

# A WEIGHTED RANDOM WALK MODEL, WITH APPLICATION TO A GENETIC ALGORITHM

C. DOMBRY,\* *Université Claude Bernard Lyon 1*

## Abstract

We consider a weighted random walk model defined as follows. An  $n$ -step random walk on the integers with distribution  $P_n$  is weighted by giving the path  $S = (S_0, \dots, S_n)$  a probability proportional to  $\prod_{k=1}^n f(S_k) P_n(S)$ , where the function  $f$  is the so-called fitness function. In the case of power-type fitness, we prove the convergence of the renormalized path to a deterministic function with exponential speed. This function is a solution to a variational problem. In the case of the simple symmetric random walk, explicit computations are done. Our result relies on large deviations techniques and Varadhan's integral lemma. We then study an application of this model to mutation–selection dynamics on the integers where a random walk operates the mutation. This dynamics is the infinite-population limit of that of mutation–selection genetic algorithms. We prove that the population grows to  $\infty$  and make explicit its growth speed. This is a toy model for modelling the effect of stronger selection at  $\infty$  for genetic algorithms taking place in a noncompact space.

*Keywords:* Random walk; functional large deviations principle; genetic algorithm

2000 Mathematics Subject Classification: Primary 60G50

Secondary 60F10

## 1. Introduction and motivations

### 1.1. The weighted random walk model and its heuristic

Let  $\Omega_n$  be the set of paths of length  $n$ :

$$\Omega_n = \{S = (S_0, \dots, S_n) \in \mathbb{Z}^{n+1}\}.$$

Let  $\pi$  and  $\delta$  be probability measures on  $\mathbb{Z}$ . We suppose that  $\pi$  and  $\delta$  have finite supports. Let  $P_n$  be the probability measure on  $\Omega_n$  corresponding to the random walk with initial distribution  $\pi$  and independent increments identically distributed according to  $\delta$ . The probability of  $S \in \Omega_n$  is given by

$$P_n(S \in \Omega_n, S_0 = s_0, \dots, S_n = s_n) = \pi(s_0) \prod_{k=0}^{n-1} \delta(s_{k+1} - s_k).$$

We denote by  $E_n$  the associated expectation. Let  $f : \mathbb{Z} \rightarrow [0, \infty)$  be a fitness function. The weight of a path  $S \in \Omega_n$  is defined by

$$\Pi_n^f(S) = \prod_{k=1}^n f(S_k).$$

Received 15 April 2005; revision received 12 October 2006.

\* Postal address: Institut Camille Jordan, Université Claude Bernard Lyon 1, 43 Boulevard du 11 Novembre 1918, 69622 Villeurbanne cedex, France. Email address: dombry@math.univ-lyon1.fr

The object of our study is the probability measure  $P_n^f$  on  $\Omega_n$  defined by

$$P_n^f(S) = \frac{1}{Z_n^f} \Pi_n^f(S) P_n(S),$$

where

$$Z_n^f = E_n[\Pi_n^f(S)]$$

is a partition function that we suppose to be nonzero. For any  $\phi: \Omega_n \rightarrow \mathbb{R}$ , the associated expectation,  $E_n^f$ , satisfies

$$E_n^f[\phi] = \frac{1}{Z_n^f} E_n[\phi(S)\Pi_n^f(S)].$$

In other words, the weighted random walk model gives to a path  $S$  a probability  $P_n^f(S)$  proportional to  $\Pi_n^f(S) P_n(S)$ .

We are interested in the asymptotic properties of  $P_n^f$  as the length  $n$  goes to infinity, i.e. its convergence (law of large numbers) and its fluctuations (central limit theorem). Note that the sequence of probability measures  $P_n^f$ ,  $n \geq 1$ , is well defined if the following assumption holds.

**Assumption 1.** For every  $n \geq 1$ ,  $Z_n^f$  is nonzero.

This assumption is satisfied if  $f$  is positive. In the case where  $f$  vanishes on the set of negative integers, it is satisfied if there exist  $n_1 > 0$  and  $n_2 \geq 0$  such that  $\pi(n_1) > 0$  and  $\delta(n_2) > 0$ .

**Remark 1.** The system  $(\Omega_n, P_n^f)$  is not projective, in the sense that there is no measure on the set of infinite paths whose finite-dimensional distributions are  $P_n^f$ . Furthermore, if the fitness function  $f$  vanishes on the set of nonpositive integers, then the probability  $P_n^f$  forces the paths to remain positive.

We now give some heuristics and simple examples. In the case of the constant fitness function  $f \equiv c > 0$ , all paths have the same weight and  $P_n^f = P_n$ . Suppose that the increments have drift  $m$  (i.e.  $\delta$  has mean  $m$ ) and finite variance  $\sigma^2$ . Then the distribution of  $S_n$  has mean  $E[S_0] + nm$  and variance  $\text{var}(S_0) + n\sigma^2$ . In the zero-drift case the random walk is said to be *diffusive*; otherwise it is said to be *ballistic*. In the case of the fitness function  $f(x) = \mathbf{1}_{(0,\infty)}(x)$ , the effect of the weight is conditioning on strictly positive paths: let

$$\Omega_n^+ = \{S \in \Omega_n : S_k > 0 \text{ for all } k \in \{1, \dots, n\}\};$$

then the weighted random walk distribution is  $P_n^f(\cdot) = P_n(\cdot \mid \Omega_n^+)$ . Intuitively, if the drift  $m$  is nonpositive, then the weighted random walk will have diffusive behaviour, with  $S_n/n$  going to 0 as  $n$  goes to  $\infty$ . If the drift is positive, then the weighted random walk will have ballistic behaviour with speed  $m$ , i.e.  $S_n/n$  has limit  $m$  as  $n$  goes to  $\infty$ . In the sequel, we focus on nondecreasing fitness functions. If the fitness is nondecreasing, then higher paths have greater weight and hence are enhanced in the weighted random walk model. We are interested in the case of fitness functions which are stronger at infinity, i.e. such that  $f(x) \rightarrow \infty$  as  $x \rightarrow \infty$ . The central question is: how does a stronger fitness function modify the limit speed of the walk? This will depend on how fast the fitness function goes to  $\infty$ , according to the following definition.

**Definition 1.** Let  $f : \mathbb{Z} \rightarrow [0, \infty)$  be a nondecreasing fitness function such that  $\lim_{x \rightarrow \infty} f(x) = \infty$ . We say that

- $f$  is a subpower fitness function if, for all  $\gamma > 0$ ,  $f(x)x^{-\gamma} \rightarrow 0$  as  $x \rightarrow \infty$ ; that
- $f$  is a power fitness function of type  $\beta \in (0, \infty)$  if there exists some  $l > 0$  such that  $f(x)x^{-\beta} \rightarrow l$  as  $x \rightarrow \infty$ ; and that
- $f$  is an overpower fitness function if, for all  $\gamma > 0$ ,  $f(x)x^{-\gamma} \rightarrow \infty$ , as  $x \rightarrow \infty$ .

The heuristic is the following. In the subpower case the fitness is not great enough at infinity to have an important effect, and its effect will be negligible. In the overpower case its effect will be very strong, and it will force all the increments to be maximal. Finally, the power case will be the interesting one, where an equilibrium is found between the entropy of the random walk and the force of the weighting.

**Example 1.** For the sake of clarity, throughout the paper we develop an example with different explicit fitness functions. We focus on the case of the symmetric simple random walk starting from point 1, i.e. in which  $\delta = \frac{1}{2}(\delta_1 + \delta_{-1})$  and  $\pi = \delta_1$ .

- The case of a power fitness function of type  $\beta > 0$  will be illustrated by the reference fitness function  $f_\beta$  defined by

$$f_\beta(x) = \begin{cases} x^\beta & \text{if } x \geq 0, \\ 0 & \text{if } x < 0. \end{cases}$$

We will show in the sequel that the normalized weighted random walk  $(1/n)S_{[nt]}$  converges on  $[0, 1]$  to a deterministic function  $\psi_\beta$  that we compute explicitly (see Theorem 1 and Proposition 1).

