

Adaptive Landscape with Singularity in Evolutionary Processes

Shuyun Jiao, Song Xu, and Ping Ao

Abstract Adaptive landscape, proposed by Sewall Wright, has been used to find optimized solutions of a system. The optimized solution of an evolutionary system is when evolution maximizes or minimizes the value of some function of the trait under consideration, thus providing an absolute measure of fixation for a biological process in a probabilistic sense. We survey the role of adaptive landscape and give some general results concerning the question of infinite potential escaping. The results presented include complex dynamical behaviors manifested by adaptive landscape with singularity in all parameters regimes. In addition, both metaphoric and quantitative description of many complex biological phenomena is provided by adaptive landscape, such as the rare event of transition between different stable states.

1 Introduction

The concept of adaptive landscape was first proposed by Sewall Wright [1], in attempting to visualize his shifting balance theory in evolutionary dynamics. Since then, the metaphoric and visualizing part of this concept has been widely used in population genetics and evolutionary biology [2–4]. Wright’s original landscape

S. Jiao

Shanghai Center for Systems Biomedicine, Department of Mathematics,
Shanghai Jiao Tong University, Shanghai, China

Xinyang Normal University, Henan, China
e-mail: jiaoshuyun@gmail.com

S. Xu

Department of Biomathematics, University of California, Los Angeles, USA
e-mail: sxu11@ucla.edu

P. Ao (✉)

Shanghai Center for Systems Biomedicine, Shanghai Jiao Tong University, Shanghai, China
e-mail: aoping@sjtu.edu.cn; aoping@u.washington.edu

can be interpreted as the surface of selective fitness [5]. Its definition has two main derivatives [6]: project the individual fitness over space of allele type or genotype or phenotype, or project the mean fitness of a population over the frequencies of allele type or genotype or phenotype. The primary goal of adaptive landscape in modern science is understanding, i.e., creating orderly pictures of the world compatible with our observations and can provide new predictions.

Evolutionary processes are highly nonlinear. We need to understand these processes. Adaptive landscape, when used in visualizing evolutionary processes, is remarkably suitable grasping the nature of nonlinearity conveniently.

The appeal of this approach is its analogy with a physical landscape, whose gradient predicts a rolling marbles spatial trajectory [7]. But it is complex that required to fully integrate the genetics with Darwinian natural selection. Geometric representations of microevolutionary trajectories over adaptive landscape of one or another are motivated by the desire to provide some heuristic intuition into the process of microevolution. If we can extract the population genetics by investigating the contours of such a topographic surface, we can visualize how an evolving population will behave without always resorting to the algebraic heavy lifting. So the notion of a landscape implies the existence of an exact potential function projected over some space, its predictive utility comes from the implication that at each instant, the system will shift its configuration by following the steepest gradient on the potential function [7]. The predictive capability is the basis for the hypothetico-deductive method [8].

Though the landscape has such good predictive property, it has suffered certain conceptual and theoretical problems [9]. Biologically, there is an ongoing argument about the heuristic value of Wright's landscape diagrams [10]. Some think there might not be anything like peaks separated by valleys at all [10], the problem of how a population crosses an adaptive valley in its way from one adaptive peak to another may be nonexistent [11]. Since then such a landscape has been known as the fitness landscape in some parts of literatures. However, there are a considerable amount of confusions on the definitions of fitness [12, 13]. Mathematically, Pigliucci and Kaplan argue that in the end it may be impossible to articulate the metaphor in a way that is both conceptually coherent and practically fruitful [14].

To get further at the predictive side we need a more realistic handle on adaptive landscape. Biologists [9, 15] noticed a framework was needed to construct adaptive landscape. Recently considerable progress has been made on the evolutionary dynamics of transiting the fitness valley for finite populations in the presence of diverse evolutionary forces. de Vladar and Barton [15] used information entropy to analyze the process of population evolution. It is pointed out that the method works only for high mutation rates and breaks down for low mutation rates [15]. One of us proposed adaptive landscape can be quantified as potential function [13]. He identified gradual parameter changes that preserve the stationary distribution as being reversible in the thermodynamic sense, this line of thinking does not lead to any constraint on the increase in mean fitness that would correspond to the constraint identified by Carnot in classical thermodynamics [16]. These theoretical progress makes the utility of adaptive landscape to complex biological processes and understanding these processes.

In this chapter we address how the singular adaptive landscape indicates the evolutionary dynamics and predicts how long the population stays at an adaptive state. There are two reasons for embarking on such an effort. The first one is practical. We want to manifest the reach of conceptually coherent and practically fruitful adaptive landscape. In general, adaptive landscape appears working well, notwithstanding its basically singularity. The next reason comes into play when we want to know the escape problems from the infinite potential.

2 Technical Preliminaries

The tools for handling a dynamical system are diverse. Here we only base on Fokker–Planck equation and some relevant knowledge about diffusion approximation, which is a widely used theory; however, there are some shortcomings about diffusion theory such as addresses by [17].

2.1 Diffusion Approximation

In general, we first consider the deterministic effects of certain biological factors for an evolutionary process. We expect to get some useful information for an evolutionary process with infinite population size. But randomness always exists, and population size is also finite, these result in the information lacked. We have to consider the effect of stochastic factors. We briefly outline the diffusion approximation according to [18] in the following.

We only address one-dimensional process. That is, we only consider populations with haploid individuals and one locus with two allele A and a and focus on the evolution of allele A . At generation t the frequency of allele A is i/N , after driving by deterministic and stochastic factors, at generation $t + 1$ the allele frequency becomes j/N . Here the probability that allele frequency becomes j/N is

$$P_j(t + 1) = \sum_{i=0}^N W_{ij} P_i(t) \quad (1)$$

where W_{ij} is the transition probability from i/N to j/N . We replace the allele frequency i/N by real number x , $0 \leq x \leq 1$. And $P(x, t + 1)$ denotes the probability of allele frequency x after $t + 1$ generations. Given that A starts out at gene frequency x_0 and additional Markov property, then

$$P(x, t + \delta t) = \sum_{\delta x} W(x, t + \delta t | x - \delta x, t) P(x - \delta x, t). \quad (2)$$

We assume that $P(x, t)$ admits a density function $\rho(x, t)$, that is,

$$\frac{\partial P(x, t)}{\partial x} = \rho(x, t). \tag{3}$$

In order to derive forward Fokker–Planck equation, we assume the following properties hold.

1)

$$\lim_{\delta t \rightarrow 0} \frac{1}{\delta t} p(z, t + \delta t | x, t) = 0 \tag{4}$$

uniformly in x, z , and t for $|x - z| \geq \varepsilon$.

2)

$$\lim_{\delta t \rightarrow 0} \frac{1}{\delta t} \int dz (z - x) p(z, t + \delta t | x, t) = M(x, t) + o(\varepsilon) \tag{5}$$

3)

$$\lim_{\delta t \rightarrow 0} \frac{1}{\delta t} \int dz (z - x)^2 p(z, t + \delta t | x, t) = V(x, t) + o(\varepsilon) \tag{6}$$

Equations (5) and (6) are uniform in x, z , and t .

4) All higher-order coefficients are $o(\varepsilon)$.

Consider the time evolution of the expectation of a function $f(z)$ which is twice continuously differentiable. Thus,

$$\partial_t \int dx f(x) p(x, t | y, t') \tag{7}$$

$$= \lim_{\delta t \rightarrow 0} \frac{1}{\delta t} \left\{ \int dx f(x) [p(x, t + \delta t | y, t') - p(x, t | y, t')] \right\} \tag{8}$$

$$= \lim_{\delta t \rightarrow 0} \frac{1}{\delta t} \left\{ \int dx \int dz f(x) p(x, t + \delta t | z, t) p(z, t | y, t') - \int dz f(z) p(z, t | y, t') \right\} \tag{9}$$

Expand $f(x)$ at position z :

$$f(x) = f(z) + f'(z)(x - z) + \frac{1}{2} f''(z)(x - z)^2 + o((x - z)^2) \tag{10}$$

