n-2)! (k-1)!}{(p-1)! N^k \ln^n N} \left(1 + \frac{n \left(\frac{\Gamma'(k)}{\Gamma(k)} + \gamma_E - \ln \ln N \right) - (k-1)}{\ln N} + \dots \right) \quad (\text{A10})$$

as expected, $j_{p,k,n}$ has an amplitude proportional to $1/\ln^n N$. To compute $r_p(k)$ for $k < p$ to order $1/\ln^2 N$, one only needs in (A6) the terms $n=1$ and $n=2$ (the term $n=0$ gives a contribution only for $k=p$),

$$r_p(k) \simeq \frac{N^k}{k!} \left(k \frac{p!}{(p-k+1)!} J_{p,k,1}^{p-k+1} + \frac{k(k-1)}{2} \sum_{p_1=2}^{p-k} \frac{p!}{p_1!(p-k+2-p_1)!} J_{p,k,2}^{p_1, p-k+2-p_1} + \dots \right). \quad (\text{A11})$$

After some algebra, one gets, for $k < p$,

$$r_p(k) = \frac{p}{(p-k+1)(p-k) \ln N} \left\{ 1 + \frac{1}{\ln N} \left[\sum_{n=1}^{k-1} \frac{1}{n} + \frac{2(k-1)}{p-k+2} \left(\sum_{n=1}^{p-k-1} \frac{1}{n} - \frac{3}{2} \right) - \ln \ln N \right] + \dots \right\} \quad (\text{A12})$$

[we used, among other things, $\Gamma'(k)/\Gamma(k) + \gamma_E = 1 + \frac{1}{2} + \dots + \frac{1}{k-1}$].

We can now compute the $\langle T_k \rangle$. From the recurrence

$$\langle T_p \rangle = 1 + \sum_{k=1}^p r_p(k) \langle T_k \rangle, \quad (\text{A13})$$

we get, using $\sum_k r_p(k) = 1$ and $\langle T_1 \rangle = 0$,

$$\langle T_p \rangle = \frac{1 + \sum_{k=2}^{p-1} r_p(k) \langle T_k \rangle}{\sum_{k=1}^{p-1} r_p(k)}. \quad (\text{A14})$$

For the first values of p , we obtain

$$\langle T_2 \rangle = \ln N + \ln \ln N + o(1),$$

$$\langle T_3 \rangle = \frac{5}{4} (\ln N + \ln \ln N) + o(1),$$

$$\langle T_4 \rangle = \frac{25}{18} (\ln N + \ln \ln N) - \frac{1}{54} + o(1). \quad (\text{A15})$$

APPENDIX B: THE PARISI BROKEN REPLICA SYMMETRY

The replica trick is a powerful approach to calculate the typical free energy of a sample in the theory of disordered systems. In the replica trick, one considers n replicas of the same random sample, one averages the product of their partition functions, and at the end of the calculation one takes the limit $n \rightarrow 0$. In some cases, the n dependence of this averaged product is simple enough for the analytic continuation $n \rightarrow 0$ to be unique leading to the desired free energy.

In the case of mean-field spin glasses, the situation is more complicated: the symmetry between the replicas gets broken as n takes noninteger values ($n < 1$) and remains broken in the limit $n \rightarrow 0$. In this appendix we recall the statistical properties of the trees predicted by the Parisi theory of the broken replica symmetry [42,45,46,48].

One starts with an integer $n = n_0$ number of replicas. These replicas are grouped into n_0/n_1 groups of n_1 replicas. Each of these groups of n_1 replicas is decomposed into n_1/n_2 groups of n_2 replicas and so on: each group of n_i replicas is formed of n_i/n_{i+1} groups of n_{i+1} replicas each. When this hierarchy consists of k levels, it is characterized by $k+1$ integers

$$n = n_0 > n_1 > n_2 > \dots > n_k = 1. \quad (\text{B1})$$

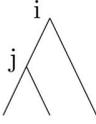
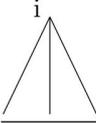
At level i , there are a total of n/n_i groups of size n_i . Therefore, the probability that m distinct individuals chosen at random belong to the same group at level i (without specifying whether they belong or not to the same group at level $i+1$) is

$$Q_m = \frac{\frac{n}{n_i} \binom{n_i}{m}}{\binom{n}{m}} = \frac{n(n_i-1)(n_i-2) \dots (n_i-m+1)}{n(n-1) \dots (n-m+1)}. \quad (\text{B2})$$

One can also associate a tree to each choice of m replicas: the m replicas are at the bottom of the tree and when two replicas belong to the same group at level i , but to different

BRUNET *et al.*PHYSICAL REVIEW E **76**, 041104 (2007)

TABLE II. All possible trees of three replicas, their probabilities and degeneracies.

	$\frac{n(n_i - n_{i+1})(n_j - n_{j+1})}{n(n-1)(n-2)}$	three cases
	$\frac{n(n_i - n_{i+1})(n_i - 2n_{i+1})}{n(n-1)(n-2)}$	one case

groups at level $i+1$, their branches merge at level i .

The various possible trees which might occur for three replicas or four replicas are shown in Tables II and III with their probabilities. For example, for the first tree of Table II, the probability that two branches merge at level j and the remaining branches merge at level i is

$$\frac{n(n_i - n_{i+1})(n_j - n_{j+1})}{n(n-1)(n-2)}, \quad (\text{B3})$$

as there are n possible choices for the leftmost replica, $n_i - n_{i+1}$ choices for the rightmost replica, and $n_j - n_{j+1}$ choices for the replica at the center of the figure. The degeneracy factor is simply the number of different ways of permuting the roles of the replicas at the bottom of the tree.

In the Parisi ansatz, all the calculations are done as if all the n_i 's and all the ratios n_i/n_{i+1} were integers. At the end of the calculation, however, one takes the limit $n \rightarrow 0$ and one reverses the inequality (B1) into

$$n = n_0 < n_1 < n_2 < \dots < n_k = 1. \quad (\text{B4})$$

One then takes a continuous limit ($k \rightarrow \infty$), where n_i becomes a continuous variable x ,

$$n_i = x. \quad (\text{B5})$$

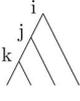
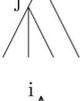
In the spin-glass theory [45,46], there is an ultrametric distance between pairs of replicas, related to the overlap $q_{\alpha,\beta}$. (The distance is a decreasing function of the overlap.) This overlap $q_{\alpha,\beta}$ depends on the level at which the branches of these two replicas merge: this means that at each level i of the hierarchy, one associates a value q_i of the overlap and that $q_{\alpha,\beta} = q_i$ if the two replicas α and β belong to the same group at level i and to different groups at level $i+1$. (q_i is an increasing function of i with $q_0=0$ and $q_k=1$.) In the limit $k \rightarrow \infty$, when n_i becomes a continuous variable (B5), the overlap q_i becomes an increasing function $q(x) = q(n_i) = q_i$ with $q(0)=0$ and $q(1)=1$.

The probability that two replicas have an overlap $q_{\alpha,\beta} < q_i$ is

$$\begin{aligned} & \text{Prob}(q_{\alpha,\beta} = q_0) + \text{Prob}(q_{\alpha,\beta} = q_1) + \dots + \text{Prob}(q_{\alpha,\beta} = q_{i-1}) \\ & = 1 - Q_2(n_i) = \frac{n - n_i}{n - 1}. \end{aligned} \quad (\text{B6})$$

Therefore, in the $n \rightarrow 0$ limit, the probability $P(q)$ that the

TABLE III. All possible trees of four replicas, their probabilities and degeneracies.

	$\frac{n(n_i - n_{i+1})(n_j - n_{j+1})(n_k - n_{k+1})}{n(n-1)(n-2)(n-3)}$	twelve cases
	$\frac{n(n_i - n_{i+1})(n_j - n_{j+1})(n_k - n_{k+1})}{n(n-1)(n-2)(n-3)}$	six cases
	$\frac{n(n_i - n_{i+1})(n_i - 2n_{i+1})(n_j - n_{j+1})}{n(n-1)(n-2)(n-3)}$	six cases
	$\frac{n(n_i - n_{i+1})(n_j - n_{j+1})(n_j - 2n_{j+1})}{n(n-1)(n-2)(n-3)}$	four cases
	$\frac{n(n_i - n_{i+1})(n_i - 2n_{i+1})(n_i - 3n_{i+1})}{n(n-1)(n-2)(n-3)}$	one case

overlap $q_{\alpha,\beta}$ between two replicas α and β takes the value q is then given by

$$\int_0^{q(x)} P(q') dq' = \lim_{n \rightarrow 0} (1 - Q_2) = x \quad (\text{B7})$$

and this leads to the famous relation [42] between the function $q(x)$ and the probability distribution of the overlap

$$P(q) = \frac{dx}{dq}. \quad (\text{B8})$$

In our models, the coalescence time between a pair of individuals in the population defines, clearly, an ultrametric distance. In order to see whether the statistics predicted by the replica approach remain valid for the trees of the exponential model discussed in the present paper, one needs to relate the overlap $q(x)$ or the parameter x (which indexes the height of the hierarchy) to the coalescence time T by a function $T(x)$. It turns out that this can be achieved by identifying the probability e^{-T} that the coalescence time between two individuals is larger than T [see $R_2(T)$ in (63)] with the probability that two replicas belong to different groups at level i . In other words,

$$e^{-T} = 1 - Q_2 = \frac{n(n - n_i)}{n(n - 1)}, \quad (\text{B9})$$

which leads in the $n \rightarrow 0$ limit to

$$e^{-T} = x. \quad (\text{B10})$$

With this identification, if one assumes that the statistics of the trees are given by Parisi's theory, one can compute all the statistical properties of the coalescence times of trees. For

example, by taking the $n \rightarrow 0$ limit of (B2), one gets that the probability Q_m that m individuals have a coalescence time $T_m < T$ is given by

$$Q_m \rightarrow \frac{\Gamma(m-x)}{(m-1)!\Gamma(1-x)} \quad (\text{B11})$$

which, by taking the derivative with respect with T , gives

$$\begin{aligned} \langle (T_m)^p \rangle &= \int_0^1 dx T(x)^p \frac{dQ_m}{dx} \\ &= \int_0^\infty dT T^p \frac{d}{dT} \frac{\Gamma(m-e^{-T})}{(m-1)!\Gamma(1-e^{-T})}. \end{aligned} \quad (\text{B12})$$

This coincides with the result of the direct calculation (62) of the moments of the T_m and shows that the statistics of the trees in the exponential model are the same as the ones predicted by the mean-field theory of spin glasses.

APPENDIX C: THE NEUTRAL MODEL

In this appendix we recall some well-known results on the statistical properties of the coalescence times in neutral models [37,41] and derive (66).

We consider a population of fixed size N with nonoverlapping generations. Each individual i at a given generation g has $k_i(g)$ offspring at the next generation. We assume that the $k_i(g)$ are random and independent, and we call p_k the probability that $k_i(g)=k$. The total number M of offspring is therefore given by

$$M = \sum_{i=1}^N k_i. \quad (\text{C1})$$

To keep the size of the population constant we choose N individuals at random among these M individuals.

The probability q_n that n individuals have the same parent at the previous generation is

$$q_n = \left\langle \frac{\sum_i \binom{k_i}{n}}{\binom{M}{n}} \right\rangle = \left\langle \frac{\sum_i k_i(k_i-1)\cdots(k_i-n+1)}{M(M-1)\cdots(M-n+1)} \right\rangle. \quad (\text{C2})$$

For a population of large size, if p_k decays fast enough with k for the moments of k to be finite, the law of large numbers

gives that the denominator is approximately equal to $(N\langle k \rangle)^n$ and

$$q_n \simeq \frac{1}{N^{n-1}\langle k \rangle^n} \left\langle \frac{\Gamma(k+1)}{\Gamma(k-n+1)} \right\rangle. \quad (\text{C3})$$

We see (when the moments of k are finite) that q_2 is much larger than all the other q_n when the size N of the population is large, and therefore in the ancestry of a finite number n of individuals, branches coalesce only by pairs. Similarly, the probability that two or more pairs of individuals coalesce at the same generation is negligible.

Let $T_n(g)$ be the age of the most recent common ancestor of a group of n individuals at generation g . As for large N only coalescences by pairs may occur from one generation to the previous one, one has

$$T_n(g+1) = \begin{cases} T_n(g)+1 & \text{with probability } 1 - \frac{1}{2}n(n-1)q_2, \\ T_{n-1}(g)+1 & \text{with probability } \frac{1}{2}n(n-1)q_2. \end{cases} \quad (\text{C4})$$

In the steady state [49], this implies that

$$\langle T_n^p \rangle = \left(1 - \frac{n(n-1)}{2}q_2 \right) \langle (1+T_n)^p \rangle + \frac{n(n-1)}{2}q_2 \langle (1+T_{n-1})^p \rangle \quad (\text{C5})$$

and using the fact that $T_1(g)=0$, one gets

$$\langle T_n \rangle = \left(2 - \frac{2}{n} \right) \frac{1}{q_2}. \quad (\text{C6})$$

We see that all the times T_n scale like N (since $q_2 \sim N^{-1}$) and that

$$\frac{\langle T_3 \rangle}{\langle T_2 \rangle} = \frac{4}{3}, \quad \frac{\langle T_4 \rangle}{\langle T_2 \rangle} = \frac{3}{2}, \quad \dots, \quad \frac{\langle T_n \rangle}{\langle T_2 \rangle} = \frac{2(n-1)}{n}. \quad (\text{C7})$$

One can also calculate from (C5) higher moments of the T_n 's or their generating functions

$$\frac{\langle (T_2)^2 \rangle}{\langle T_2 \rangle^2} = 2, \quad \frac{\langle (T_3)^2 \rangle}{\langle T_3 \rangle^2} = \frac{13}{8}. \quad (\text{C8})$$

These distributions of the T_n as well as their correlations are universal (in the sense that they do not depend on the details of the distribution of the p_k 's).

- [1] L. Peliti, e-print arXiv:cond-mat/9712027.
 [2] D. A. Kessler, H. Levine, D. Ridgway, and L. Tsimring, J. Stat. Phys. **87**, 519 (1997).
 [3] L. S. Tsimring, H. Levine, and D. A. Kessler, Phys. Rev. Lett. **76**, 4440 (1996).
 [4] P. Bak and K. Sneppen, Phys. Rev. Lett. **71**, 4083 (1993).
 [5] M. Kloster and C. Tang, Phys. Rev. Lett. **92**, 038101 (2004).

- [6] M. Kloster, Phys. Rev. Lett. **95**, 168701 (2005).
 [7] R. E. Snyder, Ecology **84**, 1333 (2003).
 [8] É. Brunet, B. Derrida, A. H. Mueller, and S. Munier, Phys. Rev. E **73**, 056126 (2006).
 [9] É. Brunet, B. Derrida, A. H. Mueller, and S. Munier, Europhys. Lett. **76**, 1 (2006).
 [10] R. A. Fisher, Ann. Eugen. **7**, 355 (1937).

BRUNET *et al.*PHYSICAL REVIEW E **76**, 041104 (2007)

- [11] A. Kolmogorov, I. Petrovsky, and N. Piscounov, *Moscow Univ. Math. Bull.* **1**, 1 (1937).
- [12] W. van Saarloos, *Phys. Rep.* **386**, 29 (2003).
- [13] B. Derrida and H. Spohn, *J. Stat. Phys.* **51**, 817 (1988).
- [14] É. Brunet and B. Derrida, *Phys. Rev. E* **70**, 016106 (2004).
- [15] H.-P. Breuer, W. Huber, and F. Petruccione, *Europhys. Lett.* **30**, 69 (1995).
- [16] C. R. Doering, C. Mueller, and P. Smereka, *Physica A* **325**, 243 (2003).
- [17] A. Lemarchand and B. Nowakowski, *J. Chem. Phys.* **111**, 6190 (1999).
- [18] E. Moro, *Phys. Rev. Lett.* **87**, 238303 (2001).
- [19] P. L. Krapivsky and S. N. Majumdar, *Phys. Rev. Lett.* **85**, 5492 (2000).
- [20] E. Iancu, A. H. Mueller, and S. Munier, *Phys. Lett. B* **606**, 342 (2005).
- [21] S. Munier and R. Peschanski, *Phys. Rev. Lett.* **91**, 232001 (2003).
- [22] A. H. Mueller and A. I. Shoshi, *Nucl. Phys. B* **692**, 175 (2004).
- [23] C. Mueller and R. B. Sowers, *J. Funct. Anal.* **128**, 439 (1995).
- [24] D. Panja, *Phys. Rep.* **393**, 87 (2004).
- [25] J. G. Conlon and C. R. Doering, *J. Stat. Phys.* **120**, 421 (2005).
- [26] C. Escudero, *Phys. Rev. E* **70**, 041102 (2004).
- [27] E. Moro, *Phys. Rev. E* **69**, 060101(R) (2004).
- [28] J. Mai, I. M. Sokolov, and A. Blumen, *Phys. Rev. Lett.* **77**, 4462 (1996).
- [29] D. A. Kessler, Z. Ner, and L. M. Sander, *Phys. Rev. E* **58**, 107 (1998).
- [30] E. Moro, *Phys. Rev. E* **70**, 045102(R) (2004).
- [31] L. Pechenik and H. Levine, *Phys. Rev. E* **59**, 3893 (1999).
- [32] A. Lemarchand, A. Lesne, and M. Mareschal, *Phys. Rev. E* **51**, 4457 (1995).
- [33] H. P. McKean, *Commun. Pure Appl. Math.* **28**, 323 (1975).
- [34] M. D. Bramson, *Commun. Pure Appl. Math.* **31**, 531 (1978).
- [35] C. Bender and S. Orszag, *Advanced Mathematical Methods for Scientists and Engineers* (McGraw-Hill, New York, 1978).
- [36] J. F. C. Kingman, *Stochastic Proc. Appl.* **13**, 235 (1982).
- [37] J. F. C. Kingman, *J. Appl. Probab.* **19A**, 27 (1982).
- [38] J. Pitman, *Ann. Probab.* **27**, 1870 (1999).
- [39] J. Schweinsberg, *Electron. J. Probab.* **5**, 1 (2000).
- [40] E. Bolthausen and A.-S. Sznitman, *Commun. Math. Phys.* **197**, 247 (1998).
- [41] S. Tavaré, D. J. Balding, R. C. Griffiths, and P. Donnelly, *Genetics* **145**, 505 (1997).
- [42] G. Parisi, *Phys. Rev. Lett.* **50**, 1946 (1983).
- [43] É. Brunet and B. Derrida, *Phys. Rev. E* **56**, 2597 (1997).
- [44] É. Brunet and B. Derrida, *J. Stat. Phys.* **103**, 269 (2001).
- [45] M. Mézard, G. Parisi, N. Sourlas, G. Toulouse, and M. A. Virasoro, *J. Phys. (Paris)* **45**, 843 (1984).
- [46] M. Mézard, G. Parisi, and M. A. Virasoro, *Spin Glass Theory and Beyond* (World Scientific, Singapore, 1987).
- [47] D. Ruelle, *Commun. Math. Phys.* **108**, 225 (1987).
- [48] G. Parisi, *J. Phys. A* **13**, 1101 (1980).
- [49] D. Simon and B. Derrida, *J. Stat. Mech.: Theory Exp.* 2006 P05002.

Statistics at the tip of a branching random walk and the delay of traveling waves

É. BRUNET^(a) and B. DERRIDA^(b)

*Laboratoire de Physique Statistique, École Normale Supérieure, UPMC Paris 6, Université Paris Diderot, CNRS
 24 rue Lhomond, 75005 Paris, France, EU*

received 24 July 2009; accepted in final form 16 September 2009

published online 12 October 2009

PACS 02.50.-r – Probability theory, stochastic processes, and statistics

PACS 05.40.-a – Fluctuation phenomena, random processes, noise, and Brownian motion

PACS 89.75.Hc – Networks and genealogical trees

Abstract – We study the limiting distribution of particles at the frontier of a branching random walk. The positions of these particles can be viewed as the lowest energies of a directed polymer in a random medium in the mean-field case. We show that the average distances between these leading particles can be computed as the delay of a traveling wave evolving according to the Fisher-KPP front equation. These average distances exhibit universal behaviors, different from those of the probability cascades studied recently in the context of mean-field spin-glasses.

Copyright © EPLA, 2009

The interest for branching random walks has a long history in mathematics [1–3], physics and biology. In biology they are commonly used to model the genealogies of evolving populations, the spread of an advantageous gene or of an infection, the combined effects of selection and mutations [4–6]. In Physics they also appear in many contexts such as reaction-diffusion models [7,8], particle physics [9,10], or the theory of disordered systems [11,12].

In one dimension, the right frontier of a branching random walk is the region located near its rightmost particle. An interesting question is what does the branching random walk look like when seen from this frontier. For example one can try to determine the position of the second, the third, ... or the n -th rightmost particle in the frame of the first rightmost particle. The statistical properties of these positions depend on time and have a well-defined long-time limit [3] which we study in this letter using traveling-wave equations of the Fisher-KPP type [4,13,14]

$$\frac{\partial h}{\partial t} = \frac{\partial^2 h}{\partial x^2} + h - h^2. \quad (1)$$

The fluctuating distances between these rightmost particles allows one to understand why directed polymers in a random medium [11] have non-self-averaging properties similar to mean-field spin-glasses [15]. Their study is

also motivated by the growing interest for the statistics of extreme events [12,16–22] which dominate a number of physical processes [23,24]. The last two decades have seen the emergence of universal statistical properties of the probability cascades describing the energies of the low lying states of several spin-glass models [22,25–30]. Somewhat surprisingly, as shown below, the distribution of the distances between the extreme positions of particles in a branching random walk (which are nothing but the energies of the low lying states in the mean-field version of directed polymer problem [11]) is different from the predictions of the probability cascades [22,25–30].

To start with a simple case, we consider a continuous time branching Brownian motion in one dimension. At time $t=0$ there is a single particle at the origin $x=0$; this particle diffuses (for convenience we normalize the variance of its displacement during time t to be $2t$) and branches at rate 1. (This means that during every infinitesimal time interval dt , the displacement of the particle is a random variable η such that $\langle \eta \rangle = 0$ and $\langle \eta^2 \rangle = 2dt$, and that there is a probability dt that the particle splits into two particles.) Whenever a branching event occurs, the offspring become themselves independent branching Brownian motions which diffuse and branch with the same rates.

The number of particles in the system grows exponentially with time and they occupy a region which grows linearly with time [1,2]. It has been known for a

^(a)E-mail: Eric.Brunet@lps.ens.fr

^(b)E-mail: Bernard.Derrida@lps.ens.fr

long time [1,2,12] that the probability distribution of the rightmost particle of a branching Brownian motion can be determined by solving a traveling-wave equation: the probability $Q_0(x, t)$ that, at time t , there is no particle at the right of x , satisfies

$$\frac{\partial Q_0}{\partial t} = \frac{\partial^2 Q_0}{\partial x^2} + Q_0^2 - Q_0. \quad (2)$$

(The derivation of (2) is standard: one decomposes the time interval $(0, t+dt)$ into two intervals $(0, dt)$ and $(dt, t+dt)$, and write that $Q_0(x, t+dt) = Q_0(x, t)^2 dt + \langle Q_0(x-\eta, t) \rangle_\eta (1-dt)$ where the first term represents the contribution of a branching event and η in the second term the displacement due to diffusion during the first time interval $(0, dt)$. With our normalization $\langle \eta^2 \rangle = 2dt$.)

Up to the change $h = 1 - Q_0$, eq. (2) is the Fisher-KPP equation (1). Since at time $t = 0$ there is a single particle at the origin, the initial condition is simply

$$Q_0(x, 0) = 1 \quad \text{for } x > 0, \quad Q_0(x, 0) = 0 \quad \text{for } x < 0. \quad (3)$$

If $Q_n(x, t)$ is the probability that there are exactly n particles on the right of x , one can see, as for Q_0 , that the generating function $\psi_\lambda(x, t)$, defined as

$$\psi_\lambda(x, t) = \sum_{n \geq 0} \lambda^n Q_n(x, t), \quad (4)$$

evolves according to the same eq. (2), the only difference being that the initial condition is replaced by

$$\psi_\lambda(x, 0) = 1 \quad \text{for } x > 0, \quad \psi_\lambda(x, 0) = \lambda \quad \text{for } x < 0. \quad (5)$$

We are now going to see that the knowledge of $\psi_\lambda(x, t)$ allows one to obtain the average distances between the rightmost particles of the system. If $p_n(x, t)$ is the probability of finding the n -th rightmost particle at position x , it is easy to see that

$$\frac{\partial Q_0}{\partial x} = p_1(x, t) \quad \text{and} \quad \frac{\partial Q_n}{\partial x} = p_{n+1}(x, t) - p_n(x, t). \quad (6)$$

The average position $\langle X_n(t) \rangle$ of the n -th rightmost particle and the average distance $\langle d_{n,n+1}(t) \rangle$ between the n -th and $(n+1)$ -th rightmost particles can then be defined by

$$\langle X_n(t) \rangle = \int x p_n(x, t) dx, \quad (7)$$

$$\langle d_{n,n+1}(t) \rangle = \langle X_n(t) \rangle - \langle X_{n+1}(t) \rangle. \quad (8)$$

(One should notice that the normalization of $p_n(x, t)$ is not 1 but $\int p_n(x, t) dx = (1 - e^{-t})^{n-1}$ due to the events for which the total number of particles at time t is still less than n . One could prefer to use different definitions of the positions or of the distances, for example by conditioning on the fact that there are at least $n+1$ particles in the system, but any such definition would coincide with (7), (8) up to contributions which decay

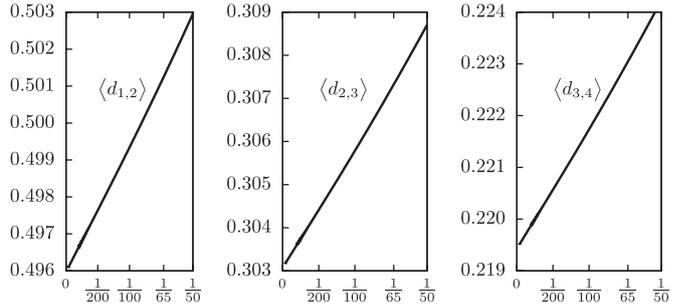


Fig. 1: The average distances between the first rightmost particles $\langle d_{1,2}(t) \rangle$, $\langle d_{2,3}(t) \rangle$ and $\langle d_{3,4}(t) \rangle$ of a branching random walk vs. $1/t$, for t up to 3000.

exponentially with time and disappear in the long-time limit that we study below.)

With the definition (8) we obtain from (4), (6), (7) that

$$\sum_{n \geq 1} \lambda^n \langle d_{n,n+1}(t) \rangle = \int x \left[\frac{\partial Q_0}{\partial x} - \frac{\partial \psi_\lambda}{\partial x} \right] dx, \quad (9)$$

which relates the distances $\langle d_{n,n+1}(t) \rangle$ between the rightmost particles to the solution $\psi_\lambda(t)$ of the partial differential equation (2) with the initial condition (5).

We have integrated numerically the equations satisfied by $\psi_\lambda(x, t)$ and its derivatives with respect to λ to measure the distances $\langle d_{n,n+1}(t) \rangle$ between the n -th and $(n+1)$ -th rightmost particles. In our numerical integration, we had to discretize space and time; we checked that our results shown in fig. 1 were stable when we decreased our integration steps.

One can remark that, in contrast to standard Monte Carlo simulations, where all the branching events would be simulated and for which the maximum reachable time would be $t \sim 20$ (with a number of particles e^t of order 10^9), the integration of (2) or of its derivatives allows one to achieve much larger times. One can also notice in fig. 1 that the distances converge like $1/t$ to well-defined values. We will see that this $1/t$ convergence is consistent with our analytic expression (24) below.

We did not find an analytic theory to predict the limiting values that we measured as in fig. 1:

$$\begin{aligned} \langle d_{1,2} \rangle &\simeq 0.496, & \langle d_{2,3} \rangle &\simeq 0.303, & \langle d_{3,4} \rangle &\simeq 0.219, \\ \langle d_{4,5} \rangle &\simeq 0.172, & \langle d_{5,6} \rangle &\simeq 0.142, & \langle d_{6,7} \rangle &\simeq 0.121. \end{aligned} \quad (10)$$

As shown below (17), we can, however, predict their large n behavior.

Before doing so, it is interesting to compare our results (10) to the expected values of the gaps between the low lying energies of spin-glass models such as the REM and the GREM [31,32]. In these models one can show that these energies are given by probability cascades [22,25,27–29] and that the energy gaps at the leading edge are the same as those of a Poisson process on the line with an exponential density. For such a process,

with density $e^{-\alpha x}$, the probability distribution of the positions is $p_n(x) = \exp[-n\alpha x - e^{-\alpha x}/\alpha]/[\alpha^{n-1}(n-1)!]$ from which one gets through (7), (8)

$$\langle d_{n,n+1} \rangle_{\text{GREM}} = \frac{1}{\alpha n}. \quad (11)$$

Clearly, there is no choice of α for which our numerical results (10) are compatible with (11).

It is well known [2,14] that the solution $Q_0(x, t)$ of the Fisher-KPP equation (2) with the step initial condition (3) becomes, for large t , a traveling wave of the form

$$Q_0(x, t) \simeq F[x - \langle X_1(t) \rangle], \quad (12)$$

where the shape $F(z)$ of the front ($F(z) \rightarrow 1$ as $z \rightarrow \infty$ and $F(z) \rightarrow 0$ as $z \rightarrow -\infty$) is time-independent and its position, which can be defined as the average position $\langle X_1(t) \rangle$ of the rightmost particle, has the following long-time behavior [2,14,33]:

$$\langle X_1(t) \rangle = 2t - \frac{3}{2} \ln t + \mathcal{O}(1). \quad (13)$$

$\psi_\lambda(x, t)$ is also the solution of the Fisher-KPP equation (2), but with the initial condition (5). As $\psi_\lambda(x, 0)$ decays fast enough [14], one expects the same large t behavior as (12), (13), up to a λ -dependent delay $f(\lambda)$ due to the change of initial condition:

$$\psi_\lambda(x, t) \simeq F[x - \langle X_1(t) \rangle + f(\lambda)]. \quad (14)$$

From (9), (12), (14) we see that the translation $f(\lambda)$ is nothing but the generating function of the average distances

$$f(\lambda) = \lim_{t \rightarrow \infty} \sum_{n \geq 1} \lambda^n \langle d_{n,n+1}(t) \rangle. \quad (15)$$

We were not able to find an analytic expression of the delay function $f(\lambda)$ for arbitrary λ . For λ close to 1, however, we are going to show that

$$f(\lambda) = \tau_\lambda - \ln \tau_\lambda + \mathcal{O}(1) \quad \text{with } \tau_\lambda = -\ln(1 - \lambda). \quad (16)$$

This implies (15) that the distances have the following large n asymptotics:

$$\langle d_{n,n+1}(t) \rangle_{t \rightarrow \infty} \simeq \frac{1}{n} - \frac{1}{n \ln n} + \dots \quad (17)$$

Compared with (11), we see that there is a correction, which we believe to be universal as discussed below. (Note that the same asymptotic distances would be obtained for uncorrelated particles distributed according to a Poisson point process with a density $-xe^{-x}$ for negative x .)

For λ close to 1, the time τ_λ in (16) is the characteristic time it takes $\psi_\lambda(-\infty, t)$ to reach a value close to 0. The most naive idea to derive (16) would be to say that it takes this time τ_λ for $\psi_\lambda(x, t)$ to look like the step function $Q_0(x, 0)$, and then to start moving like $Q_0(x, t)$. As the asymptotic velocity is 2 (see (13)) this would lead to a

delay $f(\lambda) \simeq 2\tau_\lambda$ which is wrong (see (16)) by a factor 2. The problem with this idea is that, while $\psi_\lambda(x, t)$ evolves to approach 0 on the negative x axis, a tail builds up on the positive x axis which has a strong influence on the dynamics later on.

To derive (16), we need to understand the shape of $\psi_\lambda(x, t)$ for $t > \tau_\lambda$. Let Y_t be the position where $\psi_\lambda(Y_t, t) = 1/2$. We have checked both numerically and analytically that the following picture holds for t and τ_λ large, with a given ratio t/τ_λ larger than 1.

In the range where $x - Y_t$ is of order 1

$$\psi_\lambda(x, t) \simeq \phi_{v(t)}(x - Y_t), \quad (18)$$

where $v(t) = \dot{Y}_t$ is the instantaneous velocity of the front and ϕ_v is the solution of the Fisher-KPP equation moving at a constant velocity v , *i.e.* the solution of

$$\phi_v'' + v\phi_v' + \phi_v^2 - \phi_v = 0, \quad (19)$$

with $\phi_v(-\infty) = 0$ and $\phi_v(+\infty) = 1$. (The same form (18) is used in [34].) For definiteness, we normalize such that $\phi_v(0) = 1/2$. This determines a unique solution which has, if $v > 2$, the following asymptotics for $z \rightarrow +\infty$:

$$1 - \phi_v(z) \simeq B_\gamma e^{-\gamma z} + o(e^{-\gamma z}), \quad (20)$$

where γ is the smallest solution of

$$v = \gamma + \gamma^{-1}. \quad (21)$$

On the other hand, in the range $x - Y_t \gg 1$, $\psi_\lambda(x, t)$ is accurately given by the solution of the equation obtained by linearizing (2) around 1:

$$1 - \psi_\lambda(x, t) \simeq \frac{(1 - \lambda)e^t}{\sqrt{\pi}} \int_{x/\sqrt{4t}}^{\infty} e^{-u^2} du. \quad (22)$$

Then, using the asymptotics of the error function $\int_X^\infty \exp(-u^2) du \simeq \exp(-X^2)/(2X)$ and requiring that (18), (20) and (22) match in the range $1 \ll x - Y_t \ll \sqrt{t}$, one gets that $\gamma(t)$ and Y_t should satisfy

$$B_{\gamma(t)} e^{-\gamma(t)(x - Y_t)} \simeq \frac{(1 - \lambda)e^t \sqrt{t}}{Y_t \sqrt{\pi}} e^{-\frac{Y_t^2}{4t} - \frac{Y_t(x - Y_t)}{2t}}. \quad (23)$$

To match the dependence in $x - Y_t$, we need $Y_t \simeq 2t\gamma(t)$ to first order. Then matching the prefactors leads to

$$Y_t \simeq 2\sqrt{t(t - \tau_\lambda)} - \frac{\ln t}{2\gamma(t)} - \frac{\ln[2\sqrt{\pi}\gamma(t)B_{\gamma(t)}]}{\gamma(t)}, \quad (24)$$

$$\text{with } \gamma(t) \simeq \sqrt{\frac{t - \tau_\lambda}{t}}. \quad (25)$$

Note that the relation (21) is indeed satisfied to leading order as $\dot{Y}_t \simeq \gamma(t) + \gamma(t)^{-1}$. In fig. 2 we see that the agreement between the leading term in (24) and the position obtained by integrating numerically (2) with the initial condition (5) is quite good. One could also see

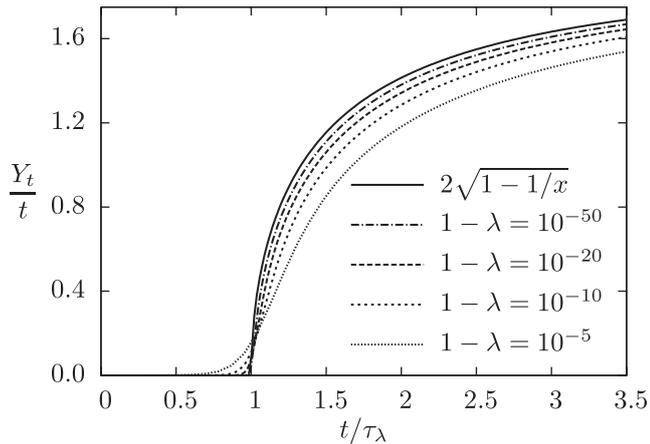


Fig. 2: The prediction (24) to the leading order for the position of the front (full line) is compared to the position measured by integrating (2) with the initial condition (5). As in (16), $\tau_\lambda = -\ln(1-\lambda)$.

in (24) a $1/t$ convergence of the λ -dependent delay¹ which is consistent with the numerical results of fig. 1.

For t/τ_λ large, $\gamma(t) \rightarrow 1$ and $v(t) \rightarrow 2$. For $v=2$, the solution of (19) satisfies $1-\phi_2(z) \simeq Aze^{-z}$ for large z [14,33]. For v slightly larger than 2, the next term in the large z expansion (20) is $1-\phi_v(z) \simeq B_\gamma e^{-\gamma z} + C_\gamma e^{-z/\gamma} + o(e^{-z/\gamma})$. For consistency in the limit $v \rightarrow 2$, one has

$$B_\gamma \simeq -C_\gamma \simeq \frac{A}{2(1-\gamma)} \quad \text{as } \gamma \rightarrow 1, \quad (26)$$

so that $B_{\gamma(t)} \simeq At/\tau_\lambda$ from (25), (26). Thus, (24) becomes $Y_t \simeq 2t - \tau_\lambda - (3/2) \ln t + \ln \tau_\lambda + \mathcal{O}(1)$, which gives (13), (14), (16).

One can repeat everything if, instead of starting with a single particle at the origin, one starts with K particles at positions y_1, \dots, y_K . One simply needs to replace $\psi_\lambda(x, t)$ defined in (4) by $\prod_{1 \leq i \leq K} \psi_\lambda(x - y_i, t)$, with a similar change for $Q_0(x, t) \equiv \psi_0(x, t)$. As a result, in the long-time limit, the delay function $f(\lambda)$, and therefore the distances between the rightmost particles remain unchanged. This property is remarkable: whatever the positions of the initial particles are (as long as there are a finite number of them) the limiting average distances and probably the whole limiting measure seen from the rightmost particle are the same.

We can also extend all our calculations to more general branching random walks. For example one may consider a discrete time case where at each time step, every particle splits into K new particles, and the position of each new particle is shifted from its parent by a random amount ϵ

¹There is another way of understanding this $1/t$ convergence. In the expression (13) of the position of the front, it is known [34] that the bounded term $\mathcal{O}(1)$ can be written as a series in powers of $1/\sqrt{t}$, where all the coefficients depend on the initial condition except the coefficient of $1/\sqrt{t}$, which is universal. As our method is equivalent to measuring the difference in positions of two fronts with different initial conditions, the t , $\ln t$ and $1/\sqrt{t}$ terms in the position cancel out, leaving a constant contribution ($f(\lambda)$), with a $1/t$ correction and a $1/t^{3/2}$ second-order correction.

drawn from a given distribution $\rho(\epsilon)$. Apart from a few changes, such as (21) which is replaced by

$$v = \gamma^{-1} \ln \left[K \int \rho(\epsilon) e^{\gamma \epsilon} d\epsilon \right], \quad (27)$$

τ_λ in (16) which becomes $-\ln(1-\lambda)/\ln K$ or $\gamma(t)$ in (25) which becomes the solution of

$$\gamma^2 \frac{dv}{d\gamma} = -\frac{\ln(1-\lambda)}{t}, \quad (28)$$

everything remains unchanged. In particular (16), (17) are simply divided by the value γ_0 of γ which minimizes the expression (27) of v . Thus, the asymptotics of both the delay (16) and the distances (17) look universal, up to a scale factor γ_0 .

In the present letter we have seen that the distances between the rightmost particles at the frontier of a branching random walk have statistical properties (10), (17) which can be understood as the delay (14), (15), (16) of a traveling wave. Other properties, such as the correlations of these distances or even their whole probability distribution can also be understood in terms of the delay of a traveling wave. For example if $R_{n,m}(x, y, t)$ is the probability that there are n particles at the right of x and m particles at the right of y , one can show that

$$\langle d_{n,n+1}(t) d_{m,m+1}(t) \rangle = \int dx x \int dy y \frac{\partial^2 R_{n,m}(x, y, t)}{\partial x \partial y},$$

while the generating function $\sum_{n,m} \lambda^n \mu^m R_{n,m}(x, x+c, t)$ defined as in (4) evolves according to the Fisher-KPP equation (2) with a new initial condition.

A surprising aspect of the present work is that the statistics of the leading particles, in the long-time limit, do not depend on the positions or on the number of particles we start with, as long as there is a finite number of them. This means that the limiting measure has the following stability property: if one takes two realizations of the leading particles according to this measure and shifts one of them by an arbitrary amount, then the superimposition of these two realizations gives a new realization of the same measure, up to a translation.

REFERENCES

- [1] MCKEAN H. P., *Commun. Pure Appl. Math.*, **28** (1975) 323.
- [2] BRAMSON M. D., *Commun. Pure Appl. Math.*, **31** (1978) 531.
- [3] LALLEY S. P. and SELLEKE T., *Ann. Probab.*, **15** (1987) 1052.
- [4] FISHER R. A., *Ann. Eugen.*, **7** (1937) 355.
- [5] KESSLER D. A., LEVINE H., RIDGWAY D. and TSIMRING L., *J. Stat. Phys.*, **87** (1997) 519.
- [6] GOLDING I., KOZLOVSKY Y., COHEN I. and BEN-JACOB E., *Physica A*, **260** (1998) 510.
- [7] DOERING C. R., MUELLER C. and SMERKA P., *Physica A*, **325** (2003) 243.

-
- [8] DE MASI A., FERRARI P. and LEBOWITZ J., *J. Stat. Phys.*, **44** (1986) 589.
- [9] MUNIER S., *Phys. Rep.*, **473** (2009) 1.
- [10] PESCHANSKI R., *Nucl. Phys. B*, **805** (2008) 377.
- [11] DERRIDA B. and SPOHN H., *J. Stat. Phys.*, **51** (1988) 817.
- [12] DEAN D. S. and MAJUMDAR S. N., *Phys. Rev. E*, **64** (2001) 046121.
- [13] KOLMOGOROV A., PETROVSKY I. and PISCOUNOV N., *Bull. Moskov. Univ. A*, **1** (1937) 1.
- [14] VAN SAARLOOS W., *Phys. Rep.*, **386** (2003) 29.
- [15] MÉZARD M., PARISI G., SOURLAS N., TOULOUSE G. and VIRASORO M. A., *J. Phys. (Paris)*, **45** (1984) 843.
- [16] TRACY C. A. and WIDOM H., *Commun. Math. Phys.*, **159** (1994) 151.
- [17] BOUCHAUD J.-P. and MÉZARD M., *J. Phys. A*, **30** (1997) 7997.
- [18] BURKHARDT T. W., GYÖRGYI G., MOLONEY N. R. and RACZ Z., *Phys. Rev. E*, **76** (2007) 041119.
- [19] GYÖRGYI G., MOLONEY N. R., OZOGANY K. and RACZ Z., *Phys. Rev. Lett.*, **100** (2008) 210601.
- [20] MAJUMDAR S. N. and KRAPIVSKY P. L., *Phys. Rev. E*, **65** (2002) 036127.
- [21] SABHAPANDIT S. and MAJUMDAR S. N., *Phys. Rev. Lett.*, **98** (2007) 140201.
- [22] RUZMAIKINA A. and AIZENMAN M., *Ann. Probab.*, **33** (2005) 82.
- [23] IGLOI F. and MONTHUS C., *Phys. Rep.*, **412** (2005) 277.
- [24] KADANOFF L. P., *J. Stat. Phys.*, **122** (2006) 1293.
- [25] RUELLE D., *Commun. Math. Phys.*, **108** (1987) 225.
- [26] BOLTHAUSEN E. and SZNITMAN A.-S., *Commun. Math. Phys.*, **197** (1998) 247.
- [27] AIZENMAN M., SIMS R. and STARR S. L., *Contemp. Math. Ser.*, **437** (2007) 1.
- [28] ARGUIN L. P., *J. Stat. Phys.*, **126** (2007) 951.
- [29] BOVIER A. and KURKOVA I., *Commun. Math. Phys.*, **263** (2006) 535.
- [30] BOVIER A. and KURKOVA I., *J. Stat. Phys.*, **126** (2007) 933.
- [31] DERRIDA B., *Phys. Rev. B*, **24** (1981) 2613.
- [32] DERRIDA B., *J. Phys. Lett.*, **46** (1985) L401.
- [33] BRUNET É. and DERRIDA B., *Phys. Rev. E*, **56** (1997) 2597.
- [34] EBERT U. and VAN SAARLOOS W., *Phys. Rev. Lett.*, **80** (1998) 1650; *Physica D* **146** (2000) 1.

A Branching Random Walk Seen from the Tip

Éric Brunet · Bernard Derrida

Received: 22 November 2010 / Accepted: 17 March 2011 / Published online: 2 April 2011
© Springer Science+Business Media, LLC 2011

Abstract We show that all the time-dependent statistical properties of the rightmost points of a branching Brownian motion can be extracted from the traveling wave solutions of the Fisher-KPP equation. The distribution of all the distances between the rightmost points has a long time limit which can be understood as the delay of the Fisher-KPP traveling waves when the initial condition is modified. The limiting measure exhibits the surprising property of superposability: the statistical properties of the distances between the rightmost points of the union of two realizations of the branching Brownian motion shifted by arbitrary amounts are the same as those of a single realization. We discuss the extension of our results to more general branching random walks.