- The case of a subpower fitness function will be illustrated by the reference fitness function  $f_0$  defined by

$$f_0(x) = \begin{cases} \log(1+x) & \text{if } x \geq 0, \\ 0 & \text{if } x < 0. \end{cases}$$

In this case, the author believes that the normalized weighted random walk  $(1/n)S_{[nt]}$  converges on  $[0, 1]$  to the deterministic function  $\psi_0 \equiv 0$  (see the simulations in Subsection 1.2). This is reminiscent of the behaviour of the simple random walk conditioned to remain positive. The effect of the weight vanishes.

- The case of an overpower fitness function will be illustrated by the reference fitness function  $f_\infty$  defined by

$$f_\infty(x) = \begin{cases} \exp(x) - 1 & \text{if } x \geq 0, \\ 0 & \text{if } x < 0. \end{cases}$$

In this case, the author believes that the normalized weighted random walk  $(1/n)S_{[nt]}$  converges on  $[0, 1]$  to the deterministic function  $\psi_\infty(t) = t$  (see the simulations in Subsection 1.2). The weight is very strong and forces almost all the increments to equal 1.

It is worth noting that, intuitively, the subpower case corresponds to the limit of the power fitness case as  $\beta \rightarrow 0$ : we can prove that, for any  $t \in [0, 1]$ ,  $\psi_\beta(t)$  has limit  $\psi_0(t) = 0$  as  $\beta$  goes to 0. In the same way, the overpower case corresponds to the limit of the power fitness case as  $\beta \rightarrow \infty$ : we can prove that, for any  $t \in [0, 1]$ ,  $\psi_\beta(t)$  has limit  $\psi_\infty(t) = t$  as  $\beta$  goes to  $\infty$ . Hence, the parameter  $\beta$  is understood to represent the strength of the weight in the model.

**1.2. Simulation**

Let  $n, \pi, \delta,$  and  $f$  be fixed. The aim of this section is to explain how to sample from the distribution  $P_n^f$ . We make use of the Hasting–Metropolis algorithm. We denote by  $\Delta_k = S_k - S_{k-1}, 1 \leq k \leq n,$  the  $k$ th increment of the path  $S$ . We represent a path by its initial value and its increments, so that  $S \cong (S_0, \Delta_1, \dots, \Delta_n)$  is the path defined by  $S_k = S_0 + \Delta_1 + \dots + \Delta_k$ . The algorithm runs as follows.

1. Start from any configuration  $S^0$  such that  $P_n^f(S^0) > 0$ .
2. With  $S^t \cong (S_0^t, \Delta_1^t, \dots, \Delta_n^t),$  construct  $\tilde{S}^t$  as follows.
  - (a) Choose  $k \in \{0, \dots, n\}$  uniformly.
  - (b) For  $k = 0,$  draw  $\tilde{S}_0$  according to distribution  $\pi$  and let  $\tilde{S}^t$  be the path obtained by replacing  $S_0^t$  by  $\tilde{S}_0$ .
  - (c) For  $k \geq 1,$  draw  $\tilde{\Delta}$  according to distribution  $\delta$  and let  $\tilde{S}^t$  be the path obtained by replacing  $\Delta_k^t$  by  $\tilde{\Delta}$ .

3. Compute

$$p = \frac{P_n^f(\tilde{S}^t)}{P_n^f(S^t)} = \frac{\pi(\tilde{S}_0^t)}{\pi(S_0^t)} \prod_{k=1}^n \frac{f(\tilde{S}_k^t)\delta(\tilde{S}_k^t - \tilde{S}_{k-1}^t)}{f(S_k^t)\delta(S_k^t - S_{k-1}^t)}.$$

4. Draw  $u$  randomly in  $(0, 1)$ . If  $u \leq p$  then let  $S^{t+1}$  be equal to  $\tilde{S}^t$ . Otherwise, let  $S^{t+1}$  be equal to  $S^t$ .

The sequence  $S^t$  forms an irreducible Markov chain on the state space  $\{S \in \Omega_n : P_n^f(S) > 0\}$ . This Markov chain admits the measure  $P_n^f$  as a reversible measure. Hence,  $P_n^f$  is the unique stationary distribution of the Markov chain, and  $S^t$  converges in distribution to this stationary distribution as  $t$  goes to  $\infty$ . Thus, to sample from distribution  $P_n^f,$  just run this Markov chain for a long time. What ‘a long time’ actually means is not so clear and is an active subject of research in the theory of Markov chains. There are also more complex methods based on the Propp–Wilson algorithm which could give perfect simulations. This point will not be discussed further here.

1.2.1. *Numerical simulations.* In order to illustrate the previous discussion about the heuristics of the model, we here present some simulations obtained using the Hasting–Metropolis algorithm. The settings are those described in Example 1. We focus on four cases: the linear case (with fitness function  $f_1$ ; see Figure 1), the square root case (with fitness function  $f_{1/2}$ ; see Figure 2), the overpower case (with fitness function  $f_\infty$ ; see Figure 3), and the subpower case (with fitness function  $f_0$ ; see Figure 4). In each case we plot three samples of the weighted random walk (in the left-hand graph) and the expected limit (in the right-hand graph). We draw paths with  $n = 1000$  steps. To obtain the samples, we ran the Hasting–Metropolis algorithm for  $10^5$  iterations.

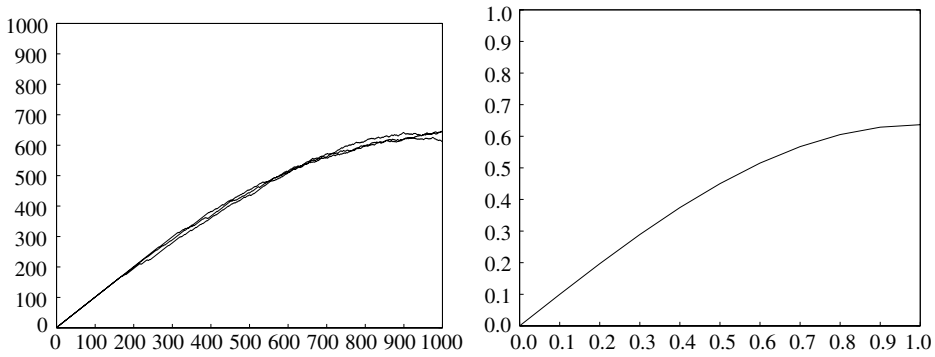


FIGURE 1: The power fitness function of type  $\beta = 1$  (fitness  $f = f_1$ ): three samples from the weighted random walk (*left*) and the deterministic limit  $\psi_1(t) = (2/\pi) \sin(\pi t/2)$  (*right*).

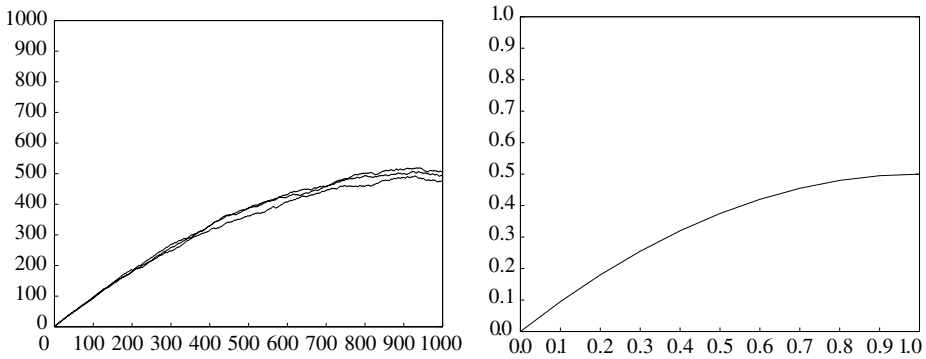


FIGURE 2: The power fitness function of type  $\beta = \frac{1}{2}$  (fitness  $f = f_{1/2}$ ): three samples from the weighted random walk (*left*) and the deterministic limit  $\psi_{1/2}(t) = t - \frac{1}{2}t^2$  (*right*).