Now substitute Eq. (10) with Eq. (9):

$$\begin{aligned}
& \lim_{\delta t \rightarrow 0} \frac{1}{\delta t} \left\{ \int dx dz \left[f(z) + f'(z)(x-z) + \frac{1}{2} f''(z)(x-z)^2 \right] \right. \\
& \quad \left. \times p(x, t + \delta t | z, t) p(z, t | y, t') - \int dz f(z) p(z, t | y, t') \right\} \\
&= \lim_{\delta t \rightarrow 0} \frac{1}{\delta t} \left\{ \int dx dz \left[f(z) + f'(z)(x-z) + \frac{1}{2} f''(z)(x-z)^2 \right] \right. \\
& \quad \left. \times p(x, t + \delta t | z, t) p(z, t | y, t') - \int dx dz f(z) p(x, t + \delta t | z, t) p(z, t | y, t') \right\} \\
&= \int dz \left\{ \lim_{\delta t \rightarrow 0} \frac{1}{\delta t} \int dx \left[f'(z)(x-z) + \frac{1}{2} f''(z)(x-z)^2 \right] \right. \\
& \quad \left. \times p(x, t + \delta t | z, t) \right\} p(z, t | y, t') \\
&= \int dz \left[f'(z) M(z, t) + \frac{1}{2} f''(z) V(z, t) \right] p(z, t | y, t') \\
&= \int dz f(z) \left\{ -\partial_z [M(z, t) p(z, t | y, t')] + \frac{1}{2} \partial_{zz} [V(z, t) p(z, t | y, t')] \right\} \\
&\stackrel{z \rightarrow x}{=} \int dx f(x) \left\{ -\frac{\partial}{\partial x} [M(x, t) p(x, t | y, t')] + \frac{1}{2} \partial_{xx} [V(x, t) p(x, t | y, t')] \right\}. \quad (11)
\end{aligned}$$

Then, we obtain the Kolmogorov forward equation (Fokker–Planck equation)

$$\partial_t p(x, t | y, t') = -\frac{\partial}{\partial x} [M(x, t) p(x, t | y, t')] + \frac{1}{2} \frac{\partial^2}{\partial x^2} [V(x, t) p(x, t | y, t')]. \quad (12)$$

Define $M(x)$ as the probability that x increases by systematic force that includes mutation and selection. And define $V(x)$ as the probability that x changes because of random drift, either decreasing by amount δx with the probability $V(x)/2$ or increasing by the amount δx with the probability $V(x)/2$.

$$M(x) = \lim_{\delta t \rightarrow 0} \frac{1}{\delta t} \int_0^1 \delta x W(x, t + \delta t | x - \delta x, t) d(\delta x), \quad (13)$$

$$V(x) = \lim_{\delta t \rightarrow 0} \frac{1}{\delta t} \int_0^1 (\delta x)^2 W(x, t + \delta t | x - \delta x, t) d(\delta x). \quad (14)$$

2.2 Relation for the Solutions of Between Forward Fokker–Planck Equation and Backward Fokker–Planck Equation

In most physical applications of Kolmogorov diffusion equations the coefficient $V(x)$ is essentially positive. However, in certain applications we encounter equations such that $V(x)$ vanishes at one (or possibly both) of the boundaries or one of the coefficients has no finite limit. Equations with coefficients such as described are called singular diffusion equations.

The classification of boundaries depends on the Lebesgue integrability of the function

$$g_1(x) = \exp \left\{ - \int_{x'}^x \frac{M(z)}{V(z)} dz \right\} \quad (15)$$

where $x' \in (r_1, r_2)$ is fixed, and related functions on a prescribed open interval contained in (r_1, r_2) . The above function was introduced by Feller [19]. Let I_i ($i = 1, 2$) denote the interval (x', r_i) . The function $g_1(x)$ is Lebesgue integrable on I_i (written $g_1(x) \in L(I_i)$) if

$$\int_{I_i} g_1(x) dx < \infty \quad (16)$$

i.e., the integral of $g_1(x)$ over the interval I_i is bounded. Before giving the classification criteria, we introduce the following functions:

$$g_2(x) = \frac{1}{V(x)g_1(x)} \quad h(x) = g_1(x) \int_{x'}^x g_2(z) dz \quad (17)$$

Then Feller classifies the boundaries as the following

1. The boundary r_i is *regular* if $g_1(x) \in L(I_i)$ and $g_2(x) \in L(I_i)$.
2. The boundary r_i is an *exit* boundary if $g_2(x) \notin L(I_i)$ and $h(x) \in L(I_i)$.
3. The boundary r_i is an *entrance* boundary if $g_2(x) \in L(I_i)$ and $g_2(x) \int_{x'}^x g_1(z) dz \in L(I_i)$.
4. In all other cases the boundary is called *natural*.

By utilizing the theory of semigroups, Feller has obtained the following results which relate the existence and uniqueness problem of diffusion equations to that of classifying the boundaries:

1. When none of the boundaries is regular, there exists exactly one fundamental solution (or Green's function) common to the forward and backward diffusion equations, even though the initial value problem as such may have many solutions.

2. When one boundary is regular, or when both boundaries are regular, there exist infinitely many common fundamental solutions.
3. When both boundaries are natural, the initial value problem for both forward and backward diffusion equations is uniquely determined, and the solutions are generated by a common fundamental solution.
4. When r_1 is a natural boundary and r_2 is an exit boundary, the initial value problem for the backward diffusion equation has infinitely many solutions, but that for the forward diffusion equations is uniquely determined.
5. When r_1 is a natural boundary but r_2 is a regular boundary, there exist infinitely many solutions for the initial value problem for both the forward and backward diffusion equations.
6. When neither boundary is natural, there are two sources for nonuniqueness, and in these cases two lateral conditions must be imposed.

3 Construction of Adaptive Landscape

Diffusion theory is an elegant approximation for analyzing population evolution. To construct adaptive landscape, we start from one-dimensional forward Fokker-Planck equation:

$$\frac{\partial}{\partial t} \rho(x, t) = -\frac{\partial}{\partial x} [M(x)\rho(x, t)] + \frac{1}{2} \frac{\partial^2}{\partial x^2} [V(x)\rho(x, t)]. \quad (18)$$

Among this $M(x)$ is the symbol for the average change in allele frequency [20, 21] that occurs due to systematic force. The function $V(x)$ is the average square change in allele frequency.

The diffusion process can also be expressed by the following symmetric equation

$$\partial_t \rho(x, t) = \partial_x [\epsilon D(x) \partial_x - f(x)] \rho(x, t) \quad (19)$$

with

$$\begin{aligned} f(x) &= M(x) - \epsilon D'(x), \\ 2\epsilon D(x) &= V(x). \end{aligned} \quad (20)$$

With a prime denoting differentiation of a function with respect to its argument such as $D'(x) \equiv \partial_x D(x)$ where $M(x)$ and $V(x)$ is from Eq. (18). Adaptive landscape is directly given when we consider natural boundary as Feller's classification. It is

$$\Phi(x) = \int^x dx' \frac{f(x')}{D(x')}. \quad (21)$$

The symmetric Eq.(19) has two advantages. On the one hand, the adaptive landscape is directly read out when the detailed balance is satisfied. On the other hand, the constructive method is dynamical, independent of existence and normalization of stationary distribution. We call $f(x)$ directional transition rate, integrating the effects of $M(x)$ and the derivative of $V(x)$. Directional transition rate can give equilibrium states when it appears in linear form.

The adaptive landscape not only avoids the mismatch of the fixed points of the force and those of extremals of steady state distribution, but also can be related to the dynamical behaviors.

The stationary distribution for the diffusion approximation satisfying natural boundary condition is given by

$$\rho(x, t = \infty) = \frac{1}{Z} \exp\left(\frac{\Phi(x)}{\epsilon}\right). \quad (22)$$

$$Z = \int_{-\infty}^{+\infty} dx \exp\left(\frac{\Phi(x)}{\epsilon}\right). \quad (23)$$

It has the form of Boltzmann–Gibbs distribution [12], so the scalar function $\Phi(x)$ naturally acquires the meaning of potential energy [22]. The value of Z determines the normalization of $\rho(x, t = \infty)$ from the perspective of probability, and the finite value of Z manifests the normalization of $\rho(x, t = \infty)$. The stationary distribution is not true in the face of infinite Z . It demonstrates absorbing phenomenon occurs at the boundary. Together with the flux at the boundary, the true stationary distribution could be got. The constant ϵ holds the same position as temperature of Boltzmann–Gibbs distribution in statistical mechanics. But it does not hold the nature of temperature in Boltzmann–Gibbs distribution.

4 Two Applications

As the concept and their technical implementation described, below we shall show the singular adaptive landscape describes internal equilibria for constant population size and variable biological factors such as original Wright–Fisher process and Muller ratchet.