Keywords Branching random walk · Branching Brownian motion · Extreme value statistics · Traveling waves

1 Introduction

A branching random walk is a collection of points which, starting from a single point, diffuse and branch independently of the time, of their positions or of the other points, as in Fig. 1.

Branching random walks appear in many contexts ranging from Mathematics [1, 6, 11, 14, 25, 29, 33] to Biology [21, 23, 27]. They can for example be used to describe how a growing population invades a new environment. In the one dimensional case, see Fig. 1, there is, at a given time t , a rightmost individual at position $X_1(t)$, a second rightmost at $X_2(t)$ and so on. (Note that the rightmost $X_1(t')$ at a time $t' > t$ is not necessarily a descendant of the rightmost $X_1(t)$ at time t .) The expected position $m_t = \langle X_1(t) \rangle$ of the rightmost individual as well as the probability distribution of its position $X_1(t)$ around m_t are

É. Brunet (✉) · B. Derrida
Laboratoire de Physique Statistique, École Normale Supérieure, UPMC, Université Paris Diderot,
CNRS, 24 rue Lhomond, 75005 Paris, France
e-mail: Eric.Brunet@lps.ens.fr

B. Derrida
e-mail: Bernard.Derrida@lps.ens.fr

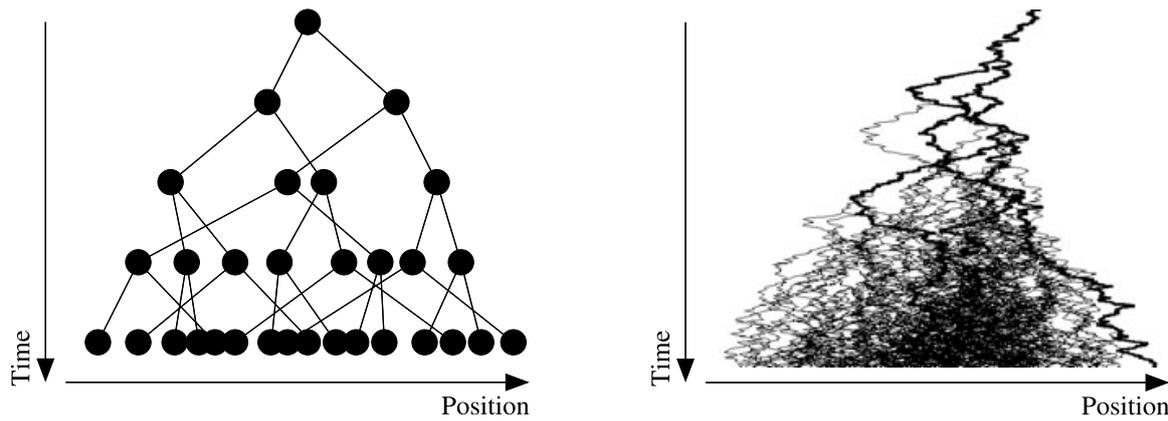


Fig. 1 Two examples of branching random walks. *Left*: a branching random walk with discrete time where each point splits into two points at each time step. *Right*: a continuous version called branching Brownian motion where points diffuse as in a Brownian motion and branch with a constant rate

well understood [11, 33]; the goal of the present paper is to describe the statistical properties of the positions of all the rightmost points in the system, in particular the distribution of the distances between the two rightmost points, the average density of points at some fixed distance from the rightmost $X_1(t)$, etc.

One motivation for studying these distances is that the problem belongs to the broader context of extreme value statistics [8, 13, 16, 22, 24, 32, 38, 40, 41]: Trying to understand the statistical properties of the rightmost points in a random set of points on the line is a problem common to the studies of the largest eigenvalues of random matrices [41], of the extrema of random signals [7, 9, 17, 18, 26, 37], or of the low lying states of some disordered systems such as spin glasses [2, 3, 10, 34, 38]. In fact, the points generated after some time t by a branching random walk can be viewed as the energies of the configurations of a directed polymer in a random medium [15, 19, 25, 35], and the distances between the rightmost points as the gaps between the low lying energy states.

The most studied example of branching random walk is the branching Brownian motion: one starts with a single point at the origin which performs a Brownian motion and branches at a given fixed rate (right part of Fig. 1). Whenever a branching event occurs, the point is replaced by two new points which then evolves as two independent branching Brownian motions. While the number of points generated after some time t grows exponentially with time, the expected position m_t of the rightmost point increases only linearly with time [11, 33]. In one dimension, McKean [33] and Bramson [11] have shown that the probability distribution of the rightmost point is given by the traveling wave solution of the Fisher-KPP equation, with a step initial condition. Here we will see that all the statistical properties of the rightmost points can be understood in terms of solutions to the Fisher-KPP equation with appropriate initial conditions [12]. We will also show that the distribution of the distances between these rightmost points has a long time limit which exhibits the striking property of superposability: the distances between the rightmost points of the union of two realizations of the branching Brownian motion have the same statistics as those of a single realization.

This paper is organized as follows: in Sect. 2 we introduce some generating functions useful to study random sets of points on the line, from which one can obtain all the properties of these random sets. In Sect. 3 we show that, for the branching Brownian motion, all these generating functions are solutions of the Fisher-KPP equation. We also show that the distribution of all the rightmost points as seen from m_t or, alternatively, as seen from $X_1(t)$, has a long time limit which can be computed as the delay of Fisher-KPP traveling waves. This distribution has the property of superposability. In Sect. 4, we present results, mostly

numerical, on some specific aspects of the limiting distribution of points in the branching Brownian motion, namely the distribution of the distance between the two rightmost points and the average density seen from the rightmost point. In Sect. 5 we explain how the results on the branching Brownian motion can be extended to more general branching random walks. Finally, we study in Sect. 6 the distribution of all the rightmost points in a specific frame which depends on the realization and which was introduced by Lalley and Sellke [29].

2 Statistics of Point Measures on the Line

In this section, we introduce some useful quantities (generating functions) to characterize random sets of points on the line such that the number $n(x)$ defined as

$$n(x) = (\text{the number of points on the right of position } x) \tag{1}$$

is finite and vanishes for x large enough.

2.1 The Generating Functions

The first generating function one can define is

$$\psi_\lambda(x) = \langle \lambda^{n(x)} \rangle. \tag{2}$$

From the knowledge of this function, one can extract the probability distribution function $p_i(x)$ of the position x of the i -th rightmost point. Indeed, by definition (2) of ψ_λ ,

$$\psi_\lambda(x) = \sum_{i \geq 0} Q_i(x) \lambda^i, \tag{3}$$

where $Q_i(x)$ is the probability that there are exactly i points on the right of x . One can notice that $Q_0(x) + Q_1(x) + \dots + Q_{i-1}(x)$ is the probability to have less than i points on the right of x . Assuming $|\lambda| < 1$, the generating function of these sums is, from (3),

$$\frac{\lambda}{1 - \lambda} \psi_\lambda(x) = Q_0(x) \lambda + [Q_0(x) + Q_1(x)] \lambda^2 + [Q_0(x) + Q_1(x) + Q_2(x)] \lambda^3 + \dots. \tag{4}$$

But $Q_0(x) + Q_1(x) + \dots + Q_{i-1}(x)$ is also the probability that the i -th rightmost point, if it exists, is on the left of x . Therefore,

$$\frac{\lambda}{1 - \lambda} \partial_x \psi_\lambda(x) = \sum_{i \geq 1} p_i(x) \lambda^i, \tag{5}$$

where $p_i(x) dx$ is the probability that the i -th rightmost point exists and is in the interval $[x, x + dx]$. (Note that $\int p_i(x) dx \leq 1$ is the probability that there are at least i points on the line.)

The knowledge of $\psi_\lambda(x)$ gives in particular the average distances between the points: from (5), one can see that

$$\begin{aligned} \int dx x \partial_x \psi_\lambda(x) &= (1 - \lambda)[\langle X_1 \rangle + \lambda \langle X_2 \rangle + \lambda^2 \langle X_3 \rangle + \dots], \\ &= \langle X_1 \rangle - \lambda[\langle X_1 \rangle - \langle X_2 \rangle] - \lambda^2[\langle X_2 \rangle - \langle X_3 \rangle] - \dots, \end{aligned}$$

where $\langle X_i \rangle = \int x p_i(x) dx$ is the average position of the i -th point (with the convention that $X_i = 0$ if there are less than i points in the system). Therefore

$$\int dx x [\partial_x \psi_0(x) - \partial_x \psi_\lambda(x)] = \sum_{i \geq 1} \langle d_{i,i+1} \rangle \lambda^i, \tag{6}$$

where $\langle d_{i,i+1} \rangle = \langle X_i \rangle - \langle X_{i+1} \rangle$ is the average distance between the i -th and the $(i + 1)$ -th point.

To obtain the correlations between the positions of pairs of points, one can start, for $y < x$, from the generating function

$$\psi_{\lambda,\mu}(x, y) = \langle \lambda^{n(x)} \mu^{n(y)} \rangle. \tag{7}$$

The coefficient in front of $\lambda^i \mu^j$ in the expansion of $\psi_{\lambda,\mu}$ in powers of λ and μ is the probability that there are exactly i points on the right of x and j points on the right of y . As in (4), the coefficient of $\lambda^i \mu^j$ in the expansion of $\lambda/(1 - \lambda) \times \mu/(1 - \mu) \times \psi_{\lambda,\mu}(x, y)$ is the probability that there are less than i points on the right of x and less than j points on the right of y , which is also the probability that the i -th rightmost point (if it exists) is on the left of x and the j -th rightmost point (if it exists) is on the left of y . Thus, for $y < x$,

$$\frac{\lambda}{1 - \lambda} \frac{\mu}{1 - \mu} \partial_x \partial_y \psi_{\lambda,\mu}(x, y) = \sum_{\substack{i \geq 1 \\ j > i}} p_{ij}(x, y) \lambda^i \mu^j, \tag{8}$$

where $p_{ij}(x, y) dx dy$ is the probability that both the i -th and j -th rightmost points exist and lie respectively in the intervals $[x, x + dx]$ and $[y, y + dy]$.

One can generalize (2), (7) by defining, for $x_0 > x_1 > \dots > x_k$, the generating functions

$$\psi_{\lambda_0, \dots, \lambda_k}(x_0, \dots, x_k) = \langle \lambda_0^{n(x_0)} \dots \lambda_k^{n(x_k)} \rangle \tag{9}$$

of the numbers $n(x_0), \dots, n(x_k)$ of points on the right of positions x_0, \dots, x_k , and get as in (5), (8) all the higher correlation functions. In that way, all the statistical properties of the measure can be derived from the knowledge of the generating functions (9).

2.2 The Measure Seen from the Rightmost Point

In the following we will often try to characterize the random set of points as seen from the rightmost point (i.e. in the frame where the rightmost point is at the origin). To do so, let us define the generating functions of the numbers $m(z)$ of points at the right of z in the frame of the rightmost point. (Note that if X_1 is the position of the rightmost, then $m(z) = n(X_1 + z)$ and one has $m(z) \geq 1$ for $z < 0$ and $m(z) = 0$ for $z > 0$.)

$$\chi_{\lambda_1, \dots, \lambda_k}(z_1, \dots, z_k) = \langle \lambda_1^{m(z_1)} \dots \lambda_k^{m(z_k)} \rangle. \tag{10}$$

(As in (9), we assume $z_1 > z_2 > \dots > z_k$.) These generating functions, as in Sect. 2.1, allow one to calculate all the statistical properties of the measure in the frame of the rightmost point (in particular the distribution of the relative distances between the points). They can be determined from the knowledge of the generating functions $\psi_{\lambda_0, \dots, \lambda_k}(x_0, \dots, x_k)$ defined in (9) by

$$\chi_{\lambda_1, \dots, \lambda_k}(z_1, \dots, z_k) = \int dx \partial_{x_0} \psi_{0, \lambda_1, \dots, \lambda_k}(x, x + z_1, \dots, x + z_k). \tag{11}$$

In Sect. 4 we will calculate the density of probability $P_{12}(a)$ that the two rightmost points are separated by a distance a (and that there are at least two points on the line) and the average density $\rho(a)$ at a distance a from the rightmost point. From (10) one can see that

$$P_{12}(a) = -\partial_a \partial_\mu \chi_\mu(-a)|_{\mu=0} = -\partial_a \int dx \partial_\mu \partial_{x_0} \psi_{0\mu}(x, x - a)|_{\mu=0}. \tag{12}$$

Then using that $\partial_{x_0} \psi_{0\mu}(x, x - a) = (\partial_x + \partial_a) \psi_{0\mu}(x, x - a)$, one gets

$$P_{12}(a) = -\partial_a^2 \int dx \partial_\mu \psi_{0\mu}(x, x - a)|_{\mu=0}. \tag{13}$$

By a similar calculation one can show that the average density $\rho(a)$ of points at distance a from the rightmost point is

$$\rho(a) = \partial_a^2 \int dx \partial_\mu \psi_{0\mu}(x, x - a)|_{\mu=1}. \tag{14}$$

2.3 Examples

We now describe a few examples of such measures.

2.3.1 A Poisson Process with an Arbitrary Density $r(x)$

Our first example is a Poisson process on the line with a density $r(x)$. We assume that $r(x)$ decays fast enough to the right so that a rightmost point exists, and that $\int r(x) dx = \infty$ so that there are infinitely many points on the line.

By definition of a Poisson process, each infinitesimal interval $[x, x + dx]$ is occupied by a point with probability $r(x) dx$ and empty with probability $1 - r(x) dx$, and the occupation numbers of disjoint intervals are uncorrelated. The probability $Q_i(x)$ that there are exactly i points on the right of x is given by

$$Q_i(x) = \frac{R(x)^i e^{-R(x)}}{i!} \quad \text{where } R(x) = \int_x^\infty r(z) dz. \tag{15}$$

From this, we obtain $\psi_\lambda(x)$ from (2), (3) and $\psi_{\lambda\mu}(x, y)$ from (7) in the Poisson process:

$$\psi_\lambda(x) = e^{-(1-\lambda)R(x)}, \quad \psi_{\lambda\mu}(x, y) = e^{-\mu(1-\lambda)R(x) - (1-\mu)R(y)}. \tag{16}$$

Using (6), the generating function of the average $\langle d_{i,i+1} \rangle$ between the i -th and $(i + 1)$ -th points is

$$\sum_{i \geq 1} \lambda^i \langle d_{i,i+1} \rangle = \int_{-\infty}^\infty dx [e^{-(1-\lambda)R(x)} - e^{-R(x)}]. \tag{17}$$

The probability distribution function $P_{12}(a)$ that the distance $d_{1,2}$ is equal to a and the average density $\rho(a)$ seen at a distance a from the rightmost point are given by

$$P_{12}(a) = \int_{-\infty}^\infty dx r(x + a)r(x)e^{-R(x)}, \quad \rho(a) = \int_{-\infty}^\infty dx r(x - a)r(x)e^{-R(x)}. \tag{18}$$

These expressions can be understood directly from the definition of the Poisson process or, with a little more algebra, from (13), (14). One can notice that $P_{12}(a)$ and $\rho(a)$ are given by the same expression with a replaced by $-a$ and are therefore analytic continuations of each other whenever $r(x)$ is analytic.

2.3.2 A Poisson Process with an Exponential Density $e^{-\alpha x}$

In the special case where the density of the Poisson process is an exponential $r(x) = \exp(-\alpha x)$, one can simply replace $R(x)$ in the previous expressions by $\exp(-\alpha x)/\alpha$. This gives

$$\begin{aligned} \psi_\lambda(x) &= \exp\left[-(1-\lambda)\frac{e^{-\alpha x}}{\alpha}\right] = \exp\left[-e^{-\alpha\left(x-\frac{\ln(1-\lambda)}{\alpha}\right)}\right], \\ \psi_{\lambda\mu}(x, y) &= \exp\left[-\mu(1-\lambda)\frac{e^{-\alpha x}}{\alpha} - (1-\mu)\frac{e^{-\alpha y}}{\alpha}\right], \end{aligned} \tag{19}$$

so that from (6)

$$\sum_{i \geq 1} \lambda^i \langle d_{i,i+1} \rangle = -\frac{\ln(1-\lambda)}{\alpha}, \tag{20}$$

and thus [12]

$$\langle d_{i,i+1} \rangle = \frac{1}{\alpha i}. \tag{21}$$

One also has from (18)

$$P_{12}(a) = \alpha e^{-\alpha a}, \quad \rho(a) = \alpha e^{\alpha a}. \tag{22}$$

2.3.3 Decorated Measures

Start with a collection of points $\{u_i\}$ distributed according to some measure ν_1 and, independently for each point u_i , replace it by a realization of another measure ν_2 shifted by u_i . We say that the points u_i are *decorated* by the measure ν_2 and call the resulting measure as ν_1 decorated by ν_2 .

If the functions $\psi_\lambda(x), \psi_{\lambda\mu}(x, y), \dots$ for the measure ν_2 are known, the decorated measure is characterized by functions $\Psi_\lambda(x), \Psi_{\lambda\mu}(x, y), \dots$ given by

$$\Psi_\lambda(x) = \left\langle \prod_i \psi_\lambda(x - u_i) \right\rangle_{u_i}, \quad \Psi_{\lambda\mu}(x, y) = \left\langle \prod_i \psi_{\lambda\mu}(x - u_i, y - u_i) \right\rangle_{u_i}, \tag{23}$$

where the average is over all realizations $\{u_i\}$ of the measure ν_1 . For instance, if ν_1 is a Poisson process of density $r(u)$, then

$$\begin{aligned} \Psi_\lambda(x) &= \prod_u [1 - r(u)du + r(u)\psi_\lambda(x - u)du] = \exp\left[\int [\psi_\lambda(x - u) - 1]r(u) du\right], \\ \Psi_{\lambda\mu}(x, y) &= \exp\left[\int [\psi_{\lambda\mu}(x - u, y - u) - 1]r(u) du\right]. \end{aligned} \tag{24}$$

2.3.4 Ruelle Cascades

For a decorated measure where the decoration ν_2 is a Poisson process of density $e^{-\alpha x}$, the average over the u_i 's in (23) leads in general to complicated expressions for $\Psi_\lambda(x)$ or $\Psi_{\lambda\mu}(x, y)$. The expressions for $P_{12}(a)$ and $\rho(a)$ are however the same as in (22) for the pure Poisson process of density $e^{-\alpha x}$. In fact, all the statistical properties of the distances

between the rightmost points are the same as those in a Poisson process with an exponential density.

This can be understood from the following reason: decorating the points u_1, \dots, u_k, \dots by independent realizations of a Poisson process of density $e^{-\alpha x}$ is equivalent to drawing a single realization of a Poisson process of density $\sum_k e^{-\alpha(x-u_k)} = e^{-\alpha x} \sum_k e^{\alpha u_k}$, which is just the same as one realization of a Poisson process of density $e^{-\alpha x}$ shifted by the random variable $\ln(\sum_k e^{\alpha u_k})/\alpha$.

The same argument applies to Ruelle cascades, which can be defined as follows [7, 9, 10, 37]: take an increasing sequence of positive numbers $\alpha_1 < \alpha_2 < \dots$ and start with a Poisson process of density $e^{-\alpha_1 x}$. At each step $k > 1$, each point in the system is decorated by a Poisson process of density $e^{-\alpha_k x}$. At step k , the measure of points in the system is simply, from the previous argument, a Poisson process of density $e^{-\alpha_k x}$ globally shifted by a random variable which depends on the positions of the points at step $k - 1$. Therefore, the statistics of the distances of the rightmost points is the same as for the Poisson process of density $e^{-\alpha_k x}$.

3 The Branching Brownian Motion and Fisher-KPP Fronts

3.1 The Fisher-KPP Equation

We are now going to see how the generating functions (2), (7), (9) can be determined when the random set of points on the line are the points generated at time t by a branching Brownian motion.

To define the branching Brownian motion we start at time $t = 0$ with a single point at the origin. This point diffuses and branches, and its offspring do the same. After some time t , a realization of the process consists of a finite number of points located at positions $X_i(t)$ for $i = 1, 2, 3, \dots$. Then, during the next time interval $dt \ll 1$, each point, independently of what the others do, moves a random distance $X_i(t + dt) - X_i(t) = \eta_i(t)\sqrt{2dt}$ with $\langle \eta_i(t) \rangle = 0$ and $\langle \eta_i(t)^2 \rangle = 1$, and, with probability dt , is replaced by two new points located at the same position $X_i(t)$.

For any function ϕ one can define the generating function $H_\phi(x, t)$ by

$$H_\phi(x, t) = \left\langle \prod_i \phi[x - X_i(t)] \right\rangle, \tag{25}$$

where the $X_i(t)$ for $i = 1, \dots, N_t$ are the positions of the N_t points of the branching Brownian motion at time t and $\langle \cdot \rangle$ denotes an average over all the possible realizations.

By analyzing what happens during the very first time interval dt , one can see that the evolution of $H_\phi(x, t)$ satisfies

$$H_\phi(x, t + dt) = (1 - dt) \langle H_\phi(x - \eta\sqrt{2dt}, t) \rangle_\eta + dt H_\phi(x, t)^2. \tag{26}$$

The first term in the right hand side represents the motion of the initial point during the first time interval dt and the second term represents the branching event which occurs with probability dt during this first time interval. Taking dt to zero, one gets

$$\partial_t H_\phi = \partial_x^2 H_\phi + H_\phi^2 - H_\phi, \tag{27}$$

which is the Fisher-KPP equation [21, 28, 33]. (The Fisher-KPP equation is often written as $\partial_t h = \partial_x^2 h + h - h^2$, but this is identical to (27) by the change of variable $h = 1 - H_\phi$.) Because there is a single point at the origin at time $t = 0$, the initial condition is simply, from (25),

$$H_\phi(x, 0) = \phi(x). \tag{28}$$

The generating function (9) at time t

$$\psi_{\lambda_0, \dots, \lambda_k}(x_0, \dots, x_k) = \langle \lambda_0^{n(x_0)} \dots \lambda_k^{n(x_k)} \rangle \tag{29}$$

can be written, for $0 > z_1 > \dots > z_k$, as

$$\psi_{\lambda_0, \dots, \lambda_k}(x, x + z_1, \dots, x + z_k) = \left\langle \prod_i \phi[x - X_i(t)] \right\rangle = H_\phi(x, t), \tag{30}$$

where the function $\phi(x)$ is given by

$$\phi(x) = \lambda_0^{1-\theta(x)} \lambda_1^{1-\theta(x+z_1)} \dots \lambda_k^{1-\theta(x+z_k)}, \tag{31}$$

and where $\theta(x)$ is the Heaviside step function defined by

$$\theta(x) = \begin{cases} 1 & \text{for } x \geq 0, \\ 0 & \text{for } x < 0. \end{cases} \tag{32}$$

See Fig. 2 for the general shape of (31).

With the choice (31) of ϕ , the generating function (9) and, therefore, all the properties of the point measure in the branching Brownian motion at time t can be obtained as solutions of the Fisher-KPP equation with the initial condition (28).

In the special case $k = 0$ and $\lambda_0 = 0$ of (31), i.e. for the initial condition $\phi(x) = \theta(x)$, one gets

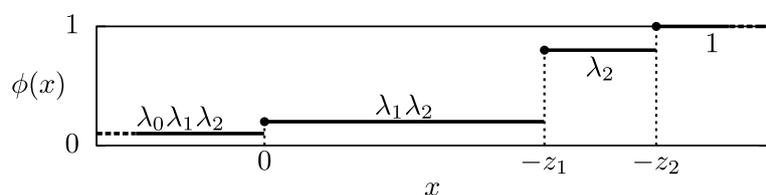
$$H_\theta(x, t) = \text{Proba}(\text{There is no point at time } t \text{ on the right of } x), \tag{33}$$

and one recovers the well-known fact [11, 33] that the solution $H_\theta(x, t)$ of the Fisher-KPP equation with a step initial condition is the cumulative distribution function of the position of the rightmost point.

In Sect. 4 we will choose $\phi = \phi_1$ and $\phi = \phi_2$, other special cases of (31), given by

$$\phi_1(x) = \begin{cases} 1 & \text{for } x \geq 0, \\ \lambda & \text{for } x < 0, \end{cases} \quad \phi_2(x) = \begin{cases} 1 & \text{for } x \geq a, \\ \mu & \text{for } 0 \leq x < a, \\ \lambda\mu & \text{for } x < 0, \end{cases} \tag{34}$$

Fig. 2 The function (31) for $k = 2$



to calculate the generating functions (2), (7) at time t

$$\psi_\lambda(x) = H_{\phi_1}(x, t), \quad \psi_{\lambda,\mu}(x, x - a) = H_{\phi_2}(x, t) \tag{35}$$

needed to determine the distribution $P_{12}(a)$ and the density $\rho(a)$ defined at the end of Sect. 2.

3.2 The Branching Brownian Motion Seen from the Rightmost Point

The Fisher-KPP equation (27) has two homogeneous solutions: $H_\phi = 1$, which is unstable, and $H_\phi = 0$, which is stable. When the initial condition $\phi(x)$ is given by the step function $\theta(x)$, see (32), the solution $H_\theta(x, t)$ of (27) becomes a traveling wave with the phase $H_\theta = 0$ invading the phase $H_\theta = 1$ [11, 21, 28]. As the front is an extended object, one can define its position m_t in several ways; for example one could define m_t as the solution of $H_\theta(m_t, t) = \alpha$ for some $0 < \alpha < 1$. Here it will be convenient to use the following definition

$$m_t = \int dx x \partial_x H_\theta(x, t). \tag{36}$$

One can see using (33) that m_t defined by (36) is the average position of the rightmost point in the branching Brownian motion.

If the initial condition (28) is not a step function but is such that $\phi(x) = 1$ for all large enough x and $\phi(x)$ is a constant smaller than 1 for all large negative x , as in (31), (34), the solution $H_\phi(x, t)$ of (27) becomes also a traveling wave. Its position $m_t^{(\phi)}$ can be defined as in (36) by

$$m_t^{(\phi)} = \int dx x \partial_x H_\phi(x, t). \tag{37}$$

We are now going to show that the whole measure seen from the rightmost point can be written in terms of this position $m_t^{(\phi)}$: one can rewrite (11) as

$$\begin{aligned} \chi_{\lambda_1, \dots, \lambda_k}(z_1, \dots, z_k) &= \int dx (\partial_x - \partial_{z_1} - \dots - \partial_{z_k}) \psi_{0, \lambda_1, \dots, \lambda_k}(x, x + z_1, \dots, x + z_k), \\ &= \psi_{0, \lambda_1, \dots, \lambda_k}(x, x + z_1, \dots, x + z_k) \Big|_{x=-\infty}^{x=+\infty} \\ &\quad - \int dx (\partial_{z_1} + \dots + \partial_{z_k}) \psi_{0, \lambda_1, \dots, \lambda_k}(x, x + z_1, \dots, x + z_k), \\ &= 1 + (\partial_{z_1} + \dots + \partial_{z_k}) \int dx x \partial_x \psi_{0, \lambda_1, \dots, \lambda_k}(x, x + z_1, \dots, x + z_k). \end{aligned} \tag{38}$$

Then from (30) and (37) one gets

$$\chi_{\lambda_1, \dots, \lambda_k}(z_1, \dots, z_k) = 1 + (\partial_{z_1} + \dots + \partial_{z_k}) m_t^{(\phi)}, \tag{39}$$

where ϕ is the function (31) with $\lambda_0 \rightarrow 0$.

Therefore, with the definition (37) of the position of the front, the whole information about the measure in the frame of the rightmost point, at any time t , can be extracted from the ϕ dependence of $m_t^{(\phi)}$.

3.3 The Limiting Measure and the Delays

In the long time limit, it is known [11, 39] that the traveling wave solution $H_\theta(x, t)$ of (27), with the initial condition (32), takes an asymptotic shape, $F(x)$. This means that

$$H_\theta(m_t + x, t) \xrightarrow[t \rightarrow \infty]{} F(x), \tag{40}$$

where $F(x)$ satisfies

$$F'' + 2F' + F^2 - F = 0, \quad F(-\infty) = 0, \quad F(\infty) = 1, \quad \int dx x \partial_x F(x) = 0. \tag{41}$$

It is also known, since the work of Bramson [11], that, in the long time limit, the traveling wave moves at a velocity 2 and that its position (36) is given by

$$m_t = 2t - \frac{3}{2} \ln t + \text{Constant} + o(1). \tag{42}$$

If the function $\phi(x)$ is not the step function but is of the form (31), (34), the solution $H_\phi(x, t)$ of (27) becomes also a traveling wave with the same shape $F(x)$. This wave is centered around the position $m_t^{(\phi)}$, defined in (37), and one has

$$H_\phi(m_t^{(\phi)} + x, t) \xrightarrow[t \rightarrow \infty]{} F(x). \tag{43}$$

For large times $m_t^{(\phi)}$ is still given by (42), but with a different constant [11]. This means that

$$m_t - m_t^{(\phi)} \xrightarrow[t \rightarrow \infty]{} f[\phi], \tag{44}$$

where $f[\phi]$ is the long time delay in the position of the front due to the modified initial condition, as compared to a front starting with a step function. Taken together, (43) and (44) give

$$H_\phi(m_t + x, t) \xrightarrow[t \rightarrow \infty]{} F(x + f[\phi]). \tag{45}$$

Using (30), this becomes

$$\psi_{\lambda_0, \dots, \lambda_k}(m_t + x, m_t + x + z_1, \dots, m_t + x + z_k) \xrightarrow[t \rightarrow \infty]{} F(x + f[\phi]), \tag{46}$$

which shows that the measure of $\{X_1(t) - m_t, X_2(t) - m_t, \dots\}$ (the rightmost points in the branching Brownian motion seen from the m_t frame) does converge when $t \rightarrow \infty$ to a limiting point measure characterized by the functions $F(x)$ and $f[\phi]$.

The measure of $\{X_2(t) - X_1(t), X_3(t) - X_1(t), \dots\}$ (the rightmost points in the branching Brownian motion seen from the $X_1(t)$ frame) also has a well-defined limit when $t \rightarrow \infty$. Indeed, using (39) and (44), one gets

$$\chi_{\lambda_1, \dots, \lambda_k}(z_1, \dots, z_k) \xrightarrow[t \rightarrow \infty]{} 1 - (\partial_{z_1} + \dots + \partial_{z_k}) f[\phi]. \tag{47}$$

Therefore, in the long time limit, all the information on the distribution of the rightmost points seen from $X_1(t)$ is contained in the ϕ dependence of the delay $f[\phi]$.

Note that, in contrast to (39) which requires the position to be defined by (37), the delay $f[\phi]$ in (46) or (47) depends only on ϕ : it would not change if we had chosen another definition of the front position.

3.4 The Superposability of Branching Brownian Motions

Let us now consider M independent branching Brownian motions starting at $t = 0$ at positions u_1, \dots, u_M . Following the same argument as in Sect. 3.1, the generating function (9) of the union of the points at time t of these M branching Brownian motions is given by the following generalization of (30)

$$\psi_{\lambda_0, \dots, \lambda_k}(x, x + z_1, \dots, x + z_k) = \prod_{\alpha=1}^M H_\phi(x - u_\alpha, t), \tag{48}$$

where $H_\phi(x, t)$ is the same solution of (27) with the initial condition (31) as in the case of a single branching Brownian motion starting at the origin. In the long time limit, using (45),

$$\psi_{\lambda_0, \dots, \lambda_k}(m_t + x, m_t + x + z_1, \dots, m_t + x + z_k) \xrightarrow[t \rightarrow \infty]{} \prod_{\alpha=1}^M F(x + f[\phi] - u_\alpha). \tag{49}$$

This means that here again, there is a limiting measure when $t \rightarrow \infty$ for the rightmost points in the m_t frame. This measure is not the same as before (when one starts with a single point at the origin), as can be seen by comparing (49) and (46). In particular, the distribution of the rightmost point is different.

We now consider the distribution of points in the frame of the rightmost one, in the long time limit. The integral in the last line of (38) can be written

$$\begin{aligned} & \int dx x \partial_x \psi_{0, \lambda_1, \dots, \lambda_k}(x, x + z_1, \dots, x + z_k) \\ &= m_t - f[\phi] + \int dx x \partial_x \psi_{0, \lambda_1, \dots, \lambda_k}(m_t + x - f[\phi], m_t + x - f[\phi] + z_1, \dots, \\ & \quad m_t + x - f[\phi] + z_k). \end{aligned} \tag{50}$$

(We made the change of variable $x \rightarrow m_t + x - f[\phi]$ and used the fact that ψ is 1 for $x = +\infty$ and 0 for $x = -\infty$.) From (49), the last term in (50) converges at long times to $\int dx x \partial_x \prod_{\alpha} F(x - u_\alpha)$ which does not depend on the parameters z_i . The front position m_t does not depend either on the z_i , so that only the term $-f[\phi]$ survives in the differentiation in the last line of (38). Finally,

$$\chi_{\lambda_1, \dots, \lambda_k}(z_1, \dots, z_k) \xrightarrow[t \rightarrow \infty]{} 1 - (\partial_{z_1} + \dots + \partial_{z_k}) f[\phi], \tag{51}$$

as in (47).

It is remarkable that the generating function χ does not depend neither on the number M of starting points nor on their positions u_α . The picture which emerges is that if we superpose several branching Brownian motions, starting at arbitrary positions, the limiting measure in the frame of the rightmost point is, when $t \rightarrow \infty$, the same as for a single branching Brownian motion.

We will say that, in the long time limit, the measure of the distances between the rightmost points in a branching Brownian motion becomes *superposable*: the union of two (or more) realizations of the process (even moved by arbitrary translations u_α) leads to the same measure in the frame of the rightmost point as for a single branching Brownian motion.

As a remark, it is easy to check, that the Poisson process with an exponential density $r(x) = e^{-\alpha x}$, see Sect. 2.3.2, is an example of a superposable measure: the superposition of

M such Poisson processes translated by arbitrary amounts u_1, \dots, u_M is identical to a single Poisson process with an exponential distribution translated by $\alpha^{-1} \ln(e^{\alpha u_1} + \dots + e^{\alpha u_M})$. One can also check that, for the same reason, all the decorated measures of Sect. 2.3.3 are superposable when ν_1 is a Poisson process with an exponential density.

In Sect. 6, we will state a stronger version of the superposability property of the branching Brownian motion.

4 Some Quantitative Properties of the Branching Brownian Motion Seen from the Rightmost Point

In this section we obtain, by integrating numerically (27) with the appropriate initial condition, some statistical properties of the limiting measure seen from the rightmost point.

4.1 Average Distances Between Consecutive Points

The analytic calculation of the delay $f[\phi]$ is in general not easy. For $\phi = \phi_1$ given by (34), however, it was possible to show [12] that when $1 - \lambda \ll 1$, the delay is given by

$$f[\phi_1] \simeq -\ln(1 - \lambda) - \ln[-\ln(1 - \lambda)] + \mathcal{O}(1), \tag{52}$$

and, from this, one could deduce that, in the long time limit, the average of the distance $d_{i,i+1}$ between the i -th and the $(i + 1)$ -th rightmost points is given for large i by

$$\langle d_{i,i+1} \rangle \simeq \frac{1}{i} - \frac{1}{i \ln i}. \tag{53}$$

In [12], the numerical values of the distances between the rightmost points were also obtained by integrating the Fisher-KPP equation with the initial condition ϕ_1 in (34) and by using (6) (in practice we integrated numerically the equations satisfied by the coefficients of the expansion of $\psi_\lambda(x)$ in powers of λ). It was found that

$$\begin{aligned} \langle d_{1,2} \rangle &\simeq 0.496, & \langle d_{2,3} \rangle &\simeq 0.303, & \langle d_{3,4} \rangle &\simeq 0.219, \\ \langle d_{4,5} \rangle &\simeq 0.172, & \langle d_{5,6} \rangle &\simeq 0.142, & \langle d_{6,7} \rangle &\simeq 0.121. \end{aligned} \tag{54}$$

The results (53), (54) gave evidence that the distances between the rightmost points of the branching Brownian motion were different from those of a Poisson process with an exponential density (21).

4.2 Distribution of the Distance Between the Two Rightmost Points

According to (13), to obtain the distribution $P_{12}(a)$ of the distance between the two rightmost points, one needs to calculate $\psi_{0\mu}(x, x - a)$ to first order in μ . We first remark that at time t for $a > 0$, $\psi_{00}(x, x - a) = \text{Proba}[n(x) = 0 \text{ and } n(x - a) = 0] = \text{Proba}[n(x - a) = 0] = H_\theta(x - a, t)$ where $H_\theta(x, t)$ is the standard Fisher-KPP front with the step initial condition (it is also easy to see from the definition (34) of ϕ_2 .) Then writing at time t that

$$\psi_{0\mu}(x, x - a) = H_\theta(x - a, t) + \mu R_a(x - a, t) + \mathcal{O}(\mu^2) \tag{55}$$

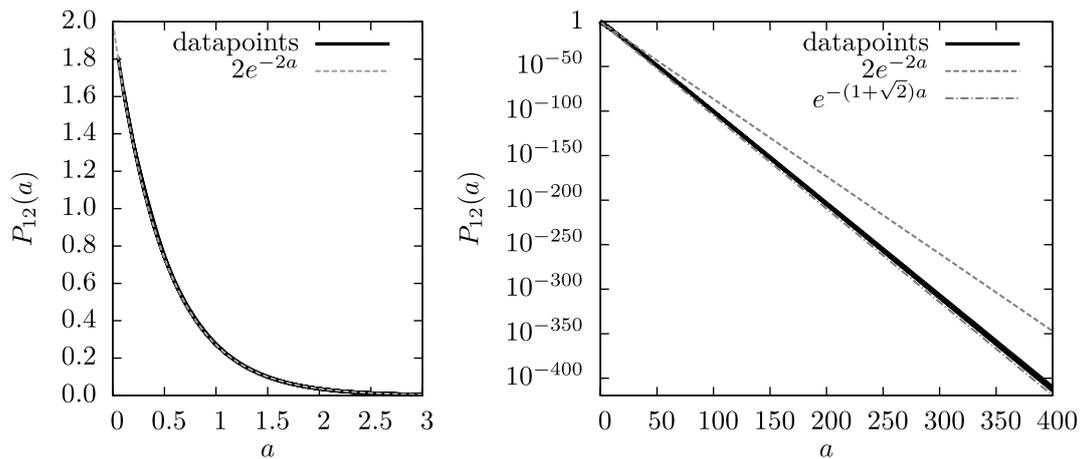


Fig. 3 The density of probability $P_{12}(a)$ of observing a distance a between the two rightmost points in the $t \rightarrow \infty$ limit, as a function of a . For a small (*left part*), the distribution is very close to $2e^{-2a}$. For larger values of a , one observes a faster exponential decay of order $e^{-(1+\sqrt{2})a}$

is solution of the Fisher-KPP equation, and using the initial condition ϕ_2 in (34), one gets

$$\partial_t R_a = \partial_x^2 R_a - R_a + 2H_\theta R_a; \quad R_a(x, 0) = \begin{cases} 1 & \text{for } -a \leq x < 0, \\ 0 & \text{otherwise.} \end{cases} \quad (56)$$

Then, from (13) one gets

$$P_{12}(a) = -\partial_a^2 \int dx R_a(x, t). \quad (57)$$

Figure 3 shows our numerical result for the distribution $P_{12}(a)$ of the distance between the two rightmost points in the long time limit. More details on our numerical procedure are given in Appendix A.

We see that $P_{12}(a)$ is very close to $2e^{-2a}$ for the values of a which have a significant probability of occurring. This is of course consistent with an average distance (54) close to $1/2$. For large a (events with a small probability), however, the exponential decay is faster. We now present a simple argument leading to the following prediction, which is consistent with our numerical data,

$$P_{12}(a) \sim e^{-(1+\sqrt{2})a} \quad \text{for large } a. \quad (58)$$

In the long time limit, the right frontier of the branching Brownian motion moves at velocity $v = 2$. Let us assume that a large distance a between the two rightmost points is produced by the following scenario: by a rare fluctuation, the rightmost point escapes and, without branching, goes significantly ahead while the rest of the points go on as usual, forming a frontier moving at velocity $v = 2$. Such an event leads to the distance a between the two rightmost points if, during a time τ , the rightmost point moves (by diffusion alone) by a distance $a + 2\tau$ without branching. The probability of such a scenario is

$$\text{Proba}(X_1 - X_2 \simeq a \text{ after an escape time } \tau) \sim \exp\left[-\frac{(a + 2\tau)^2}{4\tau}\right] \times e^{-\tau}. \quad (59)$$

The first term is the probability of diffusing over a distance $a + 2\tau$ during time τ , and the second term is the probability of not branching. The probability to observe a large distance

$a \gg 1$ is then dominated by the events with τ chosen to maximize (59), that is

$$\tau_{\text{optimal}} = \frac{a}{2\sqrt{2}}, \tag{60}$$

and this leads to (58) in good agreement with the numerical data of Fig. 3.

There is a remarkable relation between the decay rate in (58) and the shape of the traveling wave solution of (41). Around the *stable* region $F \simeq 0$, (41) can be linearized and one has

$$F(z) \simeq \text{Cste} \times e^{rz} \text{ for } z \rightarrow -\infty, \text{ with } r^2 + 2r - 1 = 0. \tag{61}$$

We emphasize that this is a linear analysis of the *stable* region, which is usually uninteresting (in contrast to the *unstable* region which determines the velocity). The solutions of (61) for r are

$$r = -1 \pm \sqrt{2}. \tag{62}$$

$r = -1 + \sqrt{2}$ is the correct root as it is the only positive solution and $F(-\infty)$ has to vanish. The other solution $r = -1 - \sqrt{2}$ (the wrong root) coincides (up to the sign) with the decay rate of the distribution $P_{12}(a)$ for the distance a between the two rightmost points (58).

As explained in Appendix B, this coincidence exists in a broad class of branching processes: each variant of the branching Brownian motion is linked to a variant of the Fisher-KPP equation, and the wrong root in the linear analysis of the stable region always gives the asymptotic decay rate of $P_{12}(a)$.

4.3 Average Density Seen from the Rightmost Point

To obtain the average density of points at a distance a on the left of the rightmost point, one needs, according to (14), to calculate $\psi_{0\mu}(x, x - a)$ for μ close to 1. As in Sect. 4.2, we first remark, from the definition (7), that $\psi_{01}(x, x - a) = \text{Proba}[n(x) = 0] = H_\theta(x, t)$ is the standard Fisher-KPP front with the step initial condition. Then, writing at time t that

$$\psi_{0\mu}(x, x - a) = H_\theta(x, t) - (1 - \mu)\tilde{R}_a(x, t) + \mathcal{O}[(1 - \mu)^2] \tag{63}$$

is solution of the Fisher-KPP equation, and using the initial condition ϕ_2 in (34), one gets

$$\partial_t \tilde{R}_a = \partial_x^2 \tilde{R}_a - \tilde{R}_a + 2H_\theta \tilde{R}_a; \quad \tilde{R}_a(x, 0) = \begin{cases} 1 & \text{for } 0 \leq x < a, \\ 0 & \text{otherwise.} \end{cases} \tag{64}$$

It is the same equation as for R_a in (56), but with a different initial condition. Then, from (14) one gets

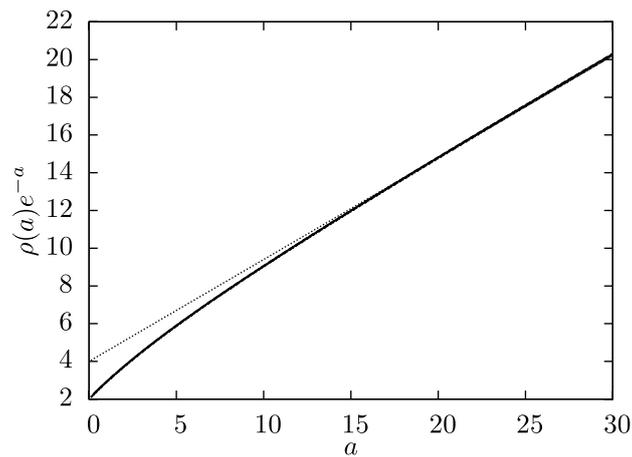
$$\rho(a) = \partial_a^2 \int dx \tilde{R}_a(x, t). \tag{65}$$

One can notice the great similarity between the expressions for the average density $\rho(a)$ of points at a distance a from the rightmost (64), (65) and the probability distribution $P_{12}(a)$ for the distance between the two rightmost points (56), (57): one goes from one to the other by simple changes of signs, as in the example of a Poisson process (18).