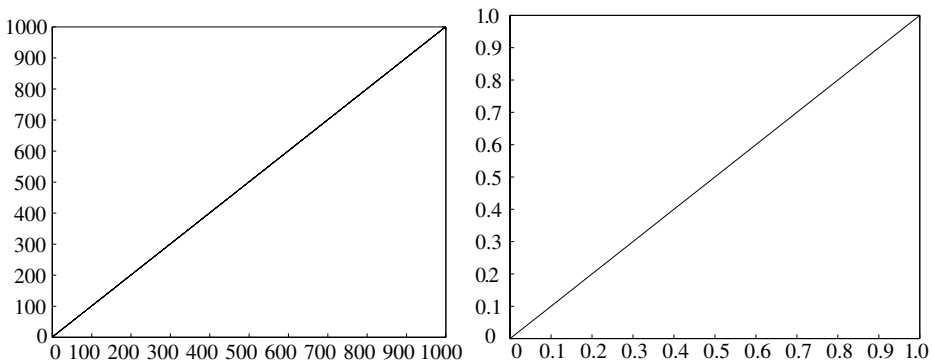


FIGURE 3: The overpower fitness function (fitness  $f = f_\infty$ ): three samples from the weighted random walk (*left*) and the deterministic limit  $\psi_\infty(t) = t$  (*right*).

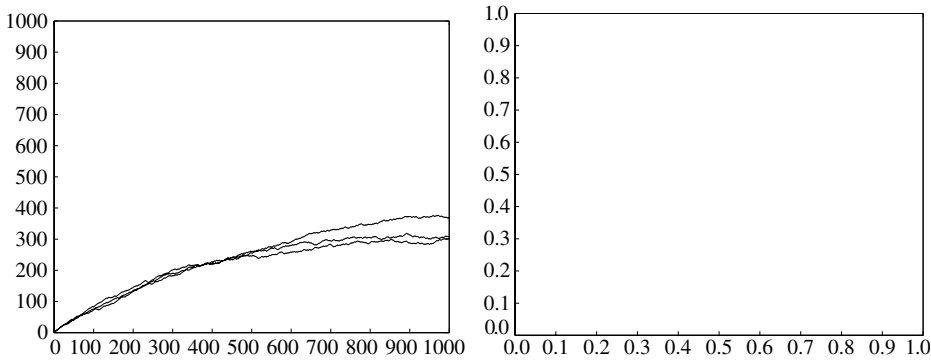


FIGURE 4: The subpower fitness function (fitness  $f = f_0$ ): three samples from the weighted random walk (left) and the deterministic limit  $\psi_0 \equiv 0$  (right).

In the case of a power fitness function (Figures 1 and 2), it seems that the three trajectories are not very far from the expected deterministic functions. However, small fluctuations appear, and it would be very nice to control them via a functional central limit theorem (like Donsker’s theorem). The author is presently working on this question. In the case of an overpower fitness function (Figure 3), it seems that the convergence is very fast: we can hardly distinguish the three paths from the deterministic function  $\psi_1$ . On the other hand, in the case of a subpower fitness function (Figure 4), the paths are very far from the expected limit function  $\psi_0$ . We believe the convergence to be very slow. A heuristic argument is that, for  $n = 1000$ ,  $\log(n)$  is equal to  $n^\beta$  with  $\beta = \log(\log(n))/\log(n) \approx 0.28$ . Hence, in the graph the regime  $\beta \approx 0$  is not reached.

**1.3. Theoretical results**

In the present work, we propose a functional convergence theorem. The proof is based on a functional large deviations principle for the weighted random walk.

Let  $\Omega$  be the space of càdlàg, real-valued functions defined on  $[0, 1]$ , endowed with the topology of the uniform norm  $\| \cdot \|_\infty$ , and let  $\mathcal{AC}$  be the subspace of absolutely continuous functions. Let  $\Psi_n : \Omega_n \rightarrow \Omega$  with  $S \mapsto \Psi_n(S)$  be the renormalization map defined by

$$\Psi_n(S) : t \mapsto \begin{cases} \frac{1}{n} S_{[nt]+1} & \text{if } 0 \leq t < 1, \\ \frac{1}{n} S_n & \text{if } t = 1. \end{cases}$$

Let  $\tilde{P}_n = P_n \circ \Psi_n^{-1}$  and  $\tilde{P}_n^f = P_n^f \circ \Psi_n^{-1}$  be the image probability measures on  $\Omega$ .

Let us state the main result of this paper. Let  $F : \Omega \rightarrow [-\infty, \infty)$  be the functional defined by

$$F(\phi) = \begin{cases} \int_0^1 \log(\phi(t)) dt & \text{if } \phi \text{ is almost everywhere positive,} \\ -\infty & \text{otherwise.} \end{cases}$$

We define the Laplace transform,  $\Lambda$ , of the increment distribution  $\delta$  by

$$\Lambda(t) = \sum_{x \in \mathbb{Z}} e^{tx} \delta(x).$$

We believe that our results hold if  $\Lambda$  is finite on a neighbourhood of 0. However, at the moment, our results are proven under the following stronger hypothesis.

**Assumption 2.**  $\delta$  has finite support.

(Note that Assumption 2 is needed only in the proof of Lemma 6, and that we are thus highly interested in a proof of this lemma in the general case.)

We denote by  $\Lambda^* : \mathbb{R} \rightarrow [0, \infty]$  the Cramer transform of the distribution  $\delta$ , defined by

$$\Lambda^*(y) = \sup_{t \in \mathbb{R}} \{ty - \log(\Lambda(t))\}.$$

According to Cramer’s theorem, under  $P_n$  the sequence  $S_n/n$  satisfies a large deviations principle with good rate function  $\Lambda^*$ . Define the functional  $I : \Omega \rightarrow [0, \infty]$  by

$$I(\phi) = \begin{cases} \int_0^1 \Lambda^*(\dot{\phi}(t)) \, dt & \text{if } \phi \in \mathcal{AC} \text{ and } \phi(0) = 0, \\ \infty & \text{otherwise.} \end{cases}$$

According to Mogulskii’s theorem, the sequence  $\tilde{P}_n = P_n \circ \Psi_n^{-1}$  satisfies a functional large deviations principle in  $\Omega$  with good rate function  $I$ .

**Theorem 1.** *Let  $f$  be a power fitness function of type  $\beta \in (0, \infty)$  and suppose that Assumptions 1 and 2 are satisfied.*

1. *The sequence of probability measures  $\tilde{P}_n^f$  satisfies the following large deviation upper bound: for any closed set  $A \subset \Omega$ ,*

$$\limsup_{n \rightarrow \infty} \frac{1}{n} \log(\tilde{P}_n^f(A)) \leq - \inf_{\phi \in A} \left\{ I - \beta F - \inf_{\phi \in \Omega} \{I - \beta F\} \right\}. \tag{1}$$

2. *The functional  $I - \beta F$  has a unique minimizer on  $\Omega$ , denoted by  $\psi_\beta$ . The function  $\psi_\beta$  is absolutely continuous, nonnegative, and concave. It vanishes at 0, is strictly positive on  $(0, 1)$ , and satisfies  $\dot{\psi}_\beta(0) > 0$ .*

3. *The sequence of probability measures  $\tilde{P}_n^f$  converges exponentially fast to  $\delta_{\psi_\beta}$ . More precisely, for all  $\varepsilon > 0$  there exists a  $\delta = \delta(\varepsilon) > 0$  such that, for all sufficiently large  $n$ ,*

$$\tilde{P}_n^f(\|\phi - \psi_\beta\|_\infty > \varepsilon) \leq e^{-n\delta}.$$

In the following proposition we characterize the minimizer  $\psi_\beta$  by the Euler–Lagrange equation associated with the minimization of the functional  $I - \beta F$ . As an application, we do exact computations in the case of the simple symmetric random walk, i.e. where  $\delta = \frac{1}{2}(\delta_1 + \delta_{-1})$ , and we make explicit the limit function  $\psi_\beta$ .