4.1 Wright–Fisher Process

The simplest stochastic model of an evolving population dates from the 1930s and was introduced independently by Ronald Fisher [23] and Sewall Wright [24]. In the model, constant organisms from one generation to the next generation are considered, and each instance of a gene in one generation is an exact copy of one

randomly chosen with replacement from previous generation [25]. Let us consider first of all diploid organisms which carry two copies of each gene and reproduce sexually. A new generation could be formed by each of the organisms producing two infinite sets of gametes before dying. Among this process mutation, selection and random drift could occur. Then random fixed size samples of these gametes survive to become offspring organisms. A random mating diploid population evolving in this way is often referred to as an ideal population. In reality, individuals do not mate at random. Some factors such as individuals' geographical locations and age may also lead to deviations from an ideal population. Nevertheless, in some cases it is found that the predictions of the ideal model are relevant [25].

Suppose we have an ideal population of individuals satisfying (i) Generations are taken to be discrete, so that the population evolves by a discrete-step Markov process. (ii) The population size is taken to be fixed, so that alleles compete only against other alleles and not against an external environment. (iii) Random mating is assumed. Concretely we consider a population of diploid sexual individuals with population size being N and factors such as one locus with two alleles A_1, A_2 . The start point in a generation is taken to be the adult stage, after all mutation and selection has occurred and selection immediately prior to reproduction. The regulation of the population number is supposed to occur through nonoverlapping generations and randomly picking N individuals from the population, then there are always $2N$ alleles in the allele pool in any generation. The treatment is consistent with [26].

We focus on the evolution of allele A_1 . Assume p_i to represent the proportion or relative frequency of allele A_1 in generation t , that is $p_i = i/2N, i = 0, \dots, 2N$. Mathematically under general diffusion approximation, frequency p_i are treated as continuous quantity x , and this also leads to the distribution of the frequency of considered allele A_1 being probability density.

4.2 Dynamics of Original Wright–Fisher Process

Adaptive landscape can be quantified as potential function. Potential function can give both quantitative and qualitative description of behaviors near steady states or metastable states. Great efforts have been spent to find such a potential function [13]. We here give specific potential function corresponding to diverse biological factors. It is clear how our potential function is related to the dynamical trajectories.

4.2.1 Pure Drift Balance

In the past, people also come into notice the effect of random drift, for example, [27] studied the effect through assuming that population fixed for a single genotype. Krakauer and Plotkin [28] studied the effect of random drift through analyzing small perturbations from the deterministic equilibrium. de Vladar and Barton [15] studied

the same process, but the method of maximizing entropy cannot work because of non-normalized stationary distribution. Waxman [29] studied the random drift process by assuming Dirac function at the boundary. We here do not focus on the boundary condition. To avoid the difficult boundary problem, here we show our method is independent of the normalization of stationary distribution and can directly get the population behaviors by adaptive landscape. We could address it from the perspective of adaptive landscape. For random sampling variance being

$$V(x) = \frac{x(1-x)}{2N}, \quad (24)$$

at the same time the directional force reads

$$f(x) = \frac{2x-1}{4N}. \quad (25)$$

In addition, we are interested in the dynamical property of adaptive landscape, so we treat Φ and Φ/ϵ no difference in this respect, that is, for convenience we can take $\epsilon = 1$ of $\epsilon D(x)$. So adaptive landscape from Eq. (21) is expressed as

$$\Phi(x) = -\ln x(1-x). \quad (26)$$

For biological understanding, no mutation equals to no driving force, without driving force the stationary distribution should be plain. That is, each state has the same chance to be ultimate state as neutral evolution, but in fact because of inertia of driving force, there exists the effect of force in the system, this makes the system move on. This is consistent with the expression of directional force $f(x)$ in Eq. (20). The inertia of driving force can just balance out the effect of random drift and let the population move on. Here from the expression of adaptive landscape, two singular points with allele frequency $x = 0$ and $x = 1$ exist. The singularity means that the population lies at a stable or unstable state. Positive infinity means the state with corresponding allele frequency stable, while negative infinity means the state with corresponding allele frequency unstable. Here the values of adaptive landscape for two singular points go to positive infinity. These demonstrate that there are indeed two stable states in the process. The adaptive landscape has U-shape.

Directional force actually gives the effects of driving force and random drift. Two stable states with infinite potential occur in the process. The fixed point with allele frequency $x = 1/2$ could be got through letting directional force being zero directly. It is the critical point to decide the eventual stable state of an initial state.

4.2.2 Mutation-Drift Balance

So far we have been considering situations where the change in composition of populations is caused by pure random drift. We now include the effect of mutation:

an A_1 allele may mutate with a probability μ to an A_2 allele, and an A_2 allele may mutate with a probability ν to an A_1 allele. These parameters are probabilities per generation. In the process we assume there is no difference on viability because of the absence of selection, the evolutionary process is only under the condition of mutation and random drift with unchanged environment. We focus on the evolution of the allele A_1 . Previous neutral evolution regarded these evolutionary processes as neutral.

To include mutation, one chooses two alleles from the current population to die and replaces it with two types chosen. That is, when we pick alleles to be the parent of a child in the next generation, the offspring can mutate with the probability μ or ν . For example, if an allele A_1 is chosen, there is a probability $1 - \mu$ that replacement allele is also an A_1 and of μ that it is an A_2 . So with mutation, in the offspring generation the frequency of allele A_1 is

$$p_{t+1} = (1 - \mu)p_t + \nu(1 - p_t). \quad (27)$$

After diffusion approximation like pure random drift process, the average change of allele frequency x : $M(x)$ and random sampling variance $V(x)$ are, respectively,

$$M(x) = -\mu x + \nu(1 - x), \quad (28)$$

$$V(x) = \frac{x(1 - x)}{2N}, \quad (29)$$

then $f(x)$, $D(x)$ in Eq. (20) are

$$\begin{aligned} f(x) &= -\mu x + \nu(1 - x) - \frac{1 - 2x}{4N} \\ &\equiv F(x - a), \end{aligned} \quad (30)$$

$$\epsilon D(x) = \frac{x(1 - x)}{4N}, \quad (31)$$

with

$$F = \frac{1 - 2N\mu - 2N\nu}{2N}, \quad a = \frac{1 - 4N\nu}{2 - 4N\mu - 4N\nu}. \quad (32)$$

Here we also take $\epsilon = 1$. So the adaptive landscape corresponding to Eq. (21) reads

$$\Phi(x) = 4N\nu \ln x + 4N\mu \ln(1 - x) - \ln x(1 - x). \quad (33)$$

There are three elements composed of adaptive landscape. The first and the second elements represent the effect due to mutation. The last element is the effect

due to random drift. There exist two singular points with allele frequency $x = 0$ and $x = 1$ from the expression of adaptive landscape. Singularity indicates the state stable or unstable. We show below that the evolutionary behavior can still be explored in terms of adaptive landscape. This method is insensitive to the singular expression. Fixed points are derived by

$$\Phi'(x) = 0. \quad (34)$$

That is

$$(4N\mu - 1 + 4N\nu - 1)x - (4N\nu - 1) = 0. \quad (35)$$

Then we get only one fixed point with allele frequency $x = a$ (its expression is Eq. (32)), and it is also the zero point of $f(x)$. In another word, when directional force is expressed in linear form, the fixed point can be read from its form directly.

We could investigate the dynamics by analyzing the relative position of the fixed point and singular points. In the following we give the detailed information.

(i) $a < 0$

There are two regimes holding the same dynamical behavior. In one of the regimes parameters satisfy $\mu \in (0, 1/4N)$ and $\nu \in (1/4N, (1 - 2N\mu)/2N)$. In another regime parameters satisfy $\mu \in (1/4N, 1)$ and $\nu \in ((1 - 2N\mu)/2N, 1/4N)$. There are two unstable states and one stable state in the process. The states at the fixed point with allele frequency $x = a$ and at the singular point with allele frequency $x = 0$ are unstable while the state with allele frequency $x = 1$ is stable. This means the population tends to fix at the state with allele frequency $x = 1$. Alleles composed of the population are likely to be allele A_1 .

