Escape from infinite adaptive peak

Song Xu*, Shuyun Jiao†‡, Pengyao Jiang§, Bo Yuan* and Ping Ao†

*Department of Computer Science and Engineering, Shanghai Jiao Tong University, 200240, Shanghai, China
†Shanghai Center for Systems Biomedicine, Key Laboratory of Systems Biomedicine of Ministry of Education,
Shanghai Jiao Tong University, 200240, Shanghai, China
‡Department of Mathematics, Xinyang Normal University, 464000, Xinyang, Henan, China
§Department of Ecology and Evolution, the University of Chicago, 1101 E 57th Street, Chicago, IL 60637, USA

Abstract—We study the transition time between different meta-stable states in the continuous Wright-Fisher (diffusion) model. We construct an adaptive landscape for describing the system both qualitatively and quantitatively. When strong genetic drift and weak mutation generate infinite adaptive peaks, we calculate the expected time to escape from such peak states. We find a new way to analytically approximate the escape time, which extends the application of Kramer's classical formulae to the cases of non-Gaussian equilibrium distribution and bridges previous results in two limits. Our adaptive landscape, compared to the classical fitness landscape or other scalar functions, is directly related to system's middle-and-long-term dynamics and is self-consistent in the whole parameter space. Our work provides a complete description for the bi-stabilities in the present model.

I. INTRODUCTION

The issue of multi-stabilities and state-transition probabilities in biological evolution have been long concerned by population geneticists [1][2][3][4][5][6]. Theoretical results are widely used for phenomenal studies of biological speciation [7], evolutionary robustness [8], and genetic substitutions [9].

Viewing the biological evolution as a stochastic process, one may classify the dynamics into two fundamentally different types of movements: toward the local stable (or meta-stable) state or between different stable states [10]. Theoreticians are strongly interested in the peak-to-peak transitions, which is often considered of stochastic essence [11][12]. On the qualitative side, mechanisms for such rare events are variously defined [13] (one example is the random genetic drift [14]). From the quantitative side, it has been frequently asked how to calculate the expected rate/time of peak-shifting events under the interaction of different biological factors.

A widely used approach to calculate the escape time is derived from the diffusion process, first introduced by [11]. It applies when the equilibrium distribution of probability density, near one or other peak, is approximately Gaussian. The escape time or first passage time was also calculated using the backward diffusion equation in population genetics and other areas under similar approximations [12][15][16]. The results estimate an exponential dependency of the escape time on the energy gap, called a "Arrehnius factor". In the present model, however, this classical results cannot be applied as strong genetic drift would generate infinite peaks on the equilibrium distribution. The Gaussian approximation is invalid as the local maximum would diverge to infinity.

The second approach for estimating the transition rate in population genetics is by calculating the probability of fixation of a single mutant in a finite population [3][4]. The rate of substitution (of a new mutation), then, is obtained as the multiplication of this probability and the average number of mutations that enter the population in each generation [9]. This approach is valid under weak forward mutation, but is not generally applicable for more complex dynamics (e.g. when the backward mutation rate is considerable). On the other hand, it provides a rather simple expression for the substitution rate, but the result is precise only under certain limiting conditions (e.g. when the forward mutation is very weak).

A third approach proposed by [6] is to calculate the leading (non-zero) eigenvalue of the diffusion equation, which gives the flux between different equilibria. Their method can be used when the forward mutation is weak, and allows the calculation under the joint effects of other factors (like the backward mutation). However, their method fails when the selection is so weak that there only exists one peak on their "deterministic equilibrium". The deterministic equilibrium fails to show the (long-term) bi-stability of the system in certain cases as the effect of genetic drift was not considered.

In the present work, we make attempts to solve this infinity-escape problem in one of the most classical model in population genetics, the (continuous) Wright-Fisher model. We are motivated to construct an adaptive landscape for an accurate and self-consistent description for the system's middle-and-long term dynamics. The concept of classical fitness/adaptive landscape has been under controversy since Wright's first proposition in 1932 [14]. Recently, one the present authors provided a new way for landscape constructions in general dynamical systems [17][18]. Results have been applied in the studies of phage-lambda genetic switch [19], Darwinian evolution [20] and cancer [21]. We will show in the process of population evolution, how genetic drift and other factors would drive a population to shift between alternative peaks on the adaptive landscape.