Figure 4 presents our numerical results for $\rho(a)$ in the long time limit. We see that $\rho(a)$ increases as

$$\rho(a) \simeq \text{Cste} \times ae^a \text{ for large } a. \tag{66}$$

Fig. 4 The average density $\rho(a)$ of points at a distance a of the rightmost in the long time limit grows like ae^a . When the data is multiplied by e^{-a} , as shown in the figure, the linear prefactor is clearly visible



Note that a Poisson process with such a density would lead to asymptotic distances between points given by (53). The branching Brownian motion is however not a Poisson process as the points are correlated, at least near the tip.

5 Generalizations to Other Branching Processes

All the results of Sects. 3 and 4 can be generalized to other branching processes on the line where points move and branch independently of the positions and of the motions of the other points. In such systems, the function $H_\phi(x, t)$ defined in (25) is also solution of an equation similar to the Fisher-KPP equation (27). Here are four examples:

- (A) The points perform Brownian motions and branch as before, but at each branching event there is a probability p to branch into three points and $1 - p$ to branch into two. Then $H_\phi(x, t)$ evolves according to

$$\partial_t H_\phi = \partial_x^2 H_\phi + p H_\phi^3 + (1 - p) H_\phi^2 - H_\phi. \tag{67}$$

- (B) Time is discrete with steps of duration δ ; at each time step, a point at position x branches into two points at positions $x + \epsilon_1$ and $x + \epsilon_2$, where the ϵ_i take independent random values distributed according to some given $\rho(\epsilon)$. The evolution of $H_\phi(x, t)$ is then given by

$$H_\phi(x, t + \delta) = \left[\int d\epsilon \rho(\epsilon) H_\phi(x - \epsilon, t) \right]^2. \tag{68}$$

In this example, the positions of the points can be thought of as the possible energies of a directed polymer on a Caley tree with independent random energies ϵ on the edges of the tree [19, 35].

- (C) Time is continuous but space is discrete with steps 1; during dt , each point at position x has a probability dt of being removed and replaced by two points at position $x + 1$. The equation satisfied by $H_\phi(x, t)$ is

$$\partial_t H_\phi(x, t) = H_\phi(x - 1, t)^2 - H_\phi(x, t). \tag{69}$$

This example is relevant to the theory of binary search trees [31, 32, 36].

(D) Time and space are discrete with steps s for space and δ for time; in a given time step, a point at position x has a probability δ of branching into two points at position x , a probability δ/s^2 of jumping to the left, δ/s^2 of jumping to the right, and $1 - \delta - 2\delta/s^2$ of doing nothing. Then:

$$H_\phi(x, t + \delta) = H_\phi(x, t) + \delta \left[\frac{H_\phi(x - s, t) + H_\phi(x + s, t) - 2H_\phi(x, t)}{s^2} - H_\phi(x, t) + H_\phi(x, t)^2 \right], \tag{70}$$

which is of course a discretized version of the original Fisher-KPP equation. (70) is actually the equation we used in our numerical simulations, see Appendix A.

In all cases, these equations have $H_\phi = 1$ as an unstable fixed point, and $H_\phi = 0$ as a stable fixed point. For initial conditions $H_\phi(x, 0) = \phi(x)$ of the type (31), (32), (34), the function $H_\phi(x, t)$ develops into a traveling wave moving at a specific velocity v^* . We recall briefly the procedure to determine the asymptotic velocity v^* of the front (which is also, through (33), the velocity of the rightmost point in the branching process). One looks for traveling wave solutions moving at velocity v of the form $H_\phi(x, t) = F(x - vt)$ and solve the linearized equation around the unstable fixed point by writing $1 - F(x) \simeq \epsilon e^{-\gamma x}$. This leads to a relation between γ and v , and the minimal value v^* of v reached at some γ^* is the velocity selected by the front [39]. (We only consider here cases where the function $v(\gamma)$ has a minimum.) For our four examples

$$\begin{aligned} \text{(A)} \quad & v = \gamma + \frac{1+p}{\gamma}; \quad v^* = 2\sqrt{1+p}, \\ \text{(B)} \quad & v = \frac{1}{\gamma\delta} \ln \left[2 \int d\epsilon \rho(\epsilon) e^{\gamma\epsilon} \right], \\ \text{(C)} \quad & v = \frac{2e^\gamma - 1}{\gamma}; \quad v^* \simeq 4.311, \\ \text{(D)} \quad & v = \frac{1}{\gamma\delta} \ln \left[1 + 2\delta \frac{\cosh(\gamma s) - 1}{s^2} + \delta \right]. \end{aligned} \tag{71}$$

Once the equation for H_ϕ of a particular branching process is written, one has access to all the generating functions $\psi_\lambda(x)$, $\psi_{\lambda,\mu}(x, x - a)$, etc., see (2), (7), (9), by choosing the appropriate initial conditions (31), (34) for the front equation. The whole measure in the frame of the rightmost point is then obtained from (39) at any finite time t . Note that to prove the existence of a long time limit to the point measure in this frame for a specific branching process, one would need a version of Bramson’s result (42) for this process which is, to our knowledge, not known in the general case.

It is natural to ask which properties of the branching Brownian motion can be extended to other branching processes. If the measure for the distances between the rightmost points has a long time limit, then the arguments of Sect. 3.4 can be easily generalized and one can show that it is superposable. We have checked that the analytical argument [12] leading to the asymptotic expression (53) for the average distances $\langle d_{i,i+1} \rangle$ at large times can be extended in case (B) for a large class of densities $\rho(\epsilon)$ and yields

$$\langle d_{i,i+1} \rangle \simeq \frac{1}{\gamma^*} \left(\frac{1}{i} - \frac{1}{i \ln i} \right) \quad \text{for large } i. \tag{72}$$

We have also checked numerically on examples (C) and (D) that the density at a distance a of the rightmost point is, as in (66),

$$\rho(a) \simeq \text{Cste} \times ae^{\gamma^* a} \quad \text{for large } a. \tag{73}$$

For the tail of the distribution $P_{12}(a)$ of the distance a between the two rightmost points, we discussed a scenario, at the end of Sect. 4.2, which can be generalized (see Appendix B) to calculate the exponential decay $P_{12}(a)$ for more general branching processes. This scenario, however, can only hold if points can move without branching, as in our examples (A) and (D); for instance, in example (A), it predicts an exponential with a decay rate equal to $\sqrt{1+p} + \sqrt{2+p}$. In examples (B) and (C), the points branch whenever they move and the tail of $P_{12}(a)$ is in general not an exponential.

Note that special care should be taken if the points are located on a discrete lattice, as in cases (C), (D) and possibly (B): quantities such as $P_{12}(a)$, see (13), become probabilities rather than densities of probability and quantities such as $\rho(a)$, see (14), become average numbers rather than average densities, and all the formulas in the previous sections need to be adapted: integrals become discrete sums, derivatives become finite differences, etc. If one interprets $n(x)$ as the number of points *strictly* on the right of x , then the generating functions $\psi_\lambda, \psi_{\lambda\mu}, \dots$ are still related to H_ϕ as in (30), (35) with the choices (31), (34) for the initial condition ϕ . Then, for instance, one can show easily that (6) becomes

$$\sum_x x[\psi_0(x) - \psi_0(x-s) - \psi_\lambda(x) + \psi_\lambda(x-s)] = \sum_{i \geq 1} \langle d_{i,i+1} \rangle \lambda^i, \tag{74}$$

where s is the lattice spacing. For these systems on the lattice, there are new properties that can be investigated. As an example, if N is the number of points on the rightmost occupied site, then it is easy to check that

$$\langle \mu^N \rangle = 1 + \sum_x [\psi_{0\mu}(x, x-s) - H_\theta(x, t)]. \tag{75}$$

The whole distribution of N can then be determined by numerical integration. In the case of our example (C), the number N corresponds to the number of leaves at the deepest level in a binary search tree [36] and we found numerically that, at large times, $\text{Proba}(N = 2) \simeq 0.50$, $\text{Proba}(N = 4) \simeq 0.23$, $\text{Proba}(N = 6) \simeq 0.11$, $\text{Proba}(N = 8) \simeq 0.06$, etc.

6 Large Time Measure in the Frame of Lalley and Sellke

We recall from the results of Sect. 3.1, see (40) and (33), that the distribution of the rightmost point is given in the long time limit by

$$\lim_{t \rightarrow \infty} \text{Proba}[X_1(t) < m_t + x] = F(x), \tag{76}$$

where m_t is the average position of the rightmost point and is asymptotically given by (42) and where F is defined in (41). The property (76) is however non-ergodic, as shown by Lalley and Sellke [29] in the sense that for a single realization of the branching Brownian motion

$$[\text{fraction of the time that } X_1(t) < m_t + x] \neq F(x). \tag{77}$$

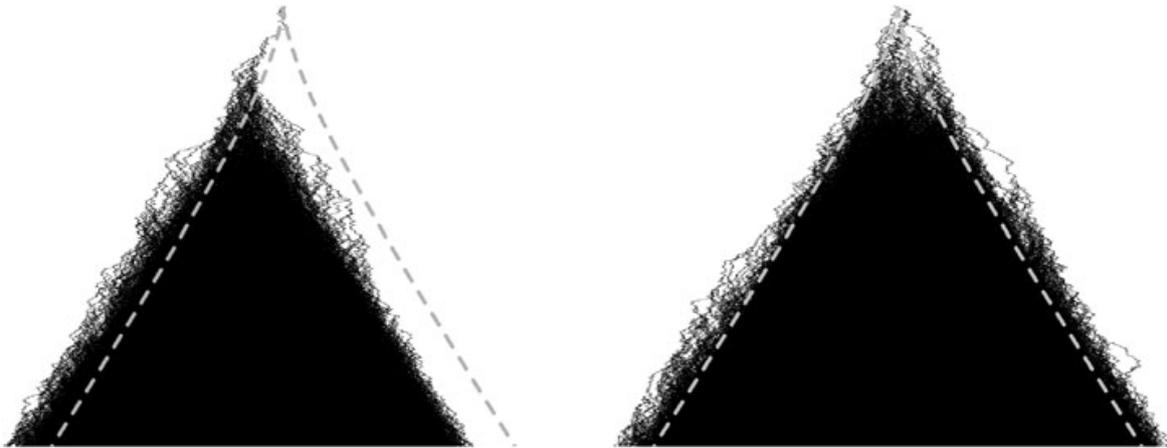


Fig. 5 Two realizations of the branching Brownian motion up to time $t = 20$. The horizontal direction represents space, and time increases downwards. The *dotted gray lines* are m_t and $-m_t$, the average positions of the rightmost and leftmost points, as measured from (36)

In fact, the left hand side of (77) is not self-averaging and depends on the realization. This is illustrated in Fig. 5: for the realization on the left, there were few branching events at early times and the first points wandered to the left, leading at larger times to an asymmetric picture. For the realization on the right of Fig. 5, there were many branching events early. For the right realization, the rightmost point is almost always on the right of m_t while it is almost always on the left of m_t in the left realization.

Visually, these strong memory effects of the early stages of the branching Brownian motion do not seem to decay with time, and it looks like the fluctuating right frontier of the system settles at some random fixed distance C from m_t :

$$X_1(t) = m_t + C + \eta_1(t) \quad \text{for large } t, \tag{78}$$

where C would depend on the realization but not the time, and where $\eta_1(t)$ would be a time-dependent random number centered around zero. A natural question is whether it is possible to define C for each realization in such a way that the distribution of $\eta_1(t)$ becomes in the long time limit independent of C and t , the idea being that the branching Brownian motion at long times seen from $m_t + C$ would “look the same” for any realization, whatever is the value of C .

A related question was addressed by Lalley and Sellke [29] in the following way: for each realization of the branching Brownian motion, define Z as

$$Z = \lim_{t \rightarrow \infty} Z_t \quad \text{where } Z_t = \sum_i [2t - X_i(t)] e^{X_i(t) - 2t}. \tag{79}$$

(The sum is over all the points $X_i(t)$ present at time t .) As shown in [29], Z_t has a limit Z for almost every realization; that limit is finite and positive. Lalley and Sellke prove then a limit theorem for the frontier of the branching Brownian motion which we interpret as follows:

$$\lim_{t \rightarrow \infty} \text{Proba}(X_1(t) < m_t + x | Z) = \exp(-AZe^{-x}) = \exp(-e^{-[x - \ln(AZ)]}), \tag{80}$$

where A is a constant related to the large x behavior of $F(x)$, see (102). This means that, if one considers only the realizations of the branching Brownian motion with a given value

of Z , then the large time distribution of the rightmost point is given by a Gumbel located around $m_t + \ln(AZ)$. To make the link with (78),

$$C = \ln(AZ), \quad \text{Proba}(\eta_1 < x) = \exp(-e^{-x}). \tag{81}$$

In Appendix C, we recall more precisely the theorem stated by Lalley and Sellke and we argue that (80) should be equivalent to their result.

A natural extension to Lalley’s and Sellke’s result is to write for all the points i

$$X_i(t) = m_t + \ln(AZ) + \eta_i(t), \tag{82}$$

as in (78), and ask whether the joint distribution of $\eta_1(t), \eta_2(t), \eta_3(t), \dots$ reaches a long time limit which is independent of the value of Z . We show in Appendix C that Lalley’s and Sellke’s result can be extended to all the generating functions H_ϕ . Our interpretation of this extension is

$$\lim_{t \rightarrow \infty} \left\langle \prod_i \phi(m_t + x - X_i(t)) \middle| Z \right\rangle = \exp(-AZe^{-x-f[\phi]}) = \exp(-e^{-[x-\ln(AZ)+f[\phi]]}), \tag{83}$$

where the delay function $f[\phi]$ is the same as in (45). By choosing $\phi = \theta$ (the step function), (83) reduces to (80). By choosing ϕ as in (31), one sees from (83) that the distribution of points at the right of the branching Brownian motion conditioned by Z reaches a long time limit where Z only appears through the global shift $\ln(AZ)$. This means that at large times, the distribution of the rightmost points in a branching Brownian motion has a well defined measure *independent of Z* located around $m_t + \ln(AZ)$.

As an example, if one chooses the function ϕ_1 defined by (34), one can easily show from (83) and (52) that, in the $m_t + \ln(AZ)$ frame, the average density of points at any position diverges in the long time limit.

6.1 Superposability Property

If one considers two branching Brownian motions a and b starting at arbitrary positions, then the points in a at large time will be characterized by a random value $Z^{(a)}$ and a realization of the point measure described by (83); idem for the points in b . If one considers the union of these two branching Brownian motions, one gets from (83)

$$\lim_{t \rightarrow \infty} \left\langle \prod_i \phi(m_t + x - X_i(t)) \middle| Z^{(a)}, Z^{(b)} \right\rangle = \exp(-AZe^{-x-f[\phi]}) = \exp(-e^{-[x-\ln(AZ)+f[\phi]]}), \tag{84}$$

with $Z = Z^{(a)} + Z^{(b)}$. This means that the point measure reached in the long time limit in the $m_t + \ln(AZ)$ frame is the same whether one started initially with one, two or, by extension, any finite number of initial points at arbitrary positions on the line. What does depend on the initial number of points is only the law of the random number Z , not the positions around $m_t + \ln(AZ)$. This is to be related to the discussion in Sect. 3.4, where we showed that, in the long time limit, the measure seen from m_t depends on the initial number of points while the measure seen from $X_1(t)$ does not.

Furthermore, the large time measure of the points in the $m_t + \ln(AZ)$ frame has the following property:

Starting with two realizations $\{\eta_i^{(a)}\}$ and $\{\eta_i^{(b)}\}$, then for any pair of real numbers α and β , the set of points $\{\eta_i^{(a)} + \alpha\} \cup \{\eta_i^{(b)} + \beta\}$ is another realization of the same measure shifted by $\ln(e^\alpha + e^\beta)$. (85)

(Think of $\{\eta_i^{(a)}\}$ as the offspring of a in the $m_t + \ln[AZ^{(a)}]$ frame and of α as $\ln Z^{(a)}$; idem for b . The shifts α and β are arbitrary because $Z^{(a)}$ and $Z^{(b)}$ are unbounded independent random numbers.)

The property (85) of the point measure in the $m_t + \ln(AZ)$ frame is a stronger version of the superposability property discussed in Sect. 3.4: clearly, it implies that the distribution of distances between the rightmost points is invariant by superposition, but it gives more information on the measure as it encompasses the position of the rightmost point. In particular, one can check that, in any such measure, the rightmost point is Gumbel distributed.

The simplest point measure with the property (85) is the Poisson process with an exponential density Ke^{-x} , for an arbitrary K . Furthermore, all the decorated measures of Sect. 2.3.3 when ν_1 is a Poisson process with an exponential density Ke^{-x} are also superposable measures. A natural question is then: can any superposable point measure be constructed as a decorated exponential Poisson process for a well chosen decoration measure? As shown recently [30], the answer is yes. In particular, the distribution of the rightmost points in a branching Brownian motion seen in the $m_t + \ln(AZ)$ frame converges to a decorated exponential Poisson process, and the decoration has a simple interpretation: it is known [4, 15, 19] that if one considers two points among the rightmost points in a branching Brownian motion at a large time t , then the time one needs to go back to find their most recent common ancestor is either very short (of order 1) or very long (of the order of the age t of the system). This means that one can group the rightmost points into families where two points belong to the same family if the branching event that generated them occurred recently. The branching Brownian motion can then be seen as a decorated exponential Poisson process, where the positions of all the families are distributed according to the Poisson process [4, 5] and where the members of a given family form the decoration. This interpretation helps to understand a question raised by our previous work [12]: the distances between points in the branching Brownian motion are given by (53), (54), but if one keeps for each family only its leader, then the distances between these leaders are given by (21) with $\alpha = 1$, as in the GREM or in the Ruelle cascade, see Sect. 2.3.4.

7 Conclusion

In the present work, we have shown that all the statistical properties of the rightmost points in a branching Brownian motion can be obtained by solving a front equation with a well-chosen initial condition. The distribution of the positions of the rightmost points seen in the frame m_t (the average position of the rightmost) has a long time limit. The properties of the limiting distribution can be expressed as the long time delays of the traveling wave solution of the Fisher-KPP equation when one varies the initial condition. This limiting distribution is however modified if one considers the union of several branching Brownian motions.

If one considers, however, only the distances between the points, for example if one looks at the distribution of all the positions of the rightmost points seen in the frame of the

rightmost one, one obtains at large times another limiting distribution which does not depend on the initial positions of the branching Brownian motions (as long as there are finitely many of them). We called this property superposability.

In Sect. 4 we have measured a few properties of this limiting distribution, and in Sect. 5 we explained how our results can be extended to more general branching random walks. Lastly, in Sect. 6, we argued that in the Lalley and Sellke frame, the branching Brownian motion satisfies a stronger version (85) of the superposability property, and that the distribution of points can be described by a decorated exponential.

In the future, it would be interesting to characterize more precisely the limiting measure of the branching Brownian motion and of the branching random walks to see whether some universal properties emerge. For example, we believe that the average density seen from the rightmost point should always grow as in (73). It would also be interesting to determine the properties of the decorating measure in Lalley’s and Sellke’s frame. The question of ergodicity, raised at the beginning of Sect. 6, is also an interesting open question.

Acknowledgements We would like to thank Julien Berestycki and Simon Harris for interesting discussions.

Appendix A: Numerical Simulations

We performed the numerical simulations by discretizing the branching Brownian motion in space (with a grid length δ) and time (with steps s) as in the example (D) of Sect. 5. The corresponding front equation is given by (70), which is of course the most straightforward discretization of the Fisher-KPP equation (27). The solutions to the discrete equation (70) converge to the solutions of the Fisher-KPP equation (27), if $s \rightarrow 0$ with δ/s^2 held constant and small enough.

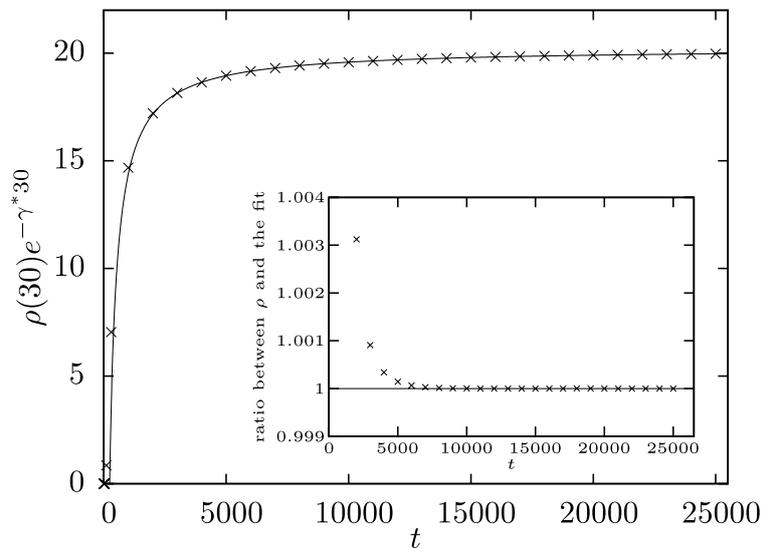
We used three sets of values for s and δ and computed the exact asymptotic velocity v^* of the front and the decay rate γ^* of the asymptotic shape $F(x)$ by minimizing numerically the function $v(\gamma)$ given in (71D). The decay rate β for the probability of observing a distance a between the two rightmost points, see (58), was computed using the recipe given at the end of Sect. 4.2 and explained in Appendix B. All these values are presented in Table 1.

The simulations were made on a finite but large domain centered around the position of the front; typically it extended to about a distance 1000 ahead and behind the center of the front (respectively 4000, 10000 or 20000 lattice sites depending on s). The values at $x = \pm\infty$ were exactly computed and used for the boundaries of the domain. Whenever the front moved by more than one unit space lattice, the whole data set was recentered. The simulations were performed up to large times of order 10000 (0.8 to 24 millions of time steps) and the data was extrapolated to obtain a value at $t = \infty$. To do this extrapolation, we

Table 1 Values of v^* , γ^* and β for our discretized branching processes, compared to the values in the Fisher-KPP case

s	δ	v^*	γ^*	β
0.25	$s^2/5 = 0.0125$	1.980480133	1.004581693	2.387337826
0.10	$s^2/5 = 0.002$	1.996840367	1.000747277	2.409772891
0.05	$s^2/6 \simeq 0.000416667$	1.999375296	1.000104046	2.412897517
	Fisher-KPP	2	1	$\sqrt{2} + 1 = 2.414213562$

Fig. 6 Average density at a distance 30 from the rightmost point as a function of time (symbols), fitted for $t \geq 5000$ by the function $A + B/t + C/t^{3/2}$ (line). The inset shows the quality of the fit by displaying the ratio between the data points and the fitting function



used a more precise asymptotic expansion of the position of the front than (42): according to [20],

$$m_t = v^*t - \frac{3}{2\gamma^*} \ln t + \text{Constant} + \frac{a_{1/2}}{t^{1/2}} + \frac{a_1}{t} + \frac{a_{3/2}}{t^{3/2}} + \dots, \tag{86}$$

where the number $a_{1/2}$ does not depend on the initial condition. As we measure the delay $m_t - m_t^{(\phi)}$, many terms cancel and one gets

$$m_t - m_t^{(\phi)} = f[\phi] + \frac{\delta a_1}{t} + \frac{\delta a_{3/2}}{t^{3/2}} + \dots. \tag{87}$$

All the quantities measured are derivatives of $f[\phi]$, see (47), and have therefore the same large time expansion as (87). Thus, we extrapolated our numerical data to the large time limit by fitting it with the function $A + B/t + C/t^{3/2}$ for times larger than (typically) 5000, see Fig. 6, and by using A as the end result.

On Fig. 3, the data points for the three values of s are shown together. On Fig. 4, we have drawn together for each value of s the function $\rho(a)e^{-\gamma^*a}$ using in each case the value of γ^* of Table 1. In both cases, the agreement between the three values of s is very good, and so we expect that on the scales of the figure, the curves would not change noticeably for smaller values of s and δ .

Appendix B: Distribution of the Distance Between the Two Rightmost Points

In this appendix we generalize, to any branching random walk, the argument leading to the asymptotic decay (58) of the distribution of the distance between the two rightmost points in the branching Brownian motion.

We consider a generic branching random walk in discrete space (with spacing s) and time (with intervals δ) defined by the following family of functions

$$p_n(r_1, \dots, r_n) = \left(\begin{array}{l} \text{The probability that a point at position } x \text{ branches} \\ \text{during a time step into } n \text{ points located at positions} \\ x + r_1, \dots, x + r_n. \end{array} \right). \tag{88}$$

We assume that $p_0 = 0$, so that there is no extinction. Then $p_1(r)$ can be thought as the probability that the point does not branch but moves by a distance r . The continuous time and/or space cases can be obtained as suitable $s \rightarrow 0$ and/or $\delta \rightarrow 0$ limits.

Let $\exp[tg(\beta)]$ be the generating function of the displacement during time t of one point *conditioned on the fact that this point does not branch*:

$$e^{tg(\beta)} = \sum_r e^{\beta r} \text{Proba}(\text{the point moves a distance } r \text{ without branching during time } t). \tag{89}$$

As the time steps are independent, the function $g(\beta)$ can be computed during the time interval δ which gives

$$e^{\delta g(\beta)} = \sum_r p_1(r) e^{\beta r}. \tag{90}$$

Note that $g(0) < 0$ as soon as the branching probability is non-zero. We want now to evaluate the probability that a point moves a distance r , without branching, during time t . For large t , it takes the form

$$\text{Proba}(\text{the point moves a distance } r \text{ without branching during time } t) \sim \exp\left[tf\left(\frac{r}{t}\right)\right], \tag{91}$$

where $f(c)$ is a large deviation function. Using (89), one finds that $f(c)$ and $g(\beta)$ are related by a Legendre transform

$$\begin{cases} \beta = -f'(c), \\ g(\beta) = f(c) + \beta c. \end{cases} \tag{92}$$

Now, assuming as in Sect. 4.2 that the events which contribute most to a large distance a between the two rightmost points are those where the rightmost point moves, without branching, a distance a ahead of the frontier of the branching Brownian motion, one gets

$$P_{12}(a) \sim \max_{\tau} \left\{ \exp\left[\tau f\left(\frac{a + v^* \tau}{\tau}\right)\right] \right\}, \tag{93}$$

where v^* is the velocity of the front. For large a , the optimal τ is also large and it satisfies, by derivation,

$$f\left(\frac{a}{\tau} + v^*\right) - \frac{a}{\tau} f'\left(\frac{a}{\tau} + v^*\right) = 0. \tag{94}$$

Let $c = a/\tau + v^*$. Using (92), (94) becomes

$$g(\beta) = \beta v^*. \tag{95}$$

Remarkably, this equation does not depend on a . Replacing into (93) gives

$$P_{12}(a) \sim e^{-\beta a}. \tag{96}$$

The asymptotic decay rate of the probability distribution function of the distance between the two rightmost points is therefore simply the positive solution β of (95) with $g(\beta)$ given by (90). In the branching Brownian motion, $g(\beta) = \beta^2 - 1$, $v^* = 2$, so that (95) gives indeed $\beta = 1 + \sqrt{2}$.

As can be checked easily from (90), the function $g(\beta)$ is convex. Therefore, as $g(0) < 0$, (95) has at most one positive solution and at most one negative solution. The positive solution is the relevant one here.

We are now going to show that the negative solution of (95), if it exists, gives the asymptotic shape of the associated traveling wave. We write the front equation associated to the branching point process (88). Using the same method as in Sect. 3, we find

$$H_\phi(x, t + \delta) = \sum_{n \geq 1} \sum_{r_1, \dots, r_n} p_n(r_1, \dots, r_n) \prod_{i=1}^n H_\phi(x - r_i). \tag{97}$$

We look at the shape F of the traveling wave solution for a step initial condition, which moves asymptotically at the velocity v^* :

$$H_\phi(x, t) = F(x - v^*t). \tag{98}$$

Using (97), we see that in the *stable* region $F(x) \ll 1$ one has

$$F(x - v^*\delta) = \sum_r p_1(r)F(x - r) + \mathcal{O}(F^2). \tag{99}$$

We look for an exponential solution to this linearized equation: $F(x) \simeq e^{\lambda x}$ with $\lambda > 0$, as $F(-\infty) = 0$. (Note that a periodic modulation of this exponential could occur as r takes only discrete values.) Inserting into (99) and using (90), one finds that the equation for λ is

$$g(-\lambda) = -\lambda v^*, \tag{100}$$

which is the same equation as (95) for $\lambda = -\beta$.

To summarize, a *positive* solution to (95) gives the exponential decay rate of the probability distribution of the distance between the two rightmost points, see (96), while a *negative* solution gives the coefficient $-\lambda$ governing the shape of the front $F(x)$ in the stable region $F(x) \ll 1$.

Appendix C: Lalley’s and Sellke’s Result

Lalley’s and Sellke’s theorem [29] is

$$\lim_{s \rightarrow \infty} \lim_{t \rightarrow \infty} \text{Proba}(X_1(t) < m_t + x | \{X_i(s)\}) = \exp(-AZe^{-x}), \tag{101}$$

where A is the constant appearing [11] in the large x expansion of the function $F(x)$ defined in (41)

$$F(x) \simeq 1 - (Ax + B)e^{-x} \quad \text{for large } x, \tag{102}$$

and Z is defined in (79). In words, given the positions $\{X_i(s)\}$ at time s , there is a $t \rightarrow \infty$ limit to the probability that the rightmost is on the left of $m_t + x$ which depends, obviously, on the $\{X_i(s)\}$ and is as such a random variable. As s goes to infinity, this random variable converges almost surely to the Gumbel distribution around $\ln(AZ)$.

This result can be extended into the following: for any suitable function ϕ (see Sect. 3.2), one has

$$\lim_{s \rightarrow \infty} \lim_{t \rightarrow \infty} \left\langle \prod_i \phi[m_t + x - X_i(t)] \middle| \{X_i(s)\} \right\rangle = \exp(-AZe^{-x-f[\phi]}), \tag{103}$$

where $f[\phi]$ is the delay function (45). For $\phi = \theta$, (103) reduces to (101).

We now give an outline of Lalley’s and Sellke’s proof applied to the case (103). Given the positions $X_i(s)$ of the points at time s , the system as time $t > s$ can be seen as a collection of independent branching Brownian motions at time $t - s$ starting from the $X_i(s)$. Therefore

$$\left\langle \prod_i \phi[x - X_i(t)] \middle| \{X_i(s)\} \right\rangle = \prod_i H_\phi(x - X_i(s), t - s), \tag{104}$$

where the product in the right hand side is made on all the points present at time s .

We replace x by $m_t + x$, to center around the position of the front, and suppose t large. It is easy to see from Bramson’s formula (42) that $m_t = m_{t-s} + 2s + o(1)$ as t becomes large, so that

$$\left\langle \prod_i \phi[m_t + x - X_i(t)] \middle| \{X_i(s)\} \right\rangle = \prod_i H_\phi(m_{t-s} + 2s + x - X_i(s) + o(1), t - s), \tag{105}$$

and, using (45),

$$\lim_{t \rightarrow \infty} \left\langle \prod_i \phi[m_t + x - X_i(t)] \middle| \{X_i(s)\} \right\rangle = \prod_i F(2s + x - X_i(s) + f[\phi]). \tag{106}$$

We now take s large. Of all the points present at time s , the rightmost is around $2s - \frac{3}{2} \log s$, see (42). Therefore, $2s - X_i(s)$ diverges for all i . Using (102),

$$\begin{aligned} & \lim_{t \rightarrow \infty} \left\langle \prod_i \phi[m_t + x - X_i(t)] \middle| \{X_i(s)\} \right\rangle \\ & \simeq \exp\left(-\sum_i [A(2s + x - X_i(s) + f[\phi]) + B] e^{-2s - x + X_i(s) - f[\phi]}\right). \end{aligned} \tag{107}$$

Following Lalley and Sellke, we introduce the quantities

$$Y_s = \sum_i e^{-2s + X_i(s)}, \quad Z_s = \sum_i [2s - X_i(s)] e^{-2s + X_i(s)}, \tag{108}$$

see (79), so that

$$\lim_{t \rightarrow \infty} \left\langle \prod_i \phi[m_t + x - X_i(t)] \middle| \{X_i(s)\} \right\rangle \simeq \exp(-[AZ_s + (Ax + Af[\phi] + B)Y_s] e^{-x - f[\phi]}). \tag{109}$$

Finally, the most technical part of Lalley’s and Sellke’s proof is that Y_s and Z_s are martingales converging when $s \rightarrow \infty$ to $\lim_{s \rightarrow \infty} Y_s = 0$ and $\lim_{s \rightarrow \infty} Z_s = Z > 0$ respectively, which leads to (103). We do not reproduce this part of the proof here as it does not concern our extension with the function ϕ and it works in (103) exactly as in (101).

In (103), the average is made on all the realizations with a given set $\{X_i(s)\}$ of points at a large time s but the only relevant quantity appearing in the generating function (109) is Z_s . One would obviously have reached the same result if one had conditioned by Z_s instead of by the $\{X_i(s)\}$. Furthermore, as Z_s converges quickly to Z , as illustrated on Fig. 5, we argue that conditioning by Z_s at a large time s or directly conditioning by Z should be equivalent, hence (80), (83).

References

1. Aïdékon, E.: Convergence in law of the minimum of a branching random walk. [arXiv:1101.1810](https://arxiv.org/abs/1101.1810) [math.PR] (2011)
2. Aizenman, M., Sims, R., Starr, S.L.: Mean-field spin glass models from the cavity-ROSt perspective. *Contemp. Math.* **437**, 1–30 (2007)
3. Arguin, L.P.: Spin glass computations and Ruelle’s probability cascades. *J. Stat. Phys.* **126**, 951–976 (2007)
4. Arguin, L.P., Bovier, A., Kistler, N.: The genealogy of extremal particles of branching Brownian motion. [arXiv:1008.4386](https://arxiv.org/abs/1008.4386) [math.PR] (2010)
5. Arguin, L.P., Bovier, A., Kistler, N.: Poissonian statistics in the extremal process of branching Brownian motion. [arXiv:1010.2376](https://arxiv.org/abs/1010.2376) [math.PR] (2010)
6. Bachmann, M.: Limit theorems for the minimal position in a branching random walk with independent logconcave displacements. *Adv. Appl. Probab.* **32**, 159–176 (2000)
7. Bolthausen, E., Sznitman, A.S.: On Ruelle’s probability cascades and an abstract cavity method. *Commun. Math. Phys.* **197**, 247–276 (1998)
8. Bouchaud, J.P., Mézard, M.: Universality classes for extreme-value statistics. *J. Phys. A* **30**, 7997–8015 (1997)
9. Bovier, A., Kurkova, I.: A tomography of the GREM: Beyond the REM conjecture. *Commun. Math. Phys.* **263**, 535–552 (2006)
10. Bovier, A., Kurkova, I.: Local energy statistics in spin glasses. *J. Stat. Phys.* **126**, 933–949 (2007)
11. Bramson, M.D.: Convergence of Solutions of the Kolmogorov Equation to Traveling Waves. *Mem. Am. Math. Soc.* vol. 44(285) (1983)
12. Brunet, É., Derrida, B.: Statistics at the tip of a branching random walk and the delay of traveling waves. *Europhys. Lett.* **87**, 60010 (2009)
13. Burkhardt, T.W., Györgyi, G., Moloney, N.R., Racz, Z.: Extreme statistics for time series: distribution of the maximum relative to the initial value. *Phys. Rev. E* **76**, 041119 (2007)
14. Chauvin, B., Rouault, A.: Supercritical branching Brownian motion and K-P-P Equation in the critical speed-area. *Math. Nachr.* **149**, 41–59 (1990)
15. Chauvin, B., Rouault, A.: Boltzmann-Gibbs weights in the branching random walk. *IMA Vol. Math. Its Appl.* **84**, 41–50 (1997)
16. Dean, D.S., Majumdar, S.N.: Extreme-value statistics of hierarchically correlated variables deviation from Gumbel statistics and anomalous persistence. *Phys. Rev. E* **64**, 046121 (2001)
17. Derrida, B.: Random-energy model: an exactly solvable model of disordered systems. *Phys. Rev. B* **24**, 2613–2626 (1981)
18. Derrida, B.: A generalization of the random energy model which includes correlations between energies. *J. Phys. Lett.* **46**, L401–L407 (1985)
19. Derrida, B., Spohn, H.: Polymers on disordered trees, spin glasses and traveling waves. *J. Stat. Phys.* **51**(5/6), 817–840 (1988)
20. Ebert, U., van Saarloos, W.: Front propagation into unstable states: universal algebraic convergence towards uniformly translating pulled fronts. *Physica D* **146**, 1–99 (2000)
21. Fisher, R.A.: The wave of advance of advantageous genes. *Ann. Eugen.* **7**, 355–369 (1937)
22. Fyodorov, Y.V., Bouchaud, J.-P.: Freezing and extreme-value statistics in a random energy model with logarithmically correlated potential. *J. Phys. A* **41**(37), 372001–372012 (2008)
23. Golding, I., Kozlovsky, Y., Cohen, I., Ben-Jacob, E.: Studies of bacterial branching growth using reaction-diffusion models for colonial development. *Physica A* **260**, 510–554 (1998)
24. Györgyi, G., Moloney, N.R., Ozogany, K., Racz, Z.: Finite-size scaling in extreme statistics. *Phys. Rev. Lett.* **100**, 210601 (2008)
25. Hu, Y., Shi, Z.: Minimal position and critical martingale convergence in branching random walks, and directed polymers on disordered trees. *Ann. Probab.* **37**(2), 742–789 (2009)
26. Igloi, F., Monthus, C.: Strong disorder RG approach of random systems. *Phys. Rep.* **412**, 277–431 (2005)
27. Kessler, D.A., Levine, H., Ridgway, D., Tsimring, L.: Evolution on a smooth landscape. *J. Stat. Phys.* **87**(3/4), 519–544 (1997)
28. Kolmogorov, A., Petrovsky, I., Piscounov, N.: Étude de l’équation de la diffusion avec croissance de la quantité de matière et son application à un problème biologique. *Bull. Univ. État Mosc. A* **1**(6), 1–25 (1937)
29. Lalley, S.P., Sellke, T.: A conditional limit theorem for the frontier of a branching Brownian motion. *Ann. Probab.* **15**, 1052–1061 (1987)
30. Maillard, P.: A characterisation of superposable random measures. [arXiv:1102.1888](https://arxiv.org/abs/1102.1888) [math.PR] (2011)
31. Majumdar, S.N., Dean, D.S., Krapivsky, P.L.: Understanding search trees via statistical physics. *Pramana J. Phys.* **64**(6) (2005)

32. Majumdar, S.N., Krapivsky, P.L.: Extreme value statistics and traveling fronts: application to computer science. *Phys. Rev. E* **65**, 036127 (2002)
33. McKean, H.P.: Applications of Brownian motion to the equation of Kolmogorov-Petrovski-Piscounov. *Commun. Pure Appl. Math.* **28**, 323–331 (1975)
34. Mézard, M., Parisi, G., Sourlas, N., Toulouse, G., Virasoro, M.A.: Replica symmetry-breaking and the nature of the spin-glass phase. *J. Phys.* **45**, 843–854 (1984)
35. Monthus, C., Garel, T.: On the critical weight statistics of the random energy model and of the directed polymer on the Cayley tree. *Phys. Rev. E* **75**(5), 051119 (2007)
36. Roberts, M.I.: Almost sure asymptotics for the random binary search tree. [arXiv:1002.3896](https://arxiv.org/abs/1002.3896) [math.PR] (2010)
37. Ruelle, D.: A mathematical reformulation of Derrida's REM and GREM. *Commun. Math. Phys.* **108**, 225–239 (1987)
38. Ruzmaikina, A., Aizenman, M.: Characterization of invariant measures at the leading edge for competing particle systems. *Ann. Probab.* **33**, 82–113 (2005)
39. van Saarloos, W.: Front propagation into unstable states. *Phys. Rep.* **386**(2–6), 29–222 (2003)
40. Sabhapandit, S., Majumdar, S.N.: Density of near-extreme events. *Phys. Rev. Lett.* **98**, 140201 (2007)
41. Tracy, C.A., Widom, H.: Level-spacing distributions and the Airy kernel. *Commun. Math. Phys.* **159**, 151–175 (1994)



An Exactly Solvable Travelling Wave Equation in the Fisher–KPP Class

Éric Brunet¹ · Bernard Derrida²

Received: 22 June 2015 / Accepted: 5 August 2015 / Published online: 8 September 2015
© Springer Science+Business Media New York 2015

Abstract For a simple one dimensional lattice version of a travelling wave equation, we obtain an exact relation between the initial condition and the position of the front at any later time. This exact relation takes the form of an inverse problem: given the times t_n at which the travelling wave reaches the positions n , one can deduce the initial profile. We show, by means of complex analysis, that a number of known properties of travelling wave equations in the Fisher–KPP class can be recovered, in particular Bramson’s shifts of the positions. We also recover and generalize Ebert–van Saarloos’ corrections depending on the initial condition.

Keywords Fisher–KPP · Front equation · Travelling wave

1 Introduction

The study of the solutions of partial differential equations describing a moving interface from a stable to an unstable medium is a classical subject [1–5] in mathematics, theoretical physics and biology [6–9]. The prototype of such equations is the Fisher–KPP equation (after Fisher [10] and Kolmogorov–Petrovskii–Piskunov [11])

$$\frac{\partial u}{\partial t} = \frac{\partial^2 u}{\partial x^2} + f(u), \quad (1)$$

where the field u satisfies $0 \leq u(x, t) \leq 1$ and where $f(u) \geq 0$. The unstable medium corresponds to $u = 0$ (i.e. $f(0) = 0$ and $f'(0) > 0$) and the stable one to $u = 1$ (i.e. $f(1) = 0$ and $f'(1) < 0$).

✉ Éric Brunet
eric.brunet@lps.ens.fr

Bernard Derrida
bernard.derrida@lps.ens.fr

¹ LPS-ENS, UMR 8550, CNRS, UPMC Univ Paris 06, Sorbonne Universités, 75005 Paris, France

² Collège de France, 75005 Paris, France

One can show that equations of type (1) exhibit a continuous family W_v of travelling wave solutions

$$u(x, t) = W_v(x - vt) \tag{2}$$

indexed by their velocities v . Explicit expressions of the travelling waves are in general not known except for particular velocities [12]. The best known example, due to Ablowitz and Zeppetella [13], is $u = [1 + c \exp[(x - vt)/\sqrt{6}]]^{-2}$ for the Fisher–KPP equation (1) with $f(u) = u - u^2$ and $v = 5/\sqrt{6}$.

Apart from describing the shapes of these travelling wave solutions (2), a central question is to understand how the long time behavior of the solutions of (1) depends on the initial condition $u(x, 0)$. In general this asymptotic regime is controlled by the rate of the exponential decay of this initial condition. A brief review of the properties of the travelling wave solutions of (1) and on the way the position and the asymptotic velocity of the solution depend on the initial condition is given in Sect. 2.

In the present paper we study a simple one dimensional lattice version of a travelling wave equation. In this lattice version we associate to each lattice site $n \in \mathbb{Z}$ a positive number $h_n(t)$ which plays the role of the field $u(x, t)$ and these $h_n(t)$ evolve according to

$$\frac{dh_n(t)}{dt} = \begin{cases} ah_{n-1}(t) + h_n(t) & \text{if } 0 \leq h_n(t) < 1, \\ 0 & \text{if } h_n(t) \geq 1. \end{cases} \tag{3}$$

We see that the evolution of $h_n(t)$ is linear except for the saturation at $h_n(t) = 1$ which is the only non-linearity in the problem. This saturation simply means that whenever $h_n(t)$ reaches the value 1, it keeps this value forever. The evolution (3) therefore combines linear growth, spreading (or diffusion) because of the coupling between neighboring sites, and saturation, very much like in Fisher–KPP equation (1).

The aim of this paper is to show that the evolution (3) leads to behaviors very similar to those expected for the usual Fisher–KPP equation (1). Moreover a number of properties of the solutions of (3) are easier to determine than for the original Fisher–KPP equation (1). Our approach is essentially based on the exact relation (32) derived in Sect. 3 which relates the times t_n at which $h_n(t)$ reaches 1 for the first time to the initial condition $h_n(0)$. We show in Sect. 4 that from (32) one can obtain a precise description of the shape of the travelling wave solutions, in particular explicit formulas for their asymptotic decay. We also show in Sect. 5 that (3) shares with the Fisher–KPP equation most of the properties expected for the dependence of the position of the front on the initial condition. Our results are summarized in Sect. 6.

2 Some Known Properties of the Fisher–KPP Class

In this section we briefly recall some properties of the Fisher–KPP equation.

