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. Munier[3]

[1]*Laboratoire de Physique Statistique, École Normale Supérieure, 24 rue Lhomond, 75231 Paris cedex 05, France*
[2]*Department of Physics, Columbia University, New York, New York, 10027, USA*
[3]*Centre de Physique Théorique, École Polytechnique, CNRS, 91128 Palaiseau, France*

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.

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 ith 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 jth 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 probabiliy $\psi(\epsilon)d\epsilon$, 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
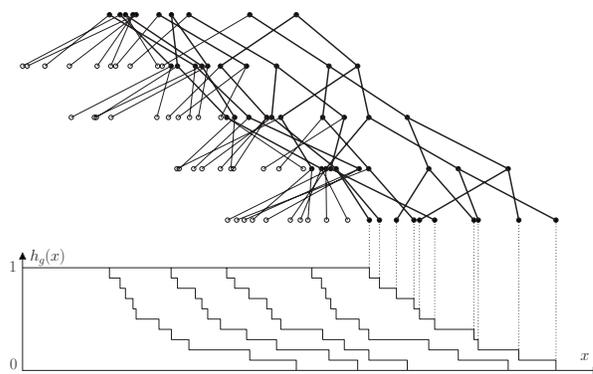

FIG. 1. Numerical simulation of the evolution of model A, with $k=2$ and $\rho(\epsilon)$ uniform between $-\frac{1}{2}$ and $\frac{1}{2}$ for $N=10$. Upper plot: The filiation between each individual and its two offspring is shown. At each generation, the N rightmost survive. Lower plot: The noisy traveling wave front $h_g(x)$, constructed as in (1), is shown for the five generations of the upper plot.
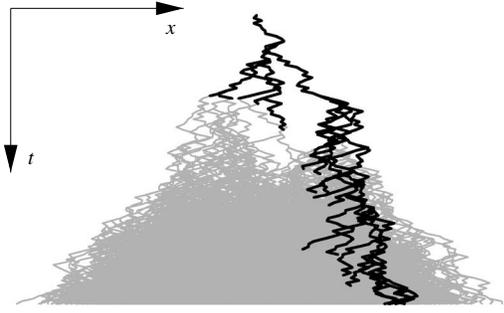
FIG. 2. A branching process for which the size N of the population is limited to five. Each time the number of walks reaches six, the leftmost walk is eliminated. Time goes downwards and the horizontal direction represents space. The actual population is represented in black, while the grey lines represent 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,

$$h_g(x) = \frac{1}{N} \int_x^\infty dz \sum_{i=1}^N \delta[z - x_i(g)]. \tag{1}$$

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

$$h_{g+1}(x) = \min[1, h_{g+1}^*(x)]. \tag{2}$$

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

$$Nh_{g+1}^*(x) = \sum_{i=1}^N n_{g+1}^{(i)}(x), \tag{3}$$

where $n_{g+1}^{(i)}(x)$ is the total number of offspring before selection of the ith 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

$$\overline{n_{g+1}^{(i)}(x)} = k \int_x^\infty d\epsilon \, \rho[\epsilon - x_i(g)], \tag{4}$$

$$\mathrm{Var}[n_{g+1}^{(i)}(x)] = k \int_x^\infty d\epsilon \, \rho[\epsilon - x_i(g)]$$
$$\times \left(1 - \int_x^\infty d\epsilon \, \rho[\epsilon - x_i(g)]\right).$$

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 $n^{(i)}_{g+1}(x)$. For the average, one has

$$N\overline{h^*_{g+1}(x)} = k \int_x^\infty d\epsilon \sum_i \rho[\epsilon - x_i(g)]$$

$$= -k \int_x^\infty d\epsilon \int dz\, \rho(\epsilon - z) Nh'_g(z), \quad (5)$$

where we used, from (1),

$$\sum_{i=1}^N \delta[x - x_i(g)] = -Nh'_g(x). \quad (6)$$

Simplifying, and doing the same transformation for the variance, one finally gets for model A

$$\overline{h^*_{g+1}(x)} = k \int d\epsilon\, h_g(x - \epsilon)\rho(\epsilon), \quad (7a)$$

$$\text{Var}[h^*_{g+1}(x)] = \frac{k}{N} \int d\epsilon\, h_g(x - \epsilon)\rho(\epsilon)$$

$$\times \left(1 - 2\int_\epsilon^\infty dz\, \rho(z)\right). \quad (7b)$$

[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

$$N\overline{h^*_{g+1}(x)} = \text{Var}[Nh^*_{g+1}(x)] = \int_x^\infty d\epsilon\, \Psi(\epsilon). \quad (9)$$

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,

$$h_{g+1}(x) = \min\left[1, k\int d\epsilon\, h_g(x - \epsilon)\rho(\epsilon)\right.$$

$$\left. + \frac{\eta_g(x)}{\sqrt{N}}\sqrt{k\int d\epsilon\, h_g(x - \epsilon)\rho(\epsilon)\left(1 - 2\int_\epsilon^\infty dz\, \rho(z)\right)}\right] \quad (12a)$$

and, for, model B,

$$h_{g+1}(x) = \min\left[1, \int d\epsilon\, h_g(x - \epsilon)\psi(\epsilon)\right.$$

$$\left. + \frac{\eta_g(x)}{\sqrt{N}}\sqrt{\int d\epsilon\, h_g(x - \epsilon)\psi(\epsilon)}\right]. \quad (12b)$$

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 $\overline{h_g(x)}$ 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 $A \to 2A$ (rate 1) and $2A \to A$ (rate $1/N$) or, more simply, is an approximate equation, valid for large N, describing the fraction of A in the chemical reaction $A + B \to 2A$ 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,...) and the $-h_g(x)^2$ term] which makes $h_g(x)=1$ a stable fixed point. So, ignoring the noise terms ($N \to \infty$), 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 $N \to \infty$, 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 $h_g(x) \simeq \exp[-\gamma(x-vg)]$, 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

$$Q_{g+1}(x) = \prod_y [1 - \psi(y)dy + \psi(y)dy Q_g(x-y)]$$

$$= \exp\left(\int dy \, \psi(y)[Q_g(x-y)-1] \right). \quad (15)$$

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 $x_{N+1} < x_N < \cdots < x_1$,

$$\Psi(x_{N+1})dx_{N+1}\Psi(x_N)dx_N \cdots \Psi(x_1)dx_1 \exp\left(-\int_{x_{N+1}}^\infty \Psi(z)dz \right). \quad (17)$$