(ii) $a = 0$

There are two regimes having the same dynamical behavior. In one of the regime mutation rates satisfy $\mu \in (0, 1/4N)$ and $\nu = 1/4N$. In another regime mutation rates satisfy $\mu \in (1/4N, 1)$ and $\nu = 1/4N$. Among these regimes the fixed point mixed the singular point with allele frequency $x = 0$. Among these cases the factor which can dominate the process is mutation rate. If parameters satisfy $\mu \in (0, 1/4N)$ and $\nu = 1/4N$, the probability of mutation from allele A_2 to allele A_1 is greater. So there is only one stable state with allele frequency $x = 1$ and one unstable state with allele frequency $x = 0$ in the process. If parameters satisfy $\mu \in (1/4N, 1)$ and $\nu = 1/4N$, the probability of mutation from allele A_1 to allele A_2 is greater. So there is one stable state with allele frequency $x = 0$ and one unstable state with allele frequency $x = 1$ in the process.

(iii) $0 < a < 1$

Two regimes have the same dynamical behavior. In one of the regimes mutation rates satisfy $\mu \in (0, 1/4N]$ and $\nu \in (0, 1/4N)$. In another regime mutation rates satisfy $\mu \in [1/4N, 1)$ and $\nu \in (1/4N, 1)$. The fixed point in the two regimes lies

at the interval $(0, 1)$. Among these cases the factor which can determine stability is mutation rate. If parameters satisfy $\mu \in (0, 1/4N]$ and $\nu \in (0, 1/4N)$, the probability of mutation from allele A_2 to allele A_1 and the probability of mutation from allele A_1 to allele A_2 are very small. There is only one unstable state with allele frequency $x = a$ and two stable states with allele frequency $x = 1$ and $x = 0$ in the process. Ultimately which state the population tends to fix is determined by the initial state. If its initial state with allele frequency is less than $x = a$, the population would move to the state with allele frequency $x = 0$. If parameters satisfy $\mu \in [1/4N, 1)$ and $\nu \in (1/4N, 1)$, the probability of mutation from allele A_2 to allele A_1 and the probability of mutation from allele A_1 to allele A_2 are very great. In one generation mutation number $2N(\mu + \nu)$ is greater than one. None of the two mutation directions can dominate the process. This results in individuals with genotype of two different genes becoming much. There is one stable state with allele frequency $x = a$ and two unstable states with allele frequency $x = 0$ and $x = 1$ in the process. Ultimately the population tends to fix at the state with allele frequency $x = a$. The population moves to the stable state with genotype $A_1 A_2$.

(iv) $a = 1$

There is only one point with the dynamical behavior. If parameters satisfy $(\mu, \nu) = (1/4N, 1/4N)$, the potential of the whole population is plain. In fact any state has equal chance to be the ultimate state. That is, the effects of mutation and random drift have no influence on the allele change. The effect of mutation is offsetted by that of random drift. The evolution has no bias on any state. We call this process new neutral evolution. Previous neutral evolution described the dynamical processes in the absence of selection. It has bias on some special states such as the case driven by unequal effects of mutation and random drift. Its description is not completely neutral.

(v) $a > 1$

There are two regimes with the same dynamical behavior. In one of the regimes parameters satisfy $\mu \in (0, 1/4N)$ and $\nu \in (1 - 2N\mu)/2N, 1]$. In another regime parameters satisfy $\mu \in (1/4N, 1/2N)$ and $\nu \in (0, (1 - 2N\mu)/2N]$. Among these cases stronger mutation rate from one of the two directions dominates the process. If parameters satisfy $\mu \in (0, 1/4N)$ and $\nu \in [1 - 2N\mu)/2N, 1)$, the probability of mutation from allele A_2 to allele A_1 is greater. So there is one stable state with allele frequency $x = 1$ and one unstable state with allele frequency $x = 0$ in the process. Ultimately the population tends to fix at the state with allele frequency $x = 1$. If parameters satisfy $\mu \in (1/4N, 1/2N)$ and $\nu \in (0, (1 - 2N\mu)/2N]$, the probability of mutation from allele A_1 to allele A_2 is greater. So there is one stable state with allele frequency $x = 0$ and one unstable state with allele frequency $x = 1$ in the process. Ultimately the population tends to fix at the state with allele frequency $x = 0$.

We theoretically analyze the dynamical behavior of the population and give the critical point in parameters space. Furthermore we can get the cause of the evolution clearly.

4.2.3 Selection-Drift Balance

The process in the presence of selection, and random drift is much complex. We only focus on some special cases to show how the process is characterized by adaptive landscape. Natural selection works when genotypes have different fitnesses [30]. But the dynamics of selection depends on relative fitness. Suppose we denote the relative fitness of selection on genotype A_1A_1 by $1 + s$, A_1A_2 by $1 + sh$, A_2A_2 by 1. Here s is called selection coefficient, a measure of the fitness of A_1A_1 relative to that of A_2A_2 and h is called heterozygote effect [30], a measure of the fitness of the heterozygote relative to the selective difference between the two homozygotes. We here assume $0 < s < 1$, A_1A_1 is fitter than A_2A_2 . Let x be the frequency of allele A_1 , then we can define mean fitness as

$$\bar{w} = x^2(1 + s) + 2x(1 - x)(1 + sh) + (1 - x)^2. \quad (36)$$

For the case of selection-random drift, the average frequency change of allele A_1 is

$$M(x) = \frac{x(1 - x)}{2\bar{w}} \frac{d\bar{w}}{dx}. \quad (37)$$

For convenience we only assume additive fitness, put differently, it depicts the important evolutionary process parameter satisfying $h = 1/2$: additive selection-random drift process. So according to Eq. (20)

$$\begin{aligned} f(x) &= M(x) - \epsilon D'(x) \\ &= \frac{sx(1 - x)}{2(1 + xs)} - \frac{1 - 2x}{4N}. \end{aligned} \quad (38)$$

Potential function, namely, adaptive landscape derived analogous to Eq. (21) ($\epsilon = 1$) reads

$$\Phi(x) = 2N \ln(1 + xs) - \ln x(1 - x). \quad (39)$$

As we can see, there are two singular points with allele frequency $x = 0$ and $x = 1$ from the expression of adaptive landscape. Singularity means the population stable in the process.

Analogous to analysis of mutation-random drift balance, we get the fixed point from

$$\Phi'(x) = 0. \quad (40)$$

They are

$$x_{1,2} = \frac{2 + (2N - 1)s \mp \sqrt{(2N - 1)^2s^2 + 4s + 4}}{4s(N - 1)}. \quad (41)$$

After some trivial calculation, we can get the relative positions of fixed points and singular points. They must satisfy the relation $0 < x_1 < 1 < x_2$ under the condition $0 < s < 1$. We can get two stable states with allele frequency $x = 0$ and $x = 1$ from the expression of adaptive landscape. In fact because of the relative positions of these points, there is only one stable state with allele frequency $x = x_1$. It is the critical state that divides the effect regimes into two stable states. This case describes that many alleles with very small effect on fitness are close to additive. That is, the heterozygote fitness is the arithmetic mean of the fitness of the two homozygotes.

4.3 *Escape Time from Infinite Potential*

Because the metastability is such an important phenomenon, it is expected to compute the life time of metastable states. It is an open question that when the metastable states become unstable based on potential function, such as represented by Kramers' escaping rate formulae [31].

For escape time, it is defined that the object leaves from one potential basin of an stable state and never comes back. If we take the initial state as a potential peak, the exit state as a potential valley. Then the mean first passage time can approximate escape time in general bell-shaped potential. Here we estimate the average escape time by computing mean first passage time for two cases. One is of mutation, random drift, the result is the same as [32]. The other is of selection and random drift.

If we assume transition process occurs in the presence of weak mutation and random drift under the condition of $4N\mu \rightarrow 0, 4N\nu \rightarrow 0$, the corresponding potential is U-shape. When first the population lies at the basin of stable state with allele frequency $x = 0$, then it arrives at the state with allele frequency $x = a$, and bounces back and forth and eventually reaches the state with allele frequency $x = 1$. Because the potential difference is infinite and adaptive landscape has U-shape, the time it climbed over the potential is very longer. General mean first passage time about a population leaving $(0, a)$ under the circumstance of initial Dirac distribution satisfies

$$(f(x) + \epsilon D'(x))\partial_x T(x) + \epsilon D(x)\partial_x^2 T(x) = -1. \quad (42)$$

Further we set

$$T'(0) = 0, \quad (43)$$

$$T(a) = 0. \quad (44)$$

Then we get

$$T(x) = \int_x^a \frac{1}{\epsilon D(y)} \exp(-\Phi(y)) dy \int_0^y \exp(\Phi(z)) dz. \quad (45)$$

Here $\Phi(x) = \int_0^x f(x')/D(x') dx' (\epsilon = 1)$ is adaptive landscape.