2.1 The Travelling Waves

For the Fisher–KPP equation (1) the shape $W_v(x)$ of the travelling wave (2) satisfies an ordinary differential equation

$$W_v'' + vW_v' + f(W_v) = 0 \tag{4}$$

with the boundary conditions $W_v(-\infty) = 1$ and $W_v(+\infty) = 0$. By linearizing (4) for small W_v (when x is large),

$$W_v'' + vW_v' + f'(0)W_v = 0, \tag{5}$$

one can see that, generically, $W_v(x)$ vanishes exponentially as $x \rightarrow \infty$

$$W_v(x) \sim e^{-\gamma x}, \tag{6}$$

with γ related to the speed v of the travelling wave by

$$v(\gamma) = \gamma + \frac{f'(0)}{\gamma}. \tag{7}$$

This relation shows that depending on v , the rate γ of the exponential decay is either real or complex, and these two regimes are separated by a critical velocity v_c where $v(\gamma)$ is minimum

$$v_c = v(\gamma_c) \quad \text{where} \quad v'(\gamma_c) = 0. \tag{8}$$

With $v(\gamma)$ given by (7), one gets $v_c = 2\gamma_c$ and $\gamma_c = \sqrt{f'(0)}$. Under certain conditions on the function $f(u)$ (such as $0 \leq f(u) \leq uf'(0)$ for all u see [4,5,14] and references therein), it is known that:

- For $0 < v < v_c$, the solutions γ of the equation $v(\gamma) = v$ are complex. The corresponding travelling waves solutions of (4) oscillate around 0 while decaying as $x \rightarrow \infty$.
- For $v > v_c$, the travelling wave is monotonically decreasing and decays for large x as

$$W_v(x) \simeq A e^{-\gamma_1 x} \quad \text{with} \quad A > 0, \tag{9}$$

where γ_1 is the smallest solution of $v(\gamma) = v$.

- For $v = v_c$, the equation $v(\gamma) = v_c$ has a double root γ_c and the travelling wave is monotonically decreasing and decays for large x as

$$W_{v_c}(x) \simeq A x e^{-\gamma_c x} \quad \text{with} \quad A > 0. \tag{10}$$

Remark 1 The facts (9) and (10) for $v \geq v_c$ are not obvious and cannot be understood from the linearized equation (5) only. These are properties of the full non-linear equation (4), which can be proved under known conditions on the non-linearity $f(u)$ (such as $0 \leq f(u) \leq uf'(0)$). Fronts which satisfy these properties are called *pulled fronts*.

For well tuned non-linearities (which fail to satisfy these conditions), travelling waves for $v \geq v_c$ might not be monotone and the asymptotics (9) and (10) might be modified; for instance in (9), depending on the value of v , one could have $A < 0$ or a decay in $\exp(-\gamma_2 x)$ where γ_2 is the largest solution of $v(\gamma) = v$. Rather than (10), one could have $A \exp(-\gamma_c x)$ without the x prefactor. In all these cases, the front equation is then said to be *pushed* [14, 15].

2.2 The Selection of the Velocity

The travelling waves W_v solutions of (4) move at a constant speed with a time independent shape. For general initial conditions $u(x, 0)$, the shape of the solution is time-dependent and the question of the selection of the speed is to predict the asymptotic shape and velocity of the solution $u(x, t)$ in the long time limit. For initial profiles decreasing from $u(-\infty, 0) = 1$ to $u(+\infty, 0) = 0$ it is known since the works of Bramson [4,5,9,14,16] under which conditions the shape of the solution $u(x, t)$ converges to a travelling wave W_v solution of (4) in the sense that one can find a displacement X_t such that

$$u(X_t + x, t) \rightarrow W_v(x), \quad \text{with} \quad \frac{X_t}{t} \rightarrow v. \tag{11}$$

In particular it is known that if the initial condition $u(x, 0)$ satisfies for large x :

- $u(x, 0) \sim e^{-\gamma x}$ with $0 < \gamma < \gamma_c$,
then the asymptotic velocity is $v(\gamma)$, the asymptotic shape is $W_{v(\gamma)}$ and

$$X_t = v(\gamma)t + \text{Cst.} \tag{12}$$

- $u(x, 0) \ll x^\alpha e^{-\gamma_c x}$ for some $\alpha < -2$ (in particular for step initial conditions),
the asymptotic velocity is $v_c = v(\gamma_c)$, the asymptotic shape is W_{v_c} and

$$X_t = v_c t - \frac{3}{2\gamma_c} \ln t + \text{Cst.} \tag{13}$$

- $u(x, 0) \sim x^\alpha e^{-\gamma_c x}$ with $\alpha \geq -2$,
the asymptotic velocity v_c and shape W_{v_c} are the same as in the previous case but the logarithmic correction to the position X_t is modified:

$$X_t = v_c t - \frac{1 - \alpha}{2\gamma_c} \ln t + \text{Cst} \quad \text{for } \alpha > -2, \tag{14}$$

$$X_t = v_c t - \frac{3}{2\gamma_c} \ln t + \frac{1}{\gamma_c} \ln \ln t + \text{Cst} \quad \text{for } \alpha = -2. \tag{15}$$

(Initial conditions decaying too slowly would not lead to a travelling wave.)

Notice that the solutions W_v of (4) can always be translated along the x axis, so the ‘‘Cst’’ in (12–15) depends on the particular solution of (4) that was chosen. It is often convenient to single out one particular solution W_v of (4): for example one may select the solution such that $W_v(0) = 1/2$ or such that $\int x W'_v(x) dx = 0$. Once a particular prescription for W_v is chosen, the ‘‘Cst’’ in the equations above is well defined. It can be computed in some cases such as (12), but its analytic expression is not known in some other cases such as (13).

2.3 Vanishing Corrections

The convergence property (11) does not allow to define the displacement X_t to better than a constant: if X_t satisfies (11), then $X_t + o(1)$ also satisfies (11). It is however quite natural to choose a particular X_t , which one might call the position of the front. A possible choice could be

$$u(X_t, t) = c, \tag{16}$$

where $c \in (0, 1)$ is a fixed given number. Another possible choice would be to interpret $-\partial u / \partial x$ as a probability density and pick X_t as its expectation:

$$X_t = - \int dx x \frac{\partial u}{\partial x}. \tag{17}$$

Either definition (16) or (17) gives a position X_t which satisfies (11). With such a precise definition for X_t as (16) or (17), it makes sense to try to improve on (12–15) and determine higher order corrections. Ebert and van Saarloos [17, 18] have claimed that for steep enough initial conditions, the first correction to (13) is of order $t^{-1/2}$ and is universal: it depends neither on the initial condition, nor on the choice of (16) or (17), nor on the value c in (16), nor on the non-linearities. They found that

$$X_t = v_c t - \frac{3}{2\gamma_c} \ln t + \text{Cst} - 3 \sqrt{\frac{2\pi}{\gamma_c^5 v''(\gamma_c)}} t^{-1/2} + \dots \tag{18}$$

2.4 The Fisher–KPP Class

The main ingredients of the Fisher–KPP equation (1) which lead to travelling waves and fronts converging to those travelling waves are a diffusive term, a growth term and a saturation term. There exist many equations with the same ingredients which share the above properties (8–18) of the Fisher–KPP equation: the equation satisfied by the travelling waves (4), the dispersion relation (7) and the values of γ_c and v_c are modified, but everything else remains the same.

To give an example which appears in the problem of directed polymers on a tree [6], let us consider an evolution equation of the type

$$G(x, t + 1) = \int G(x + \epsilon)^B \rho(\epsilon) d\epsilon. \quad (19)$$

(In the directed polymers context, B is the branching ratio on the tree and $\rho(\epsilon)$ is the distribution of the random energies associated to edges of the tree). Then $u(x, t) = 1 - G(x, t)$ satisfies a discrete time evolution equation with an unstable uniform solution $u = 0$ and a stable one $u = 1$ as in (1). Even though (1) is continuous in time while (19) is discrete, they have similar properties: travelling waves for (19) are solutions of

$$W_v(x - v) = \int W_v(x + \epsilon)^B \rho(\epsilon) d\epsilon \quad (20)$$

instead of (4). By linearizing the evolution of $G(x, t)$ around the unstable uniform solution $G = 1$ and by looking for travelling wave solutions of this linearized equation of the form $1 - G(x, t) \sim \exp[-\gamma(x - v(\gamma)t)]$, one gets a new dispersion equation which replaces (7):

$$v(\gamma) = \frac{1}{\gamma} \ln \left[B \int e^{\gamma\epsilon} \rho(\epsilon) d\epsilon \right], \quad (21)$$

but all the above behaviors (8–18) remain valid with v_c and γ_c computed from (21) and (8). For example, for $B = 2$ and a uniform $\rho(\epsilon)$ on the unit interval (i.e. $\rho(\epsilon) = 1$ for $0 < \epsilon < 1$ and $\rho(\epsilon) = 0$ elsewhere) one gets, $v(\gamma) = \frac{1}{\gamma} \ln \left[\frac{2}{\gamma} (e^\gamma - 1) \right]$ which leads to $v_c \simeq 0.815172$ and $\gamma_c \simeq 5.26208$.

Example (19) is a front equation where time is discrete. One could also consider travelling wave equations where space is discrete, say $x \in \mathbb{Z}$. For instance, one could discretize the Laplacian in (1) or take (19) with a distribution $\rho(\epsilon)$ concentrated on integer values of ϵ . When space is discrete, special care should be taken: it is clear from (2) that while the front $u(x, t)$ lives on the lattice, the travelling wave $W_v(x)$ is defined for all real values x , and even when (2) holds, the shape of the front $W_v(x - vt)$ measured on the lattice evolves periodically in time with a period $1/v$. Furthermore, the convergence (11) no longer makes any sense. One can still try to define a specific position of X_t by something like the following generalization of (17):

$$X_t = \sum_{x \in \mathbb{Z}} x [u(x, t) - u(x + 1, t)], \quad (22)$$

but with such a definition, even if the front is given by the travelling wave $W_v(x - vt)$, the difference $X_t - vt$ is no longer constant but becomes a periodic function in time because the shape of the front on the lattice evolves also periodically. Similarly, in the discrete space case, the Cst term in all the asymptotics (12–15) is in general replaced by a periodic function of time.

An alternative way to locate the front when time is continuous and space is discrete is to invert the roles of x and t : instead of defining X_t by $u(X_t, t) = c$ as in (16), one can define

t_x as the first time when the front at a given position x reaches a certain level c :

$$u(x, t_x) = c. \tag{23}$$

Note that when time and space are continuous, the functions X_t and t_x are reciprocal and one can write (12–15) as

$$t_x = \frac{x}{v(\gamma)} + \text{Cst}', \quad \text{for } u(x, 0) \sim e^{-\gamma x} \text{ with } 0 < \gamma < \gamma_c, \tag{24}$$

$$t_x = \frac{x}{v_c} + \frac{3}{2\gamma_c v_c} \ln x + \text{Cst}', \quad \text{for } u(x, 0) \ll x^\alpha e^{-\gamma_c x} \text{ for some } \alpha < -2, \tag{25}$$

$$t_x = \frac{x}{v_c} + \frac{1-\alpha}{2\gamma_c v_c} \ln x + \text{Cst}', \quad \text{for } u(x, 0) \sim x^\alpha e^{-\gamma_c x} \text{ for } \alpha > -2, \tag{26}$$

$$t_x = \frac{x}{v_c} + \frac{3}{2\gamma_c v_c} \ln x - \frac{1}{\gamma_c v_c} \ln \ln x + \text{Cst}', \quad \text{for } u(x, 0) \sim x^{-2} e^{-\gamma_c x}, \tag{27}$$

and, for steep enough initial conditions, one can write (18) as

$$t_x = \frac{x}{v_c} + \frac{1}{\gamma_c v_c} \left[\frac{3}{2} \ln x + \text{Cst}' + 3 \sqrt{\frac{2\pi v_c}{\gamma_c^3 v''(\gamma_c)}} x^{-1/2} + \dots \right]. \tag{28}$$

The main advantage of (24–28) over (12–15,18) is that they still make sense when space is discrete (with a real constant Cst' , not a periodic function of time). We will see that they remain valid for our lattice model (3).

3 The Key Formula for the Position of the Front

In this section we consider the front $h_n(t)$ defined by (3) and we establish relation (32) between the initial condition $h_n(0)$ and the first times t_n at which $h_n(t)$ reaches the value 1. Here we limit our discussion to the case $a > 0$ and to initial conditions of the form

$$h_n(0) = \begin{cases} 1 & \text{for } n \leq 0, \\ k_n & \text{for } n \geq 1, \end{cases} \tag{29}$$

where the k_n are non-negative, smaller than 1 and non-increasing, i.e.

$$1 > k_1 \geq k_2 \geq k_3 \geq \dots \geq 0. \tag{30}$$

Clearly, as $a > 0$, for a monotonic initial condition (30), the solution $h_n(t)$ of (3) remains monotonic at any later time. One can define t_n as the time when $h_n(t)$ reaches 1 for the first time (i.e. $h_n(t) = 1$ for $t \geq t_n$ while $h_n(t) < 1$ for $t < t_n$). The monotonicity (30) of the initial condition implies the monotonicity of the times t_n

$$0 < t_1 < t_2 < \dots < t_n < \dots \tag{31}$$

Most of the properties of the solutions of (3) with the initial conditions (29) discussed in this paper will be based on the following exact formula

$$\sum_{n=1}^{\infty} k_n \lambda^n = -\frac{a\lambda}{1+a\lambda} + \frac{a+1}{1+a\lambda} \sum_{n=1}^{\infty} e^{-(1+a\lambda)t_n} \lambda^n, \quad (32)$$

which relates the generating function of the initial condition $\{k_n\}$ to the times $\{t_n\}$.

Formula (32) can be derived as follows. If one defines the generating functions

$$H_m(t) = \sum_{n \geq m} h_n(t) \lambda^{n-m+1}, \quad (33)$$

one can see that for $t_{m-1} \leq t \leq t_m$ (with the convention that $t_0 = 0$) the evolution of $H_m(t)$ is given by

$$\frac{dH_m(t)}{dt} = (1+a\lambda)H_m(t) + a\lambda. \quad (34)$$

This of course can be easily solved to give

$$H_m(t) = -\frac{a\lambda}{1+a\lambda} + \Phi_m e^{(1+a\lambda)t}, \quad (35)$$

where the Φ_m 's are constants of integration. These Φ_m 's can be determined by matching the solutions at times $0, t_1, t_2, \dots$:

$$H_1(0) = \sum_{n=1}^{\infty} k_n \lambda^n, \quad H_m(t_m) = \lambda(1 + H_{m+1}(t_m)), \quad (36)$$

and one gets that for $t_{m-1} \leq t \leq t_m$

$$H_m(t) = -\frac{a\lambda}{1+a\lambda} + \frac{a+1}{1+a\lambda} \sum_{n=m}^{\infty} e^{(1+a\lambda)(t-t_n)} \lambda^{n+1-m}. \quad (37)$$

Then (32) follows by taking $m = 1$ and $t = 0$ in (37).

Remark that formula (32) appears as the solution of a kind of inverse problem: given the times t_n , one can compute the initial profile k_n by expanding in powers of λ . This gives expressions of k_n in terms of the times t_m 's for $m \leq n$. Alternatively one can determine the times t_n in terms of the initial profile k_m for $m \leq n$:

$$e^{-t_1} = \frac{a+k_1}{a+1}, \quad e^{-t_2} = \frac{ak_1+k_2}{a+1} + at_1 e^{-t_1}, \quad e^{-t_3} = \frac{ak_2+k_3}{a+1} + at_2 e^{-t_2} - \frac{(at_1)^2}{2} e^{-t_1}, \quad (38)$$

etc. Unfortunately these expressions become quickly too complicated to allow to determine how the times t_n depend asymptotically on the initial profile $\{k_n\}$ for large n . How these asymptotics can be understood from (32) will be discussed in Sect. 5.

4 Travelling Wave Solutions

4.1 The Exact Shape of the Travelling Waves

As usual, with travelling wave equations, the first solutions one can try to determine are travelling wave solutions moving at a certain velocity v . Because the $h_n(t)$ are defined on a lattice, a travelling wave solution moving at velocity v satisfies

$$h_n(t) = h_{n+1}\left(t + \frac{1}{v}\right). \quad (39)$$

Clearly this implies that the times t_n form an arithmetic progression, and by shifting the origin of time, one can choose

$$t_n = \frac{n}{v}. \tag{40}$$

This immediately gives, using (37), the generating function of the front shape at all times: for example for $0 \leq t \leq t_1 = 1/v$, one takes $m = 1$ in (37) and gets

$$\sum_{n \geq 1} h_n(t) \lambda^n = -\frac{a\lambda}{1+a\lambda} + \frac{a+1}{1+a\lambda} \times \frac{\lambda e^{(1+a\lambda)t}}{e^{(1+a\lambda)/v} - \lambda}. \tag{41}$$

Another way of determining the travelling wave solutions is to look directly for solutions of (3) of the form (39). One sees that W_v must satisfy

$$W_v(x) = 1 \quad \text{for } x \leq 0, \quad W_v(x) + aW_v(x-1) + vW'_v(x) = 0 \quad \text{for } x > 0. \tag{42}$$

These equations can be solved iteratively: for $x \leq 0$, one already knows that $W_v(x) = 1$. For $x \in [0, 1]$ one has therefore $W_v + vW'_v + a = 0$, which implies for $(x \in [0, 1])$ that $W_v(x) = (a+1)e^{-x/v} - a$ (the integration constant being fixed by continuity at $x = 0$). Knowing $W_v(x)$ for $x \in [0, 1]$, one can solve (42) for $x \in [1, 2]$ and so on.

$$W_v(x) = \begin{cases} 1 & \text{if } x \leq 0, \\ (a+1)e^{-x/v} - a & \text{if } x \in [0, 1], \\ \frac{a(1+a)}{v}(1-v-x)e^{(1-x)/v} + (1+a)e^{-x/v} + a^2 & \text{if } x \in [1, 2], \\ \dots & \dots \end{cases} \tag{43}$$

In fact one can solve directly (42) by considering for $x \in [0, 1]$ the generating function $(\lambda, x) \mapsto \sum_n \lambda^n W_n(x+n)$. Then, as can be checked directly from (41) and (42), $W_v(x)$ and $h_n(t)$ are related for all $n \in \mathbb{Z}$ and $t \geq 0$ by

$$h_n(t) = W_v(n - vt). \tag{44}$$

4.2 The Decay of the Travelling Waves

The large n behavior of the travelling wave $h_n(t)$ (or equivalently the behavior of $W_v(x)$ for large x) can be understood by analyzing the singularities in λ of the right hand side of (41). These singularities are poles located at all the real or complex zeros of

$$e^{(1+a\lambda)/v} - \lambda = 0. \tag{45}$$

(one checks there is no pole at $\lambda = -1/a$) and each pole gives rise to an exponential decay in $h_n(t)$. Using $\lambda = \exp(\gamma)$, (45) can be rewritten into

$$v(\gamma) = \frac{1 + ae^\gamma}{\gamma}. \tag{46}$$

which is the dispersion relation for (3), similar to (7) or to (21). In fact, one can obtain (46) as in Sect. 2 by looking for the velocity v compatible in (42) with an exponentially decaying travelling wave of the form $W_v(x) \sim e^{-\gamma x}$.

One is then led to distinguish three cases depending on the number of real solutions of (46). There is a critical value v_c where (46) has a double zero on the real axis. This critical velocity v_c and the corresponding decay rate γ_c are the solution of

$$a(\gamma_c - 1)e^{\gamma_c} = 1, \quad v_c = \frac{1}{\gamma_c - 1}. \tag{47}$$

1. For $v < v_c$, there is no real λ solution of (45), but there are complex roots. The large n behavior of $h_n(t)$ is then governed by the two roots λ_1 and λ_1^* of (45) closest to the origin

$$h_n(0) \simeq \frac{(a + 1)v}{(1 + a\lambda_1)(v - a\lambda_1)} \lambda_1^{-n} + \text{c.c.} \tag{48}$$

Because λ_1 and λ_1^* are complex, $h_n(t)$ changes its sign as n varies. So for $v < v_c$, as in the Fisher–KPP case, there are travelling wave solutions, but they fail to give positive profiles.

2. For $v = v_c$ given by (47), there is a double real root $\lambda_c = e^{\gamma_c}$ of (45). Then for large n the profile is of the form

$$h_n(0) \simeq \frac{2(1 + a)}{1 + v_c} \left[n + \frac{1 + 4v_c}{3(1 + v_c)} \right] e^{-\gamma_c n}. \tag{49}$$

3. For $v > v_c$ there are two real roots $1 < \lambda_1 < \lambda_2$ of (45) and the large n behavior is controlled by the smallest root:

$$h_n(0) \simeq \frac{(a + 1)v}{(1 + a\lambda_1)(v - a\lambda_1)} \lambda_1^{-n}. \tag{50}$$

We see that for all velocities, we get *explicit* expressions of the prefactors of the exponential decay of the travelling waves. These prefactors are in general not known for more traditional travelling wave equations, such as the Fisher–KPP equation (1).

In each case, corrections to (48–50) can be obtained from the contributions of the other roots of (45). For instance, in the cases $v < v_c$ or $v > v_c$ one could write

$$h_n(0) \simeq \sum_r \frac{(a + 1)v}{(1 + a\lambda_r)(v - a\lambda_r)} \lambda_r^{-n}, \tag{51}$$

where the sum is over all the complex roots λ_r of (45). In Fig. 1, we compare the exact solution (43) of (42) with the asymptotic expansion (51) truncated to a finite number of roots of (45) closest to the origin and one can see that the truncation gives a very good fit of the actual solution.

We have seen that the travelling waves for $v < v_c$ were oscillatory. For $v \geq v_c$, they decrease monotonically towards 0; this can be seen directly from equation (42) verified by $W_v(x)$: write $W_v(x) = R(x)e^{-\gamma x}$ with γ a real positive number related to v through the

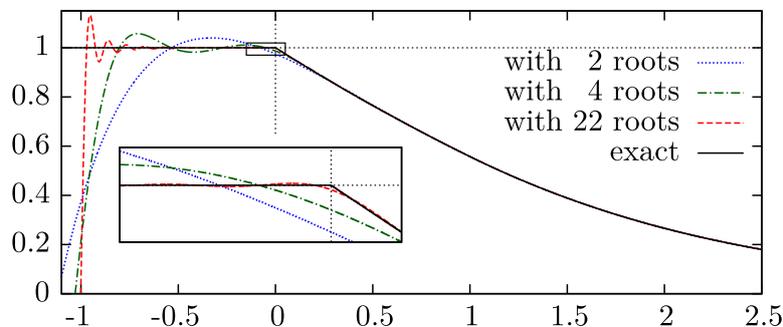


Fig. 1 The travelling wave $W_v(x)$ solution of (42) for $v = 4$ and $a = 1$ as a function of x . The *plain line* labeled “exact” is the exact small- x solution (43). The *dashed lines* are the sums (51) truncated to a given number of first terms: “with 2 roots” means only the two real roots λ_1 and λ_2 , “with 4 roots” means the two real roots and the first pair of complex conjugate roots and “with 22 roots” means the two real roots and the ten pairs of complex conjugates roots closest to the origin. The *inset* is a zoom of the small rectangle around $x = 0$ and $W_v = 1$

dispersion relation (46) (Notice that choosing such a γ is impossible if $v < v_c$). Then (42) gives

$$R(x) = e^{\gamma x} \quad \text{for } x \leq 0, \quad ae^{\gamma} [R(x - 1) - R(x)] + vR'(x) = 0 \quad \text{for } x > 0. \quad (52)$$

As $R(x)$ is strictly increasing for $x < 0$ it must be strictly increasing for all reals; otherwise, on the first local maximum x_m , one would have $R'(x_m) = 0$ and $R(x_m) > R(x_m - 1)$ which is incompatible with (52). Hence, $W_v(x)$ is positive and, from (42), strictly decreasing.

5 How the Initial Condition Determines the Asymptotic Regime

We now discuss how the position of the front at large times (or equivalently the large n asymptotics of the times t_n) depends on the initial condition.

First, by using mostly a comparison property, we will show that the final velocity of the front is determined by the large n decay of the initial condition k_n . Then, we will recover the logarithmic corrections (12–15) and sub-leading terms as in (18) by analyzing the key relation (32) between the initial profile k_n and the times t_n .

We write (32) as

$$(1 + a\lambda)K(\lambda) = -a\lambda + (a + 1)T(\lambda), \quad (53)$$

where the two functions $K(\lambda)$ and $T(\lambda)$ are defined by

$$K(\lambda) = \sum_{n=1}^{\infty} k_n \lambda^n, \quad T(\lambda) = \sum_{n=1}^{\infty} e^{-(1+a\lambda)t_n} \lambda^n. \quad (54)$$

The large n behavior of the k_n 's and of the t_n 's determines the domain of convergence of these two sums and one can try to use (53) to relate their singularities.

When $\lambda = e^{\gamma} > 1$, we will often use the following form of $T(\lambda)$ written in terms of the dispersion relation $v(\gamma)$:

$$T(e^{\gamma}) = \sum_{n=1}^{\infty} e^{\gamma[n-v(\gamma)t_n]}. \quad (55)$$

5.1 Selection of the Velocity

Let us first show that the final velocity of the front is determined by the large n behavior of the initial condition k_n in the same way as for other equations in the Fisher–KPP class. To do this, we use an obvious comparison property; considering two initial conditions $\{k_n^{(1)}\}$ and $\{k_n^{(2)}\}$ with the corresponding times $\{t_n^{(1)}\}$ and $\{t_n^{(2)}\}$, one has

$$\text{if } 0 \leq k_n^{(1)} \leq k_n^{(2)} \text{ for all } n, \text{ then } t_n^{(1)} \geq t_n^{(2)} \text{ for all } n. \quad (56)$$

To keep the discussion simple, we focus only on initial conditions $\{k_n\}$ with $k_n \geq 0$ and the following simple asymptotics:

- If $k_n \sim n^{\alpha} e^{-\gamma n}$ with $0 < \gamma < \gamma_c$.
Pick an $\epsilon > 0$ small enough so that $0 < \gamma - \epsilon$ and $\gamma + \epsilon < \gamma_c$, and consider the two travelling waves going at velocities $v(\gamma - \epsilon)$ and $v(\gamma + \epsilon)$ (they decay respectively like $e^{-(\gamma-\epsilon)n}$ and $e^{-(\gamma+\epsilon)n}$). It is clear that the initial condition $\{k_n\}$ can be sandwiched between these two travelling waves suitably shifted in space, so that, by using the comparison property one gets

$$\frac{1}{v(\gamma - \epsilon)} \leq \liminf_{n \rightarrow \infty} \frac{t_n}{n} \leq \limsup_{n \rightarrow \infty} \frac{t_n}{n} \leq \frac{1}{v(\gamma + \epsilon)}. \tag{57}$$

Now take $\epsilon \rightarrow 0$ to get

$$\lim_{n \rightarrow \infty} \frac{t_n}{n} = \frac{1}{v(\gamma)}. \tag{58}$$

- If $k_n = 0$.

It takes a time t_n to have $h_n(t) = 1$. But at time t_n , the $h_{n+m}(t)$ for $m > 0$ are positive so that, from the comparison property, one has

$$t_{n+m} \leq t_n + t_m. \tag{59}$$

The sequence $\{t_n\}$ is sub-additive and therefore t_n/n has a limit which we call $1/v$. By comparing the initial profile $k_n = 0$ to the travelling wave going at velocity v_c , one must have $1/v \geq 1/v_c$.

We are now going to show that $1/v$ cannot be strictly greater than $1/v_c$. Indeed, if we had $1/v > 1/v_c$, the series (55) defining $T(\lambda)$ would be uniformly convergent on the whole positive real axis λ because $v(\gamma)t_n/n$ would eventually be larger than $1 + \epsilon$ for some $\epsilon > 0$. One would then get

$$T'(\lambda) = \sum_{n \geq 1} \lambda^n e^{-(1+a\lambda)t_n} \left[\frac{n}{\lambda} - at_n \right] \quad \text{for all } \lambda \geq 0 \tag{60}$$

because the series (60) would also be uniformly convergent.

However, for real and large enough λ (at least for $\lambda > \min_n n/(at_n)$), one would obtain $T'(\lambda) < 0$. But, with $k_n = 0$ one has $K(\lambda) = 0$ and from (53) $T'(\lambda) = a/(a + 1)$, in contradiction with $T'(\lambda) < 0$.

We conclude that one must have

$$\lim_{n \rightarrow \infty} \frac{t_n}{n} = \frac{1}{v_c} \quad \text{if } k_n = 0 \quad \text{for } n \geq 1. \tag{61}$$

- If $k_n \sim n^\alpha e^{-\gamma_c n}$ or if $k_n = o(e^{-\gamma_c n})$.

Again, by the comparison property, the initial condition can be sandwiched between $k_n = 0$ and, for any $\epsilon > 0$, the suitably shifted travelling wave going at velocity $v(\gamma_c - \epsilon)$. This leads to conclude that

$$\lim_{n \rightarrow \infty} \frac{t_n}{n} = \frac{1}{v_c}. \tag{62}$$

The velocity selection thus works as for other equations of the Fisher–KPP class.

5.2 Sub-leading Corrections

We limit the discussion to initial conditions similar to those discussed in the previous section which lead to a front with some asymptotic velocity V :

$$\lim_{n \rightarrow \infty} \frac{t_n}{n} = \frac{1}{V}. \tag{63}$$

We also assume that $k_n \geq 0$ which implies that $V \geq v_c$ as was shown in the previous section.

If $V > v_c$, write $V = v(\gamma_1) = v(\gamma_2)$ with $0 < \gamma_1 < \gamma_c < \gamma_2$. We have seen that this velocity is reached for initial conditions such as $k_n \sim n^\alpha e^{-\gamma_1 n}$. In (55), it is then clear that the series $T(\lambda)$ is divergent for $\lambda \in (e^{\gamma_1}, e^{\gamma_2})$ and convergent for $\lambda < e^{\gamma_1}$ or $\lambda > e^{\gamma_2}$. Furthermore, in (54), the radius of convergence of $K(\lambda)$ is e^{γ_1} and, as $k_n > 0$, the function $K(\lambda)$ must have a singularity at $\lambda = e^{\gamma_1}$. We thus see that both $T(\lambda)$ and $K(\lambda)$ become

singular as λ approaches e^{γ_1} from below on the real axis. By matching the singularities of these two functions, we will obtain the sub-leading corrections to t_n for large n .

For $V = v_c$, if the initial condition is $k_n \sim n^\alpha e^{-\gamma_c n}$, the same argument applies: both $T(\lambda)$ and $K(\lambda)$ are singular when λ reaches e^{γ_c} , and one must match the singularities. But, with $V = v_c$, one could also have an initial condition which decays faster than $e^{-\gamma_c n}$ and for which the radius of convergence is larger than e^{γ_c} (even, possibly, infinite). Then, of course, $K(\lambda)$ would have no singularity at $\lambda = e^{\gamma_c}$, even though the convergence of $T(\lambda)$ would remain problematic when λ approaches e^{γ_c} . We will see that the large n behavior of t_n is tuned to “erase” the singularities in $T(\lambda)$ at e^{γ_c} to satisfy (53).

We attack the problem by assuming that the t_n are given and we try to obtain the asymptotics of the k_n . The starting point is thus to assume a velocity $V = v(\gamma_1)$ with $\gamma_1 \leq \gamma_c$, and study $T(\lambda)$ when λ gets close to e^{γ_1} . If one chooses, in all generality,

$$t_n = \frac{n}{V} + \frac{\delta_n}{\gamma_1 V}, \tag{64}$$

where $\delta_n/n \rightarrow 0$, one gets from (55)

$$T(e^\gamma) = \sum_{n=1}^{\infty} e^{\gamma n} \left[1 - \frac{v(\gamma)}{V} \right]^{n-1} e^{-\frac{\gamma v(\gamma)}{\gamma_1 V} \delta_n}. \tag{65}$$

Now we want to take $\gamma = \gamma_1 - \epsilon$ and expand for small ϵ in order to extract the nature of the singularity. Two cases arise:

- If $V > v_c$ (which means $\gamma_1 < \gamma_c$), then $v'(\gamma_1) < 0$ and to leading order

$$T(e^{\gamma_1 - \epsilon}) = \sum_{n=1}^{\infty} \exp \left[\left(\frac{\gamma_1 v'(\gamma_1)}{V} \epsilon + \dots \right) n - (1 - \mu \epsilon + \dots) \delta_n \right] \text{ for } V > v_c, \tag{66}$$

with $\mu = 1/\gamma_1 + v'(\gamma_1)/V$.

- If $V = v_c$ (which means $\gamma_1 = \gamma_c$), then $v'(\gamma_c) = 0$ and one must push the expansion further:

$$T(e^{\gamma_c - \epsilon}) = \sum_{n=1}^{\infty} \exp \left[\left(-\frac{\gamma_c v''(\gamma_c)}{2v_c} \epsilon^2 + \dots \right) n - \left(1 - \frac{1}{\gamma_c} \epsilon + \dots \right) \delta_n \right] \text{ for } V = v_c. \tag{67}$$

It is already clear that cases $V > v_c$ and $V = v_c$ need to be discussed separately. Equations (66) and (67) are the starting point of our analysis which is presented in detail in the following subsections.

We will make heavy use of the following formulas: for $\epsilon > 0$ small,

$$\sum_{n \geq 1} n^\alpha e^{-\epsilon n} \Big|_{\text{singular}} = \begin{cases} \frac{\Gamma(1+\alpha)}{\epsilon^{1+\alpha}} & \text{if } \alpha \text{ is not a negative integer,} \\ \frac{(-1)^\alpha \epsilon^{-\alpha-1} \ln \epsilon}{(-\alpha-1)!} & \text{if } \alpha \text{ is a negative integer.} \end{cases} \tag{68}$$

$$\sum_{n \geq 1} (\ln n) n^\alpha e^{-\epsilon n} \Big|_{\text{singular}} = \begin{cases} \frac{-\Gamma(1+\alpha) \ln \epsilon + \mathcal{O}(1)}{\epsilon^{1+\alpha}} & \text{if } \alpha \text{ is not a negative integer,} \\ \frac{(-\epsilon)^{-\alpha-1} \left[\frac{\ln^2 \epsilon}{2} + \mathcal{O}(\ln \epsilon) \right]}{(-\alpha-1)!} & \text{if } \alpha \text{ is a negative integer,} \end{cases} \tag{69}$$

where the meaning of “singular” for a function $F(\epsilon)$ with a singularity at 0 is that the difference between $F(\epsilon)$ and $F(\epsilon) \Big|_{\text{singular}}$ is a regular function of ϵ which can be expanded as a power series.

5.2.1 For $V > v_c$

As explained above we write $V = v(\gamma_1)$ with $\gamma_1 < \gamma_c$, and we choose t_n such that t_n/n that converges to $1/V$. If one chooses

$$t_n = \frac{n}{V} + \frac{B \ln n + C}{\gamma_1 V}, \tag{70}$$

by keeping the leading order in (66) and using (68) one gets for B not a positive integer

$$T(e^{\gamma_1 - \epsilon}) \Big|_{\text{singular}} \simeq \Gamma(1 - B)e^{-C} \left(\frac{V}{-v'(\gamma_1)\gamma_1 \epsilon} \right)^{1-B}. \tag{71}$$

It is then easy to check that matching the singularities leads to an initial condition decaying as

$$k_n \simeq \frac{(1 + a)e^{-C}}{-v'(\gamma_1)\gamma_1^2} \left[\frac{Vn}{-\gamma_1 v'(\gamma_1)} \right]^{-B} e^{-\gamma_1 n}. \tag{72}$$

Remarks: As can be easily checked, even though (71) is not valid if B is a positive integer, (72) is. One can also check that for $B = C = 0$ one recovers the asymptotics (50) of the travelling wave.

5.2.2 For $V = v_c$

if $V = v_c$, the main difference with the previous case is that $v(\gamma_c - \epsilon) - v(\gamma_c) \sim \epsilon^2$ as $\epsilon \rightarrow 0$ and one must use the expansion (67) instead of (66). As before, we choose a specific form for the times t_n which allow to easily make the comparison with the different cases (13–15) of the Fisher–KPP equation:

$$t_n = \frac{n}{v_c} + \frac{B \ln n + C}{\gamma_c v_c}. \tag{73}$$

With $\delta_n = B \ln + C$ into (67), one obtains generically (when $B \notin \{1, 3/2, 2, 5/2, 3, \dots\}$, see discussion below)

$$T(e^{\gamma_c - \epsilon}) \Big|_{\text{singular}} \simeq e^{-C} \Gamma(1 - B) \left(\frac{\gamma_c v''(\gamma_c)}{2v_c} \right)^{B-1} \epsilon^{2B-2}. \tag{74}$$

Then, using (68) again and (53), one gets

$$k_n \simeq \frac{a + 1}{\gamma_c v_c} e^{-C} \left(\frac{\gamma_c v''(\gamma_c)}{2v_c} \right)^{B-1} \frac{\Gamma(1 - B)}{\Gamma(2 - 2B)} n^{1-2B} e^{-\gamma_c n}. \tag{75}$$

We see that the asymptotics of the initial condition (75) and of the times (73) for large n are related as in the Fisher–KPP case (26) and that the constant term in (26) can be determined. As in (70), one must be careful when B is a positive integer: (74) should be modified to include the logarithmic correction of (68), but (75) is not modified as can easily be checked (the ratio of the two Gamma functions has a limit).

There is another difficulty when $B \in \{3/2, 5/2, 7/2, \dots\}$: for these values, the ratio of Gamma functions in (75) is zero. This means that an initial condition $\{k_n\}$ leading to (73) with $B = 3/2$ (for instance) must decrease faster than $n^{-2}e^{-\gamma_c n}$. For these special values of B , the right hand side of (74) is actually regular as ϵ^{2B-2} is a non-negative integer power of ϵ ; any singular part of $T(e^{\gamma_c - \epsilon})$ must come from higher order terms.

We are now going to show that no non-negative initial condition $\{k_n\}$ can lead to a time sequence $\{t_n\}$ with an asymptotic expansion starting as in (73) with $B > 3/2$. To do so, we

will show that the initial condition $k_n = 0$ leads to (73) with $B = 3/2$ (plus higher order corrections). As any non-negative initial condition must lead to times $\{t_n\}$ which are smaller than the times of the $k_n = 0$ initial condition, this will prove that B cannot be larger than $3/2$.

Consider therefore the case $k_n = 0$; one has $K(\lambda) = 0$ and, from (53), one gets $T(\lambda) = a\lambda/(a + 1)$. Obviously, $T(\lambda)$ has no singularity as λ approaches e^{γ_c} , so the right hand side of (74) must be regular, which implies that $B \in \{3/2, 5/2, 7/2, \dots\}$. We now rule out any value other than $B = 3/2$ by looking at the term of order ϵ in the expansion of $T(e^{\gamma_c - \epsilon})$. One can check that the only terms of order ϵ come from the $\epsilon\delta_n/\gamma_c$ in (67) and from (74) if $B = 3/2$, so one has

$$T(e^{\gamma_c - \epsilon}) = T(e^{\gamma_c}) + \left[\frac{1}{\gamma_c} \sum_{n=1}^{\infty} \delta_n e^{-\delta_n} + e^{-C} \Gamma(-1/2) \left(\frac{\gamma_c v''(\gamma_c)}{2v_c} \right)^{1/2} \mathbb{1}_{B=3/2} \right] \epsilon + o(\epsilon). \tag{76}$$

Notice (64) that $\delta_n \geq 0$ for the $k_n = 0$ initial condition because it is below the travelling wave at velocity v_c for which $\delta_n = 0$. The first term of order ϵ in (76) is therefore positive; on the other hand, the second term (only if $B = 3/2$) is negative. But, from $T(\lambda) = a\lambda/(a + 1)$ the term of order ϵ must be negative; therefore one must have $B = 3/2$ for the zero initial condition and, therefore, $B \leq 3/2$ for any non-negative initial condition.

To summarize, the relationship between the times (73) and the initial condition (75) we have established in this section is valid only for $B < 3/2$ because we only consider non-negative initial conditions. Furthermore, to have (73) with $B = 3/2$, one must have an initial condition decreasing faster than the $n^{-2}e^{-\gamma_c n}$ suggested by (75). No non-negative initial condition can lead to (73) with $B > 3/2$.

5.2.3 For $V = v_c$ and $B = 3/2$

The case $B = 3/2$ is of course the most delicate and it corresponds to (13, 15, 18) in the Fisher–KPP case. For the t_n given by (73) the leading singularity is not (74) but rather

$$T(e^{\gamma_c - \epsilon}) \Big|_{\text{singular}} \simeq 3e^{-C} \sqrt{\frac{2\pi v''(\gamma_c)}{\gamma_c v_c}} \epsilon^2 \ln \epsilon. \tag{77}$$

(This term comes from the first order expansion of the $\epsilon\delta_n/\gamma_c$ term in (67).) Relating this to the $\{k_n\}$ through (53), it leads through (68) to $k_n \sim n^{-3}e^{-\gamma_c n}$ with a negative prefactor. So there is no way for a non-negative initial condition to be compatible with exactly (73), without any extra term.

Therefore, we need to add some corrections to (73) when $B = 3/2$. Let us consider a correction of the form

$$t_n = \frac{n}{v_c} + \frac{\frac{3}{2} \ln n + C + Dn^{-\xi}}{\gamma_c v_c} \tag{78}$$

for some $\xi > 0$. Plugging this correction into (67) one gets

$$T(e^{\gamma_c - \epsilon}) = \sum_{n=1}^{\infty} e^{-C - \frac{\gamma_c v''(\gamma_c)}{2v_c} \epsilon^2 n} n^{-\frac{3}{2}} \left[1 + \frac{3\epsilon}{2\gamma_c} \ln n - Dn^{-\xi} + \dots \right], \tag{79}$$

where the “ \dots ” contains smaller order terms of orders $n\epsilon^3, n^{-2\xi}, \epsilon n^{-\xi}, \epsilon^2 \ln^2 n$, etc. Consider in turns the terms in the square bracket. The “1” leads to the right hand side of (74) with

$B = 3/2$, which is simply a regular term linear in ϵ . The term in $\epsilon \ln n$ gives the right hand side of (77) and the $-Dn^{-\xi}$ contribution can be computed from

$$\sum_{n=1}^{\infty} e^{-\frac{\gamma_c v''(\gamma_c)}{2v_c} \epsilon^2 n} n^{-\frac{3}{2}-\xi} \Big|_{\text{singular}} = \begin{cases} \Gamma(-\frac{1}{2} - \xi) \left(\frac{\gamma_c v''(\gamma_c)}{2v_c}\right)^{\frac{1}{2}+\xi} \epsilon^{1+2\xi} & \text{if } \xi \notin \{\frac{1}{2}, \frac{3}{2}, \frac{5}{2}, \dots\}, \\ 2\frac{\gamma_c v''(\gamma_c)}{2v_c} \epsilon^2 \ln \epsilon & \text{if } \xi = \frac{1}{2}. \end{cases} \tag{80}$$

Several subcases must be considered

- If $\xi > 1/2$ this is smaller than $\epsilon^2 \ln \epsilon$; therefore the leading singularity is still given by (77) which is incompatible with a non-negative initial condition.
- If $0 < \xi < 1/2$ the leading singularity for $T(e^{\gamma_c - \epsilon})$ is $\epsilon^{1+2\xi}$ as given by (80). This leads to

$$k_n \simeq -De^{-C} \frac{1+a}{\gamma_c v_c} \frac{\Gamma(-\frac{1}{2} - \xi)}{\Gamma(-1 - 2\xi)} \left(\frac{\gamma_c v''(\gamma_c)}{2v_c}\right)^{\frac{1}{2}+\xi} n^{-2-2\xi} e^{-n\gamma_c}. \tag{81}$$

With $0 < \xi < \frac{1}{2}$, this is positive if $D > 0$.

- If $\xi = 1/2$ the corrections from (80) and from (77) are both of order $\epsilon^2 \ln \epsilon$. This leads to

$$k_n \simeq 2\frac{1+a}{\gamma_c v_c} e^{-C} \left[D\frac{\gamma_c v''(\gamma_c)}{v_c} - 3\sqrt{\frac{2\pi v''(\gamma_c)}{\gamma_c v_c}} \right] n^{-3} e^{-\gamma_c n}, \tag{82}$$

which is positive if D is large enough. Notice also that the square bracket in (82) vanishes for

$$D = 3\sqrt{\frac{2\pi v_c}{\gamma_c^3 v''(\gamma_c)}}. \tag{83}$$

This means that initial conditions decaying faster than $n^{-3} e^{-\gamma_c n}$ (including the zero initial condition) must lead to (78) with $\xi = 1/2$ and D given by (83). This is exactly the prediction (28).