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 $x_1, \ldots, x_N$, becomes, for $k=1, \ldots, N$,

$$\frac{1}{N!} \Psi(x_{N+1})dx_{N+1}\Psi(x_N)dx_N \cdots \Psi(x_1)dx_1$$

$$\times \exp\left(-\int_{x_{N+1}}^\infty \Psi(z)dz \right) \quad \text{when } x_{N+1} < x_k. \quad (18)$$

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 $x_1, \ldots, x_N$,

$$\frac{1}{N!}\Psi(x_{N+1})dx_{N+1}\left(\int_{x_{N+1}}^{\infty}\Psi(x)dx\right)^N \exp\left(-\int_{x_{N+1}}^{\infty}\Psi(z)dz\right).$$
$$(19)$$

[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 $x_1,\ldots,x_N$ 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 $x_1(g+1),\ldots,x_N(g+1)$ are independent and one gets that, given x_{N+1}, each of the N rightmost particles is in $[x,x+dx]$ with probability

$$\frac{\Psi(x)dx}{\displaystyle\int_{x_{N+1}}^{\infty}\Psi(x)dx}\quad\text{for }x_{N+1}<x. \qquad (20)$$

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 $x_1(g+1),\ldots,x_N(g+1)$ 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

$$\Psi_{\exp}(x)=e^{-(x-X_g)}\quad\text{with }X_g=\ln(e^{x_1(g)}+e^{x_2(g)}+\cdots+e^{x_N(g)}),$$
$$(21)$$

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

$$x_{N+1}=X_g+z\quad\text{with Prob}(z)=\frac{1}{N!}\exp[-(N+1)z-e^{-z}],$$
$$(22)$$

and, once x_{N+1} has been chosen, the distribution (20) of the $x_k(g+1)$ for $k=1,\ldots,N$ becomes

$$x_k(g+1)=x_{N+1}+y_k\quad\text{with Prob}(y_k)=e^{-y_k}\quad\text{for }y_k>0.$$
$$(23)$$

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 $g\to\infty$, 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

$$\Delta X_g=X_{g+1}-X_g=z+\ln(e^{y_1}+e^{y_2}+\cdots+e^{y_N}),\qquad(24)$$

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

$$v_N=\langle\Delta X_g\rangle,\quad D_N=\langle\Delta X_g^2\rangle-\langle\Delta X_g\rangle^2. \qquad (25)$$

More generally, all cumulants of the front position at a long time g are simply g times the cumulants of ΔX_g. To compute theses cumulants, we evaluate the generating function $G(\beta)$ defined as

$$e^{G(\beta)}=\langle e^{-\beta\Delta X_g}\rangle=\int_{-\infty}^{+\infty}dz\,\text{Prob}(z)e^{-\beta z}\int_0^{+\infty}dy_1\,\text{Prob}(y_1)\cdots$$
$$\times\int_0^{+\infty}dy_N\,\text{Prob}(y_N)(e^{y_1}+\cdots+e^{y_N})^{-\beta}, \qquad (26)$$

and one obtains the cumulants by doing a small β expansion,

$$G(\beta)=\sum_{n\geq1}\frac{(-\beta)^n}{n!}\langle\Delta X_g^n\rangle_c. \qquad (27)$$

Using (22), the integral over z is easy,

$$\int_{-\infty}^{+\infty}dz\,\text{Prob}(z)e^{-\beta z}=\frac{1}{N!}\int_{-\infty}^{+\infty}dz\,\exp[-(\beta+N+1)z-e^{-z}]$$
$$=\frac{\Gamma(N+1+\beta)}{\Gamma(N+1)}. \qquad (28)$$

To calculate the integrals over y_i in (26), one can use the representation (valid for $\beta>0$)

$$Z^{-\beta}=\frac{1}{\Gamma(\beta)}\int_0^{+\infty}d\lambda\,\lambda^{\beta-1}e^{-\lambda Z} \qquad (29)$$

with $Z=e^{y_1}+\cdots+e^{y_N}$. This leads to the factorization of the integrals over $y_1,\ldots,y_N$. Replacing $\text{Prob}(y_k)$ by its explicit expression from (23), one gets for $\beta>0$ (a similar calculation can be made for $\beta>-1$),

$$e^{G(\beta)}=\frac{\Gamma(N+1+\beta)}{\Gamma(N+1)\Gamma(\beta)}\int_0^{+\infty}d\lambda\,\lambda^{\beta-1}I_0(\lambda)^N, \qquad (30)$$

where

$$I_0(\lambda)=\int_0^{+\infty}dy\,e^{-y-\lambda e^y}. \qquad (31)$$

One can rewrite $I_0(\lambda)$ in several ways,

$$I_0(\lambda) = \lambda \int_\lambda^{+\infty} \frac{du}{u^2} e^{-u}$$

$$= e^{-\lambda} + \lambda(\ln \lambda + \gamma_E) - \lambda \int_0^\lambda du \frac{1 - e^{-u}}{u}$$

$$= 1 + \lambda(\ln \lambda + \gamma_E - 1) - \sum_{k=0}^{+\infty} \frac{(-1)^k}{(k+1)(k+2)!} \lambda^{k+2},$$

$$(32)$$

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

$$\mu = \lambda N \ln N, \qquad (33)$$

one gets $I_0(\lambda)^N$ for values of μ of order 1,

$$[I_0(\lambda)]^N \simeq \exp[N\lambda(\ln \lambda + \gamma_E - 1)]$$

$$\simeq \exp\left(\frac{\mu}{\ln N}(\ln \mu - \ln N - \ln \ln N + \gamma_E - 1)\right)$$

$$\simeq e^{-\mu}\left[1 + \mu \frac{\ln \mu - \ln \ln N + \gamma_E - 1}{\ln N}\right.$$

$$\left. + \frac{1}{2}\left(\mu \frac{\ln \mu - \ln \ln N + \gamma_E - 1}{\ln N}\right)^2 + \cdots\right], \quad (34)$$

where terms of order $1/N$ have been dropped. Replacing this expression into (30) and using

$$\int_0^\infty d\mu \, \mu^{x-1} e^{-\mu} (\ln \mu)^k = \frac{d^k}{dx^k} \Gamma(x), \qquad (35)$$

one gets

$$e^{G(\beta)} \simeq \frac{\Gamma(N+1+\beta)}{\Gamma(N+1)\Gamma(\beta)} \frac{1}{(N\ln N)^\beta}\left(\Gamma(\beta)\right.$$

$$\left. + \frac{\Gamma'(\beta+1) + \Gamma(\beta+1)(-\ln \ln N + \gamma_E - 1)}{\ln N} + \cdots\right)$$

$$\simeq \frac{\Gamma(N+1+\beta)}{\Gamma(N+1)} \frac{1}{(N\ln N)^\beta}\left[1 + \frac{\beta}{\ln N}\left(\frac{\Gamma'(\beta+1)}{\Gamma(\beta+1)}\right.\right.$$

$$\left.\left. - \ln \ln N + \gamma_E - 1\right) + \cdots\right]. \qquad (36)$$

(The next order is obtained in Appendix A.) The Stirling formula allows to simplify the expression

$$\frac{\Gamma(N+1+\beta)}{\Gamma(N+1)} \frac{1}{N^\beta} = 1 + O\left(\frac{1}{N}\right). \qquad (37)$$