In general, when the landscape is sharply peaked at $x = a$, the main contribution to the integral Eq. (45) is due to a small region near $x = a$. But here there is a fat potential valley rather than a thin one. This results in the contribution to the integral Eq. (45) from a big region around $x = a$, and at the same time $\Phi'(a) = 0$, the probability from $x = 0$ to $x = a$ equals to that from $x = a$ to $x = 1$. So the escape time from $x = 0$ to $x = 1$ can be approximated to two mean first passage time from $x = 0$ to $x = a$. Because escape time is when the population leaves the basin of stable state with allele frequency $x = 0$, the escape time in the presence of weak mutation and random drift is computed as

$$\begin{aligned}
 T &\approx 2 \times T_{MFPT}(0 \rightarrow a) \\
 &= 8N \int_0^a y^{-4N\nu} (1-y)^{-4N\mu} dy \int_0^y z^{4N\nu-1} (1-z)^{4N\mu-1} dz \\
 &= \frac{2(a-0)}{\nu} + \frac{8N\mu}{\nu} \sum_{n=1}^{\infty} \frac{a^{n+1}}{n+1} \prod_{k=2}^n \left(\frac{k-1+4N\mu}{k} \right) \\
 &\quad + 8N(1-4N\mu) \sum_{n=1}^{\infty} \frac{a^{n+1}}{(n+1)(n+4N\nu)} \prod_{k=2}^n \left(\frac{k-4N\mu}{k} \right) \\
 &= \frac{1}{\nu} + \frac{4N\mu}{\nu} \sum_{n=1}^{\infty} \frac{2^{-n}}{n+1} \prod_{k=2}^n \left(\frac{k-1+4N\mu}{k} \right) \\
 &\quad + 4N(1-4N\mu) \sum_{n=1}^{\infty} \frac{2^{-n}}{(n+1)(n+4N\nu)} \prod_{k=2}^n \left(\frac{k-4N\mu}{k} \right). \quad (46)
 \end{aligned}$$

Another way to get the escape time is that approximated mean first passage time from $x = 0$ to $x = 1$. It is

$$\begin{aligned}
 \tau_1 &\approx T_{MFPT}(0 \rightarrow 1) \\
 &= 4N \int_0^1 y^{-4N\nu} (1-y)^{-4N\mu} dy \int_0^y z^{4N\nu-1} (1-z)^{4N\mu-1} dz \\
 &= \frac{1}{\nu} + \frac{4N\mu}{\nu} \sum_{n=1}^{\infty} \frac{1}{n+1} \prod_{k=2}^n \left(\frac{k-1+4N\mu}{k} \right) \\
 &\quad + 4N(1-4N\mu) \sum_{n=1}^{\infty} \frac{1}{(n+1)(n+4N\nu)} \prod_{k=2}^n \left(\frac{k-4N\mu}{k} \right). \quad (47)
 \end{aligned}$$

These results demonstrate that the stable state with infinite potential has the probability to be unstable. It gives a quantitative measure of escape with infinite potential. From Eqs. (46) and (47), we can see the results are the same if the first element $1/\nu$ approximates the escape time.

If we assume transition process occurs in the presence of weak selection and random drift under the condition of $4Ns \rightarrow 0$, we can get the escape time leaving the basin of stable state with allele frequency $x = 0$ by the same procedure. It is

$$\begin{aligned} \tau_2 &\approx T_{MFPT}(0 \rightarrow 1) \\ &= 4N \int_0^1 \frac{4N}{y(1-y)} e^{-2N \ln(1+ys) + \ln y(1-y)} dy \int_0^y e^{2N \ln(1+zs) - \ln z(1-z)} dz \\ &\rightarrow \infty. \end{aligned} \tag{48}$$

Because the integral about the variable y is not integral at $y = 0$. Here infinite potential indeed means infinite escape time. So the above two infinite potentials have different meanings. One is true infinite, the other is false infinite.

4.4 Muller Ratchet

Muller's ratchet proposed in 1964 [33] that the genome of an asexual population accumulates deleterious mutations in an irreversible manner. It is a mechanism that has been suggested as an explanation for the evolution of sex [34]. For asexually reproducing population, without recombination, chromosomes are directly passed down to offsprings. As a consequence, the deleterious mutations accumulate so that the fittest class loses. For sexually reproducing population, because of the existence of recombination between parental genomes, a parent carrying high mutational loads can have offspring with fewer deleterious mutations. The high cost of sexual reproduction is thus offsetted by the benefits of inhibiting the ratchet [35]. Muller's ratchet has been received growing attention.

Here in one-dimensional case, we consider one locus with two alleles (for example, A and a), that is, there are two classes in the haploid asexual population, one class with allele A while the other with allele a , supposed mutation from allele A to a is deleterious. We assume fixed population size of N , which means we have N alleles in all. We also assume that $N > 1$. Generations are non-overlapping. The lifecycle of the individuals in the population is from adults to juveniles, during which we consider irreversible mutation, selection, and random drift. The frequency of the allele A for generation t is p while that of allele a is $1 - p$. Let μ be the probability that an offspring of an adult with allele A is an individual with allele a , labeled by $M_{1,0}$, that is $M_{1,0} = \mu$. Analogously, $M_{0,0} = 1 - \mu$, $M_{0,1} = 0$, $M_{1,1} = 1$. The relative viability of individuals with allele A is $v_0 = 1$ while that of individuals with allele a is $v_1 = 1 - \sigma$ where σ can be treated as an effective selection coefficient associated with deleterious mutations. So the values of parameters for μ and σ are from 0 to 1. Then in generation $t + 1$, when selection and deleterious mutation are active, the probability that the offspring of a parent with allele A is chosen to be with allele a is $\mu p(1 - \sigma)$, the probability that the offspring with allele A is $(1 - \mu)p$,

the probability that the offspring of a parent with allele a is still with allele a is $(1 - \sigma)(1 - p)$. So the frequency of allele A in generation $t + 1$ is

$$p' = \frac{(1 - \mu)p}{1 - \sigma + \sigma(1 - \mu)p}. \quad (49)$$

After using diffusion approximation, frequency p is treated as continuous quantities x , and this leads to the distribution of the frequency for the allele A being the probability density. Let $\rho(x, t)$ be the probability density of the frequency for the allele A being x at time t . Then it obeys the dynamical equation

$$\frac{\partial}{\partial t} \rho(x, t) = -\frac{\partial}{\partial x} [M(x)\rho(x, t)] + \frac{1}{2} \frac{\partial^2}{\partial x^2} [V(x)\rho(x, t)], \quad (50)$$

and according to the definition of $M(x)$ and $V(x)$, the explicit expressions of them are

$$\begin{aligned} M(x) &= \frac{(1 - \mu)x}{1 - \sigma + \sigma(1 - \mu)x} - x \\ &= \frac{x[(\sigma - \mu) - \sigma(1 - \mu)x]}{1 - \sigma + \sigma(1 - \mu)x}, \end{aligned} \quad (51)$$

$$V(x) = \frac{x(1 - x)}{N}. \quad (52)$$

So according to Eq. (21) ($\epsilon = 1$), we have adaptive landscape as the following

$$\begin{aligned} \Phi(x) &= \frac{2N\mu(1 - \sigma)}{1 - \sigma\mu} \ln(1 - x) - \ln x(1 - x) \\ &\quad + \frac{2N(1 - \mu)}{1 - \sigma\mu} \ln(1 - \sigma + x\sigma(1 - \mu)). \end{aligned} \quad (53)$$

From the expression of adaptive landscape $\Phi(x)$, we may find there are two singular points with allele frequency $x = 0$ and $x = 1$ of adaptive landscape, characterized by infinite value. Infinity means the state relative stable or unstable in the system.

4.5 Dynamics of Muller Ratchet

To understand the mechanism of Muller's ratchet, a full characterization of dynamical process is a prerequisite for obtaining more accurate decaying time. Here we study the dynamical behavior by investigating the positions and relative stabilities of all fixed points as [36]. We further derive the parameter regions for all possible cases.