**Proposition 1.** (Identification of the minimizer  $\psi_\beta$ .) *Let  $\psi : [0, 1] \mapsto \mathbb{R}$  be a twice differentiable function such that  $\psi(0) = 0$  and, for any  $t > 0$ ,  $\psi(t) \neq 0$ . Suppose that  $\psi$  satisfies the differential equation*

$$\frac{d}{dt} [(\Lambda^*)'(\dot{\psi}(t))] + \frac{\beta}{\psi(t)} = 0 \tag{2}$$

on  $(0, 1]$ , with limit condition

$$(\Lambda^*)'(\dot{\psi}(1)) = 0. \tag{3}$$

Then  $\psi = \psi_\beta$  is the unique minimizer of the functional  $I - \beta F$ .

**Corollary 1.** (The simple symmetric random walk.) *Let  $\delta = \frac{1}{2}(\delta_1 + \delta_{-1})$  and let  $f$  be a power fitness function of type  $\beta \in (0, \infty)$ . Then*

$$\psi_\beta(t) = \frac{1}{G_\beta(1)} G_\beta^{-1}(G_\beta(1)t),$$

where

$$G_\beta(y) = \int_0^y \frac{dx}{\sqrt{1-x^{2\beta}}}.$$

For  $\beta = 1$  and  $\beta = \frac{1}{2}$ , the limit has simple expressions: respectively,

$$\psi_1(t) = \frac{2}{\pi} \sin\left(\frac{\pi}{2}t\right),$$

$$\psi_{1/2}(t) = t - \frac{1}{2}t^2.$$

**Remark 2.** Equation (2) is the Euler–Lagrange equation corresponding to the minimization of the functional  $I - \beta F$ .

#### 1.4. Infinite-population genetic algorithm

Finite-population genetic algorithms, introduced by Holland [10], are widely used in applications. They are relevant in many areas, for instance biology, computer science, optimization, etc. The task of a genetic algorithm is to search a fitness landscape for maximal values. A population of individuals, considered as candidate solutions to the given problem, is evolved under steps of mutation and steps of selection. The dynamics of the genetic algorithm simulates, supposedly like in natural systems, the survival of the fittest among the individuals.

Despite their numerous heuristic successes, mathematical results describing the behaviour of genetic algorithms are rather sparse. Among the exceptions are those in [3], [4], [13], [12], [5], [6], [7], [1], and [2].

Mathematically speaking, genetic algorithms are Markov chains on the product space  $E^p$ , where  $E$  is the state space and  $p \geq 1$  is the size of the running population. The dynamics is produced by the combination of two basic operators: mutation and selection. Mutation is driven by an ergodic Markov transition kernel  $Q(\cdot, \cdot)$  on  $E$ . Let  $x = (x_i)_{1 \leq i \leq p} \in E^p$  be a population. The effect of mutation on  $x$  is modelled by the random choice of a new population with probability

$$Q(x_1, \cdot) \otimes Q(x_2, \cdot) \otimes \cdots \otimes Q(x_p, \cdot).$$

Selection uses a function  $f \geq 0$  on  $E$ , usually called the fitness. Given a population  $x$ , consider the probability measure  $\pi_x$  on  $\{x_1, \dots, x_p\}$  defined by

$$\pi_x = \frac{1}{\langle f(x) \rangle} \sum_{i=1}^p f(x_i) \delta_{x_i},$$

where

$$\langle f(x) \rangle = \sum_{i=1}^p f(x_i).$$

The population after selection is randomly chosen in  $\{x_1, \dots, x_p\}^p$  with probability  $\pi_x^{\otimes p}$ .



Del Moral and Guionnet [5], [6] studied the infinite-population limit of mutation–selection algorithms, in connection with nonlinear filtering. In particular, those authors showed that the empirical law on  $E$  of  $x(n)$ , the population at time  $n$ , converges toward a deterministic measure  $\mu_n$  as the population size,  $p$ , goes to  $\infty$ . They also investigated the large deviations properties of this convergence. The sequence of limit measures  $\mu_n$  forms a measure-valued dynamical system:

$$\mu_{n+1} = T(\mu_n).$$

Here the operator  $T = W \circ M$  is the combination of the mutation operator,  $M$ , and the selection operator,  $W$ , with fitness  $f \geq 0$ . The mutation operator is defined by  $M: \mu \mapsto \mu Q$ , where  $Q$  is the mutation kernel. For a given measure  $\mu$  such that  $\mu(f) \in (0, \infty)$ , the selection operator  $W$  replaces  $\mu$  by  $W(\mu) = \hat{\mu}$ , such that, for any bounded function,

$$\hat{\mu}(g) = \frac{\mu(fg)}{\mu(f)}.$$

The link between genetic algorithms and weighted random walks comes from the following expression for the law  $\mu_n$ .

**Lemma 1.** *Let  $(Y_n)_{n \geq 0}$  be a Markov chain with transition kernel  $Q$  and initial probability measure  $\mu_0$ . Then, for all  $x \in E$ ,*

$$\mu_n(x) = \frac{E_{\mu_0}[\mathbf{1}_{\{Y_n=x\}} \Pi_n^f]}{E_{\mu_0}[\Pi_n^f]},$$

where

$$\Pi_n^f = \prod_{k=1}^n f(Y_k).$$

Hence, the random walk model is linked with the class of genetic algorithms in which the state space,  $E$ , is the integer set  $\mathbb{Z}$ , the mutation kernel,  $Q$ , is given by  $Q(x, y) = \delta(y - x)$ , and the initial distribution satisfies  $\mu_0 = \pi$ . In this case the measure at time  $n$  in the mutation–selection dynamics is the same as the distribution of the final value,  $S_n$ , in the model of weighted random walks of length  $n$ . In other words, the measure  $\mu_n$  is equal to the distribution of  $S_n$  under  $P_n^f$ .

Note that this class of genetic algorithms is a family of toy models used in studying the effect of stronger selection at  $\infty$  for genetic algorithms running in a noncompact space. Our study focuses on the effect of increasing fitness at  $\infty$ . The finite-population model is also used in biology to describe the evolution of populations of RNA viruses [14], [11]. Its long-time behaviour was studied by Bérard and Bienvenue [1], [2]. In [12] the authors investigated the infinite-population model in the case in which  $\beta = 1$  and proved a law of large numbers, a central limit theorem, and a large deviations principle. More precisely, let  $\mu_0 = \delta_1$  and let  $X_n$  be a random variable of law  $\mu_n$ ; then  $X_n/n$  converges to  $2/\pi$  almost surely, and the reduced random variable

$$\frac{1}{\sqrt{n}} \left( X_n - \frac{2}{\pi} n \right)$$

converges in law to a centred Gaussian law of variance  $4/\pi^2$ . The proof in [12] is based on a precise estimate of the Laplace transform of  $X_n$ , itself based on series calculus. This method does not extend to the case of nonlinear fitness. As a by-product of our results on the weighted random walks, we prove a law of large numbers valid for all power fitness functions.

**Theorem 2.** *In the mutation–selection dynamics on the integers with mutation based on the simple symmetric random walk and fitness of power type  $\beta \in (0, \infty)$ , the population at time  $n$  is located around the point  $v_\beta n$ . More precisely,  $X_n/n$  converges almost surely to  $v_\beta$ . The speed is given by*

$$v_\beta = \left( \int_0^1 \frac{dx}{\sqrt{1 - x^{2\beta}}} \right)^{-1}.$$

This method is robust enough to deal with more general mutation operators. We are interested in, and are working on, other mutation operators: mutation can be induced by more general random walks, namely dynamic random walks (see [9]). Dealing with other fitness functions would also be very interesting. Central limit theorems are expected to hold but have not been proved to the author’s knowledge.

**Remark 3.** The linear growth might come as a surprise, since the mean position of the random walk, with a constant selection, is of order  $\sqrt{n}$ . This clarifies the strong effect of stronger selection at  $\infty$ . In a loose sense, selection takes advantage of the fluctuations that the mutation operator yields.

**Remark 4.** Our method is robust enough to deal with other mutation operators. We can obtain a law of large numbers giving the speed. It would be of great interest to study how the speed depends on the mutation operator and on the fitness function. In the sequel, we give some numerical simulations exploring this direction.

**Remark 5.** As shown in the introduction, we expect the speed to equal 0 in the case of subpower fitness (for example logarithmic fitness). There might arise a new type of behaviour between the diffusive and the ballistic one, with a new rescaling order, between  $\sqrt{n}$  and  $n$ .