Then, one gets from (36) the following expression for the generating function:

$$G(\beta) = -\beta \ln \ln N - \frac{\beta}{\ln N}\left(\ln \ln N + 1 - \gamma_E - \frac{\Gamma'(1+\beta)}{\Gamma(1+\beta)}\right)$$

$$+ o\left(\frac{1}{\ln N}\right). \qquad (38)$$

[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),

$$v_N = \frac{\langle X_g \rangle}{g} = \langle \Delta X_g \rangle = \ln \ln N + \frac{1}{\ln N}(\ln \ln N + 1) + \cdots,$$

$$D_N = \frac{\langle X_g^2 \rangle_c}{g} = \langle \Delta X_g^2 \rangle_c = \frac{\pi^2}{3 \ln N} + \cdots$$

$$\frac{\langle X_g^n \rangle_c}{g} = \langle \Delta X_g^n \rangle_c = \frac{n!\zeta(n)}{\ln N} = \frac{n!}{\ln N}\sum_{i\geq 1}\frac{1}{i^n} + \cdots, \qquad (39)$$

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 ith individual [at position $x_i(g)$] is

$$W_i(x) = \frac{\psi[x - x_i(g)]}{\Psi(x)}. \qquad (40)$$

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

$$W_i = e^{x_i(g) - X_g} = \frac{e^{x_i(g)}}{e^{x_1(g)} + \cdots + e^{x_N(g)}} = \frac{e^{y_i}}{e^{y_1} + \cdots + e^{y_N}},$$

(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

$$q_p = \left\langle \sum_{i=1}^{N} W_i^p \right\rangle,$$

(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

$$q_p = N \langle W_1^p \rangle$$

$$= N \int_0^{+\infty} dy_1 \, e^{-y_1} \cdots \int_0^{+\infty} dy_N \, e^{-y_N} e^{py_1} (e^{y_1} + \cdots + e^{y_N})^{-p}.$$

(43)

Using the representation (29), one obtains

$$q_p = \frac{N}{(p-1)!} \int_0^{+\infty} d\lambda \, \lambda^{p-1} I_p(\lambda) I_0(\lambda)^{N-1}$$

(44)

in terms of the function $I_0(\lambda)$ introduced in (31) and of its derivatives

$$I_p(\lambda) = \int_0^{+\infty} dy \, e^{(p-1)y - \lambda e^y} = (-)^p \frac{d^p}{d\lambda^p} I_0(\lambda)$$

$$= \lambda^{1-p} \int_\lambda^{+\infty} du \, u^{p-2} e^{-u}.$$

(45)

For small λ one gets, by taking derivatives of Eq. (32), to leading order,

$$I_0(\lambda) \simeq 1 + \lambda(\ln \lambda + \gamma_E - 1), \quad I_1(\lambda) \simeq -(\ln \lambda + \gamma_E),$$

$$I_p(\lambda) \simeq \frac{(p-2)!}{\lambda^{p-1}} \quad \text{for } p \geq 2.$$

(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 $p \geq 2$,

$$q_p \simeq \frac{1}{\ln N} \frac{1}{p-1}.$$

(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 $q_2 \gg q_3 \gg q_4 \gg \cdots$.

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 reads [1]

$$r_p(k) = \binom{p}{k-1} \left\langle \sum_{i=1}^{N} W_i^{p-k+1} (1 - W_i)^{k-1} \right\rangle.$$

(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),

$$r_p(k) = \binom{p}{k-1} \sum_{j=0}^{k-1} \binom{k-1}{j} (-1)^{k-1-j} q_{p-j}.$$

(49)

Replacing (47) in (49), one gets after some algebra

$$r_p(k) \simeq \frac{1}{\ln N} \frac{p}{(p-k)(p-k+1)},$$

(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,

$$r_p(p) = 1 - \sum_{k=1}^{p-1} r_p(k) \simeq 1 - \frac{p-1}{\ln N}.$$

(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

$$P_p(g+1) = \sum_{k=2}^{p} r_p(k) P_k(g) + r_p(1) \delta_g^0.$$

(52)

Using (50) and (51), this becomes

---

[1] In 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].

$$P_p(g+1) - P_p(g) = -\frac{p-1}{\ln N}P_p(g)$$

$$+ \sum_{k=2}^{p-1} \frac{1}{\ln N}\frac{p}{(p-k)(p-k+1)}P_k(g)$$

$$+ r_p(1)\delta_g^0. \tag{53}$$

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

$$R_p(t) = \frac{1}{(p-1)!}\frac{d}{dt}\frac{\Gamma(p-e^{-t})}{\Gamma(1-e^{-t})}$$

$$= \frac{1}{(p-1)!}\frac{d}{dt}[(1-e^{-t})(2-e^{-t})\cdots(p-1-e^{-t})], \tag{62}$$

which is just a polynomial of order $p-1$ in the variable e^{-t}. More explicitly, for the first values of p, one finds

$$R_2(t) = e^{-t}, \quad R_3(t) = \frac{3}{2}e^{-t} - e^{-2t}, \tag{63}$$

$$R_4(t) = \frac{11}{6}e^{-t} - 2e^{-2t} + \frac{1}{2}e^{-3t}, \quad \cdots .$$

The average coalescence times [using (62)] are

$$\langle T_p \rangle = \sum_{g=0}^{\infty} gP_p(g) = \ln N \int_0^{+\infty} dt\, tR_p(t)$$

$$= \ln N \int_0^{\infty} dt\left[1 - (1-e^{-t})\left(1-\frac{e^{-t}}{2}\right)\cdots\left(1-\frac{e^{-t}}{p-1}\right)\right] \tag{64}$$

and one gets

$$\langle T_2 \rangle = \ln N, \quad \langle T_3 \rangle = \frac{5}{4}\langle T_2 \rangle, \quad \langle T_4 \rangle = \frac{25}{18}\langle T_2 \rangle, \quad \cdots . \tag{65}$$

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),

$$\langle T_2^{\text{neutral}} \rangle = \mathcal{O}(N), \quad \langle T_3^{\text{neutral}} \rangle = \frac{4}{3}\langle T_2^{\text{neutral}} \rangle, \tag{66}$$

$$\langle T_4^{\text{neutral}} \rangle = \frac{3}{2}\langle T_2^{\text{neutral}} \rangle, \quad \cdots .$$

(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.