According to general analysis of a dynamical system, letting

$$\Phi'(x) = 0, \quad (54)$$

we get

$$2\sigma(1 - \mu)(N - 1)x^2 + (2N(\mu - \sigma) + 3\sigma - \sigma\mu - 2)x + (1 - \sigma) = 0. \quad (55)$$

We solved Eq. (55) and found two fixed points. If we denote

$$\alpha = 2 - 3\sigma + \sigma\mu + 2N\sigma - 2N\mu, \quad (56)$$

$$\beta = 8\sigma(1 - \mu)(N - 1)(1 - \sigma). \quad (57)$$

They are

$$x_{1,2} = \frac{\alpha \mp \sqrt{\alpha^2 - \beta}}{4\sigma(1 - \mu)(N - 1)}. \quad (58)$$

For two singular points with allele frequency $x = 0$ and $x = 1$, if allele frequency x goes to 1, and parameter satisfies $\sigma \in (\mu, (2N\mu - 1)/(2N\mu - \mu))$, then the value of adaptive landscape $\Phi(1)$ goes to $-\infty$, so the population is unstable at the state with allele frequency $x = 1$. When allele frequency x moves to 1, and selection satisfies $\sigma \in ((2N\mu - 1)/(2N\mu - \mu), 1)$, the value of adaptive landscape $\Phi(1)$ tends to $+\infty$, so the population is stable at the state with allele frequency $x = 1$. For allele frequency x goes to 0, the value of adaptive landscape $\Phi(0)$ goes to $+\infty$ in almost parameters regimes except $\sigma = 1$, so the population is always stable at the state with allele frequency $x = 0$. When selection satisfies $\sigma = 1$, the viability of the suboptimal class is zero, so the population stays at the initial state.

Here we address dynamical behaviors by the positions of two real inequivalent fixed points with allele frequency $x_1 < x_2$ first.

For convenience, we denote

$$\sigma_1 = \frac{(2+2\mu-10N\mu+4N^2\mu+2N\mu^2+4(1-\mu)\sqrt{N(N-1)((2N-1)\mu-1)}}{(\mu-2N+1)^2}.$$

1. We find two different real fixed points in two regimes. When parameters satisfy $\mu \in (0, 2/(2N - 1 + 2\sqrt{N(N - 1)}))$ and $\sigma \in (\mu, 1)$, parameters satisfy $\mu \in (2/(2N - 1 + 2\sqrt{N(N - 1)}), 1)$ and $\sigma \in (\sigma_1, 1)$ two different real fixed points occur. Among them parameters regimes do not include $\mu \in ((2N - 1)/4N(N - 1), 1)$ and $\sigma = (2N\mu - 1)/(2N\mu - \mu)$.

We discuss the relative positions for the fixed points and the singular points with allele frequency $x = 0$, $x = 1$ and stabilities of them in the following.

(i) $1 < x_1 < x_2$

In one of the regimes parameters satisfy $\mu \in (1/(2N - 1), 2/(2N - 1 + 2\sqrt{N(N - 1)})]$ and $\sigma \in (\mu, (2N\mu - 1)/(2N\mu - \mu))$, in another regime parameters satisfy $\mu \in (2/(2N - 1 + 2\sqrt{N(N - 1)}), (2N - 1)/4N(N - 1))$ and $\sigma \in (\sigma_1, (2N\mu - 1)/(2N\mu - \mu))$, the fixed points satisfy $1 < x_1 < x_2$. At the same time the state with singular point $x = 1$ is unstable. There is one stable state with allele frequency $x = 0$ in the system. The population tends to evolve to the stable state.

(ii) $1 = x_1 < x_2$

In the regions parameters satisfy $\mu \in (1/(2N - 1), (2N - 1)/4N(N - 1))$ and $\sigma = (2N\mu - 1)/(2N\mu - \mu)$, the two fixed points satisfy allele frequency $x_1 = 1$, $1 < x_2$. The state with allele frequency $x = 1$ is unstable. There is only one stable state with allele frequency $x = 0$ in the system.

(iii) $0 < x_1 < 1 < x_2$

In one of the regimes parameters satisfy $\mu \in (0, 1/(2N - 1))$ and $\sigma \in (\mu, 1)$, in another regime parameters satisfy $\mu \in (1/(2N - 1), 1)$ and $\sigma \in ((2N\mu - 1)/(2N\mu - \mu), 1)$, the fixed points satisfy $0 < x_1 < 1 < x_2$. There is only one unstable state with allele frequency $x = x_1$ in the system, and two stable states with allele frequency $x = 1$ and $x = 0$ exist in the system. The population tends to evolve to which stable state dependent on the position of the initial state. If the initial state with allele frequency is greater than x_1 , the population tends to evolve to the stable state with allele frequency $x = 1$.

(iv) $0 = x_1 < 1 < x_2$

In the regime parameters satisfy $\mu \in (0, 1)$ and $\sigma = 1$, the fixed points satisfy allele frequency $x_1 = 0$, $1 < x_2$. When selection rate satisfies $\sigma = 1$, the process stays at the initial state. Because for this case the viability of the sub-fittest class is zero.

(v) $0 = x_1 < x_2 < 1$

The case $0 = x_1 < x_2 < 1$ is impossible. For $x_1 = 0$, the other parameter must satisfy $\sigma = 1$, at the same time x_2 must be greater than one.

(vi) $0 < x_1 < x_2 < 1$

In one of the regimes parameters satisfy $\mu \in ((2N - 1)/4N(N - 1), (2N - 1)/(4N - 3))$ and $\sigma \in (\sigma_1, (2N\mu - 1)/(2N\mu - \mu))$, in another regime parameters satisfy $\mu \in ((2N - 1)/(4N - 3), 1)$ and $\sigma \in (\sigma_1, (2N\mu - 1)/(2N\mu - \mu))$, the fixed points satisfy $0 < x_1 < x_2 < 1$. The state with allele frequency x_1 is unstable while that with allele frequency x_2 is stable. There are two stable states with allele frequency $x = 0$ and $x = x_2$ and two unstable states with allele frequency $x = 1$ and $x = x_1$ in the system. The population evolves to which stable states dependent on the initial position.

(vii) $x_1 < 0$ or $x_2 < 0$

The case $x_1 < 0$ is impossible, and the case $x_2 < 0$ is impossible.

2. Then we discuss the case of two equivalent real fixed points with allele frequency $x_2 = x_1$.

In the regimes of parameters satisfying $\mu \in (2/(2N - 1 + 2\sqrt{N(N - 1)}), 1)$ and $\sigma = \sigma_1$, we find two same fixed points with allele frequency

$$x_{1,2} = \frac{\alpha}{4\sigma(1 - \mu)(N - 1)}. \quad (59)$$

(i) $1 < x_{1,2}$

In the regimes of parameters satisfying $\mu \in (2/(2N - 1 + 2\sqrt{N(N - 1)}), (2N - 1)/4N(N - 1))$ and $\sigma = \sigma_1$, there are two same fixed points satisfying $1 < x_{1,2}$, and they are unstable. There is one stable state with allele frequency $x = 0$ in the process.

(ii) $1 = x_{1,2}$

At the two points of $((2N - 1)/4N(N - 1), 2N/(2N - 1)^2)$ and $((2N - 1)/(4N - 3), (4(N - 1)(3 - 6N + 4N^2) + 8(N - 1)(4N - 3)\sqrt{N(N - 1)/(4N - 3)})/(2N - 1)^2)$, there are two same fixed points satisfying $x_{1,2} = 1$, and they are unstable. There is one stable state with allele frequency $x = 0$ in the process.

(iii) $0 < x_{1,2} < 1$

In one of the regimes parameters satisfy $\mu \in ((2N - 1)/4N(N - 1), (2N - 1)/(4N - 3))$, $\sigma = \sigma_1$, in another regime parameters satisfy $\mu \in ((2N - 1)/(4N - 3), 1)$ and $\sigma = \sigma_1$, there are two same fixed points satisfying $0 < x_{1,2} < 1$, and they are unstable. There is one stable state with allele frequency $x = 0$ in the process.

3. Finally we consider two imaginary fixed points $|x_1| = |x_2|$ where the $|\cdot|$ denotes the length for an imaginary points.

In the regime of parameters satisfying $\mu \in (2/(2N - 1 + 2\sqrt{N(N - 1)}), 1)$ and $\sigma \in (\mu, \sigma_1)$, there are two imaginary fixed points in the system. There is only one stable state with allele frequency $x = 0$. The population always evolves to the stable state.