1.4.1. *Numerical simulations.* We use numerical simulations to evaluate the speed of evolution of the genetic algorithm. We consider the following two cases.

- The mutation operator is fixed and the selection operator is varying. We take as mutation operator  $Q(x, x + 1) = Q(x, x - 1) = \frac{1}{2}$ , i.e. a mutation corresponds to a step of a simple symmetric random walk. The fitness function is the power function  $f_\beta(x) = x^\beta$ . The parameter  $\beta$  can be seen as the strength of the selection operator. We evaluate the speed  $v_\beta$  as a function of  $\beta$ . The results appear in the left-hand graph in Figure 5. When  $\beta = 0$  there is no selection, and the speed is equal to 0. The vertical tangent at  $\beta = 0$  indicates the strength of selection, which is high even for small values of  $\beta$ . The limit  $v_\infty = 1$  shows that a strong selection forces almost every mutation step to be  $+1$ .
- The selection operator is fixed and the mutation operator is varying. We take as selection operator the selection with linear fitness function  $f_1$ . The mutation operator is

$$Q(x, x + 1) = 1 - Q(x, x - 1) = \frac{1 + \alpha}{2}$$

with  $\alpha \in (-1, 1)$ , i.e. a mutation corresponds to a step of a simple random walk with drift  $\alpha$ . The parameter  $\alpha$  can be seen as the strength of mutation. We evaluate the speed  $v_\alpha$  as a function of  $\alpha$ . The results appear in the right-hand graph in Figure 5. Note that for  $-1 < \alpha < 0$  the random walk is transient with negative speed, yet the strength of selection yields a positive speed  $v_\alpha > 0$ .

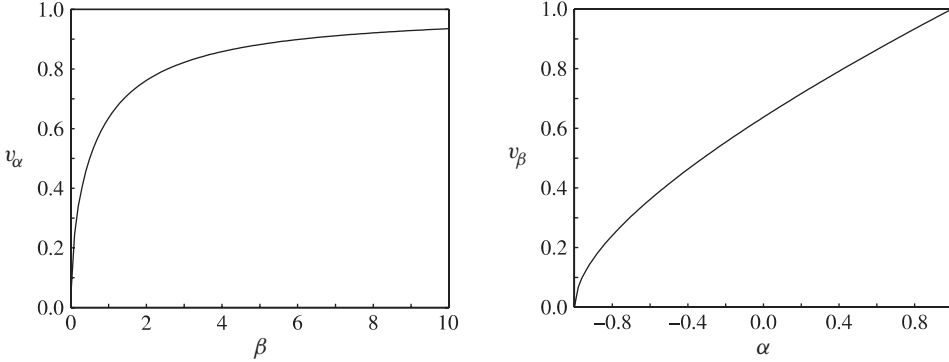


FIGURE 5: The speed  $v_\beta$  as a function of the selection strength  $\beta$  (left) and the speed  $v_\alpha$  as a function of the mutation strength  $\alpha$  (right).

**2. Proof of the theorems**

**2.1. Proof of Theorem 1**

The main result of Theorem 1 is the large deviation upper bound.

For every path  $S \in \Omega_n$ , the weight  $\Pi_n^f(S)$  can be written

$$\begin{aligned} \Pi_n^f(S) &= \prod_{i=1}^n f(S_i) = \exp\left(n \frac{1}{n} \sum_{i=1}^n \log\left(f_n\left(\frac{S_i}{n}\right)\right) + n \log(f(n))\right) \\ &= f(n)^n \exp(nF_n(\psi_n(S))) \end{aligned}$$

where  $f_n : \mathbb{R} \rightarrow \mathbb{R}$  is the function

$$f_n(x) = \frac{f([nx])}{f(n)}$$

and  $F_n : \Omega \rightarrow [-\infty, \infty)$  is the functional

$$F_n(\phi) = \int_0^1 \log(f_n(\phi(t))) dt.$$

For any Borel set  $A \subset \Omega$ , the probability  $\tilde{P}_n^f$  satisfies

$$\tilde{P}_n^f(A) = \frac{\int_A \exp(nF_n(\phi)) d\tilde{P}_n(\phi)}{\int_\Omega \exp(nF_n(\phi)) d\tilde{P}_n(\phi)}.$$

In this setting, Varadhan’s integral lemma is an adequate tool to derive the large deviations properties of the measures  $\tilde{P}_n^\beta$  from the large deviations properties of  $\tilde{P}_n$  (see [8, Section 4.3]). The difficulty comes from the irregularity of the functionals  $F_n$ , which are neither bounded nor continuous. In order to prove (1) we have to prove that, if the fitness function  $f$  is a power function of type  $\beta \in (0, \infty)$ , then both of the following inequalities hold:

$$\limsup_{n \rightarrow \infty} \frac{1}{n} \log\left(\int_A \exp(\beta n F_n(\phi)) d\tilde{P}_n(\phi)\right) \leq -\inf_{\phi \in A} \{I - \beta F\}(\phi) \tag{4}$$

for any closed set  $A \subset \Omega$  and

$$\liminf_{n \rightarrow \infty} \frac{1}{n} \log \left( \int_{\Omega} \exp(\beta n F_n(\phi)) d\tilde{P}_n(\phi) \right) \geq - \inf_{\phi \in \Omega} \{I - \beta F\}(\phi). \tag{5}$$

For the sake of clarity, we begin the proof with some preliminary lemmas. The proof of (5), which requires more technicalities, is relegated to Appendix A.

2.1.1. *Preliminaries: some properties of the functionals.* The following lemma states some convergence properties of the functions  $f_n$  and the functionals  $F_n$ .

**Lemma 2.** *Suppose that  $f$  is a fitness function of power type  $\beta \in (0, \infty)$ .*

1. *The sequence of functions  $f_n$  converges pointwise to the power function  $f_\beta$  as  $n \rightarrow \infty$ , and the convergence is uniform on the compact sets of  $\mathbb{R}$ . Furthermore, there exist real numbers  $\alpha_1$  and  $\alpha_2$  such that, for all real numbers  $t$  and eventually all integers  $n$ ,*

$$\log(f_n(t)) \leq \alpha_1 |t| + \alpha_2.$$

2. *If  $\phi_n$  is a sequence in  $\Omega$  such that  $\phi_n \rightarrow \phi$  as  $n \rightarrow \infty$ , then*

$$\limsup F_n(\phi_n) \leq \beta F(\phi).$$

*Furthermore, for all  $\phi \in \Omega$  and eventually all integers  $n$ ,*

$$F_n(\phi) \leq \alpha_1 \|\phi\| + \alpha_2.$$

*Proof.* Let  $t \in \mathbb{R}$ . If  $t \leq 0$  then, since  $f$  is nondecreasing,  $f(\lfloor nt \rfloor) \leq f(0)$ . Since  $f(n) \rightarrow \infty$  as  $n \rightarrow \infty$ , the quotient

$$f_n(t) = \frac{f(\lfloor nt \rfloor)}{f(n)} \leq \frac{f(0)}{f(n)}$$

has limit  $f_\beta(t) = 0$  as  $n \rightarrow \infty$ . The convergence is uniform on  $(-\infty, 0]$ . If  $t > 0$  then the equivalence  $f(n) \sim \ln^\beta$  implies the limit  $f_n(t) \rightarrow f_\beta(t) = t^\beta$ . More precisely, let  $\varepsilon > 0$  and  $N \in \mathbb{N}$  be such that, for every  $n \geq N$ ,

$$\left| \frac{f(n)}{\ln^\beta} - 1 \right| \leq \varepsilon.$$

If  $t \geq t_0$  then, for any  $n \geq N/t_0 + 1$ ,

$$\frac{-2\varepsilon}{1 + \varepsilon} \left( t - \frac{1}{n} \right)^\beta + \left[ \left( t - \frac{1}{n} \right)^\beta - t^\beta \right] \leq f_n(t) - t^\beta \leq \frac{2\varepsilon}{1 - \varepsilon} t^\beta.$$

This proves uniform convergence on the compact sets of  $\mathbb{R}$ .