To finish, notice that we found the first terms of the asymptotic expansion for the times t_n when the initial condition decays as $n^\alpha e^{-n\gamma_c}$ when $\alpha > -2$ (see (73) for $B < 3/2$) and when $\alpha < -2$ (it is of the form (78) with $\xi = -1 - \alpha/2$ for $-3 < \alpha < -2$ and $\xi = 1/2$ for $\alpha \leq -3$), but we did not yet considered the case where $k_n \simeq n^{-2} e^{-n\gamma_c}$. One can check that by taking, as in (27),

$$t_n = \frac{n}{v_c} + \frac{\frac{3}{2} \ln n - \ln \ln n + C}{\gamma_c v_c}, \tag{84}$$

one obtains

$$T(e^{\gamma_c - \epsilon}) \Big|_{\text{singular}} \simeq e^{-C} \sqrt{\frac{8\pi \gamma_c v''(\gamma_c)}{v_c}} \epsilon \ln \epsilon, \tag{85}$$

which leads to

$$k_n \simeq \frac{1+a}{\gamma_c v_c} e^{-C} \sqrt{\frac{8\pi \gamma_c v''(\gamma_c)}{v_c}} n^{-2} e^{-\gamma_c n}. \tag{86}$$

6 Summary

In the previous section, we have computed the initial conditions k_n as a function of the times t_n . Table 1 summarizes our results.

Table 1 Asymptotics of t_n as a function of the leading behavior of the initial condition k_n

1	$k_n \sim n^\alpha e^{-\gamma n}$ with $\gamma < \gamma_c$	$t_n \simeq \frac{n}{v(\gamma)} + \frac{1}{\gamma v(\gamma)} \left[-\alpha \ln n + C \right]$	see (70, 72)
2	$k_n \sim n^\alpha e^{-\gamma_c n}$ with $\alpha > -2$	$t_n \simeq \frac{n}{v_c} + \frac{1}{\gamma_c v_c} \left[\frac{1-\alpha}{2} \ln n + C \right]$	see (73, 75)
3	$k_n \sim n^{-2} e^{-\gamma_c n}$	$t_n \simeq \frac{n}{v_c} + \frac{1}{\gamma_c v_c} \left[\frac{3}{2} \ln n - \ln \ln n + C \right]$	see (84, 86)
4	$k_n \sim n^\alpha e^{-\gamma_c n}$ with $-3 \leq \alpha < -2$	$t_n \simeq \frac{n}{v_c} + \frac{1}{\gamma_c v_c} \left[\frac{3}{2} \ln n + C + D n^{1+\frac{\alpha}{2}} \right]$	see (78, 81, 82)
5	$k_n \ll n^\alpha e^{-\gamma_c n}$ for some $\alpha < -3$	$t_n \simeq \frac{n}{v_c} + \frac{1}{\gamma_c v_c} \left[\frac{3}{2} \ln n + C + 3 \sqrt{\frac{2\pi v_c}{\gamma_c^3 v''(\gamma_c)}} n^{-\frac{1}{2}} \right]$	see (78, 82, 83)

Table 2 Asymptotic expansion of X_t as a function of the leading behavior of the initial condition $u(x, 0)$

1	$u(x, 0) \sim x^\alpha e^{-\gamma x}$ with $\gamma < \gamma_c$	$X_t \simeq v(\gamma)t + \frac{\alpha}{\gamma} \ln t + C'$
2	$u(x, 0) \sim x^\alpha e^{-\gamma_c x}$ with $\alpha > -2$	$X_t \simeq v_c t + \frac{\alpha-1}{2\gamma_c} \ln t + C'$
3	$u(x, 0) \sim x^{-2} e^{-\gamma_c x}$	$X_t \simeq v_c t - \frac{3}{2\gamma_c} \ln t + \frac{1}{\gamma_c} \ln \ln t + C'$
4	$u(x, 0) \sim x^\alpha e^{-\gamma_c x}$ with $-3 \leq \alpha < -2$	$X_t \simeq v_c t - \frac{3}{2\gamma_c} \ln t + C' - D' t^{1+\frac{\alpha}{2}}$
5	$u(x, 0) \ll x^\alpha e^{-\gamma_c x}$ for some $\alpha < -3$	$X_t \simeq v_c t - \frac{3}{2\gamma_c} \ln t + C' - 3 \sqrt{\frac{2\pi}{\gamma_c^5 v''(\gamma_c)}} t^{-\frac{1}{2}}$

These asymptotics agree with all previously known results discussed in Sect. 2. Case 4 is a new prediction, and the domain of validity of Ebert–van Saarloos correction (18) from [17] is made precise (case 5).

The constant C can easily be computed in cases 1 to 3, but we did not manage to get a closed expression in cases 4 and 5. Similarly, we have no expression for D in case 4; in particular, for $\alpha = -3$, the D of case 4 is not given by the prefactor of $n^{-1/2}$ in case 5 because for $\alpha = -3$ the right hand side of (82) must not vanish.

The vanishing terms $n^{1+\alpha/2}$ and $n^{-1/2}$ in cases 4 and 5 depend only on the leading behavior of k_n for large n . One could compute higher order corrections in cases 1 to 3 using the same technique by looking at the next singularities in $T(e^{\gamma-\epsilon})$, but one would need to know a bit more about the asymptotic behavior of k_n : one would find that

$$\text{If } k_n = An^\alpha e^{-\gamma n} \left(1 + o\left(\frac{\ln n}{n}\right) \right),$$

$$\text{then } t_n = [\text{as above}] + \begin{cases} D \frac{\ln n}{n} & \text{in case 1 for } \alpha \neq 0, \\ D \frac{\ln n}{\sqrt{n}} & \text{in case 2 for } \alpha \notin \{-1, 0, 1\}, \\ D \frac{1}{\sqrt{n}} & \text{in case 3,} \end{cases} \quad (87)$$

where the prefactor D could be computed in each case. These vanishing corrections in cases 1 to 3 are less universal than in case 4 to 5 as they do not depend *only* on the leading behavior of k_n for large n , but also on the fact that the sub-leading behavior of k_n decays fast enough compared to the leading behavior. For case 1 with $\alpha = 0$ and case 2 with $\alpha = 1$, the initial condition behaves asymptotically as the travelling wave eventually reached by the front, and vanishing corrections might depend on the initial condition in a more complicated way. Case 2 with $\alpha = -1$ or $\alpha = 0$ are border cases with slightly different corrections.

If we conjecture that the new results (cases 4 and 5) of Table 1 hold for the whole Fisher–KPP class one can obtain, by inverting the relations between t_n and n of Table 1, the asymptotics of the position X_t for initial conditions of the form $u(x, 0) \sim x^\alpha e^{-\gamma x}$. This is done in Table 2.

7 Conclusion

The main result of the present work is the exact relation (32) between the initial condition and the positions of the front at time t for the model (3). Relating the asymptotics of the t_n 's to those of the k_n 's, using the exact relation (32) is an interesting but not easy problem of complex analysis. It allows to obtain precise expressions of the shape of the travelling waves, including prefactors which are usually not known in the usual equations of the Fisher–KPP type. It also allows one to recover the known long time asymptotics of the front position, and to get previously unknown results; in particular, we have shown how fast an initial condition should decay to exhibit the Ebert–van Saarloos correction, and that there is a range of initial conditions which exhibit the $-3/2 \ln t$ Bramson logarithmic term but for which the Ebert–van Saarloos correction is modified (See cases 4 and 5 of Tables 1 and 2).

As shown here the analysis of the asymptotics (3), using complex analysis, is tedious but rather straightforward. Higher corrections to the asymptotics of the position could be determined. One could also try to study how, depending on the initial condition, the asymptotic shape is reached. Furthermore, it would be interesting to generalize (3) to evolutions involving more than two neighboring sites, or to a non-lattice version of the model. More challenging would be to attack the noisy version of the problem [19,20].

Appendix: An Heuristic Derivation of the Positions of the Front

In this appendix we show that several expressions of the position of the front for a Fisher–KPP front can be recovered by considering a simplified version of the Fisher–KPP equation (1) where the non-linear term is replaced by an absorbing boundary. Consider the following linearized Fisher–KPP equation with a given time-dependent boundary X_t with $X_0 = 0$:

$$\begin{cases} \frac{\partial u}{\partial t} = \frac{\partial^2 u}{\partial x^2} + f'(0)u & \text{if } x > X_t, \\ u(X_t, t) = 0. \end{cases} \quad (88)$$

For a given $y > 0$, we look at the value $u(X_t + y, t)$ of the solution at a distance y from the boundary. Intuitively, if X_t increases too quickly with t , this quantity is pushed to zero. On the other hand, if X_t increases too slowly, it diverges with t . It is only for finely tuned choices of X_t that $u(X_t + y, t)$ remains of order 1.

Now we suppose that X_t is no longer given *a priori* but is instead determined by

$$u(X_t + 1, t) = 1. \quad (89)$$

It has been shown [21] that the solution of (88, 89) for compactly supported initial conditions leads to the same long time asymptotics for X_t as for the Fisher–KPP equation (see Sect. 2): one recovers the Bramson term (13) and the Ebert–van Saarloos correction (18).

For initial conditions decaying fast enough, one expects X_t to be asymptotically linear. If X_t were really linear (not only asymptotically but at all times), (88) would be very easy to solve. In this Appendix, we solve a simplified version of (88) where the boundary is replaced by a straight line. This allows us to recover the velocity and the logarithmic corrections (12–15) of the Fisher–KPP equation.

The version we actually consider is therefore the following: For each given time t , we replace the boundary by a linear boundary of slope X_t/t and solve

$$\begin{cases} \frac{\partial u}{\partial s} = \frac{\partial^2 u}{\partial x^2} + f'(0)u & \text{if } x > \frac{X_t}{t}s, \\ u\left(\frac{X_t}{t}s, s\right) = 0. \end{cases} \tag{90}$$

We then tune the value of X_t to satisfy (89) at time t .

For an initial condition $\delta(x - x_0)$ the solution to (90) is

$$g(x, s|x_0) = \frac{e^{f'(0)s}}{\sqrt{4\pi s}} \left[\exp\left(-\frac{(x - x_0)^2}{4s}\right) - \exp\left(\frac{X_t}{t}x_0 - \frac{(x + x_0)^2}{4s}\right) \right]. \tag{91}$$

Taking $s = t$ and writing $x = X_t + y$, one obtains

$$g(X_t + y, t|x_0) = \frac{1}{\sqrt{4\pi t}} \exp\left[f'(0)t - \frac{(X_t + y)^2 + x_0^2 - 2X_t x_0}{4t} \right] 2 \sinh\left(\frac{yx_0}{2t}\right). \tag{92}$$

Given a general initial condition $u(x_0, 0)$ for $x_0 > 0$ one has

$$u(X_t + y, t) = \int_0^\infty dx_0 g(X_t + y, t|x_0)u(x_0, 0), \tag{93}$$

which, after writing $X_t = ct - \delta_t$ with $\delta_t \ll t$, leads to

$$u(X_t + y, t) = \frac{1}{\sqrt{\pi t}} \exp\left[t\left(f'(0) - \frac{c^2}{4}\right) - \frac{c}{2}(y - \delta_t) - \frac{(y - \delta_t)^2}{4t} \right] \times I_t(y), \tag{94}$$

$$\text{with } I_t(y) = \int_0^\infty dx_0 u(x_0, 0) \exp\left[\frac{cx_0}{2} - \frac{\delta_t x_0}{2t} - \frac{x_0^2}{4t} \right] \sinh\left(\frac{yx_0}{2t}\right).$$

Depending on the initial condition $u(x_0, 0)$, we can now determine for which values of c and δ_t the front $u(X_t + y, t)$ remains of order 1 for y of order 1 as t increases.

- For $u(x_0, 0) \simeq Ae^{-\gamma x_0}$ with $\gamma < c/2$, one finds that the integral $I_t(y)$ is dominated by $x_0 \simeq (c - 2\gamma)t$. One obtains

$$\begin{aligned} I_t(y) &\simeq A\sqrt{4\pi t} \sinh\left(\frac{c - 2\gamma}{2}y\right) \exp\left[\left(\frac{c^2}{4} - \gamma c + \gamma^2\right)t - \frac{c - 2\gamma}{2}\delta_t\right], \\ \text{and } u(X_t + y, t) &\simeq 2A \sinh\left[\frac{c - 2\gamma}{2}y\right] \exp\left[\left(f'(0) - \gamma c + \gamma^2\right)t + \gamma\delta_t - \frac{c}{2}y\right]. \end{aligned} \tag{95}$$

Writing $u(X_t + y, t) \sim 1$ leads to $c = \gamma + f'(0)/\gamma = v(\gamma)$ and $\delta_t \simeq \text{Cst}$. The starting hypothesis $\gamma < c/2$ then translates into $\gamma < \gamma_c = \sqrt{f'(0)}$. We conclude that

$$\text{For } u(x_0, 0) \sim e^{-\gamma x_0} \text{ with } \gamma < \gamma_c, \quad X_t \simeq v(\gamma)t + C, \quad (96)$$

as in (12).

- For $u(x_0, 0) \simeq Ax_0^\alpha e^{-\gamma x_0}$ with $\gamma < c/2$, the integral $I_t(y)$ is again dominated by $x_0 \simeq (c - 2\gamma)t$. The large t expression of $u(X_t + y, t)$ has an extra term $[(c - 2\gamma)t]^\alpha$ which is canceled by taking now $\delta_t \simeq -\frac{\alpha}{\gamma} \ln t + \text{Cst.}$ (The value of c remains the same.) We conclude that

$$\text{For } u(x_0, 0) \sim x_0^\alpha e^{-\gamma x_0} \text{ with } \gamma < \gamma_c, \quad X_t \simeq v(\gamma)t + \frac{\alpha}{\gamma} \ln t + C. \quad (97)$$

- For $u(x_0, 0) \ll e^{-\gamma x_0}$ for some $\gamma > c/2$ (steep initial condition), the integral $I_t(y)$ is dominated by x_0 of order 1. This leads to

$$I_t(y) \simeq \frac{y}{2t} \int_0^\infty dx_0 u(x_0, 0)x_0 \exp\left[\frac{cx_0}{2}\right], \quad (98)$$

$$\text{and } u(X_t + y, t) \sim \frac{y}{t^{3/2}} \exp\left[t\left(f'(0) - \frac{c^2}{4}\right) - \frac{c}{2}(y - \delta_t)\right].$$

One needs to take $c = 2\sqrt{f'(0)} = v_c = 2\gamma_c$ and $\delta_t = \frac{3}{2\gamma_c} \ln t + \text{Cst.}$ The starting hypothesis $\gamma > c/2$ translates into $\gamma > \gamma_c$ and we conclude that

$$\text{For } u(x_0, 0) \ll e^{-\gamma x_0} \text{ for some } \gamma > \gamma_c, \quad X_t \simeq v_c t - \frac{3}{2\gamma_c} \ln t + C, \quad (99)$$

as in (13).

- For $u(x_0, 0) \simeq Ax_0^\alpha e^{-\frac{c}{2}x_0}$, depending on the value of α , the integral $I_t(y)$ is dominated by values of x_0 of order 1 or of order \sqrt{t} . In any case, $x_0 \ll t$ and one can simplify $I_t(y)$ into

$$I_t(y) \simeq \frac{y}{2t} \int_0^\infty dx_0 u(x_0, 0)x_0 \exp\left[\frac{c}{2}x_0 - \frac{x_0^2}{4t}\right]. \quad (100)$$

When $\alpha < -2$, this integral is dominated by x_0 of order 1, the Gaussian term can be dropped and one recovers (98) and (99).

When $\alpha \geq -2$, the integral is dominated by x_0 of order \sqrt{t} . One gets

$$I_t(y) \simeq A \frac{y}{2t} \int_1^\infty dx_0 x_0^{\alpha+1} \exp\left[-\frac{x_0^2}{4t}\right] \simeq \begin{cases} Ay2^\alpha \Gamma\left(1 + \frac{\alpha}{2}\right)t^{\frac{\alpha}{2}} & \text{if } \alpha > -2, \\ Ay \frac{\ln t}{4t} & \text{if } \alpha = -2. \end{cases} \quad (101)$$

Into (94) one must therefore take $c = 2\sqrt{f'(0)} = v_c = 2\gamma_c$ and $\delta_t = \frac{1-\alpha}{2\gamma_c} \ln t + \text{Cst}$ if $\alpha > -2$ or $\delta_t = \frac{3}{2\gamma_c} \ln t - \frac{1}{\gamma_c} \ln \ln t$ if $\alpha = -2$. We conclude that

$$\text{For } u(x_0, 0) \sim x_0^\alpha e^{-\gamma_c x_0}, \quad X_t \simeq \begin{cases} v_c t - \frac{3}{2\gamma_c} \ln t + C & \text{if } \alpha < -2, \\ v_c t - \frac{2\gamma_c}{3} \ln t + \frac{\ln \ln t}{\gamma_c} + C & \text{if } \alpha = -2, \\ v_c t - \frac{1-\alpha}{2\gamma_c} \ln t + C & \text{if } \alpha > -2, \end{cases} \quad (102)$$

as in (13–15).

References

1. Aronson, D.G., Weinberger, H.F.: Nonlinear diffusion in population genetics, combustion, and nerve propagation. Lecture Notes in Mathematics, vol. 5. Springer, Berlin (1975)
2. McKean, H.P.: Applications of brownian motion to the equation of Kolmogorov–Petrovskii–Piscounov. *Commun. Pure Appl. Math.* **28**, 323 (1975)
3. Kametaka, Y.: On the nonlinear diffusion equation of Kolmogorov–Petrovskii–Piskunov type. *Osaka J. Math.* **13**, 11 (1976)
4. Bramson, M.D.: Maximal displacement of branching Brownian motion. *Commun. Pure Appl. Math.* **31**, 531 (1978)
5. Bramson, M.D.: Convergence of solutions of the Kolmogorov equation to traveling waves. *Mem. Am. Math. Soc.* **44**, 1–190 (1983)
6. Derrida, B., Spohn, H.: Polymers on disordered trees, spin glasses and traveling waves. *J. Stat. Phys.* **51**, 817 (1988)
7. Murray, J.D.: *Mathematical biology I: an introduction*. Interdisciplinary Applied Mathematics, 3rd edn. Springer, New York (2002)
8. Meerson, B., Vilenkin, A., Sasorov, P.V.: Emergence of fluctuating traveling front solutions in macroscopic theory of noisy invasion fronts. *Phys. Rev. E* **87**(1), 012117 (2013)
9. Munier S.: Lecture notes on quantum chromodynamics and statistical physics. [arXiv:1410.6478](https://arxiv.org/abs/1410.6478) (2014)
10. Fisher, R.A.: The wave of advance of advantageous genes. *Ann. Eugen.* **7**, 355 (1937)
11. Kolmogorov A, Petrovsky I, Piscounov N.: Étude de l'équation de la diffusion avec croissance de la quantité de matière et son application à un problème biologique, *Bull. Univ. État Moscou, A* **1**, 1 (1937)
12. Ma, W., Fuchssteiner, B.: Explicit and exact solutions to a Kolmogorov–Petrovskii–Piskunov equation. *Int. J. Non-linear Mech.* **31**, 329 (1996)
13. Ablowitz, M.J., Zeppetella, A.: Explicit solutions of Fisher's equation for a special wave speed. *Bull. Math. Biol.* **41**, 835 (1979)
14. van Saarloos, W.: Front propagation into unstable states. *Phys. Rep.* **386**, 29 (2003)
15. Benguria, R.D., Depassier, M.C.: Variational characterization of the speed of propagation of fronts for the nonlinear diffusion equation. *Commun. Math. Phys.* **175**, 221 (1996)
16. Hamel, F., Nolen, J., Roquejoffre, J.M., Ryzhik, L.: A short proof of the logarithmic Bramson correction in Fisher–KPP equations. *NHM* **8**, 275 (2013)
17. Ebert, U., van Saarloos, W.: Front propagation into unstable states: universal algebraic convergence towards uniformly translating pulled fronts. *Phys. D* **146**, 1 (2000)
18. Mueller, A.H., Munier, S.: Phenomenological picture of fluctuations in branching random walks. *Phys. Rev. E* **90**(4), 042143 (2014)
19. Brunet, É., Derrida, B., Mueller, A.H., Munier, S.: A phenomenological theory giving the full statistics of the position of fluctuating pulled fronts. *Phys. Rev. E* **73**, 056126 (2006)
20. Mueller, C., Mytnik, L., Quastel, J.: Effect of noise on front propagation in reaction-diffusion equations of KPP type. *Invent. math.* **184**, 405 (2010)
21. Henderson C.: Population stabilization in branching Brownian motion with absorption. [arXiv:1409.4836](https://arxiv.org/abs/1409.4836) (2014)

Vanishing corrections for the position in a linear model of FKPP fronts

Julien Berestycki, Éric Brunet, Simon C. Harris, Matthew I. Roberts

October 13, 2015

Abstract

Take the linearised FKPP equation $\partial_t h = \partial_x^2 h + h$ with boundary condition $h(m(t), t) = 0$. Depending on the behaviour of the initial condition $h_0(x) = h(x, 0)$ we obtain the asymptotics — up to a $o(1)$ term $r(t)$ — of the absorbing boundary $m(t)$ such that $\omega(x) := \lim_{t \rightarrow \infty} h(x + m(t), t)$ exists and is non-trivial. In particular, as in Bramson's results for the non-linear FKPP equation, we recover the celebrated $-3/2 \log t$ correction for initial conditions decaying faster than $x^\nu e^{-x}$ for some $\nu < -2$.

Furthermore, when we are in this regime, the main result of the present work is the identification (to first order) of the $r(t)$ term which ensures the fastest convergence to $\omega(x)$. When $h_0(x)$ decays faster than $x^\nu e^{-x}$ for some $\nu < -3$, we show that $r(t)$ must be chosen to be $-3\sqrt{\pi}/t$ which is precisely the term predicted heuristically by Ebert-van Saarloos [EvS00] in the non-linear case (see also [MM14, BD15, Hen14]). When the initial condition decays as $x^\nu e^{-x}$ for some $\nu \in [-3, -2)$, we show that even though we are still in the regime where Bramson's correction is $-3/2 \log t$, the Ebert-van Saarloos correction has to be modified.

Similar results were recently obtained by Henderson [Hen14] using an analytical approach and only for compactly supported initial conditions.

1 Introduction

The celebrated Fisher-Kolmogorov-Petrovsky-Piscounof equation (FKPP) in one dimension for $h : \mathbb{R} \times \mathbb{R}^+ \rightarrow \mathbb{R}$ is:

$$\partial_t h = \partial_x^2 h + h - h^2, \quad h(x, 0) = h_0(x). \quad (1)$$

This equation is a natural description of a reaction-diffusion model [Fis37, KPP37, AW78]. It is also related to branching Brownian motion: for the Heaviside initial condition $h_0(x) = \mathbb{1}_{\{x < 0\}}$, $h(x, t)$ is the probability that the rightmost particle at time t in a branching Brownian motion (BBM) is to the right of x .

For suitable initial conditions where $h_0(x) \in [0, 1]$, $h_0(x)$ goes to 1 fast enough as $x \rightarrow -\infty$ and $h_0(x)$ goes to 0 fast enough as $x \rightarrow \infty$, it is known that $h(x, t)$ develops into a travelling wave: there exists a centring term $m(t)$ and an asymptotic shape $\omega_v(x)$ such that

$$\lim_{t \rightarrow \infty} h(m(t) + x, t) = \omega_v(x) \in (0, 1), \quad (2)$$

where $m(t)/t \rightarrow v$ and $\omega_v(x)$ is a travelling wave solution to (1) with velocity v : that is, the unique (up to translation) non-trivial solution to

$$\omega_v'' + v \omega_v' + \omega_v - \omega_v^2 = 0 \quad (3)$$

with $\omega_v(-\infty) = 1$ and $\omega_v(+\infty) = 0$.

In his seminal works [Bra83], Bramson showed how the initial condition h_0 (and in particular its large x asymptotic behaviour) determines $m(t)$ in (2). For the important example $h_0(x) = \mathbb{1}_{\{x < 0\}}$ corresponding to the rightmost particle in BBM, he finds

$$m(t) = 2t - \frac{3}{2} \log t + a + o(1) \quad (4)$$

for some constant a , and a limiting travelling wave with (critical) speed $v = 2$. (Here and throughout, we use the notation $f(t) = o(1)$ to mean that $f(t) \rightarrow 0$ as $t \rightarrow \infty$.)

What makes Bramson's results extremely interesting is their universality; for instance Bramson proves [Bra83] that the previous result still holds if the reaction term $h - h^2$ in (1) is replaced by $f(h)$ with $f(0) = f(1) = 0$, $f'(0) = 1$ and $f(x) \leq x$. The universality goes further than that, and for many other front equations, it is believed and sometimes known that the centring term $m(t)$ follows the same kind of behaviour as for (1): one needs to compute a function $v(\gamma)$ which has a minimum v_c at a point γ_c (in the FKPP case (1), $v(\gamma) = \gamma + 1/\gamma$, $\gamma_c = 1$, $v_c = 2$); then for an initial condition decreasing like $e^{-\gamma x}$, the front converges to a travelling wave with velocity $v(\gamma)$ if $\gamma \leq \gamma_c$ and critical velocity v_c if $\gamma \geq \gamma_c$.

When the centring term $m(t)$ is defined as in (2), it is not uniquely determined: if $m(t)$ is any suitable centring term, then $m(t) + o(1)$ is also a suitable centring term. Instead one can try to give a more precise definition for $m(t)$. For example, one could reasonably ask for

$$h(m(t), t) = \alpha \text{ for some } \alpha \in (0, 1) \quad \text{or} \quad \partial_x^2 h(m(t), t) = 0 \quad \text{or} \quad m(t) = - \int dx x \partial_x h(x, t) \quad (5)$$

in addition to (2). In the case $h_0(x) = \mathbb{1}_{\{x < 0\}}$, so that $h(x, t) = \mathbb{P}(R_t > x)$ where R_t is the position of the rightmost particle in a BBM at time t , the first definition in (5) would be the α -quantile of R_t , the second definition would be the mode of the distribution of R_t , and the third definition would be the expectation of R_t .

It has been heuristically argued [EvS00, MM14, Hen14, BD15] that any quantity $m(t)$ defined as in (5) behaves for large t as

$$m(t) = v_c t - \frac{3}{2\gamma_c} \log t + a - 3\sqrt{\frac{2\pi}{\gamma_c^5 v''(\gamma_c)}} \times \frac{1}{\sqrt{t}} + o\left(\frac{1}{\sqrt{t}}\right), \quad (6)$$

for any front equation of the FKPP type and for any initial condition that decays fast enough. In the FKPP case (1), one has $\gamma_c = 1$ and $v''(\gamma_c) = 2$ so that $m(t) = 2t - (3/2) \log t + a - 3\sqrt{\pi/t} + o(1/\sqrt{t})$.

Heuristically, the coefficient of the $1/\sqrt{t}$ term does not depend on the precise definition of $m(t)$ because the front $h(x, t)$ converges very quickly to its limiting shape in the region where h is neither very close to 0 nor very close to 1, so that the difference between any two reasonable definitions of $m(t)$ converges quickly (faster than $1/\sqrt{t}$) to some constant. Note that the constant term “ a ” is expected to be non-universal and to depend on the model, the initial condition and the precise definition of $m(t)$.

As argued in [EvS00], the reason why the “ $\log t$ ” and the “ $1/\sqrt{t}$ ” terms in (6) are so universal is that they are driven by the way the front develops very far on the right, in a region where it is exponentially small and where understanding the position $m(t)$ of the front is largely a matter of solving the linearised front equation.

However there is a catch: solving directly the linearised equation $\partial_t h = \partial_x^2 h + h$ with (for instance) a step initial condition $h_0(x) = \mathbb{1}_{\{x < 0\}}$, one finds $h_{\text{linear}}(x, t) = \frac{1}{2} e^t \operatorname{erfc}(x/\sqrt{4t})$. Defining the position $m(t)$ by $h_{\text{linear}}(m(t), t) = 1$ gives $m(t) = 2t - \frac{1}{2} \log t + a + \mathcal{O}((\log^2 t)/t)$ rather than (4); the linearised equation has the same velocity 2 as for the FKPP equation, a logarithmic correction but with a different prefactor and no $1/\sqrt{t}$ correction. The problem is that with the linearised equation, the $h_{\text{linear}}(x, t)$ increases exponentially on the left of $m(t)$ and this “mass” pushes the front forward, leading to a $-\frac{1}{2} \log t$ rather than a $-\frac{3}{2} \log t$ correction. This means that in order to recover the behaviour of $m(t)$ for the FKPP equation, one must have a front equation with some saturation mechanism on the left. The behaviour of $m(t)$ is not expected to depend on which saturation mechanism is chosen, but one must be present. For these reasons, we consider in this paper a linearised FKPP with a boundary on the left, as in [Hen14].

We emphasize that, in the present work, the FKPP equation is only a motivation: we do not attempt to establish the equivalence between the FKPP equation and the linear model with

a boundary. Our results are proved only for the linear model with boundary, and we can only conjecture that they do apply to the FKPP equation.

2 Statement of the problem and main results

We study the following linear partial differential equation with initial condition $h_0(x)$ and a given boundary $m : [0, \infty) \rightarrow \mathbb{R}$:

$$\begin{cases} \partial_t h = \partial_x^2 h + h & \text{for } x > m(t), \\ h(m(t), t) = 0, & h(x, 0) = h_0(x). \end{cases} \quad (7)$$

Observe that without loss of generality we can (and will) insist that $m(0) = 0$ since otherwise we can simply shift the reference frame by $m(0)$ by the change of coordinate $x \mapsto x - m(0)$.

The same system was studied in [Hen14] by PDE methods for compactly supported initial conditions. In this paper, we use probabilistic methods, writing the solution of the heat equation as an expectation involving Brownian motion with a killing boundary. We give more general results, in particular lifting the compactly supported hypothesis.

If the boundary is linear, $m(t) = vt$, the problem is easily solved explicitly. However, as soon as $m(t)$ is no longer linear, gaining any explicit information about the solution is known to be hard (see for instance [HE15]) and there are few available results.

Motivated by the earlier FKPP discussion about convergence to a travelling wave as in (2), we are looking for functions $m : [0, \infty) \rightarrow \mathbb{R}$ and $\omega : [0, \infty) \rightarrow [0, \infty)$ such that

$$\lim_{t \rightarrow \infty} h(m(t) + x, t) = \omega(x) \quad \text{for all } x \geq 0 \quad (8)$$

with ω non-trivial, $\omega(0) = 0$ and $\omega(x) > 0$ for all $x > 0$. Note that such a function ω necessarily satisfies

$$\omega''(x) + v\omega'(x) + \omega(x) = 0, \quad \forall x \geq 0. \quad (9)$$

In this case, the boundary condition anchors the front. Requiring the convergence of $h(m(t) + x, t)$ to a limiting shape means that $m(t)$ must increase fast enough to prevent the mass near the front from growing exponentially, but not so fast that it tends to zero. This provides a saturation mechanism, and even though it might seem very unlike FKPP fronts to have $h(m(t), t) = 0$, as discussed earlier we do expect the two systems to behave similarly.

Throughout the article we use the following notation:

- $f(x) \sim g(x)$ means $f(x)/g(x) \rightarrow 1$ as $x \rightarrow \infty$;
- $f(x) = \mathcal{O}(g(x))$ means there exists $C > 0$ such that $|f(x)| \leq C|g(x)|$ for all large x ;
- $f(x) = o(g(x))$ means $f(x)/g(x) \rightarrow 0$ as $x \rightarrow \infty$.
- A random variable G is said to have ‘‘Gaussian tails’’ if there exist two positive constants c_1, c_2 such that $\mathbb{P}(|G| > z) \leq c_1 \exp(-c_2 z^2)$ for all $z \geq 0$.

Our first theorem recovers the analogue of Bramson’s results for the system (7), (8).

Theorem 1. *For each of the following bounded initial conditions h_0 , a twice continuously differentiable function $m(t)$ such that $m(0) = 0$ and $m''(t) = \mathcal{O}(1/t^2)$ leads to a solution $h(x, t)$ to (7) with a non-trivial limit (8) if and only if $m(t)$ has the following large time asymptotics where a is an arbitrary constant:*

(a) *if $h_0(x) \sim Ax^\nu e^{-\gamma x}$ with $0 < \gamma < 1$ for large x ,*

$$m(t) = \left(\gamma + \frac{1}{\gamma}\right)t + \frac{\nu}{\gamma} \log t + a + o(1), \quad (10a)$$

$$\text{and then } \omega(x) = \alpha \left(e^{-\gamma x} - e^{-\frac{x}{\gamma}} \right) \quad \text{with } \alpha = Ae^{-\gamma a} \left(\frac{1}{\gamma} - \gamma \right)^\nu.$$

(b) if $h_0(x) \sim Ax^\nu e^{-x}$ with $\nu > -2$ for large x ,

$$m(t) = 2t - \frac{1-\nu}{2} \log t + a + o(1), \quad (10b)$$

$$\text{and then } \omega(x) = \alpha x e^{-x} \quad \text{with } \alpha = \frac{Ae^{-a}}{\sqrt{\pi}} 2^\nu \Gamma\left(1 + \frac{\nu}{2}\right).$$

(c) if $h_0(x) \sim Ax^{-2}e^{-x}$ for large x ,

$$m(t) = 2t - \frac{3}{2} \log t + \log \log t + a + o(1), \quad (10c)$$

$$\text{and then } \omega(x) = \alpha x e^{-x} \quad \text{with } \alpha = \frac{Ae^{-a}}{4\sqrt{\pi}}.$$

(d) if $h_0(x) = \mathcal{O}(x^\nu e^{-x})$ with $\nu < -2$ for large x and such that the value of α below is non-zero,

$$m(t) = 2t - \frac{3}{2} \log t + a + o(1), \quad (10d)$$

$$\text{and then } \omega(x) = \alpha x e^{-x} \quad \text{with } \alpha = \frac{e^{-a-\Delta}}{2\sqrt{\pi}} \int_0^\infty dy h_0(y) y e^y \psi_\infty(y),$$

where Δ and ψ_∞ are quantities depending on the whole function m (and not only the asymptotics) which are introduced (in (61) and (68)) in the proofs.

Remarks.

- From the probabilistic representation of $h(x, t)$ written later in the paper (21), it is clear that the solution $h(x, t)$ to (7) must be an increasing function of h_0 and a decreasing function of m (in the sense that if $m^{(1)}(t) \geq m^{(2)}(t)$ for all t , then $h^{(1)}(x, t) \leq h^{(2)}(x, t)$ for all x and t). This implies that the α given in Theorem 1 must be increasing functions of h_0 and decreasing functions of m . This was obvious from the explicit expression of α in cases (a), (b) and (c). In case (d), given the complicated expressions for Δ and ψ_∞ , it is not obvious at all from its expression that α decreases with m .
- Consider now a twice differentiable function m without the assumption that $m''(t) = \mathcal{O}(1/t^2)$. The monotonicity of $h(x, t)$ with respect to m still holds, and by sandwiching such a m between two sequences of increasingly close functions that satisfy the $\mathcal{O}(1/t^2)$ condition, one can show easily in cases (a), (b) and (c) that if m has the correct asymptotics, then $h(m(t) + x, t)$ converges as in Theorem 1. Case (d) is more difficult as both Δ and ψ_∞ might be ill defined when one does not assume $m''(t) = \mathcal{O}(1/t^2)$.

We now turn to the analogue of the Ebert-van Saarloos correction (6) for our model (7). As explained in the introduction and shown in Theorem 1, with a characterization as in (8), $m(t)$ is only determined up to $o(1)$. If we wish to improve upon Theorem 1, then we need a more precise definition for $m(t)$, analogous to (5). Natural possible definitions could be

$$h(m(t) + 1, t) = 1 \quad \text{or} \quad \partial_x h(m(t), t) = 1. \quad (11)$$

However, it is not obvious that such a function $m(t)$ even exists, would be unique or differentiable. We are furthermore interested only in the long time asymptotics of $m(t)$. Therefore, instead of requiring something like (11) we rather look, as in [Hen14], for the function $m(t)$ such that the convergence (8) is as fast as possible.

Our main result, Theorem 2, tells us how fast $h(m(t) + x, t)$ converges for suitable choices of m in case (d) of Theorem 1. This case is the most classical as it contains, for example, initial conditions with bounded support. It is the case studied by Ebert-Van Saarloos and Henderson, and is the case for which universal behaviour is expected. Theorem 2 is followed by two corollaries that highlight important consequences.

Theorem 2. *Suppose that h_0 is a bounded function such that $h_0(x) = \mathcal{O}(x^\nu e^{-x})$ for large x for some $\nu < -2$, and such that α defined in (10d) is non-zero. Suppose also that m is twice continuously differentiable with*

$$m(t) = 2t - \frac{3}{2} \log(t+1) + a + r(t) \quad (12)$$

where $r(0) = -a$, $r(t) \rightarrow 0$ as $t \rightarrow \infty$ and $r''(t) = \mathcal{O}(t^{-2-\eta})$ for large t for some $\eta > 0$. Then for any $x \geq 0$,

$$h(m(t)+x, t) = \alpha x e^{-x} \left[1 - r(t) - \frac{3\sqrt{\pi}}{\sqrt{t}} + \mathcal{O}\left(t^{1+\frac{\nu}{2}}\right) + \mathcal{O}\left(\frac{1}{t^{\frac{1}{2}+\eta}}\right) + \mathcal{O}\left(\frac{\log t}{t}\right) + \mathcal{O}(r(t)^2) \right] \quad (13)$$

with α as in (10d).

If we further assume that $h_0(x) \sim Ax^\nu e^{-x}$ for large x for some $A > 0$ and $-4 < \nu < -2$, then

$$h(m(t)+x, t) = x e^{-x} \left[\alpha \left(1 - r(t) - \frac{3\sqrt{\pi}}{\sqrt{t}} \right) - b t^{1+\frac{\nu}{2}} + o\left(t^{1+\frac{\nu}{2}}\right) + \mathcal{O}\left(\frac{1}{t^{\frac{1}{2}+\eta}}\right) + \mathcal{O}(r(t)^2) \right] \quad (14)$$

with

$$b = -\frac{A}{\sqrt{4\pi}} e^{-a} 2^{\nu+1} \Gamma\left(\frac{\nu}{2} + 1\right) > 0. \quad (15)$$

This result allows us to bound the rate of convergence $h(m(t)+x, t)$ to $\alpha x e^{-x}$: it is generically of order $\max(1/\sqrt{t}, |r(t)|, t^{1+\nu/2})$.

This also suggests that for $m(t)$ defined as in either choice of (11), one should have $r(t) \sim -3\sqrt{\pi}/\sqrt{t}$ for $\nu < -3$ and $r(t) \asymp t^{1+\nu/2}$ for $-3 \leq \nu < -2$. Note however that we are not sure that such a $m(t)$ exists and, if it exists, we do not know whether it satisfies the hypothesis on $m''(t)$ that we used in the Theorem.

In the following two corollaries we highlight the best rates of convergence of $h(m(t)+x, t) \rightarrow x e^{-x}$ that we can obtain from Theorem 2. For simplicity, we dropped the technical requirement that $m(0) = 0$ in the corollaries; the expression for α must therefore be adapted.

Corollary 3. *Suppose that h_0 is a bounded function such that $h_0(x) = \mathcal{O}(x^\nu e^{-x})$ for large x with $\nu < -3$ and such that α is non-zero. If we choose*

$$m(t) = 2t - \frac{3}{2} \log(t+1) + a + \frac{c}{\sqrt{t+1}}, \quad (16)$$

then

$$\text{if } \nu \leq -4, \quad c = -3\sqrt{\pi} \iff h(m(t)+x, t) = \alpha x e^{-x} + \mathcal{O}\left(\frac{\log t}{t}\right), \quad (17)$$

$$\text{if } -4 < \nu < -3, \quad c = -3\sqrt{\pi} \iff h(m(t)+x, t) = \alpha x e^{-x} + \mathcal{O}\left(t^{1+\frac{\nu}{2}}\right). \quad (18)$$

Note in particular that we have recovered the result of [Hen14], but with more general initial conditions ([Hen14] only considered compactly supported initial conditions).

Corollary 4. *Suppose that $h_0(x)$ is a bounded function such that $h_0(x) \sim Ax^\nu e^{-x}$ for large x with $-4 < \nu < -2$, with m , r and b as in Theorem 2. Then*

$$\text{if } -3 < \nu < -2, \quad r(t) = -\frac{b}{\alpha} t^{1+\frac{\nu}{2}} + o\left(t^{1+\frac{\nu}{2}}\right) \iff h(m(t)+x, t) = \alpha x e^{-x} + o\left(t^{1+\frac{\nu}{2}}\right),$$

$$\text{if } -4 < \nu \leq -3, \quad r(t) = -\frac{3\sqrt{\pi}}{\sqrt{t}} - \frac{b}{\alpha} t^{1+\frac{\nu}{2}} + o\left(t^{1+\frac{\nu}{2}}\right) \iff h(m(t)+x, t) = \alpha x e^{-x} + o\left(t^{1+\frac{\nu}{2}}\right).$$

Notice that for $h_0(x) \sim Ax^{-3} e^{-x}$ the position $m(t)$ still features a first order correction in $1/\sqrt{t}$ but with a coefficient $-(3\sqrt{\pi} + \frac{1}{4\alpha} A e^{-a})$ which is different from the $\nu < -3$ case.

3 Writing the solution as an expectation of a Bessel

In this section, we write the solution to (7) as an expectation of a Bessel process.

We only consider functions $m(t)$ that are twice continuously differentiable. For each given $m(t)$, (7) is a linear problem. We first study the fundamental solutions $q(t, x, y)$ defined as

$$\begin{cases} \partial_t q = \partial_x^2 q + q & \text{if } x > m(t), \\ q(t, m(t), y) = 0, & q(0, x, y) = \delta(x - y); \end{cases} \quad (19)$$

where δ is the Dirac distribution. Then

$$h(x, t) = \int_0^\infty dy q(t, x, y) h_0(y). \quad (20)$$

It is clear that $e^{-t} q(t, x, y)$ is the solution to the heat equation with boundary, and therefore

$$q(t, x, y) dx = e^t \mathbb{P}\left(B_t^{(y)} \in dx, B_s^{(y)} > m(s) \forall s \in (0, t)\right), \quad (21)$$

where $t \mapsto B_t^{(y)}$ is the Brownian motion started from $B_0^{(y)} = y$ with the normalization

$$\mathbb{E}[(B_{s+h}^{(y)} - B_s^{(y)})^2] = 2h. \quad (22)$$

Suppose $f : [0, \infty) \rightarrow \mathbb{R}$ is a continuous function, and $A_t(f)$ is a measurable functional that depends only on $f(s)$, $s \in [0, t]$. Then by Girsanov's theorem,

$$\mathbb{E}[A_t(B^{(y)})] = e^{-\frac{1}{4} \int_0^t ds m'(s)^2} \mathbb{E}\left[A_t(m + B^{(y)}) e^{-\frac{1}{2} \int_0^t m'(s) dB_s^{(y)}}\right]. \quad (23)$$

Plugging into (21) at position $m(t) + x$ instead of x , we get

$$q(t, m(t) + x, y) dx = e^{t - \frac{1}{4} \int_0^t ds m'(s)^2} \mathbb{E}\left[\mathbb{1}_{\{B_t^{(y)} \in dx\}} \mathbb{1}_{\{B_s^{(y)} > 0 \forall s \in (0, t)\}} e^{-\frac{1}{2} \int_0^t m'(s) dB_s^{(y)}}\right]. \quad (24)$$

We recall that, by the reflection principle, the probability that a Brownian path started from y stays positive and ends in dx is:

$$\mathbb{P}(B_t^{(y)} \in dx, B_s^{(y)} > 0 \forall s \in (0, t)) = \frac{1}{\sqrt{\pi t}} \sinh\left(\frac{xy}{2t}\right) e^{-\frac{x^2+y^2}{4t}} dx. \quad (25)$$

Using (25), we write (24) as a conditional expectation:

$$q(t, m(t) + x, y) = \frac{\sinh\left(\frac{xy}{2t}\right)}{\sqrt{\pi t}} e^{-\frac{x^2+y^2}{4t} + t - \frac{1}{4} \int_0^t ds m'(s)^2} \mathbb{E}\left[e^{-\frac{1}{2} \int_0^t m'(s) d\xi_s^{(t:y \rightarrow x)}}\right], \quad (26)$$

where $\xi_s^{(t:y \rightarrow x)}$, $s \in [0, t]$ is a Brownian motion (normalized as in (22)) started from y and conditioned not to hit zero for any $s \in (0, t)$ and to be at x at time t . Such a process is called a Bessel-3 bridge, and we recall some properties of Bessel processes and bridges in Section 4.