(tree diagram)	Neutral case	Exponential model
(tree diagram)	1	$\frac{3}{4}$
(tree diagram)	0	$\frac{1}{4}$

(tree diagram)	Neutral case	Exponential model
(tree diagram)	$\frac{2}{3}$	$\frac{1}{3}$
(tree diagram)	$\frac{1}{3}$	$\frac{1}{6}$
(tree diagram)	0	$\frac{1}{6}$
(tree diagram)	0	$\frac{2}{9}$
(tree diagram)	0	$\frac{1}{9}$

the velocity $v_N \to \infty$ as $N \to \infty$, 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]

$$
h_{g+1}(x) = \begin{cases} \min\left(1, \int d\epsilon\, \psi(\epsilon) h_g(x - \epsilon)\right) & \text{if that number} \\ & \text{is larger than } 1/N, \\ 0 & \text{otherwise.} \end{cases}
$$

(68)

{Note that in the exponential model $[\psi(\epsilon) = e^{-\epsilon}]$, it is easy to see that the solution to (68) is

$$
h_g(x) = \begin{cases} 1 & \text{for } x < Y_g, \\ e^{-(x - Y_g)} & \text{for } Y_g < x < Y_g + \ln N, \\ 0 & \text{for } x > Y_g + \ln N, \end{cases}
$$

(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

$$
v_{\text{cutoff}}^{\exp} = Y_{g+1} - Y_g = \ln(\ln N + 1) \simeq \ln \ln N,
$$

(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

$$
h_g(x) \propto L_0 \sin\left(\pi \frac{x - Y_g}{L_0}\right) e^{-\gamma_0(x - Y_g)}
$$

(71a)

and

$$
v_{\text{cutoff}}^{\text{F-KPP}} = Y_{g+1} - Y_g \simeq v(\gamma_0) - \frac{\pi^2 v''(\gamma_0)}{2L_0^2},
$$

(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 $x - Y_g < L_0$.

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

$$p(\delta) = C_1 e^{-\gamma_0 \delta} \tag{72}$$

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

$$R(\delta) = \frac{1}{\gamma_0} \ln\left(1 + C_2 \frac{e^{\gamma_0 \delta}}{L_0^\alpha}\right), \tag{73}$$

where C_2 is another constant and $\alpha = 3$. Finally, in [8] it was argued that

$$C_1 C_2 = \pi^2 \gamma_0 v''(\gamma_0). \tag{74}$$

[Note that (72)–(74) have ben 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,

$$\langle e^{-\beta Y_g} \rangle \sim e^{g G(\beta)} \quad \text{where} \tag{75}$$

$$G(\beta) = -\beta v_{\text{cutoff}} + \int d\delta\, p(\delta)(e^{-\beta R(\delta)} - 1).$$

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]

$$v_N = v(\gamma_0) - \frac{\pi^2 \gamma_0^2 v''(\gamma_0)}{2 \ln^2 N} + \gamma_0^2 v''(\gamma_0)\pi^2 \frac{3 \ln \ln N}{\ln^3 N} + \cdots$$

$$= v(\gamma_0) - \frac{\pi^2 \gamma_0^2 v''(\gamma_0)}{2(\ln N + 3 \ln \ln N)^2} + \cdots,$$

$$D_N = \gamma_0 v''(\gamma_0)\frac{\pi^4}{3 \ln^3 N} + \cdots,$$

$$\frac{\langle (Y_g - Y_0)^n \rangle_c}{g} = \gamma_0^{3-n} v''(\gamma_0)\frac{\pi^2 n!\zeta(n)}{\ln^3 N} + \cdots \quad \text{for } n \geq 2. \tag{76}$$

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

$$\text{Prob}(y_{\text{rightmost}}) = N(1 - e^{-y_{\text{rightmost}}})^{N-1}e^{-y_{\text{rightmost}}}$$

$$\simeq \exp[-(y_{\text{rightmost}} - \ln N) - e^{-(y_{\text{rightmost}} - \ln N)}]. \tag{77}$$

$y_{\text{rightmost}}$ 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

$$p(\delta) \simeq \exp(-\delta - e^{-\delta}) \simeq \exp(-\delta), \qquad (78)$$

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 isdecorrelated 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

$$\text{Prob}(y_k) = \frac{e^{-y_k}}{1 - e^{-l}} \quad \text{for } 0 < y_k < l. \qquad (79)$$

As in (26), we introduce the generating function of the displacement ΔX_g, given the value of l,

$$\langle e^{-\beta \Delta X_g} | l \rangle = \int dz \, \text{Prob}(z) e^{-\beta z} \int dy_1 \, \text{Prob}(y_1) \cdots \int dy_{N-1} \, \text{Prob}(y_{N-1}) (e^{y_1} + \cdots + e^{y_{N-1}} + e^l)^{-\beta}$$

$$= \frac{\Gamma(N+1+\beta)}{\Gamma(N+1)} \frac{1}{(1 - e^{-l})^{N-1}} \int_0^l dy_1 e^{-y_1} \cdots \int_0^l dy_{N-1} e^{-y_{N-1}} (e^{y_1} + \cdots + e^{y_{N-1}} + e^l)^{-\beta}, \qquad (80)$$

where (28) and (79) were used. By using the same representation (29) that led to (30), one gets

$$\langle e^{-\beta \Delta X_g} | l \rangle = \frac{\Gamma(N+1+\beta)}{\Gamma(N+1)\Gamma(\beta)} \int_0^\infty d\lambda \, \lambda^{\beta-1}$$

$$\times \left(\frac{1}{1 - e^{-l}} \int_0^l dy \, e^{-y - \lambda e^y} \right)^{N-1} e^{-\lambda e^l}, \qquad (81)$$

which, in terms of $I_0(\lambda)$ defined in (31), is the same as

$$\langle e^{-\beta \Delta X_g} | l \rangle = \frac{\Gamma(N+1+\beta)}{\Gamma(N+1)\Gamma(\beta)} \int_0^\infty d\lambda \, \lambda^{\beta-1}$$

$$\times \left(\frac{I_0(\lambda) - e^{-l} I_0(\lambda e^l)}{1 - e^{-l}} \right)^{N-1} e^{-\lambda e^l}, \qquad (82)$$

where, using (32),

$$\frac{I_0(\lambda) - e^{-l} I_0(\lambda e^l)}{1 - e^{-l}} = 1 - \lambda \frac{l}{1 - e^{-l}}$$

$$+ \sum_{k=0}^{+\infty} \frac{(-1)^k}{(k+1)(k+2)!} \lambda^{k+2} \frac{e^{l(k+1)} - 1}{1 - e^{-l}}. \qquad (83)$$