II. MODEL AND METHODS

A. Wright-Fisher model and diffusion process

The Wright-Fisher model describes the evolution of a diploid population at one locus (denote the corresponding alleles as A_1 and A_2). It assumes that the generations are

non-overlapping and the population size N is constant. If N is big enough (usually satisfying $N > 50$), the continuous approximation of the Wright-Fisher model can be described by the diffusion process equation

$$\partial_t \rho(x,t) = \partial_x \Big[\epsilon D(x)\partial_x - f(x) \Big] \rho(x,t) \ . \qquad (1)$$

Here x is the allele frequencies of A_1 (so the A_2 frequency is $1 - x$). $\rho(x,t)$ is the probability density that the A_1 frequency of the population is x at time t. $f(x)$ and $\epsilon D(x)$ consist of evolutionary factors (mutation, selection, genetics drift, etc.) being considered in specific models. One can easily find the equilibrium solution of Eq.(1) by setting its right side to be zero [12]:

$$\rho(x,t=\infty) = \frac{1}{Z} \exp\left[\int^x \frac{f(y)}{\epsilon D(y)} dy \right] \ , \qquad (2)$$

where Z is the normalization constant

$$Z = \int_0^1 \rho(x,t=\infty) dx \ . \qquad (3)$$

In the present work, we start from the simplest case of mutation and drift. The parameters in Eq.(1) are given by

$$f(x) = -\mu x + \nu(1-x) - (1-2x)/4N \ , \qquad (4)$$
$$\epsilon D(x) = x(1-x)/4N \ . \qquad (5)$$

Here μ is the rate of mutation from A_1 alleles to A_2 alleles and ν in the reverse direction. $2N$ is the total number of alleles at the interested locus in the diploid population.

B. Adaptive landscape and escape time

There are various ways to obtain the long-term dynamics of the model in Eq.(1). In the present work, we are motivated to study the above system using an adaptive landscape. We define

$$\Phi(x) = \int^x \frac{f(y)}{D(y)} dy \ , \qquad (6)$$

which, if combined with Eq.(2), can be easily shown to conform the Boltzmann-Gibbs distribution

$$\rho(x,t=\infty) \propto \exp\left[\Phi(x)/\epsilon \right] \ . \qquad (7)$$

In the mutation-drift case, we specify the adaptive landscape in Eq.(6) by substituting into Eqs.(4)(5) (we set $\epsilon = 1$ in the present construction)

$$\Phi(x) = (4N\nu - 1)\ln x + (4N\mu - 1)\ln(1-x) \ . \qquad (8)$$

From this expression, a bi-stable system emerges if we take $4N\nu, 4N\mu < 1$ (see Figure 1). Once such a "Boltzmann-Gibbs" potential energy has been defined, the escape time can usually be obtained by [11][12]

$$T_{escape} \sim |f|^{-1} \exp\left(\Delta\Phi\right) \ . \qquad (9)$$

Here $\Delta\Phi = \Phi(peak) - \Phi(valley)$, is the valley depth on a landscape. The exponential term $\exp\left(\Delta\Phi\right)$ is often called the "Arrhenius factor" [11]. However, this classical result fails in

the present model. As shown in Eq.(8), genetic drift generates two infinite adaptive peaks at $x = 0, 1$:

$$\Phi(0) = \Phi(1) = +\infty \ .$$

With finite $\Phi(x = a)$ (denote $x = a$ the valley state), we have $\Delta\Phi = \infty$. By Eq.(9), the escape time from $x = 0$ becomes infinite then, and it changes very quickly with x near 0. This is not a good estimation for the escape time. Mathematically, under $4N\nu = 4N\mu = 1$, the potential landscape becomes flat ($\Delta\Phi = 0$) and we obtain $T_{escape} \sim 1/|f|$; for any $4N\nu, 4N\mu < 1$, however, we obtain $T_{escape} = \infty$. The escape time does not change continuously with $4N\nu$ and $4N\mu$ near 1. Biologically, we expect that the evolutionary state would eventually reach an equilibrium distribution (though this process may be extremely slow), and for any finite-size system peak of the distribution would probably spread around. The question here is: Does the escape time really diverge with the infinite adaptive peak? Equivalently, does the infinity of equilibrium distribution at certain states imply the impossibility of peak-shifting events? If not, how does it depend on the parameters of the system? We try to answer these questions by first calculating the first passage time in the diffusion process.