4.6 Estimation of Single Click Time

To evaluate the single click time and show the further power of adaptive landscape, in the following we will demonstrate how the single click time from one relative stable state to another is derived for Muller ratchet. We also start from backward

Fokker–Planck equation as general treatment [37, 38], backward Fokker–Planck equation corresponding to Eq.(50) under natural boundary condition can be expressed in the following form

$$\partial_t \rho(x, t) = (f(x) + \epsilon D'(x)) \partial_x \rho(x, t) + \epsilon D(x) \partial_x^2 \rho(x, t). \quad (60)$$

General single click time dependent on initial Dirac function satisfies

$$(f(x) + \epsilon D'(x)) \partial_x T(x) + \epsilon D(x) \partial_x^2 T(x) = -1. \quad (61)$$

The general solution corresponding to Eq. (61) is

$$T(x) = \int_0^x dy \frac{1}{\epsilon D(y)} \exp(-\Phi(y)) \int_y^1 dz \exp(\Phi(z)), \quad (62)$$

here $\Phi(x) = \int^x dx' (f(x')/D(x')) (\epsilon = 1)$.

Here the evolutionary process occurs when $x \in [0, 1]$. We are more interested in the escape time between the two stable states with allele frequency $x = 0$ and $x = 1$. Here the adaptive landscape has rich structure which contains escaping from finite and infinite potential. In the escaping from finite potential process, there are two important states x^* , x_0^* . Interval $(0, 1)$ contains a potential well at x^* and a potential barrier at x_0^* . The single click time is composed of two elements, one denotes forming process of fittest class, the other describes losing process of fittest class. In general, the time spent on forming process is much smaller than that spent on losing process. So the escape time approximates to the time spent on losing process. Because we assume that near x_0^* we can write

$$\Phi(x) \approx \Phi(x_0^*) - \frac{1}{2} \left(\frac{x - x_0^*}{\alpha'} \right)^2. \quad (63)$$

and near x^*

$$\Phi(x) \approx \Phi(x^*) + \frac{1}{2} \left(\frac{x - x^*}{\beta'} \right)^2. \quad (64)$$

At the same time, if the central maximum of $\Phi(x)$ is large compared with $1/N$, then $\exp(\Phi(z))$ is sharply peaked at x_0^* , while $\exp(-\Phi(y))/D(y)$ is very small near $y = x^*$. Equation (62) is evaluated as

$$\begin{aligned} T_{1 \rightarrow 0} &\approx \int_{x^*}^0 dy \frac{1}{D(y)} \exp(-\Phi(y)) \int_1^{x_0^*} dz \exp(\Phi(z)) \\ &\approx \frac{2\pi\alpha'\beta' \exp(\Phi(x_0^*) - \Phi(x^*))}{D(x^*)} \\ &\propto \frac{1}{D(x^*)} \exp(\Phi(x_0^*) - \Phi(x^*)). \end{aligned} \quad (65)$$

From the expression of Eq. (65), the single click time is not sensitive to the boundary assumption with $x = 0$ and $x = 1$, it is closely related to the potential difference $\Phi(x_0^*) - \Phi(x^*)$.

In the higher mutation rates regime, where x_0^* equals to a stable state x_2 , x^* corresponds to the unstable state x_1 that the population lies between the adaptive states 0 and x_2 . The potential barrier $\Phi(x_2) - \Phi(x_1)$ is finite. The difference of potential is

$$\begin{aligned} \Phi(x_2) - \Phi(x_1) &= \frac{2N\mu(1-\sigma) - 1 + \sigma\mu}{1 - \sigma\mu} \ln \left(1 - \frac{x_2 - x_1}{1 - x_1} \right) - \ln \left(1 + \frac{x_2 - x_1}{x_1} \right) \\ &\quad + \frac{2N(1-\mu)}{1 - \sigma\mu} \ln \left(1 + \frac{\sigma(1-\mu)(x_2 - x_1)}{1 - \sigma + x_1\sigma(1-\mu)} \right) \\ &= \frac{2N\mu(1-\sigma)}{1 - \sigma\mu} \ln \left(1 - \frac{2\sqrt{\alpha^2 - \beta}}{\alpha - \sqrt{\alpha^2 - \beta}} \right) \\ &\quad - \ln \left(1 - \frac{4(\alpha^2 - \beta)}{(\alpha - \sqrt{\alpha^2 - \beta})^2} \right) + \frac{2N(1-\mu)}{1 - \sigma\mu} \\ &\quad \times \ln \left(1 + \frac{2\sqrt{\alpha^2 - \beta}}{4N - 6 - 6N\sigma + 7\sigma - \sigma\mu + 2N\mu + \sqrt{\alpha^2 - \beta}} \right), \end{aligned} \quad (66)$$

where α and β are the same as Eqs. (56) and (57). According to classical derivation, corresponding to Eq. (62) the single click time approximates to

$$\begin{aligned} T_{1 \rightarrow 0} &= \lim_{x \rightarrow 1} \int_0^x dy \frac{1}{\epsilon D(y)} \exp(-\Phi(y)) \int_y^1 dz \exp(\Phi(z)) \\ &\approx 2N \int_0^{x^*} \frac{(1-y)^{2N\mu(\sigma-1)/(1-\sigma\mu)}}{(1-\sigma + y\sigma(1-\mu))^{2N(1-\mu)/(1-\sigma\mu)}} dy \\ &\quad \times \int_{x_0^*}^1 z^{-1} (1-z)^{(\sigma\mu-1-2N\mu(\sigma-1))/(1-\sigma\mu)} \\ &\quad \times (1-\sigma + \sigma z(1-\mu))^{2N(1-\mu)/(1-\sigma\mu)} dz \\ &\approx \frac{1}{D(x^*)} \exp(\Phi(x_2) - \Phi(x_1)) \\ &\approx \frac{N(N-1)^2\sigma^2(1-\mu)^2}{(\alpha - \sqrt{\alpha^2 - \beta})(4\sigma(N-1)(1-\mu) - \alpha + \sqrt{\alpha^2 - \beta})}. \end{aligned} \quad (67)$$

The approximated single click time varies with mutation rates. The single click time $T_{1 \rightarrow 0}$ increases with population size N in certain regime, decreases with mutation rates μ and selection rates σ in the parameters regime $\mu \in (2N/4N(N-1), 1)$ and

$\sigma \in (\sigma_1, (2N\mu - 1)/(2N\mu - \mu))$. When mutation rate holds constant, with selection rates increasing, the difference of potential between two fixed points decreases, the viability of suboptimal class decreases, the population evolves to the fittest class. On the other hand, when selection holds constant, with deleterious mutation increasing, the population of suboptimal class increases, the difference of potential between two fixed points decreases.

For the lower mutation rates regime, where the potential barrier is infinite. The landscape has U-shape. The single click time can also be estimated under the condition of weak mutation and weak selection, x^* corresponds to the fixed point x_1 that the population lies at the lowest potential.

$$\begin{aligned}
 T_{Escape}(1 \rightarrow 0) &\approx T_{MFPT}(1 \rightarrow 0) \\
 &\approx 2N \int_0^1 \frac{(1-y)^{2N\mu(\sigma-1)/(1-\sigma\mu)}}{(1-\sigma + y\sigma(1-\mu))^{2N(1-\mu)/(1-\sigma\mu)}} dy \\
 &\quad \times \int_{x_0^*}^1 z^{-1}(1-z)^{(\sigma\mu-1-2N\mu(\sigma-1))/(1-\sigma\mu)} \\
 &\quad \times (1-\sigma + \sigma z(1-\mu))^{2N(1-\mu)/(1-\sigma\mu)} dz \\
 &\approx \frac{1-\sigma\mu}{\mu(1-\sigma)}. \tag{68}
 \end{aligned}$$

From expression of Eq. (68), the single click time goes to infinity with mutation rates tending to zero in the parameters regimes of $\mu \in (0, 1/(2N - 1))$ and $\sigma \in (\mu, 1)$. The results of the single click time are not sensitive to the population size.