Expressions (82) and (83) are valid for any value of l. We now consider a large fluctuation $l = \ln N + \delta$ with $1 \ll \delta \lesssim \ln \ln N$. 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

$$\left(\frac{I_0(\lambda) - e^{-l} I_0(\lambda e^l)}{1 - e^{-l}} \right)^{N-1} \simeq \exp\left[-\mu\left(1 + \frac{\delta}{\ln N}\right) \right.$$

$$+ \sum_{k=0}^{+\infty} \frac{(-1)^k}{(k+1)(k+2)!}$$

$$\times \left. \left(\frac{\mu}{\ln N} \right)^{k+2} e^{\delta(k+1)} \right]. \qquad (84)$$

We are only interested in the leading order in $1/\ln N$. Dropping higher order terms, one gets, in (82),

$$\langle e^{-\beta \Delta X_g} | \delta \rangle \simeq \frac{\Gamma(N+1+\beta)}{\Gamma(N+1)\Gamma(\beta)} \frac{1}{(N \ln N)^\beta} \int_0^\infty d\mu \, \mu^{\beta-1}$$

$$\times \exp\left[-\mu\left(1 + \frac{\delta + e^\delta}{\ln N}\right) \right]$$

$$\simeq \frac{1}{(\ln N)^\beta} \left(1 + \frac{e^\delta}{\ln N}\right)^{-\beta}, \qquad (85)$$

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

$$\Delta X_g(\delta) \simeq \ln \ln N + \ln\left(1 + \frac{e^\delta}{\ln N}\right) \simeq v_{\text{cutoff}} + R(\delta), \qquad (86)$$

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-
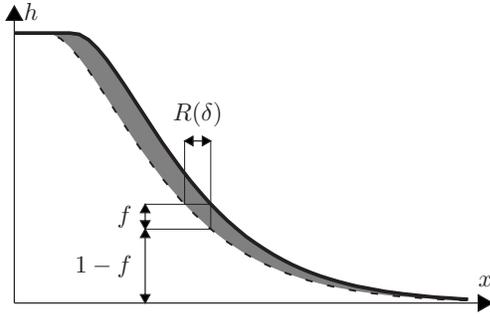
FIG. 3. Effect of a fluctuation of a front. The dashed line is the front (87) in the absence of a fluctuation. The plain line is the front (88) if a rare fluctuation occured. The grey area represents the contribution to the front from the descendants of the fluctuation. After the front has relaxed, they 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

$$h_{\text{no fluctuation}}(x,g) \propto e^{-\gamma_0(x - Y_g^{\text{no fluctuation}})} \qquad (87a)$$

with

$$Y_g^{\text{no fluctuation}} = Y_{g_0} + v_{\text{cutoff}}(g - g_0). \qquad (87b)$$

[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

$$h_{\text{fluctuation}}(x,g) \propto e^{-\gamma_0(x - Y_g^{\text{fluctuation}})} \qquad (88a)$$

with

$$Y_g^{\text{fluctuation}} = Y_{g_0} + v_{\text{cutoff}}(g - g_0) + R(\delta). \qquad (88b)$$

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

$$f = 1 - e^{-\gamma_0 R(\delta)}. \qquad (89)$$

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:

$$\text{Prob}(f) = \frac{C_1 C_2}{\gamma_0 L_0^\alpha} \frac{1}{f^2}. \qquad (90)$$

[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

$$\text{Prob}(f)$$
$$= \begin{cases} \dfrac{1}{\ln N} \dfrac{1}{f^2} & \text{for the exponential model.} \\[2ex] \dfrac{\pi^2 \gamma_0^3 v''(\gamma_0)}{\ln^3 N} \dfrac{1}{f^2} & \text{for the generic Fisher-KPP case.} \end{cases}$$
$$(91)$$

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

$$q_p = \int_0^1 df\, \text{Prob}(f) f^p = \frac{C_1 C_2}{\gamma_0 L_0^\alpha} \frac{1}{p-1} \qquad (92)$$

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

$$r_p(k) = \binom{p}{k-1} \int_0^1 df\, \text{Prob}(f) f^{p-k+1}(1-f)^{k-1}$$
$$= \frac{C_1 C_2}{\gamma_0 L_0^\alpha} \frac{p}{(p-k)(p-k+1)}, \qquad (93)$$

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

$$\langle T_2 \rangle \simeq \frac{\ln^3 N}{\pi^2 \gamma_0^3 v''(\gamma_0)}, \qquad (94)$$

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

---

[2]In 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,

$$D_N \times \langle T_2 \rangle \simeq \frac{\pi^2}{3\gamma_0^2}. \qquad (95)$$

We will test numerically this identity in Sec. V.

As a side remark, we note that if Prob(f) of (91) is replaced by Cste 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 Nth 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 $i_1, \ldots, i_p$ is simply given by $\max[T_g(i_1, i_2), T_g(i_1, i_3), \ldots, T_g(i_1, i_p)]$. 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 $\alpha_1, \ldots, \alpha_p$ individuals of the current generation, then this node is the most common ancestor of $\Sigma_{i \neq j} \alpha_i \alpha_j$ pairs of individuals. More generally, this node is the most common ancestor of $\binom{\Sigma_i \alpha_i}{q} - \Sigma_i \binom{\alpha_i}{q}$ 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: $\psi(\epsilon) = \Sigma_q \phi_q \delta(\epsilon - q)$. 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 $\Sigma_x n(x, g) \phi_{y-x}$, 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 $N \simeq 10^{100}$. To

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,

$$\text{Prob}(y \text{ comes from } x) = \frac{n(x,g)\phi_{y-x}}{\sum_{x'} n(x',g)\phi_{y-x'}}. \qquad (96)$$