Furthermore, for  $n \geq N$  we have the following majorizations:

$$f_n(t) \leq \begin{cases} 1 & \text{if } x \leq 0, \\ \max \left( 1, \frac{l + \varepsilon}{l - \varepsilon} t^\beta \right) & \text{if } x > 0. \end{cases}$$

Let  $\alpha_1 = \beta$  and  $\alpha_2 = \log((l + \varepsilon)/(l - \varepsilon))$ . For all real numbers  $t$  and all integers  $n \geq N$ ,

$$\log(f_n(t)) \leq \alpha_1|t| + \alpha_2.$$

Let  $\phi_n \rightarrow \phi$  in  $\Omega$ . Since uniform convergence implies pointwise convergence,  $\phi_n(t) \rightarrow \phi(t)$  and  $f_n(\phi_n(t)) \rightarrow f_\beta(\phi(t))$  for any  $t \in \mathbb{R}$ . This implies that  $\log(f_n(\phi_n(t))) \rightarrow \beta \log(\phi(t))$ , and we can apply Lebesgue's convergence theorem to the positive part (which is uniformly bounded) and Fatou's lemma to the negative part, to obtain

$$\limsup_{n \rightarrow \infty} \int_0^1 \log(f_n(\phi_n(t))) dt \leq \int_0^1 \limsup_{n \rightarrow \infty} \log(f_n(\phi_n(t))) dt = \beta \int_0^1 \log(\phi(t)) dt.$$

Furthermore, for all  $\phi \in \Omega$ , all  $t \in [0, 1]$ , and eventually all  $n$ ,

$$\log(f_n(\phi(t))) \leq \alpha_1|\phi(t)| + \alpha_2 \leq \alpha_1\|\phi\| + \alpha_2.$$

Integrating this inequality over  $[0, 1]$  yields

$$F_n(\phi) \leq \alpha_1\|\phi\| + \alpha_2.$$

The domain of a functional  $J: \Omega \rightarrow (-\infty, \infty]$  is the set where  $J < \infty$ .

**Lemma 3.** *The functionals  $I$ ,  $-\beta F$ , and  $I - \beta F$  are lower semicontinuous and strictly convex on their domains.*

*Proof.* Since the functions  $\Lambda^*$  and  $-\beta \log$  are strictly convex on their domains and lower semicontinuous, so are the functionals  $I$ ,  $-\beta F$ , and  $I - \beta F$ .

**Lemma 4.** *In any nonempty closed set  $A \subset \Omega$ , there exists a point  $\phi_0$  such that*

$$I(\phi_0) - \beta F(\phi_0) = \inf_{\phi \in A} \{I(\phi) - \beta F(\phi)\}.$$

*Proof.* The previous lemma states that  $\beta F$  is upper semicontinuous. Furthermore, the inequality  $\log(x) \leq x - 1$  implies that  $F(\phi) \leq \|\phi\| - 1$  for any  $\phi \in \Omega$ . We can thus apply [15, Lemma 4.2], which yields Lemma 4.

2.1.2. *Proof of Theorem 1, part 1: the large deviation upper bound.* We first prove (4). We apply [15, Theorem 4.1]. Let  $A \subset \Omega$  be a closed set, let  $F_{n,A}: \Omega \rightarrow [-\infty, \infty)$  be the functionals defined by

$$F_{n,A}(\phi) = \begin{cases} F_n(\phi) & \text{if } \phi \in A, \\ -\infty & \text{if } \phi \notin A, \end{cases}$$

and let

$$F_A(\phi) = \begin{cases} F(\phi) & \text{if } \phi \in A, \\ -\infty & \text{if } \phi \notin A. \end{cases}$$

By Lemma 2, if  $\phi_n \rightarrow \phi$  in  $\Omega$  then

$$\limsup_{n \rightarrow \infty} F_{n,A}(\phi_n) \leq F_A(\phi).$$

Furthermore, the following tail condition holds:

$$\lim_{M \rightarrow \infty} \limsup_{n \rightarrow \infty} \frac{1}{n} \log \left( \int_{\Omega} e^{nF_{n,A}(\phi)} \mathbf{1}_{\{F_{n,A}(\phi) \geq M\}} d\tilde{P}_n(\phi) \right) = -\infty$$

(this is proved as in [15] as a consequence of the inequality  $F_{n,A}(\phi) \leq \alpha_1 \|\phi\|_\infty + \alpha_2$ ). The large deviation upper bound holds for  $\tilde{P}_n$  with the good rate function  $I : \Omega \rightarrow [0, \infty]$ . Applying [15, Theorem 4.1] then yields (4).

As remarked above, the proof of (5) is relegated to Appendix A.

2.1.3. *Proof of Theorem 1, part 2: existence and unicity of the minimizer  $\psi_\beta$ .* The existence and unicity of the minimizer,  $\psi_\beta$ , of the functional  $I - \beta F$  on  $\Omega$  is a consequence of Lemmas 3 and 4. Its existence follows from Lemma 4 with  $A = \Omega$ , and its unicity is ensured by the strict convexity of  $I - \beta F$  stated in Lemma 3.

We investigate some properties of the minimizer  $\psi_\beta$ . The domain of  $I$  is included in the set of absolutely continuous functions vanishing at 0. The domain of  $-\beta F$  is included in the set of nonnegative functions. The function  $\psi_\beta$  belongs to the domain of the functional  $I - \beta F$ , which is included in the set of absolutely continuous, nonnegative functions vanishing at 0. Let  $\hat{\psi}_\beta$  be the concave function defined by

$$\hat{\psi}_\beta(t) = \sup \left\{ \int_E \psi_\beta(u) \, du : E \subset [0, 1], |E| = t \right\}.$$

Then  $I(\psi_\beta) = I(\hat{\psi}_\beta)$ , and  $\hat{\psi}_\beta \geq \psi_\beta$  implying that  $F(\hat{\psi}_\beta) \geq F(\psi_\beta)$ . This implies the equality  $\hat{\psi}_\beta = \psi_\beta$  and the concavity of  $\psi_\beta$ . If there exists some point  $u$  such that  $0 < u < 1$  and  $\psi_\beta(0) = \psi_\beta(u) = 0$ , then, because it is concave,  $\psi_\beta \leq 0$  must be nonpositive on  $[u, 1]$  and such that  $F(\psi_\beta) = -\infty$ . This contradicts the fact that  $\psi_\beta$  minimizes  $I - \beta F$ . In the same way, if  $\hat{\psi}_\beta(0) \leq 0$ , then the concavity implies that  $\psi_\beta \leq 0$  and  $F(\psi_\beta) = -\infty$ , also yielding a contradiction.

2.1.4. *Proof of Theorem 1, part 3: the law of large numbers for  $\tilde{P}_n^f$ .* Let us apply the large deviation upper bound to the closed set

$$A = \{\phi \in \Omega : \|\phi - \psi_\beta\|_\infty \geq \varepsilon\}.$$

This yields

$$\limsup_{n \rightarrow \infty} \frac{1}{n} \log(\tilde{P}_n^f(\|\phi - \psi_\beta\|_\infty \geq \varepsilon)) \leq -2\delta$$

with

$$2\delta = \inf_{\phi \in A} \left\{ I - \beta F - \inf_{\phi \in \Omega} \{I - \beta F\} \right\}.$$

Lemma 4 and the fact that  $\psi_\beta \notin A$  implies that  $\delta > 0$ . Hence, for  $n$  large enough,

$$\tilde{P}_n^f(\|\phi - \psi_\beta\|_\infty \geq \varepsilon) \leq e^{-n\delta}.$$

## 2.2. Proof of Proposition 1 and Corollary 1

2.2.1. *Proof of Proposition 1.* Let  $\psi \in \Omega$  be a twice differentiable function satisfying (2) and (3), and let  $\phi \in \Omega$ . We suppose that  $\psi(0) = \phi(0) = 0$  and that  $\psi(t) > 0$  for  $0 < t \leq 1$ . By the convexity of  $\Lambda^*$ ,

$$\Lambda^*(\dot{\phi}(t)) \geq \Lambda^*(\dot{\psi}(t)) + (\Lambda^*)'(\dot{\psi}(t))(\dot{\phi}(t) - \dot{\psi}(t)),$$

and the convexity of  $-\beta \log$ ,

$$-\beta \log(\phi(t)) \geq -\beta \log(\psi(t)) - \frac{\beta}{\psi(t)}(\phi(t) - \psi(t)).$$