It is convenient to think of the path $s \mapsto \xi_s^{(t:y \rightarrow x)}$ as the straight line $s \mapsto y + (x - y)s/t$ plus some fluctuations. This leads us to define

$$\begin{aligned} \psi_t(y, x) &:= \mathbb{E}\left[e^{-\frac{1}{2} \int_0^t m'(s) \left(d\xi_s^{(t:y \rightarrow x)} - \frac{x-y}{t} ds\right)}\right] = \mathbb{E}\left[e^{-\frac{1}{2} \int_0^t m'(s) d\xi_s^{(t:y \rightarrow x)}}\right] e^{\frac{m(t)}{2t}(x-y)}, \\ &= \mathbb{E}\left[e^{\frac{1}{2} \int_0^t m''(s) \left(\xi_s^{(t:y \rightarrow x)} - (y + \frac{x-y}{t}s)\right) ds}\right], \end{aligned} \quad (27)$$

where we have used integration by parts. With this quantity, (26) now reads

$$q(t, m(t) + x, y) = \frac{\sinh\left(\frac{xy}{2t}\right)}{\sqrt{\pi t}} e^{\frac{m(t)}{2t}(y-x) - \frac{x^2+y^2}{4t} + t - \frac{1}{4} \int_0^t ds m'(s)^2} \psi_t(y, x), \quad (28)$$

and the main part of the present work is to estimate $\psi_t(y, x)$.

4 The Bessel toolbox

Before we begin our main task, we need some fairly standard estimates on Bessel-3 processes and Bessel-3 bridges. From here on, we refer to these simply as Bessel processes and Bessel bridges; the “3” will be implicit. We include proofs for completeness.

We build most of our processes on the same probability space. We fix a driving Brownian motion $(B_s, s \geq 0)$ started from 0 under a probability measure \mathbb{P} , with the normalization $\mathbb{E}[B_t^2] = 2t$.

For each $y \geq 0$ we introduce a Bessel process $\xi^{(y)}$ started from y as the strong solution to the SDE

$$\xi_0^{(y)} = y, \quad d\xi_s^{(y)} = dB_s + \frac{2}{\xi_s^{(y)}} ds. \quad (29)$$

It is well-known that $\xi_s^{(y)}$ has the law of a Brownian motion conditioned to never hit zero.

We also introduce, for each $t \geq 0$ and $y \geq 0$

$$\xi_s^{(t:y \rightarrow 0)} = \frac{t-s}{t} \xi_{\frac{st}{t-s}}^{(y)} \quad \text{for } s \in [0, t). \quad (30)$$

This process is a Bessel bridge from y to 0 in time t , which is a Brownian motion started from y and conditioned to hit 0 for the first time at time t . One can check by direct substitution that $\xi_s^{(t:y \rightarrow 0)}$ solves

$$\xi_0^{(t:y \rightarrow 0)} = y, \quad d\xi_s^{(t:y \rightarrow 0)} = d\tilde{B}_{t,s} + \left(\frac{2}{\xi_s^{(t:y \rightarrow 0)}} - \frac{\xi_s^{(t:y \rightarrow 0)}}{t-s} \right) ds, \quad (31)$$

where for each t , $(\tilde{B}_{t,s}, s \in [0, t))$ is the strong solution to

$$\tilde{B}_{t,0} = 0, \quad d\tilde{B}_{t,s} = \frac{t-s}{t} d\left(B_{\frac{ts}{t-s}}\right), \quad (32)$$

and is thus itself a Brownian motion.

One can compute directly the law of the Brownian motion conditioned to hit zero for the first time at time t using (25) and check that this law solves the forward Kolmogorov equation (or Fokker Planck equation) associated with the SDE (or Langevin equation) (31).

Similarly, we construct the Bessel bridge from y to x in time t , the Brownian motion conditioned not to hit zero for any $s \in (0, t)$ and to be at x at time t , through

$$\xi_0^{(t:y \rightarrow x)} = y, \quad d\xi_s^{(t:y \rightarrow x)} = d\tilde{B}_{t,s} + \left(\frac{x}{t-s} \coth \frac{x\xi_s^{(t:y \rightarrow x)}}{2(t-s)} - \frac{\xi_s^{(t:y \rightarrow x)}}{t-s} \right) ds. \quad (33)$$

The advantages of constructing all the processes from a single Brownian path $s \mapsto B_s$ is that they can be compared directly, realization by realization. In particular we use the following comparisons:

Lemma 5. *For any $y \geq z \geq 0$ and $s \geq 0$,*

$$\xi_s^{(z)} \leq \xi_s^{(y)} \leq \xi_s^{(z)} + y - z \quad \text{and} \quad y + B_s \leq \xi_s^{(y)}. \quad (34)$$

Furthermore, for any $y \geq 0$, $x \geq z \geq 0$, $t \geq 0$ and $s \in [0, t]$,

$$\xi_s^{(t:0 \rightarrow 0)} \leq \xi_s^{(t:y \rightarrow 0)} \leq \xi_s^{(t:0 \rightarrow 0)} + y \frac{t-s}{t}, \quad \xi_s^{(t:y \rightarrow z)} \leq \xi_s^{(t:y \rightarrow x)} \leq \xi_s^{(t:y \rightarrow z)} + \frac{(x-z)s}{t}. \quad (35)$$

Proof. To prove (34) we make three observations.

- The processes $\xi_s^{(y)}$ and $y + B_s$ both start from y and

$$d(\xi_s^{(y)} - (y + B_s)) = \frac{ds}{\xi_s^{(y)}} > 0, \quad s > 0, \quad (36)$$

so that $\xi_s^{(y)} > y + B_s$ for all $s > 0$ and $y \geq 0$.

- $\xi_s^{(y)}$ and $\xi_s^{(z)}$ follow the same SDE (29) and $\xi_0^{(y)} \geq \xi_0^{(z)}$, so the two processes must remain ordered at all times (see for instance [Kun97]).

- We have

$$d(\xi_s^{(y)} - \xi_s^{(z)}) = \left(\frac{1}{\xi_s^{(y)}} - \frac{1}{\xi_s^{(z)}} \right) ds, \quad (37)$$

and since $\xi_s^{(y)} \geq \xi_s^{(z)}$ for all $s \geq 0$ we see that $\xi_s^{(y)} - \xi_s^{(z)}$ is decreasing, yielding $\xi_s^{(y)} - \xi_s^{(z)} \leq y - z$ for all $s \geq 0$.

The inequalities in the left part of (35) are a direct consequence of (34) through the change of time (30). We now focus on the inequalities in the right part of (35). First we assume that $z > 0$.

The fact that for $x \geq z$ we have $\xi_s^{(t:y \rightarrow x)} \geq \xi_s^{(t:y \rightarrow z)}$ follows from the fact that $x \coth(ax) \geq z \coth(az)$ for any $a > 0$ and $x \geq z$.

For the other inequality, the fact that $u(\coth u - 1)$ is decreasing yields that

$$\begin{aligned} d\xi_s^{(t:y \rightarrow x)} &= d\tilde{B}_{t,s} + \frac{2}{\xi_s^{(t:y \rightarrow x)}} \times \frac{x\xi_s^{(t:y \rightarrow x)}}{2(t-s)} \left(\coth \frac{x\xi_s^{(t:y \rightarrow x)}}{2(t-s)} - 1 \right) ds + \frac{x - \xi_s^{(t:y \rightarrow x)}}{t-s} ds \\ &\leq d\tilde{B}_{t,s} + \frac{2}{\xi_s^{(t:y \rightarrow x)}} \times \frac{z\xi_s^{(t:y \rightarrow z)}}{2(t-s)} \left(\coth \frac{z\xi_s^{(t:y \rightarrow z)}}{2(t-s)} - 1 \right) ds + \frac{x - \xi_s^{(t:y \rightarrow x)}}{t-s} ds \\ &\leq d\tilde{B}_{t,s} + \frac{z}{t-s} \left(\coth \frac{z\xi_s^{(t:y \rightarrow z)}}{2(t-s)} - 1 \right) ds + \frac{x - \xi_s^{(t:y \rightarrow x)}}{t-s} ds, \end{aligned} \quad (38)$$

so that, writing $\zeta_s := \xi_s^{(t:y \rightarrow x)} - \xi_s^{(t:y \rightarrow z)} \geq 0$ for the difference process,

$$d\zeta_s \leq \frac{x - z - \zeta_s}{t - s} ds. \quad (39)$$

But the solution to $\frac{d\phi_s}{ds} = (x - z - \phi_s)/(t - s)$ and $\phi_0 = 0$ is $\phi_s = (x - z)s/t$, implying that $\zeta_s \leq (x - z)s/t$, which concludes the proof for $z > 0$. For the case $z = 0$ the proof is the same but uses the inequalities $1 \leq u \coth u \leq 1 + u$ for $u \geq 0$. \square

We note that, intuitively, as the length of a Bessel bridge tends to infinity, on any compact time interval the bridge looks more and more like a Bessel process. Similarly, as the start point of a Bessel process tends to infinity, on any compact interval it looks more and more like a Brownian motion relative to its start position. We make this precise in the lemma below.

Lemma 6. *For all $s \geq 0$ and $y \geq 0$,*

$$\xi_s^{(t:y \rightarrow 0)} \rightarrow \xi_s^{(y)} \quad \text{as } t \rightarrow \infty. \quad (40)$$

For all $s \geq 0$

$$\xi_s^{(y)} - y \rightarrow B_s \quad \text{as } y \rightarrow \infty. \quad (41)$$

For all $s \geq 0$ and any $y_t \rightarrow \infty$ as $t \rightarrow \infty$,

$$\xi_s^{(t:y_t \rightarrow 0)} - y_t \frac{t-s}{t} \rightarrow B_s \quad \text{as } t \rightarrow \infty. \quad (42)$$

Proof. For (40), we simply recall (30) which defined

$$\xi_s^{(t:y \rightarrow 0)} = \frac{t-s}{t} \xi_{\frac{st}{t-s}}^{(y)} \quad \text{for } s \in [0, t), \quad (43)$$

so we are done by continuity of paths.

For (41), recall from Lemma 5 that $\xi_s^{(y)} - y \geq B_s$. This both gives us the required lower bound, and tells us that for any $s \geq 0$, $\inf_{u \in [0, s]} \xi_u^{(y)} \rightarrow \infty$ as $y \rightarrow \infty$. Thus

$$\xi_s^{(y)} - y = B_s + 2 \int_0^s \frac{1}{\xi_u^{(y)}} du \leq B_s + \frac{2s}{\inf_{u \in [0, s]} \xi_u^{(y)}} \rightarrow B_s \text{ as } y \rightarrow \infty. \quad (44)$$

Finally, for (42), we write

$$\xi_s^{(t:y_t \rightarrow 0)} - y_t \left(\frac{t-s}{t} \right) = \left[\xi_s^{(y_t)} - y_t \right] - \left[\frac{s}{t} (\xi_s^{(y_t)} - y_t) \right] + \left[\left(\frac{t-s}{t} \right) (\xi_{s+\frac{s^2}{t-s}}^{(y_t)} - \xi_s^{(y_t)}) \right]. \quad (45)$$

By (40), $\xi_s^{(y_t)} - y_t \rightarrow B_s$. By (34), $B_s \leq \xi_s^{(y_t)} - y_t \leq \xi_s^{(0)}$, so

$$\frac{s}{t} (\xi_s^{(y_t)} - y_t) \rightarrow 0 \quad \text{as } t \rightarrow \infty. \quad (46)$$

Using our coupling between the Bessel processes and Brownian motion we have $dB_u \leq d\xi_u^{(y_t)} \leq d\xi_u^{(0)}$ for all $u \geq 0$ and hence

$$B_{s+\frac{s^2}{t-s}} - B_s \leq \xi_{s+\frac{s^2}{t-s}}^{(y_t)} - \xi_s^{(y_t)} \leq \xi_{s+\frac{s^2}{t-s}}^{(0)} - \xi_s^{(0)} \quad (47)$$

so by continuity of paths,

$$\left(\frac{t-s}{t} \right) (\xi_{s+\frac{s^2}{t-s}}^{(y_t)} - \xi_s^{(y_t)}) \rightarrow 0 \quad \text{as } t \rightarrow \infty, \quad (48)$$

which concludes the proof of (42). \square

We need the fact that the increments of a Bessel process over time s are roughly of order $s^{1/2}$. By paying a small price on the exponent, we obtain the following uniform bounds:

Lemma 7. *For any $\epsilon > 0$ small enough, there exists a positive random variable G with Gaussian tail such that uniformly in $s \geq 0$ and $y \geq 0$,*

$$|\xi_s^{(y)} - y| \leq G \max \left(s^{\frac{1}{2}-\epsilon}, s^{\frac{1}{2}+\epsilon} \right) \quad \text{and} \quad |B_s| \leq G \max \left(s^{\frac{1}{2}-\epsilon}, s^{\frac{1}{2}+\epsilon} \right). \quad (49)$$

Furthermore, uniformly in $x \geq 0$, $y \geq 0$, $t \geq 0$ and $0 \leq s \leq t$,

$$\left| \xi_s^{(t:y \rightarrow x)} - \left(y + \frac{x-y}{t} s \right) \right| \leq G \max \left(s^{\frac{1}{2}-\epsilon}, s^{\frac{1}{2}+\epsilon} \right). \quad (50)$$

Proof. From (34) we have $B_s \leq \xi_s^{(y)} - y \leq \xi_s^{(0)}$. Also by symmetry $\mathbb{P}(|B_s| > x) = 2\mathbb{P}(B_s > x)$. Thus to prove (49), it is sufficient to show that

$$\mathbb{P} \left(\sup_{s>0} \frac{\xi_s^{(0)}}{\max(s^{1/2-\epsilon}, s^{1/2+\epsilon})} > x \right) \leq c_1 e^{-c_2 x^2} \quad (51)$$

for some positive c_1 and c_2 . The proof is elementary and we defer it to an appendix.

To prove (50), notice that from (35) we have

$$\xi_s^{(t:y \rightarrow x)} - \left(y + \frac{x-y}{t} s \right) \leq \xi_s^{(t:0 \rightarrow 0)}. \quad (52)$$

But from the change of time (30) and (49),

$$\xi_s^{(t:0 \rightarrow 0)} = \frac{t-s}{t} \xi_{\frac{st}{t-s}}^{(0)} \leq G \frac{t-s}{t} \max \left\{ \left(\frac{st}{t-s} \right)^{\frac{1}{2}-\epsilon}, \left(\frac{st}{t-s} \right)^{\frac{1}{2}+\epsilon} \right\} \leq G \max \left(s^{\frac{1}{2}-\epsilon}, s^{\frac{1}{2}+\epsilon} \right), \quad (53)$$

where the last step is obtained by pushing the $(t-s)/t$ inside the max. This provides the upper bound of (50). For the lower bound, we introduce Brownian bridges $s \mapsto B_s^{(t:y \rightarrow x)}$ started from y and conditioned to be at x at time t . We couple the Brownian bridge to the Bessel bridges by building them over the family $\tilde{B}_{t,s}$ of Brownian motions defined in (32):

$$B_0^{(t:y \rightarrow x)} = y, \quad dB_s^{(t:y \rightarrow x)} = d\tilde{B}_{t,s} + \frac{x - B_s^{(t:y \rightarrow x)}}{t-s} ds. \quad (54)$$

One can check directly that

$$B_s^{(t:y \rightarrow x)} = y + \frac{x-y}{t}s + B_s^{(t:0 \rightarrow 0)}. \quad (55)$$

Furthermore, by comparing (54) to (33), it is immediate from the fact that $\coth u \geq 1$ for all $u \geq 0$ that $\xi_s^{(t:y \rightarrow x)} \geq B_s^{(t:y \rightarrow x)}$. Therefore

$$\xi_s^{(t:y \rightarrow x)} - \left(y + \frac{x-y}{t}s \right) \geq B_s^{(t:0 \rightarrow 0)}. \quad (56)$$

Also, as in (30), we can relate B_s and $B_s^{(t:0 \rightarrow 0)}$ through a time change:

$$B_s^{(t:0 \rightarrow 0)} = \frac{t-s}{t} B_{\frac{st}{t-s}} \quad \text{for } s \in [0, t), \quad (57)$$

and, as in (53),

$$|B_s^{(t:0 \rightarrow 0)}| = \frac{t-s}{t} |B_{\frac{st}{t-s}}| \leq G \frac{t-s}{t} \max \left\{ \left(\frac{st}{t-s} \right)^{\frac{1}{2}-\epsilon}, \left(\frac{st}{t-s} \right)^{\frac{1}{2}+\epsilon} \right\} \leq G \max \left(s^{\frac{1}{2}-\epsilon}, s^{\frac{1}{2}+\epsilon} \right), \quad (58)$$

which concludes the proof. \square

5 Simple properties of $\psi_t(y, x)$ and proof of Theorem 1

As in the hypothesis of Theorem 1, we assume throughout this section that m is twice continuously differentiable with

$$m(0) = 0 \quad \text{and} \quad m''(s) = \mathcal{O}\left(\frac{1}{s^2}\right). \quad (59)$$

The large s behaviour of $m''(s)$ implies that there exists a v such that, for large s ,

$$m'(s) = v + \mathcal{O}\left(\frac{1}{s}\right) \quad \text{and} \quad m(s) = vs + \mathcal{O}(\log s). \quad (60)$$

We define

$$\Delta = \frac{1}{4} \int_0^\infty ds (m'(s) - v)^2, \quad (61)$$

which is finite because of (59).

5.1 Simple properties of $\psi_t(y, x)$

We recall from (27) that the main quantity we are interested in is

$$\psi_t(y, x) = \mathbb{E}[e^{I_t(y, x)}], \quad (62)$$

with

$$I_t(y, x) = \frac{1}{2} \int_0^t ds m''(s) \left(\xi_s^{(t:y \rightarrow x)} - \left(y + \frac{x-y}{t}s \right) \right), \quad (63)$$

where we recall that $\xi_s^{(t:y \rightarrow x)}$, $s \in [0, t]$ is a Bessel bridge from y to x over time t . We mainly need to consider $x = 0$ so we use the shorthand

$$\psi_t(y) := \psi_t(y, 0). \quad (64)$$

We also define

$$I(y) = \frac{1}{2} \int_0^\infty ds m''(s) (\xi_s^{(y)} - y) \quad (65)$$

where $\xi_s^{(y)}$, $s \geq 0$ is a Bessel process started from y .

Proposition 8. *The function $\psi_t(y, x)$ has the following properties:*

- *It is bounded away from zero and infinity: there exist two positive constants $0 < K_1 < K_2$ depending on the function $m''(s)$ such that for any x, y, t ,*

$$K_1 \leq \psi_t(y, x) \leq K_2. \quad (66)$$

- *It hardly depends on x for large times: recalling that $\psi_t(y) := \psi_t(y, 0)$,*

$$\psi_t(y, x) = \psi_t(y) \left(1 + x \mathcal{O}\left(\frac{\log t}{t}\right) \right) \quad \text{uniformly in } y \text{ and } x. \quad (67)$$

- *For fixed y , it has a finite and positive limit as $t \rightarrow \infty$:*

$$\psi_\infty(y) := \lim_{t \rightarrow \infty} \psi_t(y) = \mathbb{E}[e^{I(y)}] > 0. \quad (68)$$

- *The large time limit $\psi_\infty(y)$ has a well-behaved large y limit: for any function $t \mapsto y_t$ that goes to infinity as $t \rightarrow \infty$,*

$$\lim_{y \rightarrow \infty} \psi_\infty(y) = \lim_{t \rightarrow \infty} \psi_t(y_t) = \mathbb{E}\left[e^{\frac{1}{2} \int_0^\infty ds m''(s) B_s}\right] = e^\Delta. \quad (69)$$

Proof. For the first result, Lemma 7 tells us that

$$\left| \xi_s^{(t:y \rightarrow x)} - \left(y + \frac{x-y}{t}s \right) \right| \leq G \max\left(s^{\frac{1}{2}-\epsilon}, s^{\frac{1}{2}+\epsilon}\right), \quad (70)$$

where $G > 0$ is a random variable with Gaussian tail independent of t, y and x . Then, since $m''(s) = \mathcal{O}(1/s^2)$,

$$\left| I_t(y, x) \right| \leq \frac{1}{2} \int_0^\infty ds |m''(s)| G \max\left(s^{\frac{1}{2}-\epsilon}, s^{\frac{1}{2}+\epsilon}\right) = G\mathcal{O}(1). \quad (71)$$

For the second result, we compare paths going to x with paths going to 0: we know from Lemma 5 that $0 \leq \xi_s^{(t:y \rightarrow 0)} - \xi_s^{(t:y \rightarrow x)} + xs/t \leq xs/t$, so

$$\begin{aligned} |I_t(y, 0) - I_t(y, x)| &\leq \frac{1}{2} \int_0^t ds |m''(s)| \times \left| \xi_s^{(t:y \rightarrow 0)} - \xi_s^{(t:y \rightarrow x)} + \frac{x}{t}s \right| \\ &\leq \frac{x}{2t} \int_0^t ds |m''(s)| s = x \mathcal{O}\left(\frac{\log t}{t}\right). \end{aligned} \quad (72)$$

We now turn to the third result. For any fixed s and y , Lemma 6 tells us that $\xi_s^{(t:y \rightarrow 0)} \rightarrow \xi_s^{(y)}$ as $t \rightarrow \infty$. Thus, using (70) and (71), we can apply dominated convergence and obtain

$$I_t(y, 0) = \frac{1}{2} \int_0^t ds m''(s) \left(\xi_s^{(t:y \rightarrow 0)} - y \frac{t-s}{t} \right) \rightarrow \frac{1}{2} \int_0^\infty ds m''(s) (\xi_s^{(y)} - y) = I(y). \quad (73)$$

Furthermore, as the bound (71) is a random variable with Gaussian tails, using dominated convergence again we get

$$\lim_{t \rightarrow \infty} \mathbb{E}[e^{I_t(y, 0)}] = \mathbb{E}[e^{I(y)}]. \quad (74)$$

For the fourth statement, by Lemma 6 for any fixed s we have

$$\lim_{y \rightarrow \infty} (\xi_s^{(y)} - y) = B_s \quad \text{and} \quad \lim_{t \rightarrow \infty} \left(\xi_s^{(t:y_t \rightarrow 0)} - y_t \frac{t-s}{t} \right) = B_s. \quad (75)$$

Then, by dominated convergence using again a uniform Gaussian bound from Lemma 7,

$$\lim_{y \rightarrow \infty} \psi_\infty(y) = \lim_{t \rightarrow \infty} \psi_t(y_t) = \mathbb{E} \left[e^{\frac{1}{2} \int_0^\infty ds m''(s) B_s} \right]. \quad (76)$$

It now remains to compute the right-hand-side. Let

$$X_t := \frac{1}{2} \int_0^t ds m''(s) B_s. \quad (77)$$

By integration by parts,

$$X_t = \frac{1}{2} m'(t) B_t - \frac{1}{2} \int_0^t m'(s) dB_s = \frac{1}{2} \int_0^t (m'(t) - m'(s)) dB_s \quad (78)$$

so X_t is a time change of Brownian motion with

$$\mathbb{E} \left[e^{X_t} \right] = e^{\frac{1}{2} \text{var}(X_t)} = e^{\frac{1}{8} \int_0^t (m'(t) - m'(s))^2 ds} \rightarrow e^{\frac{1}{4} \int_0^\infty (v - m'(s))^2 ds} = e^\Delta. \quad (79)$$

Therefore, by dominated convergence as in (76), $\mathbb{E}[e^{X_\infty}] = e^\Delta$. \square

5.2 Proof of Theorem 1

Since $m(0) = 0$ and $m''(s) = \mathcal{O}(1/s^2)$, we can write $m(s) = vs + \delta(s)$ with $\delta(0) = 0$, $\delta(s) = \mathcal{O}(\log s)$, and $\delta'(s) = \mathcal{O}(1/s)$. Note that

$$\int_0^t ds m'(s)^2 = \int_0^t ds (v^2 + 2v\delta'(s) + \delta'(s)^2) = v^2 t + 2v\delta(t) + 4\Delta + \mathcal{O}\left(\frac{1}{t}\right), \quad (80)$$

where we recall that $\Delta = \frac{1}{4} \int_0^\infty ds \delta'(s)^2$. We now fix $x > 0$, so that any terms written as $\mathcal{O}(f(t))$ might depend on x ; since x is fixed this will not matter. For instance, instead of (67) we simply write that $\psi_t(y, x) = \psi_t(y) e^{\mathcal{O}(\frac{\log t}{t})}$.

We recall (28):

$$q(t, m(t) + x, y) = \frac{\sinh\left(\frac{xy}{2t}\right)}{\sqrt{\pi t}} e^{\frac{m(t)}{2t}(y-x) - \frac{x^2+y^2}{4t} + t - \frac{1}{4} \int_0^t ds m'(s)^2} \psi_t(y, x). \quad (81)$$

Substituting in the estimate above we get

$$q(t, m(t) + x, y) = \frac{1}{\sqrt{\pi t}} e^{t(1 - \frac{v^2}{4}) - \frac{v}{2}\delta(t) - \Delta - \frac{v}{2}x + \mathcal{O}(\frac{\log t}{t})} \sinh\left(\frac{xy}{2t}\right) e^{\frac{v}{2}y + \frac{\delta(t)}{2t}y} \psi_t(y) e^{-\frac{y^2}{4t}}. \quad (82)$$

Then since $h(x, t) = \int_0^\infty dy q(t, x, y) h_0(y)$ —see (20)—we have

$$h(m(t) + x, t) = \frac{1}{\sqrt{4\pi t^{3/2}}} e^{t(1-\frac{v^2}{4}) - \frac{v}{2}\delta(t) - \Delta - \frac{v}{2}x + \mathcal{O}(\frac{\log t}{t})} H(x, t), \quad (83)$$

with

$$H(x, t) = \int_0^\infty dy h_0(y) 2t \sinh\left(\frac{xy}{2t}\right) e^{\frac{v}{2}y + \frac{\delta(t)}{2t}y} \psi_t(y) e^{-\frac{y^2}{4t}}. \quad (84)$$

We now must choose v and $\delta(t)$, depending on the initial condition, such that (83) has a finite and non-zero limit as $t \rightarrow \infty$.

We use the following simple calculus lemma to evaluate $H(x, t)$. We defer the proof to the end of this section.

Lemma 9. *Let $\phi(y)$ a bounded function such that*

$$\phi(y) \sim Ay^\alpha \quad \text{as } y \rightarrow \infty \quad (85)$$

for some $A > 0$ and some α . If $\epsilon_t = o(t^{-1/2})$ then, as $t \rightarrow \infty$,

$$\int_0^\infty dy \phi(y) e^{-\frac{y^2}{4t} + \epsilon_t y} \psi_t(y) \begin{cases} \sim A 2^\alpha e^\Delta \Gamma\left(\frac{1+\alpha}{2}\right) t^{\frac{1+\alpha}{2}} & \text{if } \alpha > -1 \\ \sim \frac{A}{2} e^\Delta \log t & \text{if } \alpha = -1 \\ \rightarrow \int_0^\infty dy \phi(y) \psi_\infty(y) & \text{if } \alpha < -1. \end{cases} \quad (86a)$$

$$\int_0^\infty dy \phi(y) e^{-\frac{y^2}{4t} + \epsilon_t y} \psi_t(y) \begin{cases} \sim \frac{A}{2} e^\Delta \log t & \text{if } \alpha = -1 \end{cases} \quad (86b)$$

$$\int_0^\infty dy \phi(y) e^{-\frac{y^2}{4t} + \epsilon_t y} \psi_t(y) \begin{cases} \rightarrow \int_0^\infty dy \phi(y) \psi_\infty(y) & \text{if } \alpha < -1. \end{cases} \quad (86c)$$

If (85) is replaced by $\phi(y) = \mathcal{O}(y^\alpha)$, then (86c) remains valid, and (86a) and (86b) are respectively replaced by $\mathcal{O}(t^{(1+\alpha)/2})$ and $\mathcal{O}(\log t)$.

We now continue with the proof of Theorem 1. We distinguish two cases.

Case 1: $h_0(y) = \mathcal{O}(y^\nu e^{-\frac{v}{2}y})$ for some ν

We introduce $H_1(t)$ such that $xH_1(t)$ is the same as $H(x, t)$ with the sinh expanded to first order:

$$H_1(t) = \int_0^\infty dy \left(h_0(y) e^{\frac{v}{2}y} \right) y e^{\frac{\delta(t)}{2t}y} \psi_t(y) e^{-\frac{y^2}{4t}}. \quad (87)$$

For any $z \geq 0$, by Taylor's theorem (with the Lagrange remainder), there exists $w \in [0, z]$ such that $0 \leq \sinh(z) - z = \frac{z^3}{6} \cosh(w) \leq \frac{z^3}{6} e^z$. It follows that

$$\left| H(x, t) - xH_1(t) \right| \leq \frac{x^3}{24t^2} \int_0^\infty dy \left(|h_0(y)| e^{\frac{v}{2}y} \right) y^3 e^{\frac{x+\delta(t)}{2t}y} \psi_t(y) e^{-\frac{y^2}{4t}}. \quad (88)$$

By applying Lemma 9 to $\phi(y) = |h_0(y)| e^{\frac{v}{2}y} y^3$ with $\alpha = \nu + 3$ we obtain

$$H(x, t) - xH_1(t) = \begin{cases} \mathcal{O}(t^{\nu/2}) & \text{if } \nu > -4, \\ \mathcal{O}(t^{-2} \log t) & \text{if } \nu = -4, \\ \mathcal{O}(t^{-2}) & \text{if } \nu < -4. \end{cases} \quad (89)$$

We now apply Lemma 9 to $H_1(t)$ with $\alpha = \nu + 1$ and obtain

$$xH_1(t) \sim \begin{cases} x \frac{A}{2} e^\Delta \log t & \text{if } h_0(y) \sim Ay^{-2} e^{-\frac{v}{2}y} \text{ with } A > 0, \\ x A e^\Delta 2^{\nu+1} \Gamma\left(1 + \frac{\nu}{2}\right) t^{1+\frac{\nu}{2}} & \text{if } h_0(y) \sim Ay^\nu e^{-\frac{v}{2}y} \text{ with } A > 0 \text{ and } \nu > -2, \\ x \int_0^\infty dy h_0(y) y e^{\frac{v}{2}y} \psi_\infty(y) & \text{if } h_0(y) = \mathcal{O}(y^\nu e^{-\frac{v}{2}y}) \text{ for some } \nu < -2, \end{cases} \quad (90)$$

where we assumed that in the third case the right hand side is non-zero. As the difference (89) between $H(x, t)$ and $xH_1(t)$ is always asymptotically small compared to the values in the right hand side of (90), it follows that (90) also gives the asymptotic behaviour of $H(x, t)$.

We now plug this estimate of $H(x, t)$ into (83). To prevent $h(m(t) + x, t)$ from growing exponentially fast we need to take $\nu = 2$. Then $\delta(t)$ must be adjusted (up to a constant a) to kill the remaining time dependence. We find

$$\delta(t) = \begin{cases} -\frac{1-\nu}{2} \log t + a + o(1) & \text{if } h_0(y) \sim Ay^\nu e^{-y} \text{ with } A > 0 \text{ and } \nu > -2, \\ -\frac{3}{2} \log t + \log \log t + a + o(1) & \text{if } h_0(y) \sim Ay^{-2} e^{-y} \text{ with } A > 0, \\ -\frac{3}{2} \log t + a + o(1) & \text{if } h_0(y) = \mathcal{O}(y^\nu e^{-y}) \text{ for some } \nu < -2. \end{cases} \quad (91)$$

In (83), when $t \rightarrow \infty$, all the t -dependence disappears and what remains is $\omega(x)$ from the Theorem, with the claimed value of α . This proves cases (b), (c) and (d) of Theorem 1.

Case 2: $h_0(y) \sim Ay^\nu e^{-\gamma y}$ with $\gamma < \nu/2$

We write $h_0(y) = g_0(y)e^{-\gamma y}$ with $g_0(y) \sim Ay^\nu$ so that (84) becomes

$$H(x, t) = 2t \int_0^\infty dy g_0(y) \sinh\left(\frac{xy}{2t}\right) \psi_t(y) e^{\frac{\delta(t)}{2t}y} e^{\frac{\nu}{2}y - \gamma y - \frac{y^2}{4t}}. \quad (92)$$

The terms in the second exponential reach a maximum at $y = \lambda t$ with $\lambda = \nu - 2\gamma$. We make the change of variable $y = \lambda t + u\sqrt{t}$; after rearranging we have

$$H(x, t) = 2t^{\nu+\frac{3}{2}} e^{\frac{\lambda^2}{4}t + \lambda\frac{\delta(t)}{2}} \int_{-\lambda\sqrt{t}}^\infty du \frac{g_0(\lambda t + u\sqrt{t})}{t^\nu} \sinh\left(\frac{\lambda x}{2} + \frac{ux}{2\sqrt{t}}\right) \psi_t(\lambda t + u\sqrt{t}) e^{u\frac{\delta(t)}{2\sqrt{t}} - \frac{u^2}{4}}. \quad (93)$$

We bound each term in the integral with the goal of applying dominated convergence.

- As g_0 is bounded for small y and $g_0 \sim Ay^\nu$ for large y , we can take \tilde{A} such that $|g_0(y)| \leq \tilde{A}(y+1)^\nu$. Then

$$\frac{|g_0(\lambda t + u\sqrt{t})|}{t^\nu} \leq \tilde{A}\lambda^\nu \left(1 + \frac{u\sqrt{t} + 1}{\lambda t}\right)^\nu \leq \tilde{A}\lambda^\nu e^{\frac{|\nu|(u\sqrt{t}+1)}{\lambda t}} \leq 2\tilde{A}\lambda^\nu e^u \text{ for } t \text{ large enough.} \quad (94)$$

- We have the simple bound

$$\sinh\left(\frac{\lambda x}{2} + \frac{ux}{2\sqrt{t}}\right) \leq e^{\frac{\lambda x}{2} + \frac{ux}{2\sqrt{t}}} \leq e^{\frac{\lambda x}{2} + u} \text{ for } t \text{ large enough.} \quad (95)$$

- $\psi_t(\cdot)$ is bounded by Proposition 8.
- Finally, $\exp(u\delta(t)/(2\sqrt{t})) \leq e^u$ for t large enough.

We have bounded the integrand in (93) by a constant times $\exp(3u - u^2/4)$ for t large enough, so we can apply dominated convergence. As $t \rightarrow \infty$, the $g_0(\cdot)/t^\nu$ term converges to $A\lambda^\nu$, the $\sinh(\cdot)$ term to $\sinh(\lambda x/2)$, the $\psi_t(\cdot)$ term to e^Δ and the exponential to $e^{-u^2/4}$. We are left with some constants and the integral of $e^{-u^2/4}$, which is $\sqrt{4\pi}$, and finally:

$$H(x, t) \sim 2t^{\nu+\frac{3}{2}} e^{\frac{\lambda^2}{4}t + \lambda\frac{\delta(t)}{2}} A\lambda^\nu \sinh\left(\frac{\lambda x}{2}\right) e^\Delta \sqrt{4\pi}. \quad (96)$$

In (83), this gives

$$h(m(t) + x, t) = 2 \sinh\left(\frac{\lambda x}{2}\right) e^{-\frac{v}{2}x} \times e^{t\left(1 - \frac{v^2}{4} + \frac{\lambda^2}{4}\right) - \frac{v-\lambda}{2}\delta(t) + o(1)} t^\nu A \lambda^\nu. \quad (97)$$

Recall that $\lambda = v - 2\gamma$. To avoid exponential growth, we need $1 - v^2/4 + \lambda^2/4 = 0$, which implies $v = \gamma + 1/\gamma$ with $\gamma < 1$ because we started with the assumption $\gamma < v/2$. As $\frac{v-\lambda}{2} = \gamma$, to have convergence of $h(m(t) + x, t)$ we need $\delta(t)$ to be of the form

$$\delta(t) = \frac{\nu}{\gamma} \log t + a + o(1) \quad \text{for large } t. \quad (98)$$

Writing the $\sinh(\cdot)$ as the difference of two exponentials leads to $2 \sinh(\lambda x/2) e^{-vx/2} = e^{-\gamma x} - e^{-(1/\gamma)x}$; we then recover case (a) of Theorem 1 with the claimed value of $\omega(x)$ and α .

This completes the proof of Theorem 1, subject to proving Lemma 9. \square

Proof of Lemma 9. Recall from Proposition 8 that $\psi_t(y)$ is bounded in t and y , $\psi_\infty(y) := \lim_{t \rightarrow \infty} \psi_t(y)$ exists, $\lim_{y \rightarrow \infty} \psi_\infty(y)$ exists and equals e^Δ , and $\lim_{t \rightarrow \infty} \psi_t(t^\alpha) = e^\Delta$ for any $\alpha > 0$.

For $\alpha < -1$, the result is obtained with dominated convergence by noticing that $e^{-y^2/(4t) + \epsilon_t y}$ is bounded by $e^{t\epsilon_t^2}$ (value obtained at $y = 2t\epsilon_t$). With $\epsilon_t = o(t^{-1/2})$, this is bounded by a constant.

For $\alpha > -1$, cut the integral at $y = 1$. The integral from 0 to 1 is bounded, and in the integral from 1 to ∞ we make the substitution $y = u\sqrt{t}$:

$$\int_0^\infty dy \phi(y) e^{-\frac{y^2}{4t} + \epsilon_t y} \psi_t(y) = \mathcal{O}(1) + t^{\frac{1+\alpha}{2}} \int_{\frac{1}{\sqrt{t}}}^\infty du \frac{\phi(u\sqrt{t})}{t^{\alpha/2}} e^{-\frac{u^2}{4} + \sqrt{t}\epsilon_t u} \psi_t(u\sqrt{t}). \quad (99)$$

A simple application of dominated convergence then leads to

$$\int_0^\infty dy \phi(y) e^{-\frac{y^2}{4t} + \epsilon_t y} \psi_t(y) = \mathcal{O}(1) + t^{\frac{1+\alpha}{2}} \left(\int_0^\infty du A u^\alpha e^{-\frac{u^2}{4}} e^\Delta + o(1) \right), \quad (100)$$

and the substitution $t = u^2/4$ gives (86a).

For $\alpha = -1$, we cut the integral at $y = \sqrt{t}$ and again make the change of variable $y = u\sqrt{t}$ in the second part:

$$\int_0^\infty dy \phi(y) e^{-\frac{y^2}{4t} + \epsilon_t y} \psi_t(y) = \int_0^{\sqrt{t}} dy \phi(y) e^{-\frac{y^2}{4t} + \epsilon_t y} \psi_t(y) + \int_1^\infty du \sqrt{t} \phi(u\sqrt{t}) e^{-\frac{u^2}{4} + \sqrt{t}\epsilon_t u} \psi_t(u\sqrt{t}). \quad (101)$$

Again by dominated convergence, the second integral has a limit; we simply write it as $\mathcal{O}(1)$. For the first, the integrand is bounded so the integral from 0 to 1 is certainly $\mathcal{O}(1)$, and we may concentrate on the integral from 1 to \sqrt{t} . Making the substitution $y = t^x$, we have

$$\int_1^{\sqrt{t}} dy \phi(y) e^{-\frac{y^2}{4t} + \epsilon_t y} \psi_t(y) = (\log t) \int_0^{1/2} dx t^x \phi(t^x) e^{-\frac{t^{2x-1}}{4} + \epsilon_t t^x} \psi_t(t^x). \quad (102)$$

The integrand on the right converges for each $x \in (0, 1/2)$ to Ae^Δ so by dominated convergence,

$$\int_1^t dy \phi(y) e^{-\frac{y^2}{4t} + \epsilon_t y} \psi_t(y) \sim \frac{A}{2} e^\Delta \log t, \quad (103)$$

as required. \square

6 Estimating ψ_t : finer bounds, and Proof of Theorem 2

We want to refine Proposition 8 and estimate the speed of convergence of $\psi_t(y, x)$ to its limit as $t \rightarrow \infty$. As we are only interested up to errors of order $\frac{\log t}{t}$, it suffices to consider the case $x = 0$ since by (67), $\psi_t(y, x) = \psi_t(y)e^{x\mathcal{O}(\frac{\log t}{t})}$.

Recall that

$$\psi_t(y) = \mathbb{E}\left[e^{I_t(y)}\right], \quad \psi_\infty(y) = \mathbb{E}\left[e^{I(y)}\right], \quad (104)$$

where, introducing $I_t(y) := I_t(y, 0)$,

$$\begin{aligned} I_t(y) &= \frac{1}{2} \int_0^t ds m''(s) \left(\xi_s^{(t:y \rightarrow 0)} - y \frac{t-s}{t} \right) = \frac{1}{2} \int_0^t ds m''(s) \frac{t-s}{t} \left(\xi_{\frac{st}{t-s}}^{(y)} - y \right), \\ I(y) &= \frac{1}{2} \int_0^\infty ds m''(s) (\xi_s^{(y)} - y). \end{aligned} \quad (105)$$

We have used the change of time (30) to give the second expression of $I_t(y)$. As in the hypothesis (12) of Theorem 2, we suppose that m is twice continuously differentiable and

$$m''(t) = \frac{3}{2(t+1)^2} + r''(t) \quad \text{with} \quad r''(t) = \mathcal{O}\left(\frac{1}{t^{2+\eta}}\right), \quad \eta > 0. \quad (106)$$

Our estimate of $\psi_t(y)$ is based on the following two propositions. By writing $I_t(y) = I(y) - (I(y) - I_t(y))$ in the definition of $\psi_t(y)$, and expanding the exponential in the small correction term $I(y) - I_t(y)$, we show that:

Proposition 10. *Assuming (106), the following holds uniformly in y :*

$$\psi_t(y) = \psi_\infty(y) \left(1 - \mathbb{E}[I(y) - I_t(y)] \right) + \mathcal{O}\left(\frac{\log t}{t}\right) + y\mathcal{O}\left(\frac{1}{t}\right). \quad (107)$$

Further, some straightforward computations give that:

Proposition 11. *Assuming (106), the following holds uniformly in y :*

$$\mathbb{E}[I(y) - I_t(y)] = \frac{3\sqrt{\pi}}{\sqrt{t}} + y\mathcal{O}\left(\frac{\log t}{t}\right) + \begin{cases} \mathcal{O}\left(\frac{1}{t}\right) & \text{if } \eta > 1/2, \\ \mathcal{O}\left(\frac{\log t}{t}\right) & \text{if } \eta = 1/2, \\ \mathcal{O}\left(\frac{1}{t^{1/2+\eta}}\right) & \text{if } \eta < 1/2. \end{cases} \quad (108)$$

We prove Propositions 10 and 11 in Sections 6.2 and 6.3, after some preparatory work in Section 6.1. We now show how to prove Theorem 2 from these two propositions.

Proof of Theorem 2. We assume that $m(t)$ satisfies the hypothesis (12) of Theorem 2:

$$m(t) = 2t - \frac{3}{2} \log(t+1) + a + r(t) \quad \text{with} \quad r(t) = o(1) \quad \text{and} \quad r''(t) = \mathcal{O}\left(\frac{1}{t^{2+\nu}}\right) \quad \text{for large } t. \quad (109)$$

As in the proof of Theorem 1, we recall that $h(m(t) + x, t)$ is related to $H(x, t)$ through (83) and that $H(x, t)$ is given by (84). With $\nu = 2$ and $\delta(t) = -(3/2) \log(t+1) + a + r(t)$, these two equations read:

$$h(m(t) + x, t) = \frac{1}{\sqrt{4\pi}} e^{-a-r(t)-\Delta-x+\mathcal{O}(\frac{\log t}{t})} H(x, t), \quad (110)$$

$$H(x, t) = \int_0^\infty dy \left(h_0(y) e^y \right) 2t \sinh\left(\frac{xy}{2t}\right) e^{\frac{-(3/2) \log(t+1) + a + r(t)}{2t} y - \frac{y^2}{4t}} \psi_t(y), \quad (111)$$

We compute $H(x, t)$ for an initial condition $h_0(x) = \mathcal{O}(x^\nu e^{-x})$ for some $\nu < -2$. In (87) in the proof of Theorem 1, we introduced $H_1(t)$ which is $H(x, t)/x$ with the sinh replaced by its first order expansion:

$$H_1(t) = \int_0^\infty dy \left(h_0(y) e^y \right) y e^{\frac{-(3/2) \log(t+1) + a + r(t)}{2t} y - \frac{y^2}{4t}} \psi_t(y), \quad (112)$$

and we showed in (89) that the difference between $H(x, t)$ and $xH_1(t)$ is very small. We continue to simplify the integral by introducing successive simplifications

$$\begin{aligned} H_2(t) &= \int_0^\infty dy \left(h_0(y) e^y \right) y e^{-\frac{y^2}{4t}} \psi_t(y), \\ H_3(t) &= \int_0^\infty dy \left(h_0(y) e^y \right) y e^{-\frac{y^2}{4t}} \psi_\infty(y), \\ H_4 &= \int_0^\infty dy \left(h_0(y) e^y \right) y \psi_\infty(y), \end{aligned} \quad (113)$$

and by writing

$$\begin{aligned} H(x, t) &= \left(H(x, t) - xH_1(t) \right) + x \left(H_1(t) - H_2(t) \right) + x \left(H_2(t) - \left[1 - \frac{3\sqrt{\pi}}{\sqrt{t}} \right] H_3(t) \right) \\ &\quad + x \left[1 - \frac{3\sqrt{\pi}}{\sqrt{t}} \right] \left(H_3(t) - H_4 \right) + x \left[1 - \frac{3\sqrt{\pi}}{\sqrt{t}} \right] H_4. \end{aligned} \quad (114)$$

We now bound the successive differences in the above expression, as we did in (89), for the first one.