[Compare to (40).] Then, one obtains that

$$\bar{T}_{g+1}(y,y') = 1 + \sum_{x,x'} \text{Prob}(y \text{ comes from } x)$$

$$\times \text{Prob}(y' \text{ comes from } x')\bar{T}_g(x,x')$$

$$\times \left(1 - \frac{\delta_x^{x'}}{n(x,g)}\right). \qquad (97)$$

(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

$$\frac{1}{N(N-1)}\sum_{x,x'} \bar{T}_g(x,x')n(x,g)n(x',g)\left(1 - \frac{\delta_x^{x'}}{n(x,g)}\right). \qquad (98)$$

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
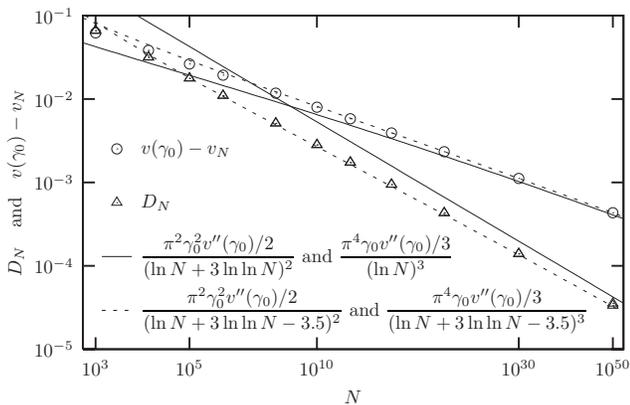

FIG. 4. Numerical simulations of model B with $\psi(\epsilon) = \frac{1}{4}\sum_{n\leq0}\delta(\epsilon-n)$. For this model, one has $\gamma_0 = \ln(2)$, $v(\gamma_0) = -1$, and $v''(\gamma_0) = 2/\ln(2)$. The circles are the correction to the velocity and the triangles the diffusion constant, as a function 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 $\ln\ln N$.)
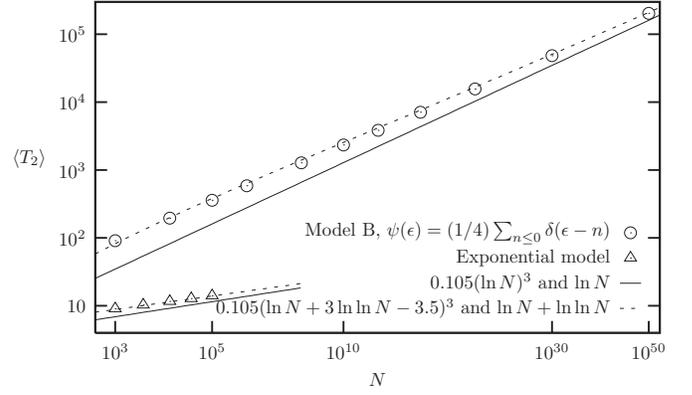

FIG. 5. Numerical simulations of $\langle T_2 \rangle$ for model B with $\psi(\epsilon) = \frac{1}{4}\sum_{n\leq0}\delta(\epsilon-n)$ (circles) and for the exponential model (triangles). The plain lines are the predictions (65) for $\langle T_2 \rangle$ and (94), while the dotted lines are the same predictions 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 $\psi(\epsilon) = \frac{1}{4}\sum_{n\leq0}\delta(\epsilon-n)$ 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
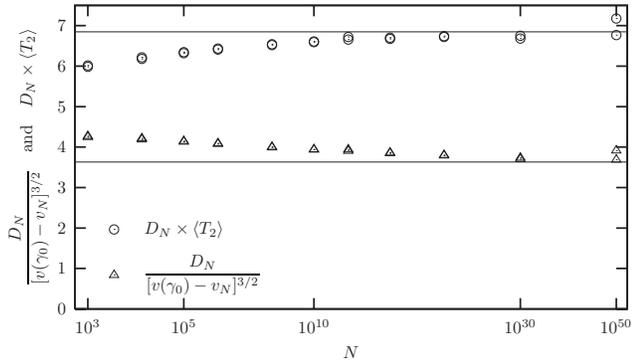
FIG. 6. Numerical simulations of model B with $\psi(\epsilon)$ $=\frac{1}{4}\Sigma_{n\leqslant 0}\delta(\epsilon-n)$. The circles represent the product $D_N\times\langle T_2\rangle$ compared to the prediction (95). The triangles are the ratio of the diffusion constant and the correction to the velocity 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) $\langle T_2\rangle\simeq\ln N$. Here again, the fit is improved by including the subleading corrections (A15) $\langle T_2\rangle\simeq\ln N+\ln\ln N$ 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
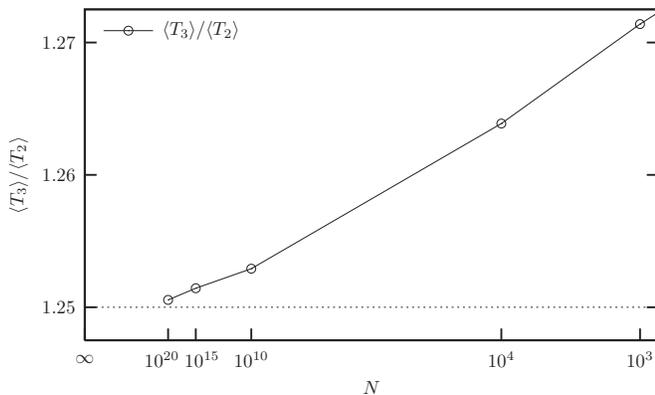

FIG. 7. Numerical simulations of model B with $\psi(\epsilon)$ $=\frac{1}{4}\Sigma_{n\leqslant 0}\delta(\epsilon-n)$. The circles represent the ratio $\langle T_3\rangle/\langle T_2\rangle$ as a function of N, compared to the result $5/4$ suggested by the phenomenological theory of Sec. IV C. (The scale 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 $N\to\infty$, the velocity of the front associated to the exponential model diverges when N $\to\infty$. 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.