Adding the preceding two inequalities and integrating over  $(0, 1]$  yields

$$(I - \beta F)(\phi) \geq (I - \beta F)(\psi) + \int_0^1 \left[ (\Lambda^*)'(\dot{\psi}(t))(\dot{\phi}(t) - \dot{\psi}(t)) - \frac{\beta}{\psi(t)}(\phi(t) - \psi(t)) \right] dt.$$

Using an integration by parts, we obtain

$$(I - \beta F)(\phi) - (I - \beta F)(\psi) \geq [(\Lambda^*)'(\dot{\psi}(t))(\phi(t) - \psi(t))]_0^1 - \int_0^1 \left( \frac{d}{dt} [(\Lambda^*)'(\dot{\psi}(t))] + \frac{\beta}{\psi(t)} \right) (\phi(t) - \psi(t)) dt.$$

By (2), the integral vanishes. By (3), the bracket also vanishes. Hence,

$$(I - \beta F)(\phi) \geq (I - \beta F)(\psi)$$

and  $\psi$  minimizes  $I - \beta F$ .

2.2.2. *Proof of Corollary 1.* After some computations, we can prove that (2) is equivalent to

$$\frac{1}{1 - \dot{\psi}(t)^2} \ddot{\psi}(t) + \frac{\beta}{\psi(t)} = 0.$$

Let us define the function

$$\psi(t) = \frac{1}{G_\beta(1)} G_\beta^{-1}(G_\beta(1)t), \quad t \in [0, 1],$$

where

$$G_\beta(y) = \int_0^y \frac{dx}{\sqrt{1 - x^{2\beta}}}.$$

We now apply Proposition 1 and prove that  $\psi$  minimizes  $I - \beta F$ . Let  $p = G_\beta(1)$ . The relevant derivatives of  $\psi$  are

$$\begin{aligned} \dot{\psi}(t) &= \sqrt{1 - (p\psi(t))^{2\beta}}, \\ \ddot{\psi}(t) &= -\beta p (p\psi(t))^{2\beta-1}. \end{aligned}$$

Using these expressions, we can show that (2) is satisfied.

Equation (3) holds since  $\dot{\psi}(1) = 0$  and  $(\Lambda^*)'(0) = 0$ .

**Remark 6.** Note that any function of the form  $\phi(t) = (1/p)G_\beta^{-1}(pt)$  is a solution to the Euler–Lagrange equation (2) and vanishes at 0. The solution with  $p = G_\beta(1)$  is the one satisfying (3).

### 2.3. Proof of Theorem 2

Theorem 2 is a direct consequence of Theorem 1. By the contraction principle, a large deviation upper bound for the sequence  $\mu_n$  can be deduced from the large deviation upper bound for  $\tilde{P}_{\mu_0, n}$ . Alternatively, and more directly,

$$\mu_n \left( \left| \frac{t}{n} - \psi_\beta(1) \right| > \varepsilon \right) = \mathbf{P}_{\mu_0, n}^\beta \left( \left| \frac{S_n}{n} - \psi_\beta(1) \right| > \varepsilon \right) = \tilde{\mathbf{P}}_{\mu_0, n}^\beta (|\phi(1) - \psi_\beta(1)| > \varepsilon).$$

Hence, by Theorem 1 there exists a  $\delta = \delta(\varepsilon) > 0$  such that, for  $n$  large enough,

$$\mu_n \left( \left| \frac{t}{n} - \psi_\beta(1) \right| > \varepsilon \right) \leq \tilde{P}_{\mu_0, n}^\beta (\|\phi - \psi_\beta\| > \varepsilon) \leq e^{-n\delta}.$$

Let  $X_n$  be a random variable of law  $\mu_n$ . The previous inequality implies that  $X_n/n$  converges almost surely to  $v_\beta$ . The speed  $v_\beta = \psi_\beta(1)$  is given explicitly by

$$v_\beta = \left( \int_0^1 \frac{dx}{\sqrt{1-x^{2\beta}}} \right)^{-1}.$$

**Appendix A. Proof of (5)**

The function  $F$  is not lower semicontinuous, so we are not able to use [8, Lemma 4.3.4] to establish the lower bound.

*Proof of (5).* Let  $B_\varepsilon = \{\phi \in \Omega : \phi > \psi_\beta \text{ on } [\varepsilon, 1]\}$ . Then

$$\begin{aligned} \tilde{E}_n[\exp(\beta n F_n)] &= \int_\Omega \exp(n F_n(\phi)) d\tilde{P}_n(\phi) \\ &\geq \int_{B_\varepsilon} \exp(n F_n(\phi)) d\tilde{P}_n(\phi) \\ &= \int_{B_\varepsilon} \exp\left(n \int_0^1 \log(f_n(\phi(t))) dt\right) d\tilde{P}_n(\phi) \\ &\geq \exp\left(n \int_\varepsilon^1 \log(f_n(\psi_\beta(t))) dt\right) \\ &\quad \times \int_{B_\varepsilon} \exp\left(n \int_0^\varepsilon \log(f_n(\phi(t))) dt\right) d\tilde{P}_n(\phi). \end{aligned} \tag{6}$$

In the last integral, let us condition on  $\phi(\varepsilon)$  and apply the Markov property. This yields

$$\int_{B_\varepsilon} \exp\left(n \int_0^\varepsilon \log(f_n(\phi(t))) dt\right) d\tilde{P}_n = \tilde{E}_n[Y_1 Y_2], \tag{7}$$

where

$$\begin{aligned} Y_1 &= \tilde{E}_n[\mathbf{1}_{B_\varepsilon} | \phi(\varepsilon)], \\ Y_2 &= \tilde{E}_n \left[ \exp\left(n \int_0^\varepsilon \log(f_n(\phi(t))) dt\right) \middle| \phi(\varepsilon) \right]. \end{aligned}$$

If  $\phi(\varepsilon) \leq \psi_\beta(\varepsilon)$  then  $Y_1 = 0$ . Let us define

$$K_n^\varepsilon = \inf_{y \geq \psi_\beta(\varepsilon)} \tilde{E}_n \left[ \exp\left(n \int_0^\varepsilon \log(f_n(\phi(t))) dt\right) \middle| \phi(\varepsilon) = y \right].$$

Then  $Y_1 Y_2 \geq Y_1 K_n^\varepsilon$  and

$$\tilde{E}_n[Y_1 Y_2] \geq K_n^\varepsilon \tilde{E}_n[Y_1] = K_n^\varepsilon \tilde{P}_n(B_\varepsilon). \tag{8}$$



Equations (6), (7), and (8) together yield

$$\begin{aligned} \liminf_{n \rightarrow \infty} \frac{1}{n} \log(\tilde{E}_n[\exp(nF_n)]) &= \liminf_{n \rightarrow \infty} \frac{1}{n} \log\left(\int_{\Omega} \exp(nF_n(\phi)) \, d\tilde{P}_n(\phi)\right) \\ &\geq \liminf_{n \rightarrow \infty} \int_{\varepsilon}^1 \log(f_n(\psi_{\beta}(t))) \, dt + \liminf_{n \rightarrow \infty} \frac{1}{n} \log(\tilde{P}_n(B_{\varepsilon})) \\ &\quad + \liminf_{n \rightarrow \infty} \frac{1}{n} \log(K_n^{\varepsilon}). \end{aligned} \tag{9}$$

The function  $\psi_{\beta}$  is positive on  $[\varepsilon, 1]$ . Lemma 2 implies that  $\log(f_n(\psi_{\beta}(t)))$  converges to  $\beta \log(\psi_{\beta}(t))$  uniformly on  $[\varepsilon, 1]$ . Hence,

$$\lim_{n \rightarrow \infty} \int_{\varepsilon}^1 \log(f_n(\psi_{\beta}(t))) \, dt = \beta \int_{\varepsilon}^1 \log(\psi_{\beta}(t)) \, dt.$$

The set  $B_{\varepsilon}$  is open, so the large deviation lower bound and Lemma 5, below, imply that

$$\liminf_{n \rightarrow \infty} \frac{1}{n} \log(\tilde{P}_n(B_{\varepsilon})) \geq -\inf_{\phi \in B_{\varepsilon}} I(\phi) \geq -I(\psi_{\beta}).$$

Letting  $\varepsilon \rightarrow 0$ , (9) yields

$$\liminf_{n \rightarrow \infty} \frac{1}{n} \log(\tilde{E}_n[\exp(nF_n)]) \geq \beta F(\psi_{\beta}) - I(\psi_{\beta}) + \limsup_{\varepsilon \rightarrow 0} \liminf_{n \rightarrow \infty} \frac{1}{n} \log(K_n^{\varepsilon}).$$

By Lemma 6, below, the term in  $K_n^{\varepsilon}$  vanishes. Hence, (5) holds.