For t large enough, $-\frac{3}{2} \log(t+1) + a + r(t) < 0$ and for $z > 0$ we have $0 \leq 1 - e^{-z} \leq z$. Thus

$$\left| H_2(t) - H_1(t) \right| \leq \frac{\frac{3}{2} \log(t+1) - a - r(t)}{2t} \int_0^\infty dy \left(|h_0(y)| e^y \right) y^2 e^{-\frac{y^2}{4t}} \psi_t(y). \quad (115)$$

An application of Lemma 9 with $\phi(y) = h_0(y) e^y y^2$ and hence $\alpha = \nu + 2$ then gives

$$H_1(t) - H_2(t) = \begin{cases} \mathcal{O} \left(t^{\frac{1+\nu}{2}} \log t \right) & \text{if } \nu > -3, \\ \mathcal{O} \left(\frac{\log^2 t}{t} \right) & \text{if } \nu = -3, \\ \mathcal{O} \left(\frac{\log t}{t} \right) & \text{if } \nu < -3. \end{cases} \quad (116)$$

For the difference involving H_2 and H_3 , we use Propositions 10 and 11 which give that uniformly in y ,

$$\psi_t(y) = \psi_\infty(y) \left(1 - \frac{3\sqrt{\pi}}{\sqrt{t}} \right) + y \mathcal{O} \left(\frac{\log t}{t} \right) + \begin{cases} \mathcal{O} \left(\frac{1}{t} \right) & \text{if } \eta > 1/2, \\ \mathcal{O} \left(\frac{\log t}{t} \right) & \text{if } \eta = 1/2, \\ \mathcal{O} \left(\frac{1}{t^{1/2+\eta}} \right) & \text{if } \eta < 1/2. \end{cases} \quad (117)$$

We get

$$\begin{aligned} H_2(t) - \left(1 - \frac{3\sqrt{\pi}}{\sqrt{t}} \right) H_3(t) &= \int_0^\infty dy \left(h_0(y) e^y \right) y e^{-\frac{y^2}{4t}} \left[\psi_t(y) - \psi_\infty(y) \left(1 - \frac{3\sqrt{\pi}}{\sqrt{t}} \right) \right], \\ &= \mathcal{O} \left(\frac{1}{t^{1/2+\eta}} \right) + \begin{cases} \mathcal{O} \left(t^{\frac{1+\nu}{2}} \log t \right) & \text{if } \nu > -3, \\ \mathcal{O} \left(\frac{\log^2 t}{t} \right) & \text{if } \nu = -3, \\ \mathcal{O} \left(\frac{\log t}{t} \right) & \text{if } \nu < -3. \end{cases} \end{aligned} \quad (118)$$

Indeed, the $y\mathcal{O}\left(\frac{\log t}{t}\right)$ gives the same correction as in (116) by another application of Lemma 9 with $\alpha = \nu + 2$. As $\int dy |h_0(y)|e^y y < \infty$ because $\nu < -2$, the contribution of the $y\mathcal{O}\left(\frac{\log t}{t}\right)$ term subsumes the other \mathcal{O} in (117) except in the case $\eta < \frac{1}{2}$.

Finally, notice that $|H_4| < \infty$ because we supposed $\nu < -2$. Recalling $\psi_\infty(y) \leq K_2$, one has

$$\begin{aligned} |H_4 - H_3(t)| &\leq \int_0^\infty dy \left(|h_0(y)|e^y \right) y \left(1 - e^{-\frac{y^2}{4t}} \right) \psi_\infty(y), \\ &\leq K_2 \int_0^{\sqrt{t}} dy \left(|h_0(y)|e^y \right) y \frac{y^2}{4t} + K_2 \int_{\sqrt{t}}^\infty dy \left(|h_0(y)|e^y \right) y, \\ &= \begin{cases} \mathcal{O}\left(t^{1+\frac{\nu}{2}}\right) & \text{if } -2 > \nu > -4, \\ \mathcal{O}\left(\frac{\log t}{t}\right) & \text{if } \nu = -4, \\ \mathcal{O}\left(\frac{1}{t}\right) & \text{if } \nu < -4, \end{cases} \end{aligned} \tag{119}$$

where we used $h_0(y)e^y = \mathcal{O}(y^\nu)$. The end result comes from the integral from 0 to \sqrt{t} ; the other integral is always $\mathcal{O}(t^{1+\nu/2})$.

Finally, collecting the differences (89), (116), (118) and (119) leads with (114) to

$$H(x, t) = xH_4 \left[1 - \frac{3\sqrt{\pi}}{\sqrt{t}} + \mathcal{O}\left(t^{1+\frac{\nu}{2}}\right) + \mathcal{O}\left(\frac{1}{t^{1/2+\eta}}\right) + \mathcal{O}\left(\frac{\log t}{t}\right) \right]. \tag{120}$$

Substituting into (110) and expanding $e^{-r(t)}$ leads to the main expression (13) of Theorem 2, with the value α given in Theorem 1.

We now turn to the second part of Theorem 2 and assume that $h_0(y) \sim Ay^\nu e^{-y}$ with $-4 < \nu < -2$. We look for an estimate of $H_4 - H_3(t)$ which is more precise than (119).

Writing $H_4 - H_3(t)$ as a single integral and doing the change of variable $y = u\sqrt{t}$ one gets

$$H_4 - H_3(t) = t^{1+\frac{\nu}{2}} \int_0^\infty du \frac{h_0(u\sqrt{t})e^{u\sqrt{t}}}{t^{\nu/2}} u \left(1 - e^{-\frac{u^2}{4}} \right) \psi_\infty(u\sqrt{t}). \tag{121}$$

A simple application of dominated convergence then gives

$$H_4 - H_3(t) \sim t^{1+\frac{\nu}{2}} Ae^\Delta \int_0^\infty du u^{\nu+1} \left(1 - e^{-\frac{u^2}{4}} \right) = -Ae^\Delta 2^{\nu+1} \Gamma\left(\frac{\nu}{2} + 1\right) t^{1+\frac{\nu}{2}}, \tag{122}$$

and (120) becomes

$$H(x, t) = xH_4 \left[1 - \frac{3\sqrt{\pi}}{\sqrt{t}} \right] + xAe^\Delta 2^{\nu+1} \Gamma\left(\frac{\nu}{2} + 1\right) t^{1+\frac{\nu}{2}} + o\left(t^{1+\frac{\nu}{2}}\right) + \mathcal{O}\left(\frac{1}{t^{\frac{1}{2}+\eta}}\right). \tag{123}$$

This leads with (110) to (13). □

6.1 Decorrelation between $I(y)$ and $\xi_s^{(y)}$

A large part of our argument relies on a statement that roughly says “ $I(y)$ and $\xi_s^{(y)}$ are almost independent for large s ”. The following proposition makes this precise.

Proposition 12. *Suppose that m is twice continuously differentiable with $m''(t) = \mathcal{O}(1/t^2)$. Define*

$$w(y, s) = \mathbb{E}\left[e^{I(y)} (\xi_s^{(y)} - y)\right] - \mathbb{E}\left[e^{I(y)}\right] \mathbb{E}[\xi_s^{(y)} - y]. \tag{124}$$

There exists a constant $C > 0$ such that

$$\begin{aligned} |w(y, s)| &\leq C \log(s+1) && \text{for all } s, y \geq 0, \\ |w(y, s)| &\leq C \left(1 + y \frac{\log(s+1)}{\sqrt{s}} \right) && \text{for all } s, y \geq 0, \\ |w(y, s+\delta) - w(y, s)| &\leq C \frac{\delta}{s+1} && \text{for all } y \geq 0, \text{ whenever } 0 \leq \delta \leq s^2. \end{aligned} \tag{125}$$

The proof of this result is quite involved. The first step is to prove two fairly accurate estimates on the difference between two bridges with different end points, the first of which is best when the starting point y is large and the second of which is more accurate when y is small.

It is well-known that a Bessel process started from y and conditioned to be at position x at time t is equal in law to a Bessel bridge from y to x in time t followed by an independent Bessel process started from x at time t . We defined $\xi_s^{(t:y \rightarrow x)}$ for $s \in [0, t]$ as a Bessel bridge from y to x in a time t . In this section, we extend the definition of $\xi_s^{(t:y \rightarrow x)}$ for $s > t$ by interpreting it as an independent Bessel started from x at time t , so that $\xi_s^{(t:y \rightarrow x)}$, $s \geq 0$ is a Bessel process conditioned to be at x at time t . We assume that the Bessel processes attached to $\xi_s^{(t:y \rightarrow x)}$ for $s \geq t$ are built for all x and t with the same noise, so that we can compare them to each other. In particular, we apply (34) and (49) to these Bessel processes.

Recall that $I(y) = \frac{1}{2} \int_0^\infty du m''(u)(\xi_u^{(y)} - y)$ and define

$$\tilde{I}_t(y, z) = \frac{1}{2} \int_0^\infty du m''(u)(\xi_u^{(t:y \rightarrow z)} - y). \quad (126)$$

Lemma 13. *If m is twice continuously differentiable with $m''(t) = \mathcal{O}(1/t^2)$, then there exists a constant c and random variables G_t with distribution independent of t and Gaussian tails such that:*

- For any t, y, z and x ,

$$|\tilde{I}_t(y, z) - \tilde{I}_t(y, x)| \leq c|z - x| \frac{\log(t+1)}{t}. \quad (127)$$

- For any t, y and z ,

$$|\tilde{I}_t(y, z) - \tilde{I}_t(y, 0)| \leq \frac{z^2}{t^{3/2}} G_t + c \left(\frac{z}{t} + \frac{z^3}{t^2} + \frac{z^2 y}{t^2} \log(t+1) \right). \quad (128)$$

Proof. Recall from (34) and (35) that $|\xi_s^{(t:y \rightarrow z)} - \xi_s^{(t:y \rightarrow x)}| \leq |z - x| \min(s/t, 1)$. Therefore

$$|\tilde{I}_t(y, z) - \tilde{I}_t(y, x)| \leq \frac{1}{2} \int_0^\infty ds |m''(s)| \left| \xi_s^{(t:y \rightarrow z)} - \xi_s^{(t:y \rightarrow x)} \right| \quad (129)$$

$$\leq \frac{1}{2} |z - x| \left(\int_0^t ds |m''(s)| \frac{s}{t} + \int_t^\infty ds |m''(s)| \right). \quad (130)$$

The first integral is $\mathcal{O}(\frac{\log t}{t})$ while the second is a $\mathcal{O}(1/t)$. Their sum can be bounded by $2c \log(t+1)/t$ for some c , which proves the simpler bound (127).

To prove (128) we consider $x = 0$ and split the integral at $t/2$ and t . For $s > t/2$, with the same simple bounds as above we have

$$\left| \int_{\frac{t}{2}}^\infty ds m''(s) \left(\xi_s^{(t:y \rightarrow z)} - \xi_s^{(t:y \rightarrow 0)} \right) \right| \leq z \left(\int_{\frac{t}{2}}^t ds |m''(s)| \frac{s}{t} + \int_t^\infty ds |m''(s)| \right) = z \mathcal{O}\left(\frac{1}{t}\right). \quad (131)$$

From 0 to $t/2$, we claim that the following bound is true:

$$0 \leq \int_0^{\frac{t}{2}} ds \frac{\xi_s^{(t:y \rightarrow z)} - \xi_s^{(t:y \rightarrow 0)}}{(1+s)^2} \leq \frac{z^3}{3t^2} + \frac{z^2}{t^{3/2}} G_t + \frac{2z^2 y}{3t^2} \log(t+1), \quad (132)$$

for some non-negative G_t with distribution independent of t and Gaussian tails. Then, as there exists some constant c' such that $|m''(s)| \leq c'/(1+s)^2$, (131) and (132) give the result (128). Therefore it only remains to prove (132).

We use the bound $\coth(x) \leq 1/x + x/3$, together with the SDEs (31) and (33). We already know from Lemma 5 that $\xi_s^{(t:y \rightarrow 0)} \leq \xi_s^{(t:y \rightarrow z)} \leq \xi_s^{(t:y \rightarrow 0)} + zs/t$ for any $s \in [0, t]$. Therefore for any $s \in [0, t]$,

$$d\xi_s^{(t:y \rightarrow z)} - d\xi_s^{(t:y \rightarrow 0)} \leq \left(\frac{z}{t-s} \coth \frac{z\xi_s^{(t:y \rightarrow z)}}{2(t-s)} - \frac{2}{\xi_s^{(t:y \rightarrow 0)}} \right) ds \quad (133)$$

$$\leq \frac{z^2 \xi_s^{(t:y \rightarrow z)}}{6(t-s)^2} ds \quad (134)$$

$$\leq \left(\frac{z^3 s}{6t(t-s)^2} + \frac{z^2 \xi_s^{(t:y \rightarrow 0)}}{6(t-s)^2} \right) ds. \quad (135)$$

By integration by parts,

$$\int_0^{\frac{t}{2}} ds \frac{\xi_s^{(t:y \rightarrow z)}}{(s+1)^2} = \int_0^{\frac{t}{2}} \frac{1}{s+1} d\xi_s^{(t:y \rightarrow z)} - \frac{\xi_{t/2}^{(t:y \rightarrow z)}}{t/2+1} + y. \quad (136)$$

Using (35), the estimate on $d\xi_s^{(t:y \rightarrow z)} - d\xi_s^{(t:y \rightarrow 0)}$ from above, and $t-s \geq t/2$ for $s \leq t/2$, we get

$$0 \leq \int_0^{\frac{t}{2}} ds \frac{\xi_s^{(t:y \rightarrow z)} - \xi_s^{(t:y \rightarrow 0)}}{(s+1)^2} \leq \int_0^{\frac{t}{2}} \frac{1}{s+1} d\xi_s^{(t:y \rightarrow z)} - \int_0^{\frac{t}{2}} \frac{1}{s+1} d\xi_s^{(t:y \rightarrow 0)} \quad (137)$$

$$\leq \int_0^{\frac{t}{2}} ds \frac{z^3 s}{6t(t-s)^2(s+1)} + \int_0^{\frac{t}{2}} ds \frac{z^2 \xi_s^{(t:y \rightarrow 0)}}{6(t-s)^2(s+1)} \quad (138)$$

$$\leq \frac{2z^3}{3t^3} \int_0^{\frac{t}{2}} ds \frac{s}{s+1} + \frac{2z^2}{3t^2} \int_0^{\frac{t}{2}} ds \frac{\xi_s^{(t:y \rightarrow 0)}}{s+1} \quad (139)$$

$$\leq \frac{z^3}{3t^2} + \frac{2z^2}{3t^2} \int_0^{\frac{t}{2}} ds \frac{y + \xi_s^{(t:0 \rightarrow 0)}}{s+1} \quad (140)$$

$$\leq \frac{z^3}{3t^2} + \frac{2z^2 y}{3t^2} \log(t+1) + \frac{2z^2 y}{3t^2} \int_0^{\frac{t}{2}} ds \frac{\xi_s^{(t:0 \rightarrow 0)}}{s}. \quad (141)$$

By the scaling property, we introduce another Bessel bridge $\tilde{\xi}^{(1:0 \rightarrow 0)}$ by setting $\xi_{tu}^{(t:0 \rightarrow 0)} = \sqrt{t} \tilde{\xi}_u^{(1:0 \rightarrow 0)}$. By adapting Lemma 7 to the new Bessel bridge, there exists a random variable G_t with distribution independent of t and Gaussian tails such that $\tilde{\xi}_u^{(1:0 \rightarrow 0)} \leq G_t u^{\frac{1}{4}}$. Hence

$$\int_0^{\frac{t}{2}} ds \frac{\xi_s^{(t:0 \rightarrow 0)}}{s} = \int_0^{\frac{1}{2}} du \frac{\xi_{tu}^{(t:0 \rightarrow 0)}}{u} \leq \sqrt{t} G_t \int_0^{\frac{1}{2}} du u^{-\frac{3}{4}} \leq 4G_t \sqrt{t}. \quad (142)$$

This bounds the last term in (141) and establishes (132), thereby completing the proof. \square

Finally, given that we are using random variables with Gaussian tails, the following trivial result is useful.

Lemma 14. *Suppose that G is a random variable with Gaussian tails. Then for any real number a and any polynomial P ,*

$$|\mathbb{E}[P(G)e^{aG}]| < \infty. \quad (143)$$

We can now prove Proposition 12.

Proof of Proposition 12. Recall the definition(126) of \tilde{I} . For any deterministic x , since $\mathbb{E}[\xi_s^{(y)}] - \mathbb{E}[\xi_s^{(y)}] = 0$ and $\mathbb{E}[e^{\tilde{I}_s(y,x)}]$ is deterministic, we have

$$w(y, s) = \mathbb{E}\left[e^{I(y)}\left(\xi_s^{(y)} - \mathbb{E}[\xi_s^{(y)}]\right)\right] \quad (144)$$

$$= \mathbb{E}\left[\left(e^{I(y)} - \mathbb{E}[e^{\tilde{I}_s(y,x)}]\right)\left(\xi_s^{(y)} - \mathbb{E}[\xi_s^{(y)}]\right)\right] \quad (145)$$

$$= \int_0^\infty \left(\mathbb{E}[e^{I(y)}|\xi_s^{(y)} = z] - \mathbb{E}[e^{\tilde{I}_s(y,x)}]\right)\left(z - \mathbb{E}[\xi_s^{(y)}]\right)\mathbb{P}(\xi_s^{(y)} \in dz) \quad (146)$$

$$= \int_0^\infty \mathbb{E}\left[e^{\tilde{I}_s(y,z)} - e^{\tilde{I}_s(y,x)}\right]\left(z - \mathbb{E}[\xi_s^{(y)}]\right)\mathbb{P}(\xi_s^{(y)} \in dz), \quad (147)$$

where we used that $\mathbb{E}[e^{I(y)}|\xi_s^{(y)} = z] = \mathbb{E}[e^{\tilde{I}_s(y,z)}]$. Then

$$|w(y, s)| \leq \int_0^\infty \mathbb{E}\left[\left|e^{\tilde{I}_s(y,z)} - e^{\tilde{I}_s(y,x)}\right|\right] \times \left|z - \mathbb{E}[\xi_s^{(y)}]\right| \times \mathbb{P}(\xi_s^{(y)} \in dz). \quad (148)$$

By the mean value theorem, $|e^a - e^b| \leq |a - b|e^{\max(a,b)} \leq |a - b|e^{b+|a-b|}$. Thus

$$|w(y, s)| \leq \int_0^\infty \mathbb{E}\left[e^{\tilde{I}_s(y,x)}\left|\tilde{I}_s(y, z) - \tilde{I}_s(y, x)\right|e^{|\tilde{I}_s(y,z) - \tilde{I}_s(y,x)|}\right] \left|z - \mathbb{E}[\xi_s^{(y)}]\right| \times \mathbb{P}(\xi_s^{(y)} \in dz) \quad (149)$$

$$\leq \int_0^\infty \mathbb{E}\left[e^{\tilde{I}_s(y,x)}\left|\tilde{I}_s(y, z) - \tilde{I}_s(y, x)\right|\right] e^{c|z-x|\frac{\log(s+1)}{s}} \left|z - \mathbb{E}[\xi_s^{(y)}]\right| \times \mathbb{P}(\xi_s^{(y)} \in dz), \quad (150)$$

where we applied (127) of Lemma 13 in the exponential. Now, by Cauchy-Schwarz,

$$|w(y, s)| \leq \mathbb{E}\left[e^{2\tilde{I}_s(y,x)}\right]^{\frac{1}{2}} \int_0^\infty \mathbb{E}\left[\left|\tilde{I}_s(y, z) - \tilde{I}_s(y, x)\right|^2\right]^{\frac{1}{2}} e^{c|z-x|\frac{\log(s+1)}{s}} \left|z - \mathbb{E}[\xi_s^{(y)}]\right| \mathbb{P}(\xi_s^{(y)} \in dz). \quad (151)$$

Decompose $\tilde{I}_s(y, x)$ in the following way:

$$\begin{aligned} 2\tilde{I}_s(y, x) &= \int_0^s du m''(u)\left(\xi_u^{(s:y \rightarrow x)} - y - (x-y)\frac{u}{s}\right) + \int_s^\infty du m''(u)\left(\xi_u^{(s:y \rightarrow x)} - x\right) \\ &\quad + \int_0^s du m''(u)(x-y)\frac{u}{s} + \int_s^\infty du m''(u)(x-y). \end{aligned} \quad (152)$$

The first integral is $2I_s(y, x)$. Using (50) it can be bounded uniformly in y, x and s by a variable with Gaussian tails. The second integral, which does not depend on y , can also be bounded uniformly in x and s using (49) by an independent variable with Gaussian tails. The third integral is $(x-y)\mathcal{O}\left(\frac{\log s}{s}\right)$ and the fourth is $(x-y)\mathcal{O}\left(\frac{1}{s}\right)$; they can be bounded together by $2c|x-y|\frac{\log(s+1)}{s}$ for some constant c . Finally, there exists a C_1 and a c such that, uniformly in s, y and x :

$$\mathbb{E}\left[e^{2\tilde{I}_s(y,x)}\right]^{\frac{1}{2}} \leq C_1 e^{c|x-y|\frac{\log(1+s)}{s}}. \quad (153)$$

Substituting back into (151), we get

$$|w(y, s)| \leq C_1 e^{c|x-y|\frac{\log(1+s)}{s}} \int_0^\infty \mathbb{E}\left[\left|\tilde{I}_s(y, z) - \tilde{I}_s(y, x)\right|^2\right]^{\frac{1}{2}} e^{c|z-x|\frac{\log(1+s)}{s}} \left|z - \mathbb{E}[\xi_s^{(y)}]\right| \mathbb{P}(\xi_s^{(y)} \in dz). \quad (154)$$

First we concentrate on showing the first line of (125), i.e. that $|w(y, s)| \leq C \log(s+1)$. Using (127) again,

$$\mathbb{E}\left[\left|\tilde{I}_s(y, z) - \tilde{I}_s(y, x)\right|^2\right]^{1/2} \leq c|z-x|\frac{\log(s+1)}{s}, \quad (155)$$

so we get, by choosing $x = \mathbb{E}[\xi_s^{(y)}]$,

$$\begin{aligned} |w(y, s)| &\leq C_1 c \frac{\log(s+1)}{s} e^{c|\mathbb{E}[\xi_s^{(y)}] - y| \frac{\log(1+s)}{s}} \int_0^\infty e^{c|z - \mathbb{E}[\xi_s^{(y)}]| \frac{\log(1+s)}{s}} (z - \mathbb{E}[\xi_s^{(y)}])^2 \mathbb{P}(\xi_s^{(y)} \in dz), \\ &= C_1 c \frac{\log(s+1)}{s} e^{c|\mathbb{E}[\xi_s^{(y)}] - y| \frac{\log(1+s)}{s}} \mathbb{E}\left[e^{c|\xi_s^{(y)} - \mathbb{E}[\xi_s^{(y)}]| \frac{\log(1+s)}{s}} (\xi_s^{(y)} - \mathbb{E}[\xi_s^{(y)}])^2\right]. \end{aligned} \quad (156)$$

It remains to bound the expectations above. Note from (34) that for all $z \geq 0$ we have $B_1 \leq \xi_1^{(z)} - z \leq \xi_1^{(0)}$ and therefore

$$B_1 - \mathbb{E}[\xi_1^{(0)}] \leq \xi_1^{(z)} - \mathbb{E}[\xi_1^{(z)}] \leq \xi_1^{(z)} - z \leq \xi_1^{(0)}, \quad (157)$$

so, with Γ the positive random variable with Gaussian tail defined by

$$\Gamma := \max\{|B_1 - \mathbb{E}[\xi_1^{(0)}]|, |\xi_1^{(0)}|\}, \quad (158)$$

we have, uniformly in z ,

$$|\xi_1^{(z)} - \mathbb{E}[\xi_1^{(z)}]| \leq \Gamma, \quad |\xi_1^{(z)} - z| \leq \Gamma. \quad (159)$$

Therefore, by the scaling property,

$$|w(y, s)| \leq C_1 c \frac{\log(s+1)}{s} e^{c\sqrt{s}\mathbb{E}[\Gamma] \frac{\log(1+s)}{s}} \mathbb{E}\left[e^{c\sqrt{s}\Gamma \frac{\log(1+s)}{s}} s\Gamma^2\right] \leq C \log(s+1), \quad (160)$$

for some constant C , where we used Lemma 14 to bound the last expectation. This is the first line of (125).

We now turn to showing the second line of (125), that $|w(y, s)| \leq C(1 + y \frac{\log(s+1)}{\sqrt{s}})$. Given that we have already proven that $|w(y, s)| \leq C \log(s+1)$, it suffices to consider $y \leq \sqrt{s}$.

Recall (128):

$$\left| \tilde{I}_s(y, z) - \tilde{I}_s(y, 0) \right| \leq \frac{z^2}{s^{3/2}} G_s + c \left(\frac{z}{s} + \frac{z^3}{s^2} + \frac{z^2 y}{s^2} \log(s+1) \right). \quad (161)$$

By Cauchy-Schwarz, if $a, b \geq 0$ and X is a non-negative random variable with finite second moment, then

$$\mathbb{E}[(aX + b)^2]^{1/2} \leq a\mathbb{E}[X^2]^{1/2} + b. \quad (162)$$

This tells us that

$$\mathbb{E}[|\tilde{I}_s(y, z) - \tilde{I}_s(y, 0)|^2]^{1/2} \leq C_2 a_{s,y,z} \quad \text{with } a_{s,y,z} = \frac{z}{s} + \frac{z^2}{s^{3/2}} + \frac{z^3}{s^2} + \frac{z^2 y}{s^2} \log(s+1) \quad (163)$$

for some constant C since the distribution of G_s does not depend on s .

Now choosing $x = 0$ in (154) and substituting (163), we get

$$|w(y, s)| \leq C_1 C_2 e^{cy \frac{\log(1+s)}{s}} \int_0^\infty a_{s,y,z} e^{cz \frac{\log(1+s)}{s}} |z - \mathbb{E}[\xi_s^{(y)}]| \mathbb{P}(\xi_s^{(y)} \in dz). \quad (164)$$

$$\leq C_3 \mathbb{E}\left[a_{s,y,\xi_s^{(y)}} e^{c\xi_s^{(y)} \frac{\log(1+s)}{s}} |\xi_s^{(y)} - \mathbb{E}[\xi_s^{(y)}]| \right], \quad (165)$$

where we used $y \leq \sqrt{s}$ to bound the factor in front of the integral by a constant. Using the scaling property, writing $\tilde{\xi}_1 = \xi_s^{(y)}/\sqrt{s}$ we have as in (159)

$$|\tilde{\xi}_1 - \mathbb{E}[\tilde{\xi}_1]| \leq \Gamma, \quad |\tilde{\xi}_1 - y/\sqrt{s}| \leq \Gamma, \quad \tilde{\xi}_1 \leq 1 + \Gamma \quad (166)$$

for some positive random variable Γ with Gaussian tails; we used $y \leq \sqrt{s}$ in the last equation. Then

$$|w(y, s)| \leq C_3 \mathbb{E} \left[a_{s,y,\sqrt{s}(1+\Gamma)} e^{c\sqrt{s}(1+\Gamma)\frac{\log(1+s)}{s}} \sqrt{s} \Gamma \right], \quad (167)$$

but

$$a_{s,y,\sqrt{s}X}\sqrt{s} = X + X^2 + X^3 + X^2 y \frac{\log(s+1)}{\sqrt{s}}, \quad (168)$$

so using Lemma 14 again we obtain $|w(y, s)| \leq C \left(1 + y \frac{\log(s+1)}{\sqrt{s}}\right)$ for some constant C , which is the second line of (125).

Finally we turn to the last line of (125) and bound the increments of $w(y, s)$. Our approach is very similar to the above, conditioning on the value of $\xi_{s+\delta}^{(y)} - \xi_s^{(y)}$ instead of $\xi_s^{(y)}$.

Let $X = \xi_{s+\delta}^{(y)} - \xi_s^{(y)}$ and $\mu = \mathbb{E}[X]$, and also define

$$\mathcal{E}(x) = \mathbb{E} \left[e^{I(y)} \middle| X = x \right]. \quad (169)$$

Directly from the definition (124) of w , since $\mathcal{E}(\mu)$ is deterministic and $\mathbb{E}[X - \mu] = 0$, we have

$$w(y, s + \delta) - w(y, s) = \mathbb{E} \left[e^{I(y)} (X - \mu) \right], \quad (170)$$

$$= \mathbb{E} \left[\left(e^{I(y)} - \mathcal{E}(\mu) \right) (X - \mu) \right], \quad (171)$$

$$= \int_{-\infty}^{\infty} (\mathcal{E}(x) - \mathcal{E}(\mu)) (x - \mu) \mathbb{P}(X \in dx). \quad (172)$$

Applying the Markov property at time s , we have

$$\begin{aligned} \mathcal{E}(x) - \mathcal{E}(\mu) &= \int_0^\infty \mathbb{P}(\xi_s^{(y)} \in dz) \mathbb{E} \left[e^{\frac{1}{2} \int_0^s du m''(u) (\xi_u^{(s:y \rightarrow z)} - y)} \right] \\ &\quad \cdot \mathbb{E} \left[e^{\frac{1}{2} \int_0^\infty du m''(s+u) (\xi_u^{(\delta:z \rightarrow z+x)} - y)} - e^{\frac{1}{2} \int_0^\infty du m''(s+u) (\xi_u^{(\delta:z \rightarrow z+\mu)} - y)} \right]. \end{aligned} \quad (173)$$

We now use the simple bound

$$|e^{\xi_u^{(\delta:z \rightarrow z+x)}} - e^{\xi_u^{(\delta:z \rightarrow z+x')}}| \leq |x - x'| \quad \text{for all } z, x, x', \delta, u \geq 0, \quad (174)$$

which follows from Lemma 5 and implies that

$$\left| \int_0^\infty du m''(s+u) (\xi_u^{(\delta:z \rightarrow z+x)} - \xi_u^{(\delta:z \rightarrow z+\mu)}) \right| \leq \int_0^\infty du |m''(s+u)| |x - \mu| \leq \frac{2c|x - \mu|}{s+1} \quad (175)$$

for some constant c . This, together with the bound $|e^a - e^b| \leq |a - b| e^{b+|a-b|}$ for any $a, b \in \mathbb{R}$, tells us that

$$\begin{aligned} &\left| e^{\frac{1}{2} \int_0^\infty du m''(s+u) (\xi_u^{(\delta:z \rightarrow z+x)} - y)} - e^{\frac{1}{2} \int_0^\infty du m''(s+u) (\xi_u^{(\delta:z \rightarrow z+\mu)} - y)} \right| \\ &\leq e^{\frac{1}{2} \int_0^\infty du m''(s+u) (\xi_u^{(\delta:z \rightarrow z+\mu)} - y)} \frac{c|x - \mu|}{s+1} e^{\frac{c|x-\mu|}{s+1}}. \end{aligned} \quad (176)$$

Substituting this into (173), we have

$$\begin{aligned} |\mathcal{E}(x) - \mathcal{E}(\mu)| &\leq \int_0^\infty \mathbb{P}(\xi_s^{(y)} \in dz) \mathbb{E} \left[e^{\frac{1}{2} \int_0^s du m''(u) (\xi_u^{(s:y \rightarrow z)} - y)} \right] \\ &\quad \cdot \mathbb{E} \left[e^{\frac{1}{2} \int_0^\infty du m''(s+u) (\xi_u^{(\delta:z \rightarrow z+\mu)} - y)} \right] \frac{c|x - \mu|}{s+1} e^{\frac{c|x-\mu|}{s+1}} \end{aligned} \quad (177)$$

$$= \mathcal{E}(\mu) \frac{c|x - \mu|}{s+1} e^{\frac{c|x-\mu|}{s+1}}. \quad (178)$$

Returning to (172), we obtain

$$|w(y, s + \delta) - w(y, s)| \leq \int_{-\infty}^{\infty} \mathcal{E}(\mu) \frac{c|x - \mu|}{s + 1} e^{\frac{c|x - \mu|}{s + 1}} |x - \mu| \mathbb{P}(X \in dx) \quad (179)$$

$$= \mathcal{E}(\mu) \mathbb{E} \left[\frac{c(X - \mu)^2}{s + 1} e^{\frac{c|X - \mu|}{s + 1}} \right]. \quad (180)$$

Finally, by scaling, conditionally on $\xi_s^{(y)} = z$ we have

$$|X - \mu| \stackrel{(d)}{=} \sqrt{\delta} \left| \xi_1^{(z/\sqrt{\delta})} - \mathbb{E} \left[\xi_1^{(z/\sqrt{\delta})} \right] \right| \leq \sqrt{\delta} \Gamma, \quad (181)$$

where Γ was defined in (158) and is a non-negative random variable with Gaussian tail. Therefore

$$\mathbb{E} \left[\frac{c|X - \mu|^2}{s + 1} e^{\frac{c|X - \mu|}{s + 1}} \right] \leq C \frac{\delta}{s + 1} \quad (182)$$

for some constant C provided $\delta \leq s^2$, and one may check similarly to (153) that $\mathcal{E}(\mu)$ is also bounded uniformly in y , s and δ . This establishes the last line of (125) and completes the proof. \square

6.2 Proof of Proposition 10

To prove Proposition 10 we proceed via three lemmas. We first write $I_t(y) = I(y) - (I(y) - I_t(y))$, and show that the correction $I(y) - I_t(y)$ is small in the following sense:

Lemma 15. *Suppose that m is twice continuously differentiable and satisfies (106). Then there exist positive random variables G and G_t with Gaussian tails, where all the G_t have the same distribution, such that uniformly in y ,*

$$I(y) = G\mathcal{O}(1) \quad \text{and} \quad I(y) - I_t(y) = G_t\mathcal{O}(t^{-\frac{1}{2}}). \quad (183)$$

Unsurprisingly, for random variables with Gaussian tails we can make series expansions rather easily:

Lemma 16. *Let G and G_t be positive random variables with Gaussian tails such that all the G_t have the same distribution. Suppose that A_t and B_t are random variables such that*

$$A_t = G\mathcal{O}(1), \quad B_t = G_t\mathcal{O}(\epsilon_t) \quad (184)$$

where $\epsilon_t \geq 0$ is a deterministic function with $\epsilon_t \rightarrow 0$ as $t \rightarrow \infty$. Then for any integer $n \geq 0$,

$$\mathbb{E}[e^{A_t + B_t}] = \sum_{p=0}^n \frac{1}{p!} \mathbb{E}[e^{A_t} B_t^p] + \mathcal{O}(\epsilon_t^{n+1}). \quad (185)$$

Taking $n = 1$, $\epsilon_t = t^{-1/2}$, $A_t = I(y)$ and $B_t = -(I(y) - I_t(y))$, we find

$$\mathbb{E}[e^{I_t(y)}] = \mathbb{E}[e^{I(y)}] - \mathbb{E}[e^{I(y)}(I(y) - I_t(y))] + \mathcal{O}\left(\frac{1}{t}\right). \quad (186)$$

The difficult part is then to show how the $I(y)$ decorrelates asymptotically from $I(y) - I_t(y)$:

Lemma 17. *Suppose that m is twice continuously differentiable with $m''(t) = \frac{3}{2(t+1)^2} + r''(t)$ where $r(t) = \mathcal{O}(t^{-2-\eta})$ for some $\eta > 0$. Then*

$$\mathbb{E}[e^{I(y)}(I(y) - I_t(y))] = \mathbb{E}[e^{I(y)}] \mathbb{E}[I(y) - I_t(y)] + \mathcal{O}\left(\frac{\log t}{t}\right) + y\mathcal{O}\left(\frac{1}{t}\right). \quad (187)$$

Of course $\psi_\infty(y) = \mathbb{E}[e^{I(y)}]$ and $\psi_t(y) = \mathbb{E}[e^{I_t(y)}]$, so these lemmas together give Proposition 10. It remains to prove the lemmas.

Proof of Lemma 15. The bound on $I(y)$ is easy by applying Lemma 7 since $|m''(s)| \leq \frac{c}{(1+s)^2}$ for all s and some constant c . We now turn to $I_t(y) - I(y)$.

Recall the expression (105) of $I_t(y)$, replace $m''(s)$ by its expression (106) and cut the integral into three pieces to obtain

$$I_t(y) = \frac{1}{2} \int_0^t ds m''(s) \frac{t-s}{t} \left(\xi_{\frac{st}{t-s}}^{(y)} - y \right), \quad (188)$$

$$= \frac{3}{4} \frac{t+1}{t} \int_0^t \frac{ds}{(s+1)^2} \left(\xi_{\frac{st}{t-s}}^{(y)} - y \right) - \frac{3}{4t} \int_0^t \frac{ds}{s+1} \left(\xi_{\frac{st}{t-s}}^{(y)} - y \right) + \frac{1}{2} \int_0^t ds r''(s) \frac{t-s}{t} \left(\xi_{\frac{st}{t-s}}^{(y)} - y \right). \quad (189)$$

Recall that, by scaling,

$$\xi_{tu}^{(y)} = \sqrt{t} \tilde{\xi}_u^{(\tilde{y})} \quad \text{with} \quad \tilde{y} = y/\sqrt{t} \quad (190)$$

where $\tilde{\xi}_u^{(\tilde{y})}$ is another, t dependent (implicit in notation), Bessel process started from \tilde{y} . We can apply Lemma 7 to the Bessel process $\tilde{\xi}_u^{(\tilde{y})}$ but, as it depends on t , the random variable G must be replaced by some other random variable \tilde{G}_t which has the same Gaussian tails as G . Then

$$|\tilde{\xi}_u^{(\tilde{y})} - \tilde{y}| \leq \tilde{G}_t \max\left(u^{\frac{1}{2}-\epsilon}, u^{\frac{1}{2}+\epsilon}\right) \quad \text{so} \quad |\xi_{tu}^{(y)} - y| \leq \tilde{G}_t \sqrt{t} \max\left(u^{\frac{1}{2}-\epsilon}, u^{\frac{1}{2}+\epsilon}\right). \quad (191)$$

In the second integral of (189), make the change of variable $u = s/t$ and use (191) to obtain

$$\left| \frac{1}{t} \int_0^t \frac{ds}{s+1} \left[\xi_{\frac{st}{t-s}}^{(y)} - y \right] \right| \leq \frac{1}{t} \int_0^1 \frac{du}{u} \tilde{G}_t \sqrt{t} \max\left\{ \left(\frac{u}{1-u}\right)^{\frac{1}{2}+\epsilon}, \left(\frac{u}{1-u}\right)^{\frac{1}{2}-\epsilon} \right\} = \tilde{G}_t \mathcal{O}(t^{-\frac{1}{2}}). \quad (192)$$

In the first integral of (189), make the change of variable $u = st/(t-s)$ to obtain

$$I_t(y) = \frac{3}{4} \frac{t+1}{t} \int_0^\infty \frac{du}{(u+1+u/t)^2} (\xi_u^{(y)} - y) + \frac{1}{2} \int_0^t ds r''(s) \frac{t-s}{t} \left(\xi_{\frac{st}{t-s}}^{(y)} - y \right) + \tilde{G}_t \mathcal{O}(t^{-\frac{1}{2}}). \quad (193)$$

We now turn to $I(y)$. In expression (105) of $I(y)$, use the expression (106) and cut the integral into the following pieces:

$$I(y) = \frac{3}{4} \int_0^\infty \frac{ds}{(s+1)^2} (\xi_s^{(y)} - y) + \frac{1}{2} \int_0^t ds r''(s) \frac{t-s}{t} (\xi_s^{(y)} - y) + \frac{1}{2} \int_0^t ds r''(s) \frac{s}{t} (\xi_s^{(y)} - y) + \frac{1}{2} \int_t^\infty ds r''(s) (\xi_s^{(y)} - y). \quad (194)$$

Applying Lemma 7 and the fact that $r''(s)$ is bounded (since it is continuous on $[0, \infty)$ and tends to 0) with $r''(s) = \mathcal{O}(s^{-2-\eta})$ for some $\eta > 0$, it is easy to check that the third and fourth integrals are bounded in modulus by $G\mathcal{O}(t^{-1/2})$ if $\epsilon < \eta$. Using Lemma 7 again, it is also easy to check that the first terms in (193) and (194) are equal up to an error of size $G\mathcal{O}(1/t)$ which we absorb in the $G\mathcal{O}(t^{-1/2})$ that we already have. Thus we get

$$I_t(y) - I(y) = \frac{1}{2} \int_0^t ds r''(s) \frac{t-s}{t} \left(\xi_{\frac{st}{t-s}}^{(y)} - \xi_s^{(y)} \right) + \tilde{G}_t \mathcal{O}(t^{-\frac{1}{2}}) + G\mathcal{O}(t^{-\frac{1}{2}}). \quad (195)$$

We now focus on the remaining integral. The difference $\xi_{st/(t-s)}^{(y)} - \xi_s^{(y)}$ is the position at time $s^2/(t-s) = st/(t-s) - s$ of a new Bessel process started from $\xi_s^{(y)}$. It is also, by scaling, equal

to $t^{-1/2}$ times the position at time $ts^2/(t-s)$ of another Bessel process started from $\sqrt{t}\xi_s^{(y)}$. Applying Lemma 7 again to this last Bessel process, we get

$$\left| \xi_{\frac{st}{t-s}}^{(y)} - \xi_s^{(y)} \right| \leq \frac{\hat{G}_t}{\sqrt{t}} \max \left\{ \left(\frac{ts^2}{t-s} \right)^{\frac{1}{2}+\epsilon}, \left(\frac{ts^2}{t-s} \right)^{\frac{1}{2}-\epsilon} \right\} \leq \frac{\hat{G}_t}{\sqrt{t}} \times \begin{cases} \frac{t}{t-s} s^{1+2\epsilon} & \text{if } 1 < s < t, \\ \frac{t}{t-1} & \text{if } 0 < s < 1. \end{cases} \quad (196)$$

where \hat{G}_t is another t -dependent positive random variable with the same Gaussian tail as G . Since $r''(s)$ is bounded and $r''(s) = \mathcal{O}(s^{-2-\eta})$, the integral $\int_1^\infty ds r''(s) s^{1+2\epsilon}$ is finite provided $\epsilon < \eta/2$, and we obtain

$$I_t(y) - I(y) = G_t \mathcal{O}(t^{-\frac{1}{2}}), \quad (197)$$

with $G_t = \max(G, \tilde{G}_t, \hat{G}_t)$. This concludes the proof. \square

Proof of Lemma 16. With the hypothesis of the lemma, write $|A_t| \leq \alpha G$ and $|B_t| \leq \beta \epsilon_t G_t$ for some $\alpha > 0$ and $\beta > 0$. Writing

$$e^{A_t+B_t} = \sum_{p=0}^{\infty} \frac{1}{p!} e^{A_t} B_t^p, \quad (198)$$

we can apply dominated convergence—since the partial sums are dominated by $\exp(A_t + |B_t|)$ which has finite expectation—and obtain

$$\mathbb{E}[e^{A_t+B_t}] = \sum_{p=0}^{\infty} \frac{1}{p!} \mathbb{E}[e^{A_t} B_t^p]. \quad (199)$$

It only remains to show that the sum for $p \geq n+1$ is $\mathcal{O}(\epsilon_t^{n+1})$. To do this observe that

$$\left| \frac{1}{p!} \mathbb{E}[e^{A_t} B_t^p] \right| \leq \frac{\epsilon_t^p}{p!} \mathbb{E}[e^{\alpha G} (\beta G_t)^p] \leq \epsilon_t^p \mathbb{E}[e^{\alpha G + \beta G_t}], \quad (200)$$

where the last expectation is finite. Then, as soon as $\epsilon_t < 1$, we have

$$\left| \sum_{p=n+1}^{\infty} \frac{1}{p!} \mathbb{E}[e^{A_t} B_t^p] \right| \leq \frac{\epsilon_t^{n+1}}{1-\epsilon_t} \mathbb{E}[e^{\alpha G + \beta G_t}], \quad (201)$$

which concludes the proof. \square

Proof of Lemma 17. Define

$$J_t(y) = 2\mathbb{E}[e^{I(y)}(I(y) - I_t(y))] - 2\mathbb{E}[e^{I(y)}]\mathbb{E}[I(y) - I_t(y)]. \quad (202)$$

We want to show that $J_t(y) = \mathcal{O}(\frac{\log t}{t}) + y\mathcal{O}(\frac{1}{t})$. Clearly,

$$J_t(y) = 2\left(\mathbb{E}[e^{I(y)}I(y)] - \mathbb{E}[e^{I(y)}]\mathbb{E}[I(y)]\right) - 2\left(\mathbb{E}[e^{I(y)}I_t(y)] - \mathbb{E}[e^{I(y)}]\mathbb{E}[I_t(y)]\right) \quad (203)$$

$$\begin{aligned} &= \int_0^\infty ds m''(s) \left(\mathbb{E}[e^{I(y)}(\xi_s^{(y)} - y)] - \mathbb{E}[e^{I(y)}]\mathbb{E}[\xi_s^{(y)} - y] \right) \\ &\quad - \int_0^t ds m''(s) \frac{t-s}{t} \left(\mathbb{E}[e^{I(y)}(\xi_{\frac{ts}{t-s}}^{(y)} - y)] - \mathbb{E}[e^{I(y)}]\mathbb{E}[\xi_{\frac{ts}{t-s}}^{(y)} - y] \right) \end{aligned} \quad (204)$$

$$= \int_0^\infty ds m''(s) w(y, s) - \int_0^t ds m''(s) \frac{t-s}{t} w\left(y, \frac{ts}{t-s}\right), \quad (205)$$

where we recall the definition of w from (124). We now apply Proposition 12. Cut the integrals at $t/2$ and rearrange the terms:

$$J_t(y) = \int_{\frac{t}{2}}^{\infty} ds m''(s)w(y, s) + \int_0^{\frac{t}{2}} ds m''(s)\frac{s}{t}w(y, s) - \int_{\frac{t}{2}}^t ds m''(s)\frac{t-s}{t}w\left(y, \frac{ts}{t-s}\right) - \int_0^{\frac{t}{2}} ds m''(s)\frac{t-s}{t}\left(w\left(y, \frac{ts}{t-s}\right) - w(y, s)\right). \quad (206)$$

Using from Proposition 12 that $|w(y, s)| \leq C \log(s+1)$ and of course $m''(s) = \mathcal{O}(1/s^2)$, the first and third integrals are both $\mathcal{O}(\frac{\log t}{t})$, uniformly in y . Now using from Proposition 12 that $|w(y, s)| \leq C(1 + y\frac{\log(s+1)}{\sqrt{s}})$, the second integral is $y\mathcal{O}(\frac{1}{t}) + \mathcal{O}(\frac{\log t}{t})$.