C. First passage time

We come back to the standard formula for the first passage time, derived from the backward diffusion equation [12]. We later show how it can be used to estimate the escape time. We study the first passage event through some state $a < x_1 < 1$ (again $x = a$ is the valley state), starting from some state $0 < x = x_0 < a$. The average first passage time $T(x_0 \to x_1)$ (denoted here as T; it is a function of the initial state $x = x_0$) satisfies

$$\left[f(x) + \epsilon D'(x) \right] \partial_x T + \epsilon D(x)\partial_x^2 T = -1 \ .$$

To study the time to exit $(0, x_1)$ from x_1, we set the boundary conditions as ($x = 0$ reflecting and $x = x_1$ absorbing)

$$\partial_x T|_{x_0=0} = 0 \ ,$$
$$T|_{x_0=x_1} = 0 \ .$$

The solution is

$$T(x_0 \to x_1) = T \ ,$$
$$= \int_{x_0}^{x_1} \frac{1}{\epsilon D(y)} \exp\left[-\Phi(y) \right] dy \int_0^y \exp\left[\Phi(z) \right] dz \ . \qquad (10)$$

Here Φ is just our adaptive landscape in Eq.(6). We specify Eq.(10) under weak mutation ($4N\nu, 4N\mu < 1$) and drift by substituting Eqs.(5, 8), and there is

$$T(x_0 \to x_1) = 4N \int_{x_0}^{x_1} y^{-4N\nu}(1-y)^{-4N\mu} dy$$
$$\cdot \int_0^y z^{4N\nu-1}(1-z)^{4N\mu-1} dz \ . \qquad (11)$$

The second integral is just the integral of $\rho(x,t=\infty)$ in $[0,y]$, having the form of incomplete Beta function. No previous
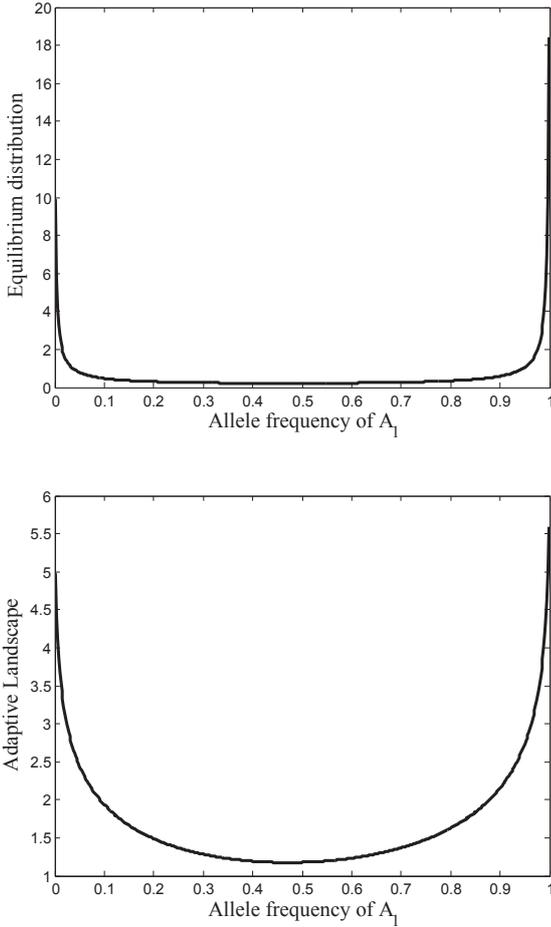
Fig. 1. Equilibrium distribution and the corresponding adaptive landscape for a typical bi-stable system under mutation and drift. The parameter is set as $1/4N > \nu > \mu$ (drift dominates mutation); the peak at $x = 1$ is higher than that at $x = 0$ on each figure. (a) As $\nu, \mu \to 0$, the probabilities would accumulate more sharply at the boundaries; when $\nu = \mu = 0$, it develops two Dirac delta functions. (b) The U-shaped landscape shows a "fat" valley at the middle states and two sharp peaks at the boundaries. A population may wander some time at the middle (heterozygotic) states, but would be absorbed quickly into the boundary (homozygotic) state once appearing near one. As $\nu, \mu \to 0$, the landscape would continuously converge to a symmetric configuration; there would be no sudden change in the limit $\nu = \mu = 0$.