**Lemma 5.** *The open set  $B_{\varepsilon} = \{\phi \in \Omega : \phi > \psi_{\beta} \text{ on } [\varepsilon, 1]\}$  is such that*

$$\inf_{\phi \in B_{\varepsilon}} I(\phi) \leq I(\psi_{\beta}).$$

*Proof.* For any  $\delta \in (0, 1)$ , consider the function  $\phi_{\delta}$  defined by

$$\phi_{\delta}(t) = \frac{1}{\delta} \psi_{\beta}(\delta t).$$

Since  $\psi_{\beta}$  is strictly concave on  $(0, 1)$ , the function  $\phi_{\delta}$  belongs to the set  $B_{\varepsilon}$  (since  $\phi_{\delta}(0) = \psi_{\beta}(0) = 0$  and  $\phi_{\delta}(t) = \psi_{\beta}(\delta t) > \psi_{\beta}(t)$  for  $t > 0$ ). Furthermore, as  $\delta \rightarrow 1$ ,  $I(\phi_{\delta}) \rightarrow I(\psi_{\beta})$ . This proves Lemma 5.

**Lemma 6.** *The quantity*

$$K_n^{\varepsilon} = \inf_{y \geq \psi_{\beta}(\varepsilon)} \tilde{E}_n \left[ \exp\left(n \int_0^{\varepsilon} \log(f_n(\phi(t))) \, dt\right) \mid \phi(\varepsilon) = y \right]$$

*satisfies*

$$\limsup_{\varepsilon \rightarrow 0} \liminf_{n \rightarrow \infty} \frac{1}{n} \log(K_n^{\varepsilon}) \geq 0.$$

*Proof.* Let  $y \geq \psi_\beta(\varepsilon)$  be such that  $\tilde{P}_n(\phi(\varepsilon) = y) > 0$ . Let us define

$$K_n^\varepsilon(y) = \tilde{E}_n \left[ \exp \left( n \int_0^\varepsilon \log(f_n(\phi(t))) dt \right) \middle| \phi(\varepsilon) = y \right].$$

Since we are dealing with discrete probabilities, this conditional expectation can be written

$$K_n^\varepsilon(y) = \frac{1}{\tilde{P}_n(\phi(\varepsilon) = y)} \sum_{\phi} \exp \left( n \int_0^\varepsilon \log(f_n(\phi(t))) dt \right) \tilde{P}_n(\phi),$$

where the sum is taken over the finite set of functions  $\phi \in \Omega$  such that  $\phi(\varepsilon) = y$  and  $\tilde{P}_n(\phi) > 0$ . The sum is underestimated by the contribution of the paths  $\phi \in U$ , where  $U$  is defined by

$$U = \left\{ \phi \in \Omega : \phi(t) \geq \frac{y}{\varepsilon}t \text{ on } [0, \varepsilon], \phi(\varepsilon) = y \right\}.$$

For all paths  $\phi \in U$ ,

$$\begin{aligned} \exp \left( n \int_0^\varepsilon \log(f_n(\phi(t))) dt \right) &\geq \exp \left( n \int_0^\varepsilon \log \left( f_n \left( \frac{y}{\varepsilon}t \right) \right) dt \right) \\ &\geq \exp \left( n \int_0^\varepsilon \log \left( f_n \left( \frac{\psi_\beta(\varepsilon)}{\varepsilon}t \right) \right) dt \right). \end{aligned}$$

The probability of the set  $U$  can be underestimated by

$$\tilde{P}_n(U) \geq uv^{[n\varepsilon]},$$

where  $u = \min(\pi(k)) > 0$  for  $k$  in the support of  $\pi$  and  $v = \min(\delta(k)) > 0$  for  $k$  in the support of  $\delta$ . The probability  $\tilde{P}_n(\phi(\varepsilon) = y)$  is less than 1. This yields the underestimation

$$\frac{1}{n} \log(K_n^\varepsilon) \geq \frac{1}{n} \log(uv^{[n\varepsilon]}) + \int_0^\varepsilon \log \left( f_n \left( \frac{\psi_\beta(\varepsilon)}{\varepsilon}t \right) \right) dt,$$

and the lemma follows.

### Acknowledgement

We gratefully acknowledge the anonymous referee for useful remarks and advice that greatly improved this article.

### References

- [1] BÉRARD, J. AND BIENVENÛE, A. (2000). Convergence of a genetic algorithm with finite population. In *Mathematics and Computer Science (Versailles, 2000)*, Birkhäuser, Basel, pp. 155–163.
- [2] BÉRARD, J. AND BIENVENÛE, A. (2003). Sharp asymptotic results for simplified mutation-selection algorithms. *Ann. Appl. Prob.* **13**, 1534–1568.
- [3] CERF, R. (1996). The dynamics of mutation-selection algorithms with large population sizes. *Ann. Inst. H. Poincaré Prob. Statist.* **32**, 455–508.
- [4] CERF, R. (1998). Asymptotic convergence of genetic algorithms. *Adv. Appl. Prob.* **30**, 521–550.
- [5] DEL MORAL, P. AND GUIONNET, A. (1998). Large deviations for interacting particle systems: applications to non-linear filtering. *Stoch. Process. Appl.* **78**, 69–95.
- [6] DEL MORAL, P. AND GUIONNET, A. (1999). On the stability of measure valued processes with applications to filtering. *C. R. Acad. Sci. Paris Sér. I Math.* **329**, 429–434.

- [7] DEL MORAL, P. AND GUIONNET, A. (2001). On the stability of interacting processes with applications to filtering and genetic algorithms. *Ann. Inst. H. Poincaré Prob. Statist.* **37**, 155–194.
- [8] DEMBO, A. AND ZEITOUNI, O. (1998). *Large Deviations Techniques and Applications (Appl. Math. New York 38)*, 2nd edn. Springer, New York.
- [9] GUILLOTIN-PLANTARD, N., PINÇON, B., DOMBRY, C. AND SCHOTT, R. (2006). Data structures with dynamical random transistions. *Random Structures Algorithms* **28**, 403–426.
- [10] HOLLAND, J. H. (1975). *Adaptation in Natural and Artificial Systems*. University of Michigan Press, Ann Arbor.
- [11] KESSLER, D. A., LEVINE, H., RIGDWAY, D. AND TSIMRING, L. (1997). Evolution on a smooth landscape. *J. Statist. Phys.* **87**, 519–544.
- [12] MAZZA, C. AND PIAU, D. (2001). On the effect of selection in genetic algorithms. *Random Structures Algorithms* **18**, 185–200.
- [13] RABINOVICH, Y. AND WIGDERSON, A. (1999). Techniques for bounding the convergence rate of genetic algorithms. *Random Structures Algorithms* **14**, 111–138.
- [14] TSIMRING, L. S., LEVINE, H. AND KESSLER, D. A. (1996). RNA virus evolution via fitness-space model. *Phys. Rev. Lett.* **76**, 4440–4443.
- [15] VARADHAN, S. R. S. (1996). Asymptotic probabilities and differential equations. *Commun. Pure Appl. Math.* **19**, 261–286.