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),

$$
e^{G(\beta)} = \frac{1}{\ln^\beta N} \frac{1}{\Gamma(\beta)} \int_0^\infty d\mu \; \mu^{\beta-1} e^{-\mu}
$$
$$
\times \left[1 + \mu \frac{\ln \mu - \ln \ln N + \gamma_E - 1}{\ln N} \right.
$$
$$
\left. + \frac{1}{2}\left(\mu \frac{\ln \mu - \ln \ln N + \gamma_E - 1}{\ln N} \right)^2 + \cdots \right]. \quad (A1)
$$

The integrals of each term can be computed using (35). One gets

$$
e^{G(\beta)} = \frac{1}{\ln^\beta N} \left[1 + \frac{\beta}{\ln N}\left(\frac{\Gamma'(\beta+1)}{\Gamma(\beta+1)} - l \right) \right.
$$
$$
\left. + \frac{\beta(\beta+1)}{2\ln^2 N}\left(\frac{\Gamma''(\beta+2)}{\Gamma(\beta+2)} - 2\frac{\Gamma'(\beta+2)}{\Gamma(\beta+2)}l + l^2 \right) + \cdots \right]
$$
$$
(A2)
$$

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,

$$
v_N = \ln \ln N + \frac{\ln \ln N + 1}{\ln N} - \frac{(\ln \ln N)^2 - 1 + \frac{\pi^2}{6}}{2\ln^2 N} + \cdots \;,
$$

$$
D_N = \frac{\pi^2}{3}\frac{1}{\ln N} - \frac{1}{\ln^2 N}\left(\frac{\pi^2}{3}\ln \ln N - \frac{\pi^2}{6} + 2\zeta(3) \right) + \cdots \;.
$$
$$
(A3)
$$

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 ith 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 ith individual of generation g. The probability distribution of the p_i is

$$
\text{Prob}(p_1, \ldots, p_N) = \frac{p!}{p_1! \cdots p_N!} \delta^p_{p_1 + \cdots + p_N} W_1^{p_1} \cdots W_N^{p_N}.
$$
$$
(A4)
$$

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

$$
r_p(k) = \binom{N}{k} \sum_{p_1 \geqslant 1, \ldots, p_k \geqslant 1} \frac{p!}{p_1! \cdots p_k!} \delta^p_{p_1 + \cdots + p_k} \langle W_1^{p_1} \cdots W_k^{p_k} \rangle.
$$
$$
(A5)
$$

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,

$$
r_p(k) = \binom{N}{k} \sum_{n \geqslant 0} \binom{k}{n} \sum_{p_1 \geqslant 2, \ldots, p_n \geqslant 2} \frac{p!}{p_1! \cdots p_n!} \delta^{p-k+n}_{p_1 + \cdots + p_n}
$$
$$
\times \langle W_1^{p_1} \cdots W_n^{p_n} W_{n+1} \cdots W_k \rangle. \quad (A6)
$$

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),

$$
J_{p,k,n}^{p_1, \ldots, p_n} = \langle W_1^{p_1} \cdots W_n^{p_n} W_{n+1} \cdots W_k \rangle
$$
$$
= \int_0^\infty dy_1 e^{-y_1} \cdots \int_0^\infty dy_N e^{-y_N} \frac{e^{p_1 y_1 + \cdots + p_n y_n + y_{n+1} + \cdots + y_k}}{(e^{y_1} + \cdots + e^{y_N})^p}.
$$
$$
(A7)
$$

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):

$$
J_{p,k,n}^{p_1, \ldots, p_n} = \frac{1}{(p-1)!} \int_0^{+\infty} d\lambda \; \lambda^{p-1} I_{p_1}(\lambda) \cdots
$$
$$
\times I_{p_n}(\lambda) I_1(\lambda)^{k-n} I_0(\lambda)^{N-k}. \quad (A8)
$$

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

$$\lambda^{p-1} I_{p_1} \cdots I_{p_n} I_1^{k-n} I_0^{N-k} \simeq \frac{(p_1-2)! \cdots (p_n-2)!}{N^{k-1} \ln^{n-1} N} \mu^{k-1} e^{-\mu} \left(1 + \frac{(k-n-\mu)(\ln \ln N - \ln \mu - \gamma_E) - \mu}{\ln N} + \cdots \right), \qquad (A9)$$

(A8) can then be evaluated using (35). One gets

$$J_{p,k,n}^{p_1,\ldots,p_n} = \frac{(p_1-2)! \cdots (p_n-2)!}{(p-1)!} \frac{(k-1)!}{N^k \ln^n N} \left(1 + \frac{n\left(\frac{\Gamma'(k)}{\Gamma(k)} + \gamma_E - \ln \ln N\right) - (k-1)}{\ln N} + \cdots \right) \qquad (A10)$$

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 $k=p$),

$$r_p(k) \simeq \frac{N^k}{k!} \left(k \frac{p!}{(p-k+1)!} J_{p,k,1}^{p-k+1} \right.$$

$$\left. + \frac{k(k-1)}{2} \sum_{p_1=2}^{p-k} \frac{p!}{p_1!(p-k+2-p_1)!} J_{p,k,2}^{p_1,p-k+2-p_1} + \cdots \right). \qquad (A11)$$

After some algebra, one gets, for $k < p$,

$$r_p(k) = \frac{p}{(p-k+1)(p-k)} \frac{1}{\ln N} \left\{ 1 + \frac{1}{\ln N} \left[\sum_{n=1}^{k-1} \frac{1}{n} \right. \right.$$

$$\left. \left. + \frac{2(k-1)}{p-k+2} \left(\sum_{n=1}^{p-k-1} \frac{1}{n} - \frac{3}{2} \right) - \ln \ln N \right] + \cdots \right\} \qquad (A12)$$

[we used, among other things, $\Gamma'(k)/\Gamma(k) + \gamma_E = 1 + \frac{1}{2} + \cdots + \frac{1}{k-1}$].

We can now compute the $\langle T_k \rangle$. From the recurrence

$$\langle T_p \rangle = 1 + \sum_{k=1}^{p} r_p(k) \langle T_k \rangle, \qquad (A13)$$

we get, using $\sum_k r_p(k) = 1$ and $\langle T_1 \rangle = 0$,

$$\langle T_p \rangle = \frac{1 + \sum_{k=2}^{p-1} r_p(k) \langle T_k \rangle}{\sum_{k=1}^{p-1} r_p(k)}. \qquad (A14)$$

For the first values of p, we obtain

$$\langle T_2 \rangle = \ln N + \ln \ln N + o(1),$$

$$\langle T_3 \rangle = \frac{5}{4}(\ln N + \ln \ln N) + o(1),$$

$$\langle T_4 \rangle = \frac{25}{18}(\ln N + \ln \ln N) - \frac{1}{54} + o(1). \qquad (A15)$$

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 $n \to 0$. In some cases, the n dependence of this averaged product is simple enough for the analytic continuation $n \to 0$ 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 $n \to 0$. 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

$$n = n_0 > n_1 > n_2 > \cdots > n_k = 1. \qquad (B1)$$

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

$$Q_m = \frac{\frac{n}{n_i}\binom{n_i}{m}}{\binom{n}{m}} = \frac{n(n_i-1)(n_i-2)\cdots(n_i-m+1)}{n(n-1)\cdots(n-m+1)}. \qquad (B2)$$

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

TABLE II. All possible trees of three replicas, their probabilities and degeneracies.
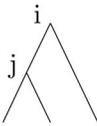
	$\dfrac{n(n_i - n_{i+1})(n_j - n_{j+1})}{n(n-1)(n-2)}$	three cases
	$\dfrac{n(n_i - n_{i+1})(n_i - 2n_{i+1})}{n(n-1)(n-2)}$	one case

TABLE III. All possible trees of four replicas, their probabilities and degeneracies.