Analogous to the derivation of $T_{Escape}(1 \rightarrow 0)$, we can calculate

$$\begin{aligned}
 T_{0 \rightarrow 1} &= \lim_{x \rightarrow 0} \int_x^1 dy \frac{1}{\epsilon D(y)} \exp(-\Phi(y)) \int_0^y dz \exp(\Phi(z)) \\
 &\approx \lim_{x \rightarrow 0} 2N \int_x^1 \frac{(1-y)^{2N\mu(\sigma-1)/(1-\sigma\mu)}}{(1-\sigma + y\sigma(1-\mu))^{2N(1-\mu)/(1-\sigma\mu)}} dy \\
 &\quad \times \int_0^y z^{-1}(1-z)^{(\sigma\mu-1-2N\mu(\sigma-1))/(1-\sigma\mu)} \\
 &\quad \times (1-\sigma + \sigma z(1-\mu))^{2N(1-\mu)/(1-\sigma\mu)} dz \\
 &\approx 2N \int_0^1 \frac{(1-y)^{2N\mu(\sigma-1)/(1-\sigma\mu)}}{(1-\sigma + y\sigma(1-\mu))^{2N(1-\mu)/(1-\sigma\mu)}} dy \\
 &\quad \times \int_0^y (z^{-1} + \dots) dz \\
 &\rightarrow \infty. \tag{69}
 \end{aligned}$$

From expression of Eq. (69), the single click time goes to infinity with mutation rates being zero.

Our results show that the expected time to loss of the fittest class can be predicted for a wide range of parameter values that appear to be biologically relevant. Furthermore our results show that the time is insensitive to population size in the presence of weak mutation and weak selection. This tells us that the ratchet process may occur in finite but not small populations.

5 Discussion

Adaptive landscape introduced by Sewall Wright, a re-emerging powerful concept in systems biology, is used as a tool to describe complex biological processes. In the past Prugel-Bennett and Shapiro, Rattray took analogies between population evolution and statistical physics and proposed maximum polymorphism to approximate polygenic system. Their method does not ensure that the arbitrary entropy measure converges to the correct stationary distribution, though it makes accurate prediction. Barton group uses information entropy measure to analyze the process of population evolution analogous to statistical thermodynamics. It points out that the method works only for high mutation rates and breaks down for low mutation rates [15], though this method ensures the convergence to the correct stationary distribution by maximizing introduced entropy. In a word, past methods of constructing adaptive landscape must need the normalization of stationary distribution. The present chapter manifests the nonlinear and complex dynamics of the evolutionary system by adaptive landscape analytically. Especially it can describe the dynamical behavior under the action of random drift. Waxman and Loewe [26] studied the same process, but they put Dirac function at boundary. This results in the solution not satisfying the equation. The theoretical results suggest our constructive potential function may be a reasonable candidate to quantify the adaptive landscape and investigate the complex biological processes.

Recently, Zhou and Qian [39] also constructs landscape function to study the complex dynamical property by discrete and diffusion Moran process. They also meet the dilemma diffusion approximation and mismatch between fixed points and extrema of adaptive landscape. But this mismatch can be explained in our constructive method. Our method does not need the existence and normalization of the stationary distribution. Our constructive method is independent of the stationary distribution. Our method investigates the long-term dynamical property of the system and reduces the complexity of calculating stochastic differential equations. Our adaptive landscape can give a new definition of neutral evolutionary, that is, the population is under the equal action of mutation and random drift. And the analytical results of mutation-random drift balance are consistent with the results in general textbook [21]. Whatever its shape and its singularity, the adaptive landscape characterizes globally the complex dynamical behaviors for a system.

Acknowledgments The critical discussion with Prof. Zhu Xiaomei is appreciated. We also thank Jiang Pengyao, Wang Yanbo, and other members in the lab for their constructive comments.

References

1. Wright S (1932) The role of mutation, inbreeding, crossbreeding and selection in evolution. *Proc Int Congr Genet* 1:356–366
2. Lande R (1976) Natural selection and random genetic drift in phenotypic evolution. *Evolution* 30:314–334
3. Arnold SJ, Pfrender ME, Jones AG (2001) The adaptive landscape as a conceptual bridge between micro and macroevolution. *Genetica* 112–113: 9–32
4. Wright S (1988) Surfaces of selective value revisited. *Am Nat* 131:115–123
5. Wright S (1967) Surfaces of selective value. *Proc Natl Acad Sci* 131:165–172
6. Provine WB (1986) Sewall Wright and evolutionary biology. University of Chicago Press, Chicago, IL
7. Weinreich DM, Sindi S, Watson RA (2013) Finding the boundary between evolutionary basins of attraction, and implications for Wright's fitness landscape analogy. *J Stat Mech Theor Exp* 01:P01001
8. Gyllenberg M, Metz JAJ, Service R (2011) When do optimisation arguments make evolutionary sense? In: Fabio ACC Chalub, Jos Francisco Rodrigues (eds) *The mathematics of Darwin's legacy*. Springer, Basel, pp 233–288
9. Ao P (2009) Global view of bionetwork dynamics: adaptive landscape. *J Genet Genom* 36: 63–73
10. Kaplan J (2008) The end of the adaptive landscape metaphor? *Biol Philos* 23:625–638
11. Gavrillets S (1997) Evolution and speciation on holey adaptive landscapes. *Trends Ecol Evol* 12:307–312
12. Ao P (2008) Emerging of stochastic dynamical equalities and steady state thermodynamics from Darwinian dynamics. *Comm Theor Phys* 49:1073–1090
13. Ao P (2005) Laws in Darwinian evolutionary theory. *Phys Life Rev* 2:117–156
14. Pigliucci M, Kaplan J (2006) *Making sense of evolution: the conceptual foundations of evolutionary theory*. University of Chicago Press, Chicago, IL
15. de Vladar HP, Barton NH (2009) Statistical mechanics and the evolution of polygenic quantitative traits. *Genetics* 181:997–1011
16. Coe JB, Barton NH (2009) On the application of statistical physics to evolutionary biology. *J Theor Biol* 259:317–324
17. Assaf M, Mobilia M (2011) Fixation of deleterious allele under mutation pressure and finite selection intensity. *J Theor Biol* 275:93–103
18. Bharucha-Reid AT (1960) *Elements of the theory of Markov processes and their applications*. McGraw-Hill, New York
19. Feller W (1954) Diffusion processes in one dimension. *Trans Am Math Soc* 77:1–31
20. Kimura M (1964) Diffusion models in population genetics. *J Appl Prob* 1:177–232
21. Ewens WJ (2004) *Mathematical population genetics*. Springer, New York
22. Ao P (2004) Potential in stochastic differential equation: novel construction. *J Phys Math Gen* 37:25–30
23. Fisher RA (1930) *The genetical theory of natural selection*. Clarendon Press, Oxford
24. Wright S (1931) Evolution in Mendelian populations. *Genetics* 16:97–159
25. Blythe RA, McKane AJ (2007) Stochastic models of evolution in genetics, ecology and linguistics. *J Stat Mech Theor Exp* 2007:P07018
26. Waxman D, Loewe L (2010) A stochastic model for a single click of Muller's ratchet. *J Theor Biol* 264:1120–1132
27. Van Nimwegen E, Crutchfield JP, Huynen M (1999) Neutral evolution of mutational robustness. *Proc Natl Acad Sci* 96:9716–9720
28. Krakauer DC, Plotkin JB (2002) Redundancy, antiredundancy, and the robustness of genomes. *Proc Natl Acad Sci* 99:1405–1409
29. Waxman D (2007) Singular solutions of the diffusion equation of population genetics. *J Theor Biol* 247:849–858

30. Gillespie JH (2004) Population genetics: a concise guide. The Johns Hopkins University Press, Baltimore
31. Kramers HA (1940) Brownian motion in a field of force and the diffusion model of chemical reactions. *Physica* 7:284–304
32. Xu S, Jiao SY, Jiang PY, Yuan B, Ao P (2012) Escape from infinite adaptive peak. In: Proceedings of Sixth International Conference on System Biology, 268–273, Xi'an, 18–20 August (2012)
33. Muller HJ (1964) The relation of recombination to mutational advance. *Mutat Res Fund Mol Mech Mutagen* 1:2–9
34. Maynard Smith J (1978) The evolution of sex. Cambridge University Press, England
35. Etheridge A, Pfaffelhuber P, Wakolbinger A (2009) How often does the ratchet click? Facts, heuristics, asymptotics. In: Bath J, Mörters P, Scheutzow M (eds) Trends in stochastic analysis. Springer, Basel, pp 233–288
36. Jiao SY, Ao P (2012) Absorbing phenomena and escaping time for Muller's ratchet in adaptive landscape. *BMC Syst Biol* S1:S10
37. Van Kampen NG (1992) Stochastic processes in physics and chemistry. North Holland, Amsterdam
38. Øksendal B (2003) Stochastic differential equations: an introduction with applications. Springer, Berlin
39. Zhou D, Qian H (2011) Redundancy, antiredundancy, and the robustness of genomes. *J Theor Biol* 99:1405–1409