We now turn to the fourth integral. Writing $\frac{st}{t-s} = s + \frac{s^2}{t-s}$ and noticing that for $s < t/2$ we have $\frac{s^2}{t-s} < s^2$ as soon as $t \geq 2$, the last part of Proposition 12 gives $|w(y, \frac{ts}{t-s}) - w(y, s)| \leq C\frac{s}{t-s}$, and therefore the fourth integral is $\mathcal{O}(\frac{\log t}{t})$, which concludes the proof. \square

6.3 Proof of Proposition 11

For $y \geq 0$ we introduce the notation $\mu(y, t) := \mathbb{E}[\xi_t^{(y)}] - y$ and observe that

$$\mu(y, tu) = \sqrt{t}\mu\left(\frac{y}{\sqrt{t}}, u\right), \quad \mu(0, s) = \frac{4}{\sqrt{\pi}}\sqrt{s}, \quad \max[0, \mu(0, s) - y] \leq \mu(y, s) \leq \mu(0, s). \quad (207)$$

(The first equality is the scaling property, and the inequalities are from (34). The second equality can be calculated directly from the probability density function for a Bessel process; see for example [RY99, page 446].)

With this notation we can rewrite

$$\mathbb{E}[I(y)] = \frac{1}{2} \int_0^{\infty} ds m''(s)\mu(y, s) \quad (208)$$

$$\mathbb{E}[I_t(y)] = \frac{1}{2} \int_0^t ds m''(s)\frac{t-s}{t}\mu\left(y, \frac{st}{t-s}\right). \quad (209)$$

As usual we use the expression (106), decomposing $\mathbb{E}[I(y) - I_t(y)]$ into terms containing $3/2(s+1)^2$ and terms containing $r''(s)$. In the former we make our usual change of time $u = st/(t-s)$, but in the latter we do not.

$$\begin{aligned} \mathbb{E}[I(y) - I_t(y)] &= \frac{3}{4} \int_0^{\infty} ds \frac{1}{(s+1)^2} \mu(y, s) - \frac{3}{4} \int_0^{\infty} du \frac{1}{\left(\frac{tu}{t+u} + 1\right)^2} \left(\frac{t}{t+u}\right)^3 \mu(y, u) \\ &\quad + \frac{1}{2} \int_0^{\infty} ds r''(s)\mu(y, s) - \frac{1}{2} \int_0^t ds r''(s)\frac{t-s}{t}\mu\left(y, \frac{st}{t-s}\right). \end{aligned} \quad (210)$$

Rearranging we get

$$\begin{aligned} \mathbb{E}[I(y) - I_t(y)] &= \frac{3}{4} \int_0^{\infty} ds \left(1 - \frac{t}{t+s}\right) \frac{1}{(s+1)^2} \mu(y, s) \\ &\quad + \frac{3}{4} \int_0^{\infty} ds \left(\frac{1}{(s+1)^2} - \frac{1}{(s+1+s/t)^2}\right) \frac{t}{t+s} \mu(y, s) \\ &\quad + \frac{1}{2} \int_0^t ds r''(s) \left(\mu(y, s) - \frac{t-s}{t}\mu\left(y, \frac{st}{t-s}\right)\right) + \frac{1}{2} \int_t^{\infty} ds r''(s)\mu(y, s), \end{aligned} \quad (211)$$

and we treat each of the four integrals on the right-hand side in turn.

The first integral in the right hand side of (211)

Making the change of variable $s = tu$ and using the first part of (207) we have

$$\int_0^\infty ds \left(1 - \frac{t}{t+s}\right) \frac{1}{(s+1)^2} \mu(y, s) = \frac{1}{\sqrt{t}} \int_0^\infty du \frac{u}{(u+1)(u+1/t)^2} \mu\left(\frac{y}{\sqrt{t}}, u\right). \quad (212)$$

We now approximate $\mu(y/\sqrt{t}, u)$ by $\mu(0, u)$, bounding the error by using the last part of (207):

$$\left| \frac{1}{\sqrt{t}} \int_0^\infty du \frac{u}{(u+1)(u+1/t)^2} \mu\left(\frac{y}{\sqrt{t}}, u\right) - \frac{1}{\sqrt{t}} \int_0^\infty du \frac{u}{(u+1)(u+1/t)^2} \mu(0, u) \right| \leq \frac{1}{\sqrt{t}} \int_0^\infty du \frac{u}{(u+1)(u+1/t)^2} \frac{y}{\sqrt{t}}. \quad (213)$$

The right hand side is $y\mathcal{O}\left(\frac{\log t}{t}\right)$, and using the second part of (207), we have

$$\frac{1}{\sqrt{t}} \int_0^\infty du \frac{u}{(u+1)(u+1/t)^2} \mu(0, u) = \frac{4\sqrt{\pi}}{\sqrt{t}} + \mathcal{O}(1/t). \quad (214)$$

We therefore conclude that

$$\int_0^\infty ds \left(1 - \frac{t}{t+s}\right) \frac{1}{(s+1)^2} \mu(y, s) = \frac{4\sqrt{\pi}}{\sqrt{t}} + y\mathcal{O}\left(\frac{\log t}{t}\right) + \mathcal{O}\left(\frac{1}{t}\right). \quad (215)$$

The second integral in the right hand side of (211)

We note that

$$\frac{1}{(s+1)^2} - \frac{1}{(s+1+s/t)^2} = \frac{1}{(s+1)^2} \mathcal{O}\left(\frac{1}{t}\right), \quad (216)$$

and $t/(t+s) \leq 1$, so using the bound $\mu(y, s) \leq \mu(y, 0) = 4\sqrt{s}/\sqrt{\pi}$ from (207), we easily see that the second integral is $\mathcal{O}(1/t)$ uniformly in y .

The third integral in the right hand side of (211)

We use the following result: for any $\delta > 0$,

$$0 \leq \mu(y, s+\delta) - \mu(y, s) \leq \mu(0, s+\delta) - \mu(0, s) = \frac{4}{\sqrt{\pi}} (\sqrt{s+\delta} - \sqrt{s}). \quad (217)$$

This follows from the Markov property plus (207). Then

$$\frac{t-s}{t} \mu\left(y, \frac{st}{t-s}\right) - \mu(y, s) = \frac{t-s}{t} \left[\mu\left(y, \frac{st}{t-s}\right) - \mu(y, s) \right] - \frac{s}{t} \mu(y, s), \quad (218)$$

so that

$$-\frac{4}{\sqrt{\pi}} \frac{s^{3/2}}{t} \leq \frac{t-s}{t} \mu\left(y, \frac{st}{t-s}\right) - \mu(y, s) \leq \frac{4}{\sqrt{\pi}} \frac{t-s}{t} \left[\left(\frac{st}{t-s}\right)^{1/2} - s^{1/2} \right] \quad (219)$$

But $\left(\frac{st}{t-s}\right)^{1/2} = s^{1/2} \left(1 + \frac{s}{t-s}\right)^{1/2} \leq s^{1/2} \left(1 + \frac{s}{2(t-s)}\right)$ so the right hand side of the previous equation is at most $(4/\sqrt{\pi}) \times s^{3/2}/(2t)$. We conclude that

$$\left| \int_0^t ds r''(s) \left[\mu(y, s) - \frac{t-s}{t} \mu\left(y, \frac{st}{t-s}\right) \right] \right| \leq \frac{4}{\sqrt{\pi}t} \int_0^t ds s^{3/2} |r''(s)| = \begin{cases} \mathcal{O}\left(\frac{1}{t}\right) & \text{if } \eta > \frac{1}{2}, \\ \mathcal{O}\left(\frac{\log t}{t}\right) & \text{if } \eta = \frac{1}{2}, \\ \mathcal{O}\left(\frac{1}{t^{2+\eta}}\right) & \text{if } \eta < \frac{1}{2}, \end{cases} \quad (220)$$

uniformly in y .

The fourth integral in the right hand side of (211)

Since $r''(s) = O(s^{-2-\eta})$ for some $\eta > 0$, using (207) again we have

$$\left| \int_t^\infty ds r''(s) \mu(y, s) \right| \leq \frac{4}{\sqrt{\pi}} \int_t^\infty ds |r''(s)| \sqrt{s} = \mathcal{O}\left(\frac{1}{t^{\frac{1}{2}+\eta}}\right) \quad (221)$$

uniformly in y .

Putting together the results from the four integrals give the proposition.

Appendix

Lemma 18. *For any $\epsilon > 0$, the non-negative random variable*

$$G := \sup_{s>0} \frac{\xi_s^{(0)}}{\max(s^{1/2-\epsilon}, s^{1/2+\epsilon})} \quad (222)$$

has Gaussian tail under \mathbb{P} .

Proof. We do this in two parts, first considering the supremum over $s \in (0, 1]$. We have

$$\mathbb{P}\left(\sup_{s \in (0,1]} \frac{\xi_s^{(0)}}{s^{1/2-\epsilon}} > z\right) \leq \sum_{n=2}^{\infty} \mathbb{P}\left(\sup_{s \in (\frac{1}{n}, \frac{1}{n-1}]} \frac{\xi_s^{(0)}}{s^{1/2-\epsilon}} > z\right). \quad (223)$$

By scaling, this equals

$$\sum_{n=2}^{\infty} \mathbb{P}\left(\sup_{s \in (1, \frac{n}{n-1}]} \frac{\xi_s^{(0)}}{s^{1/2-\epsilon}} > zn^\epsilon\right) \leq \sum_{n=2}^{\infty} \mathbb{P}\left(\sup_{s \in (1,2]} \xi_s^{(0)} > zn^\epsilon\right). \quad (224)$$

Now note that there exist $c_3 > 0$ and $c_4 > 0$ such that $\mathbb{P}(\sup_{s \in (1,2]} \xi_s^{(0)} > z) \leq c_3 \exp[-c_4 z^2]$ for all $z > 0$, so

$$\mathbb{P}\left(\sup_{s \in (0,1]} \frac{\xi_s^{(0)}}{s^{1/2-\epsilon}} > z\right) \leq c_3 \sum_{n=2}^{\infty} e^{-c_4 z^2 n^{2\epsilon}}, \quad (225)$$

and it is an easy exercise to show that there exist c_1 and c_2 (with c_1 depending on ϵ) such that $c_3 \sum_{n=2}^{\infty} e^{-c_4 z^2 n^{2\epsilon}} \leq c_1 e^{-c_2 z^2}$.

Similarly for $s \in (1, \infty)$,

$$\mathbb{P}\left(\sup_{s \in (1,\infty)} \frac{\xi_s^{(0)}}{s^{1/2+\epsilon}} > z\right) \leq \sum_{n=1}^{\infty} \mathbb{P}\left(\sup_{s \in (n, n+1]} \frac{\xi_s^{(0)}}{s^{1/2+\epsilon}} > z\right). \quad (226)$$

By scaling, this equals

$$\sum_{n=1}^{\infty} \mathbb{P}\left(\sup_{s \in (1, \frac{n+1}{n}]} \frac{\xi_s^{(0)}}{s^{1/2+\epsilon}} > zn^\epsilon\right) \leq \sum_{n=1}^{\infty} \mathbb{P}\left(\sup_{s \in (1,2]} \xi_s^{(0)} > zn^\epsilon\right). \quad (227)$$

and the end of the argument is the same as in the previous case. \square

References

- [AW78] Donald G Aronson and Hans F Weinberger, *Multidimensional nonlinear diffusion arising in population genetics*, *Advances in Mathematics* **30** (1978), no. 1, 33–76.
- [BD15] Éric Brunet and Bernard Derrida, *An exactly solvable travelling wave equation in the Fisher–KPP class*, *Journal of Statistical Physics* (2015), 1–20.
- [Bra83] Maury Bramson, *Convergence of solutions of the Kolmogorov equation to travelling waves*, *Mem. Amer. Math. Soc.* **44** (1983), no. 285, iv+190.
- [EvS00] Ute Ebert and Wim van Saarloos, *Front propagation into unstable states: universal algebraic convergence towards uniformly translating pulled fronts*, *Physica D: Nonlinear Phenomena* **146** (2000), no. 1, 1–99.
- [Fis37] R. A. Fisher, *The advance of advantageous genes*, *Ann. Eugenics* **7** (1937), 355–369.
- [HE15] Samuel Herrmann and Tanré; Etienne, *The first-passage time of the brownian motion to a curved boundary: an algorithmic approach*, arXiv preprint arXiv:1501.07060 (2015).
- [Hen14] Christopher Henderson, *Population stabilization in branching Brownian motion with absorption*, arXiv preprint arXiv:1409.4836 (2014).
- [KPP37] A. N. Kolmogorov, I. Petrovski, and N. Piscounov, *Étude de l'équation de la diffusion avec croissance de la quantité de matière et son application à un problème biologique*, *Mosc. Univ. Bull. Math.* **1** (1937), 1–25, Translated and reprinted in Pelce, P., *Dynamics of Curved Fronts* (Academic, San Diego, 1988).
- [Kun97] Hiroshi Kunita, *Stochastic flows and stochastic differential equations*, vol. 24, Cambridge university press, 1997.
- [MM14] AH Mueller and S Munier, *Phenomenological picture of fluctuations in branching random walks*, *Physical Review E* **90** (2014), no. 4, 042143.
- [RY99] Daniel Revuz and Marc Yor, *Continuous martingales and brownian motion*, vol. 293, Springer Science & Business Media, 1999.

Bibliography

- [Aïd13] Elie Aïdékon. “Convergence in law of the minimum of a branching random walk”. In: *The Annals of Probability* **41.3A** (May 2013), pp. 1362–1426. DOI: 10.1214/12-aop750 (cited on pages 11, 20, 33).
- [ABBS12] Elie Aïdékon, Julien Berestycki, Éric Brunet, and Zhan Shi. “Branching Brownian motion seen from its tip”. In: *Probability Theory and Related Fields* (Nov. 2012), pp. 405–451. DOI: 10.1007/s00440-012-0461-0 (cited on pages 25, 31).
- [ABK11] Louis-Pierre Arguin, Anton Bovier, and Nicola Kistler. “Genealogy of extremal particles of branching Brownian motion”. In: *Communications on Pure and Applied Mathematics* **64.12** (Dec. 2011), pp. 1647–1676. DOI: 10.1002/cpa.20387 (cited on pages 25, 31).
- [ABK12] Louis-Pierre Arguin, Anton Bovier, and Nicola Kistler. “Poissonian statistics in the extremal process of branching Brownian motion”. In: *The Annals of Applied Probability* **22.4** (Aug. 2012), pp. 1693–1711. DOI: 10.1214/11-aap809 (cited on pages 25, 31).
- [ABK13a] Louis-Pierre Arguin, Anton Bovier, and Nicola Kistler. “An ergodic theorem for the frontier of branching Brownian motion”. In: *Electronic Journal of Probability* **18**, 53 (2013). DOI: 10.1214/ejp.v18-2082 (cited on page 28).
- [ABK13b] Louis-Pierre Arguin, Anton Bovier, and Nicola Kistler. “The extremal process of branching Brownian motion”. In: *Probability Theory and Related Fields* **157.3** (Dec. 2013), pp. 535–574. DOI: 10.1007/s00440-012-0464-x (cited on pages 25, 31).
- [ABK15] Louis-Pierre Arguin, Anton Bovier, and Nicola Kistler. “An ergodic theorem for the extremal process of branching Brownian motion”. In: *Annales de l’Institut Henri Poincaré, Probabilités et Statistiques* **51.2** (May 2015), pp. 557–569. DOI: 10.1214/14-aihp608 (cited on page 29).
- [BDL08] Rafael D. Benguria, M. Cristina Depassier, and Michael Loss. “Upper and lower bounds for the speed of pulled fronts with a cut-off”. In: *The European Physical Journal B* **61.3** (Feb. 2008), pp. 331–334. DOI: 10.1140/epjb/e2008-00069-1 (cited on page 35).
- [BG10] Jean Bérard and Jean-Baptiste Gouéré. “Brunet-Derrida behavior of branching-selection particle systems on the line”. In: *Communications in Mathematical Physics* **298.2** (Sept. 2010), pp. 323–342. DOI: 10.1007/s00220-010-1067-y (cited on pages 16, 36).
- [BG11] Jean Bérard and Jean-Baptiste Gouéré. “Survival probability of the branching random walk killed below a linear boundary”. In: *Electronic Journal of Probability* **16**, 14 (2011), pp. 396–418. DOI: 10.1214/ejp.v16-861 (cited on page 41).
- [BN12] Henri Berestycki and Grégoire Nadin. “Spreading speeds for one-dimensional monostable reaction-diffusion equations”. In: *Journal of Mathematical Physics* **53**, 115619 (Nov. 2012). DOI: 10.1063/1.4764932 (cited on page 17).
- [BBS11] Julien Berestycki, Nathanaël Berestycki, and Jason Schweinsberg. “Survival of near-critical branching Brownian motion”. In: *Journal of Statistical Physics* **143.5** (May 2011), pp. 833–854. DOI: 10.1007/s10955-011-0224-9 (cited on page 41).
- [BBS13] Julien Berestycki, Nathanaël Berestycki, and Jason Schweinsberg. “The genealogy of branching Brownian motion with absorption”. In: *The Annals of Probability* **41.2** (Mar. 2013), pp. 527–618. DOI: 10.1214/11-aop728 (cited on pages 42, 51).
- [BB16] Julien Berestycki and Éric Brunet. “A note of the convergence of the Fisher-KPP front centred around its α -level” (Mar. 2016). URL: <http://arxiv.org/abs/1603.06005> (cited on page 24).
- [BBHHR15] Julien Berestycki, Éric Brunet, John W. Harris, Simon C. Harris, and Matthew I. Roberts. “Growth rates of the population in a branching Brownian motion with an inhomogeneous breeding potential”. In: *Stochastic Processes and their Applications* **125.5** (May 2015), pp. 2096–2145. DOI: 10.1016/j.spa.2014.12.008 (cited on page 1).

- [BBHR15] Julien Berestycki, Éric Brunet, Simon C. Harris, and Matthew I. Roberts. “Vanishing corrections for the position in a linear model of FKPP fronts” (Oct. 2015). URL: <http://arxiv.org/abs/1510.03329> (cited on pages 22, 24). Reproduced on page 136.
- [BBS14] Julien Berestycki, Éric Brunet, and Zhan Shi. “Accessibility percolation with backsteps” (Jan. 2014). URL: <http://arxiv.org/abs/1401.6894> (cited on page 1).
- [BBS16] Julien Berestycki, Éric Brunet, and Zhan Shi. “The number of accessible paths in the hypercube”. In: *Bernoulli* **22.2** (May 2016), pp. 653–680. DOI: 10.3150/14-bej641 (cited on page 1).
- [Ber09] Nathanaël Berestycki. “Recent progress in coalescent theory”. In: *Ensaaios Matemáticos* **16** (2009), pp. 1–193. URL: <http://ensaios.sbm.org.br/contents/> (cited on page 46).
- [BP88] Andrzej Bialas and Robi Peschanski. “Random cascading models and the link between long-range and short-range interactions in multiparticle production”. In: *Physics Letters B* **207.1** (June 1988), pp. 59–63. DOI: 10.1016/0370-2693(88)90887-8 (cited on page 18).
- [BS98] Erwin Bolthausen and Alain-Sol Sznitman. “On Ruelle’s probability cascades and an abstract cavity method”. In: *Communications in Mathematical Physics* **197.2** (Oct. 1998), pp. 247–276. DOI: 10.1007/s002200050450 (cited on page 46).
- [Bra78] Maury D. Bramson. “Maximal displacement of branching Brownian motion”. In: *Communications on Pure and Applied Mathematics* **31.5** (Sept. 1978), pp. 531–581. DOI: 10.1002/cpa.3160310502 (cited on pages 8, 20).
- [Bra83] Maury D. Bramson. “Convergence of solutions of the Kolmogorov equation to travelling waves”. In: *Memoirs of the American Mathematical Society* **44.285** (July 1983). DOI: 10.1090/memo/0285 (cited on pages 8, 9, 12, 19, 20).
- [BDZ14] Maury D. Bramson, Jian Ding, and Ofer Zeitouni. “Convergence in law of the maximum of nonlattice branching random walk” (Apr. 2014). URL: <http://arxiv.org/abs/1404.3423> (cited on pages 11, 20).
- [BD97] Éric Brunet and Bernard Derrida. “Shift in the velocity of a front due to a cutoff”. In: *Physical Review E* **56.3** (Sept. 1997), pp. 2597–2604. DOI: 10.1103/PhysRevE.56.2597 (cited on pages 20, 34–36).
- [BD01] Éric Brunet and Bernard Derrida. “Effect of microscopic noise on front propagation”. In: *Journal of Statistical Physics* **103.1** (Apr. 2001), pp. 269–282. DOI: 10.1023/a:1004875804376 (cited on page 38).
- [BD04] Éric Brunet and Bernard Derrida. “Exactly soluble noisy traveling-wave equation appearing in the problem of directed polymers in a random medium”. In: *Physical Review E* **70**, 016106 (July 2004). DOI: 10.1103/PhysRevE.70.016106 (cited on page 43).
- [BD09] Éric Brunet and Bernard Derrida. “Statistics at the tip of a branching random walk and the delay of traveling waves”. In: *Europhysics Letters* **87**, 60010 (Sept. 2009). DOI: 10.1209/0295-5075/87/60010 (cited on pages 25, 27). Reproduced on page 84.
- [BD11] Éric Brunet and Bernard Derrida. “A branching random walk seen from the tip”. In: *Journal of Statistical Physics* **143.3** (May 2011), pp. 420–446. DOI: 10.1007/s10955-011-0185-z (cited on pages 25, 27, 29, 30). Reproduced on page 89.
- [BD13] Éric Brunet and Bernard Derrida. “Genealogies in simple models of evolution”. In: *Journal of Statistical Mechanics: Theory and Experiment* **2013**, P01006 (Jan. 2013). DOI: 10.1088/1742-5468/2013/01/p01006 (cited on page 46).
- [BD15] Éric Brunet and Bernard Derrida. “An exactly solvable travelling wave equation in the Fisher-KPP class”. In: *Journal of Statistical Physics* **161.4** (Nov. 2015), pp. 801–820. DOI: 10.1007/s10955-015-1350-6 (cited on pages 23, 24). Reproduced on page 116.
- [BDMM06a] Éric Brunet, Bernard Derrida, Alfred H. Mueller, and Stéphane Munier. “Phenomenological theory giving the full statistics of the position of fluctuating pulled fronts”. In: *Physical Review E* **73**, 056126 (May 2006). DOI: 10.1103/PhysRevE.73.056126 (cited on pages 36, 39, 40). Reproduced on page 55.
- [BDMM06b] Éric Brunet, Bernard Derrida, Alfred H. Mueller, and Stéphane Munier. “Noisy traveling waves: effect of selection on genealogies”. In: *Europhysics Letters* **76.1** (Oct. 2006), pp. 1–7. DOI: 10.1209/ep1/i2006-10224-4 (cited on pages 42, 46).
- [BDMM07] Éric Brunet, Bernard Derrida, Alfred H. Mueller, and Stéphane Munier. “Effect of selection on ancestry: an exactly soluble case and its phenomenological generalization”. In: *Physical Review E* **76**, 041104 (4 Oct. 2007). DOI: 10.1103/PhysRevE.76.041104 (cited on pages 17, 42, 46). Reproduced on page 64.

- [BRW08] Éric Brunet, Igor M. Rouzine, and Claus O. Wilke. “The stochastic edge in adaptive evolution”. In: *Genetics* **179.1** (May 2008), pp. 603–620. DOI: 10.1534/genetics.107.079319 (cited on pages 1, 52).
- [CD05] Joseph G. Conlon and Charles R. Doering. “On travelling waves for the stochastic Fisher-Kolmogorov-Petrovsky-Piscunov equation”. In: *Journal of Statistical Physics* **120.3** (Aug. 2005), pp. 421–477. DOI: 10.1007/s10955-005-5960-2 (cited on page 50).
- [Cor16] Aser Cortines. “The genealogy of a solvable population model under selection with dynamics related to directed polymers”. In: *Bernoulli* **22.4** (Nov. 2016), pp. 2209–2236. DOI: 10.3150/15-bej726 (cited on page 46).
- [DMZ08] Youri Davydov, Ilya Molchanov, and Sergei Zuyev. “Strictly stable distributions on convex cones”. In: *Electronic Journal of Probability* **13**, 11 (2008), pp. 259–321. DOI: 10.1214/ejp.v13-487 (cited on page 30).
- [DS07] Bernard Derrida and Damien Simon. “The survival probability of a branching random walk in presence of an absorbing wall”. In: *Europhysics Letters* **78**, 60006 (June 2007). DOI: 10.1209/0295-5075/78/60006 (cited on page 41).
- [DS88] Bernard Derrida and Herbert Spohn. “Polymers on disordered trees, spin glasses, and traveling waves”. In: *Journal of Statistical Physics* **51.5** (June 1988), pp. 817–840. DOI: 10.1007/BF01014886 (cited on pages 13, 14, 18, 31).
- [DF07] Michael M. Desai and Daniel S. Fisher. “Beneficial mutation selection balance and the effect of linkage on positive selection”. In: *Genetics* **176.3** (July 2007), pp. 1759–1798. DOI: 10.1534/genetics.106.067678 (cited on page 52).
- [DWF13] Michael M. Desai, Aleksandra M. Walczak, and Daniel S. Fisher. “Genetic diversity and the structure of genealogies in rapidly adapting populations”. In: *Genetics* **193.2** (Feb. 2013), pp. 565–585. DOI: 10.1534/genetics.112.147157 (cited on page 52).
- [DMS03] Charles R. Doering, Carl Mueller, and Peter Smereka. “Interacting particles, the stochastic Fisher-Kolmogorov-Petrovsky-Piscunov equation, and duality”. In: *Physica A: Statistical Mechanics and its Applications* **325.1-2** (July 2003), pp. 243–259. DOI: 10.1016/s0378-4371(03)00203-6 (cited on pages 17, 18, 50).
- [DPK07] Freddy Dumortier, Nikola Popović, and Tasso J. Kaper. “The critical wave speed for the Fisher-Kolmogorov-Petrovskii-Piscunov equation with cut-off”. In: *Nonlinearity* **20.4** (Mar. 2007), pp. 855–877. DOI: 10.1088/0951-7715/20/4/004 (cited on page 35).
- [DR11] Rick Durrett and Daniel Remenik. “Brunet-Derrida particle systems, free boundary problems and Wiener-Hopf equations”. In: *The Annals of Probability* **39.6** (Nov. 2011), pp. 2043–2078. DOI: 10.1214/10-aop601 (cited on page 16).
- [ES00] Ute Ebert and Wim van Saarloos. “Front propagation into unstable states: universal algebraic convergence towards uniformly translating pulled fronts”. In: *Physica D: Nonlinear Phenomena* **146.1-4** (Nov. 2000), pp. 1–99. DOI: 10.1016/s0167-2789(00)00068-3 (cited on pages 18–20).
- [Fis30] Ronald A. Fisher. *The genetical theory of natural selection*. Oxford: The Clarendon Press, 1930. URL: <https://archive.org/details/geneticaltheoryo031631mbp> (cited on pages 16, 45).
- [Fis37] Ronald A. Fisher. “The wave of advance of advantageous genes”. In: *Annals of Eugenics* **7.4** (June 1937), pp. 355–369. DOI: 10.1111/j.1469-1809.1937.tb02153.x (cited on page 7).
- [GHS11] Nina Gantert, Yueyun Hu, and Zhan Shi. “Asymptotics for the survival probability in a killed branching random walk”. In: *Annales de l’Institut Henri Poincaré, Probabilités et Statistiques* **47.1** (Feb. 2011), pp. 111–129. DOI: 10.1214/10-aihp362 (cited on page 41).
- [GD13] Benjamin H. Good and Michael M. Desai. “Fluctuations in fitness distributions and the effects of weak linked selection on sequence evolution”. In: *Theoretical Population Biology* **85** (May 2013), pp. 86–102. DOI: 10.1016/j.tpb.2013.01.005 (cited on page 52).
- [Gou13] Jean-Baptiste Gouéré. “Le mouvement brownien branchant vu depuis une extrémité”. Séminaire Bourbaki 1067 (Mar. 2013). URL: <http://arxiv.org/abs/1305.4396> (cited on page 25).
- [Hal11] Oskar Hallatschek. “The noisy edge of traveling waves”. In: *Proceedings of the National Academy of Sciences of the USA* **108.5** (Feb. 2011), pp. 1783–1787. DOI: 10.1073/pnas.1013529108 (cited on pages 44, 52).
- [HG15] Oskar Hallatschek and Lukas Geyrhofer. “Collective fluctuations in models of adaptation” (June 2015). URL: <http://arxiv.org/abs/1506.08683> (cited on page 52).

- [HH07] John W. Harris and Simon C. Harris. “Survival probabilities for branching Brownian motion with absorption”. In: *Electronic Communications in Probability* **12**, 10 (2007), pp. 81–92. DOI: 10.1214/ecp.v12-1259 (cited on page 41).
- [Hen14] Christopher Henderson. “Population stabilization in branching Brownian motion with absorption” (Sept. 2014). URL: <http://arxiv.org/abs/1409.4836> (cited on page 22).
- [HS09] Yueyun Hu and Zhan Shi. “Minimal position and critical martingale convergence in branching random walks, and directed polymers on disordered trees”. In: *The Annals of Probability* **37.2** (Mar. 2009), pp. 742–789. DOI: 10.1214/08-aop419 (cited on pages 9, 13, 20).
- [IMM05] Edmond Iancu, Alfred H. Mueller, and Stéphane Munier. “Universal behavior of QCD amplitudes at high energy from general tools of statistical physics”. In: *Physics Letters B* **606.3-4** (Jan. 2005), pp. 342–350. DOI: 10.1016/j.PhysLetB.2004.12.009 (cited on page 18).
- [INW68] Nobuyuki Ikeda, Masao Nagasawa, and Shinzo Watanabe. “Branching Markov processes II”. In: *Journal of Mathematics of Kyoto University* **8.3** (1968), pp. 365–410. URL: <http://projecteuclid.org/euclid.kjm/1250524059> (cited on page 6).
- [Kes78] Harry Kesten. “Branching Brownian motion with absorption”. In: *Stochastic Processes and their Applications* **7.1** (Mar. 1978), pp. 9–47. DOI: 10.1016/0304-4149(78)90035-2 (cited on page 40).
- [Kin82] John F. C. Kingman. “On the genealogy of large populations”. In: *Journal of Applied Probability* **19** (1982), pp. 27–43. DOI: 10.2307/3213548 (cited on page 45).
- [KPP37] Andrei Kolmogorov, Ivan Petrovsky, and N. Piscounov. “Étude de l’équation de la diffusion avec croissance de la quantité de matière et son application à un problème biologique”. In: *Bull. Univ. État Moscou, A* **1.6** (1937), pp. 1–25. DOI: 10.1007/978-94-011-3030-1_38 (cited on page 7).
- [LS87] Steven P. Lalley and Thomas Sellke. “A conditional limit theorem for the frontier of a branching Brownian motion”. In: *The Annals of Probability* **15.3** (July 1987), pp. 1052–1061. DOI: 10.1214/aop/1176992080 (cited on pages 28, 29).
- [Mad15] Thomas Madaule. “Convergence in law for the branching random walk seen from its tip”. In: *Journal of Theoretical Probability* (Sept. 2015). DOI: 10.1007/s10959-015-0636-6 (cited on page 25).
- [Mai13] Pascal Maillard. “A note on stable point processes occurring in branching Brownian motion”. In: *Electronic Communications in Probability* **18**, 5 (2013). DOI: 10.1214/ECP.v18-2390 (cited on page 30).
- [Mai16] Pascal Maillard. “Speed and fluctuations of N -particle branching Brownian motion with spatial selection”. In: *Probability Theory and Related Fields* (Apr. 2016). DOI: 10.1007/s00440-016-0701-9 (cited on page 51).
- [McK75] Henry P. McKean. “Application of Brownian motion to the equation of Kolmogorov-Petrovskii-Piskunov”. In: *Communications on Pure and Applied Mathematics* **28.3** (May 1975), pp. 323–331. DOI: 10.1002/cpa.3160280302 (cited on page 7).
- [Mor58] Patrick A. P. Moran. “Random processes in genetics”. In: *Mathematical Proceedings of the Cambridge Philosophical Society* **54.01** (Jan. 1958), pp. 60–71. DOI: 10.1017/s0305004100033193 (cited on page 45).
- [MM14] Alfred H. Mueller and Stéphane Munier. “Phenomenological picture of fluctuations in branching random walks”. In: *Physical Review E* **90**, 042143 (Oct. 2014). DOI: 10.1103/PhysRevE.90.042143 (cited on pages 20, 24).
- [MT02] Alfred H. Mueller and Dionysis N. Triantafyllopoulos. “The energy dependence of the saturation momentum”. In: *Nuclear Physics B* **640.1-2** (Sept. 2002), pp. 331–350. DOI: 10.1016/s0550-3213(02)00581-3 (cited on page 18).
- [MMQ08] Carl Mueller, Leonid Mytnik, and Jeremy Quastel. “Small noise asymptotics of traveling waves”. In: *Markov Processes and Related Fields* **14.3** (2008), pp. 333–342. URL: <http://math-prf.org/journal/articles/id1157/> (cited on page 50).
- [MMQ11] Carl Mueller, Leonid Mytnik, and Jeremy Quastel. “Effect of noise on front propagation in reaction-diffusion equations of KPP type”. In: *Inventiones mathematicae* **184.2** (May 2011), pp. 405–453. DOI: 10.1007/s00222-010-0292-5 (cited on page 50).
- [MS95] Carl Mueller and Richard B. Sowers. “Random travelling waves for the KPP equation with noise”. In: *Journal of Functional Analysis* **128.2** (Mar. 1995), pp. 439–498. DOI: 10.1006/jfan.1995.1038 (cited on page 50).

- [MT95] Carl Müller and R. Tribe. “Stochastic p.d.e.’s arising from the long range contact and long range voter processes”. In: *Probability Theory and Related Fields* **102.4** (Dec. 1995), pp. 519–545. DOI: 10.1007/bf01198848 (cited on page 17).
- [Mun09] Stéphane Munier. “Quantum chromodynamics at high energy and statistical physics”. In: *Physics Reports* **473.1-4** (Apr. 2009), pp. 1–49. DOI: 10.1016/j.PhysRep.2009.02.001 (cited on page 18).
- [MP03] Stéphane Munier and Robi Peschanski. “Geometric scaling as traveling waves”. In: *Physical Review Letters* **91**, 232001 (Dec. 2003). DOI: 10.1103/PhysRevLett.91.232001 (cited on page 18).
- [MP04a] Stéphane Munier and Robi Peschanski. “Traveling wave fronts and the transition to saturation”. In: *Physical Review D* **69**, 034008 (Feb. 2004). DOI: 10.1103/PhysRevD.69.034008 (cited on page 18).
- [MP04b] Stéphane Munier and Robi Peschanski. “Universality and tree structure of high-energy QCD”. In: *Physical Review D* **70**, 077503 (Oct. 2004). DOI: 10.1103/PhysRevD.70.077503 (cited on page 18).
- [Nad15] Grégoire Nadin. “How does the spreading speed associated with the Fisher-KPP equation depend on random stationary diffusion and reaction terms?” In: *Discrete and Continuous Dynamical Systems - Series B* **20.6** (Aug. 2015), pp. 1785–1803. DOI: 10.3934/dcdsb.2015.20.1785 (cited on page 17).
- [NH13] Richard A. Neher and Oskar Hallatschek. “Genealogies of rapidly adapting populations”. In: *Proceedings of the National Academy of Sciences* **110.2** (Jan. 2013), pp. 437–442. DOI: 10.1073/pnas.1213113110 (cited on page 52).
- [Nev88] Jacques Neveu. “Multiplicative martingales for spatial branching processes”. *Seminar on Stochastic Processes, 1987*. Vol. 15. Birkhäuser Boston, 1988, pp. 223–242. DOI: 10.1007/978-1-4684-0550-7_10 (cited on page 28).
- [Pai16] Michel Pain. “Velocity of the L -branching Brownian motion”. In: *Electronic Journal of Probability* **21**, 28 (2016). DOI: 10.1214/16-ejp4639 (cited on page 51).
- [PSK10] Su-Chan Park, Damien Simon, and Joachim Krug. “The speed of evolution in large asexual populations”. In: *Journal of Statistical Physics* **138.1** (Feb. 2010), pp. 381–410. DOI: 10.1007/s10955-009-9915-x (cited on page 52).
- [PL99] Leonid Pechenik and Herbert Levine. “Interfacial velocity corrections due to multiplicative noise”. In: *Physical Review E* **59.4** (Apr. 1999), pp. 3893–3900. DOI: 10.1103/PhysRevE.59.3893 (cited on page 50).
- [Pit99] Jim Pitman. “Coalescents with multiple collisions”. In: *Annals of Probability* **27.4** (Oct. 1999), pp. 1870–1902. DOI: 10.1214/aop/1022874819 (cited on page 46).
- [Rob13] Matthew I. Roberts. “A simple path to asymptotics for the frontier of a branching Brownian motion”. In: *The Annals of Probability* **41.5** (Sept. 2013), pp. 3518–3541. DOI: 10.1214/12-aop753 (cited on pages 8, 9).
- [RES00] Andrea Rocco, Ute Ebert, and Wim van Saarloos. “Subdiffusive fluctuations of “pulled” fronts with multiplicative noise”. In: *Physical Review E* **62.1** (July 2000), R13–R16. DOI: 10.1103/PhysRevE.62.r13 (cited on page 17).
- [RBW08] Igor M. Rouzine, Éric Brunet, and Claus O. Wilke. “The traveling-wave approach to asexual evolution: Muller’s ratchet and speed of adaptation”. In: *Theoretical Population Biology* **73.1** (Feb. 2008), pp. 24–46. DOI: 10.1016/j.tpb.2007.10.004 (cited on pages 1, 52).
- [Sag99] Serik Sagitov. “The general coalescent with asynchronous mergers of ancestral lines”. In: *Journal of Applied Probability* **36.4** (Dec. 1999), pp. 1116–1125. DOI: 10.1239/jap/1032374759 (cited on page 46).
- [Sch15a] Jason Schweinsberg. “Rigorous results for a population model with selection I: evolution of the fitness distribution” (July 2015). URL: <http://arxiv.org/abs/1507.00393> (cited on page 52).
- [Sch15b] Jason Schweinsberg. “Rigorous results for a population model with selection II: genealogy of the population” (July 2015). URL: <http://arxiv.org/abs/1507.00394> (cited on page 52).
- [SU86] Tokuzo Shiga and Kohei Uchiyama. “Stationary states and their stability of the stepping stone model involving mutation and selection”. In: *Probability Theory and Related Fields* **73.1** (Mar. 1986), pp. 87–117. DOI: 10.1007/bf01845994 (cited on page 18).

- [SD08] Damien Simon and Bernard Derrida. “Quasi-stationary regime of a branching random walk in presence of an absorbing wall”. In: *Journal of Statistical Physics* **131.2** (Apr. 2008), pp. 203–233. DOI: 10.1007/s10955-008-9504-4 (cited on page 41).
- [Wri31] Sewall Wright. “Evolution in Mendelian populations”. In: *Genetics* **16.2** (Mar. 1931), pp. 97–159. URL: <http://www.genetics.org/content/16/2/97> (cited on pages 16, 45).
- [YEC10] Feng Yu, Alison Etheridge, and Charles Cuthbertson. “Asymptotic behavior of the rate of adaptation”. In: *The Annals of Applied Probability* **20.3** (June 2010), pp. 978–1004. DOI: 10.1214/09-aap645 (cited on page 52).

Some aspects of the Fisher-KPP equation and the branching Brownian motion

Abstract

The Fisher-Kolmogorov, Petrovski, Piscounov equation (FKPP) is a deterministic partial differential equation. It describes the evolution of an invasion front from a stable phase into an unstable phase. Branching Brownian motion (BBM) is a stochastic Markov process where particles diffuse and duplicate. Both the FKPP equation and the BBM can be seen as modelling the evolution of a population, but the former is deterministic and with saturation, while the latter is stochastic and without saturation. They are however directly related to each other by McKean's duality.

In this dissertation, after a brief review of classical and essential results concerning the FKPP equation and the BBM, I present some of the contributions my collaborators and I have made to this field.

A first set of results concerns the asymptotic position of the FKPP front; on two well-chosen models in the FKPP class, I present two different ways to recover the classical results of Bramson and the prediction by Ebert and van Saarloos. I also make a prediction for the next order term.

A second set of results concerns the limiting distribution of the rightmost particles in the BBM. As we found out, they are distributed according to a so-called "randomly shifted σ -decorated exponential Poisson point process", which we define and characterize. These results were mostly obtained by using the duality between the BBM and the FKPP equation.

A last set of results concerns the behaviour of noisy FKPP fronts in the limit of a weak noise. I present a phenomenological theory which allows to compute, to leading order, all the cumulants of the position. Furthermore, in models for which it makes sense, the genealogical tree of the population is given by a rescaled Bolthausen-Sznitman coalescent.

Quelques aspects de l'équation Fisher-KPP et du mouvement brownien branchant

Résumé

L'équation de Fisher-Kolmogorov, Petrovski, Piscounov (FKPP) est une équation déterministe aux dérivées partielles. Elle décrit l'évolution d'un front avec une phase stable qui envahit une phase instable. Le mouvement brownien branchant (BBM) est un processus aléatoire de Markov avec des particules qui diffusent et se reproduisent. L'équation FKPP et le BBM peuvent tous deux être vus comme une modélisation de l'évolution d'une population, mais la première est déterministe et avec saturation, alors que le second est aléatoire et sans saturation. Ils sont néanmoins tous les deux reliés par la dualité de McKean.

Dans ce mémoire, après un rappel rapide de résultats classiques et essentiels concernant l'équation FKPP et le BBM, je présente plusieurs des résultats que mes collaborateurs et moi avons obtenus.

Une première série de résultats concerne la position asymptotique d'un front FKPP ; sur deux modèles bien choisis dans la classe FKPP, je présente deux méthodes différentes pour retrouver les résultats classiques de Bramson et la prédiction d'Ebert et de van Saarloos. Je fais également une prédiction pour le terme d'ordre suivant.

Une deuxième série de résultats concerne la distribution limite des particules les plus à droite dans le BBM. Comme nous l'avons mis en évidence, elles sont distribuées selon ce qu'on peut appeler un « processus de Poisson exponentiel σ -décoré et aléatoirement décalé », que nous définissons et caractérisons. Ces résultats ont essentiellement été obtenus en utilisant la dualité entre le BBM et l'équation FKPP.

La dernière série de résultats concerne le comportement des fronts FKPP bruités dans la limite des faibles bruits. Je présente une théorie phénoménologique qui permet de calculer, à l'ordre dominant, tous les cumulants de la position. De plus, dans les modèles pour lesquels la question fait sens, l'arbre généalogique de la population est donné sur la bonne échelle de temps par le coalescent de Bolthausen-Sznitman.