	$\dfrac{n(n_i - n_{i+1})(n_j - n_{j+1})(n_k - n_{k+1})}{n(n-1)(n-2)(n-3)}$	twelve cases
	$\dfrac{n(n_i - n_{i+1})(n_j - n_{j+1})(n_k - n_{k+1})}{n(n-1)(n-2)(n-3)}$	six cases
	$\dfrac{n(n_i - n_{i+1})(n_i - 2n_{i+1})(n_j - n_{j+1})}{n(n-1)(n-2)(n-3)}$	six cases
	$\dfrac{n(n_i - n_{i+1})(n_j - n_{j+1})(n_j - 2n_{j+1})}{n(n-1)(n-2)(n-3)}$	four cases
	$\dfrac{n(n_i - n_{i+1})(n_i - 2n_{i+1})(n_i - 3n_{i+1})}{n(n-1)(n-2)(n-3)}$	one case

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

$$\frac{n(n_i - n_{i+1})(n_j - n_{j+1})}{n(n-1)(n-2)}, \tag{B3}$$

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 $n \to 0$ and one reverses the inequality (B1) into

$$n = n_0 < n_1 < n_2 < \cdots < n_k = 1. \tag{B4}$$

One then takes a continuous limit ($k \to \infty$), where n_i becomes a continuous variable x,

$$n_i = x. \tag{B5}$$

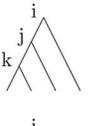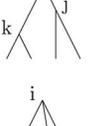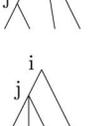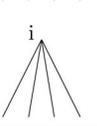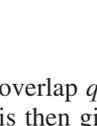In the spin-glass theory [45,46], there is an ultrametric distance between pairs of replicas, related to the overlap $q_{\alpha,\beta}$. (The distance is a decreasing function of the overlap.) This overlap $q_{\alpha,\beta}$ depends on the level at which the branches of these two replicas merge: this 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 $k \to \infty$, when n_i becomes a continuous variable (B5), the overlap q_i becomes a 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

$$\text{Prob}(q_{\alpha,\beta} = q_0) + \text{Prob}(q_{\alpha,\beta} = q_1) + \cdots + \text{Prob}(q_{\alpha,\beta} = q_{i-1})$$

$$= 1 - Q_2(n_i) = \frac{n - n_i}{n-1}. \tag{B6}$$

Therefore, in the $n \to 0$ limit, the probability $P(q)$ that the

overlap $q_{\alpha,\beta}$ between two replicas α and β takes the value q is then given by

$$\int_0^{q(x)} P(q')dq' = \lim_{n \to 0}(1 - Q_2) = x \tag{B7}$$

and this leads to the famous relation [42] between the function $q(x)$ and the probability distribution of the overlap

$$P(q) = \frac{dx}{dq}. \tag{B8}$$

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,

$$e^{-T} = 1 - Q_2 = \frac{n(n - n_i)}{n(n-1)}, \tag{B9}$$

which leads in the $n \to 0$ limit to

$$e^{-T} = x. \tag{B10}$$

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 $n \to 0$ limit of (B2), one gets that the probability Q_m that m individuals have a coalescence time $T_m < T$ is given by

$$Q_m \to \frac{\Gamma(m-x)}{(m-1)!\Gamma(1-x)} \tag{B11}$$

which, by taking the derivative with respect with T, gives

$$\langle (T_m)^p \rangle = \int_0^1 dx \, T(x)^p \frac{dQ_m}{dx}$$

$$= \int_0^\infty dT \, T^p \frac{d}{dT} \frac{\Gamma(m-e^{-T})}{(m-1)!\Gamma(1-e^{-T})}. \tag{B12}$$

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

$$M = \sum_{i=1}^N k_i. \tag{C1}$$

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

$$q_n = \left\langle \frac{\sum_i \binom{k_i}{n}}{\binom{M}{n}} \right\rangle = \left\langle \frac{\sum_i k_i(k_i-1)\cdots(k_i-n+1)}{M(M-1)\cdots(M-n+1)} \right\rangle. \tag{C2}$$

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 approximatively equal to $(N\langle k \rangle)^n$ and

$$q_n \simeq \frac{1}{N^{n-1}\langle k \rangle^n} \left\langle \frac{\Gamma(k+1)}{\Gamma(k-n+1)} \right\rangle. \tag{C3}$$

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 coalesences by pairs may occur from one generation to the previous one, one has

$$T_n(g+1)$$

$$= \begin{cases} T_n(g)+1 & \text{with probability } 1 - \frac{1}{2}n(n-1)q_2, \\ T_{n-1}(g)+1 & \text{with probability } \frac{1}{2}n(n-1)q_2. \end{cases} \tag{C4}$$

In the steady state [49], this implies that

$$\langle T_n^p \rangle = \left(1 - \frac{n(n-1)}{2}q_2\right)\langle (1+T_n)^p \rangle + \frac{n(n-1)}{2}q_2\langle (1+T_{n-1})^p \rangle \tag{C5}$$

and using the fact that $T_1(g)=0$, one gets

$$\langle T_n \rangle = \left(2 - \frac{2}{n}\right)\frac{1}{q_2}. \tag{C6}$$

We see that all the times T_n scale like N (since $q_2 \sim N^{-1}$) and that

$$\frac{\langle T_3 \rangle}{\langle T_2 \rangle} = \frac{4}{3}, \quad \frac{\langle T_4 \rangle}{\langle T_2 \rangle} = \frac{3}{2}, \quad \cdots, \quad \frac{\langle T_n \rangle}{\langle T_2 \rangle} = \frac{2(n-1)}{n}. \tag{C7}$$

One can also calculate from (C5) higher moments of the T_n's or their generating functions

$$\frac{\langle (T_2)^2 \rangle}{\langle T_2 \rangle^2} = 2, \quad \frac{\langle (T_3)^2 \rangle}{\langle T_3 \rangle^2} = \frac{13}{8}. \tag{C8}$$

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).

[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. Donelly, 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.