expansion was taken on this divergent equilibrium distribution, though [6] claimed that their eigenvalue method can be applied in the Gamma-like divergent equilibrium (see Section III.B). We describe our expansion of Eq.(11) as follows. Under $0 < z < y < x_1$, there is $0 < 1 - x_1 < 1 - y < 1 - z < 1$; the expansions on the exponential terms of $1 - y$ and $1 - z$ would converge near 0. First we expand $(1 - z)^{4N\mu-1}$ with respect to z in the second integral of of Eq.(11) (denoted as $I(y)$):

$$I(y) = \int_0^y z^{4N\nu-1}(1-z)^{4N\mu-1}dz ,$$
$$= \int_0^y z^{4N\nu-1}\left[1 + \sum_{n=1}^{\infty}\prod_{k=1}^{n}\left(\frac{k-4N\mu}{k}\right)z^n\right]dz . \quad (12)$$

The convergence is obvious given $0 < y < x_1 < 1$. Substitute $I(y)$ and expend $(1-y)^{-4N\mu}$ in the first integral of $T(x_0 \to x_1)$, we obtain the expansion result

$$T(x_0 \to x_1)$$
$$= 4N\int_{x_0}^{x_1}\left[1 + \sum_{n=1}^{\infty}\prod_{k=1}^{n}\left(\frac{k-1+4N\mu}{k}\right)y^n\right]I(y)dy ,$$
$$= \frac{x_1 - x_0}{\nu} + \frac{4N\mu}{\nu}\sum_{n=1}^{\infty}\prod_{k=2}^{n}\left(\frac{k-1+4N\mu}{k}\right)\frac{x_1^{n+1}-x_0^{n+1}}{n+1}$$
$$+ 4N(1-4N\mu)\sum_{n=1}^{\infty}\prod_{k=2}^{n}\left(\frac{k-4N\mu}{k}\right)\frac{x_1^{n+1}-x_0^{n+1}}{(n+1)(n+4N\nu)} .$$
$$(13)$$

With $x_1, x_0 < 1$, the convergence of this expansion is obvious. We discuss its convergence under two limiting conditions:

(1) Under $x \to 0$ and $\nu \to 0$. For any non-zero ν, our result for escape time is not sensitive to x_0 near 0, and we can take $x_0 \to 0$. It is of the essence that the incomplete Beta function in Eq.(12) is of near-linear dependence on the integral interval near the boundary $y = 0$, given $\nu > 0$. For $\nu = 0$, however, the expansion in Eq.(12) will not be valid. We have instead

$$I(y) = \ln y + \sum_{n=1}^{\infty}\prod_{k=1}^{n}\left(\frac{k-4N\mu}{k}\right)\frac{y^n}{n} .$$

The leading term changes from a polynomial of order $\sim y^{-1}$ to the logarithmic scale ($\sim \ln y$). The result becomes sensitive to the value of x_0 near 0 and approaches infinity at $x_0 = 0$. To ensure the convergence of the escape time as $x_0 \to 0$ under weak mutation, we should first take $x_0 \to 0$, and then we can set $\nu \to 0$. Nevertheless, as discussed above, the result $T(x_0 \to a) \to \infty$ is consistent for any way of taking $\nu \to 0$.

(2) Under $x \to 1$ and $\mu \to 1/4N$. When $4N\mu = 1$, the expansion of $(1 - y)^{-4N\mu}$ in Eq.(13) would not converge for $x_1 = 1$, as the series

$$\sum_{n=2}^{\infty}\prod_{k=2}^{n}\left(\frac{k-1+4N\mu}{k}\right)\frac{x_1^{n+1}}{n+1} \quad (14)$$

would then become a harmonic series (with logarithmic divergence rate) in maths. We return to Eq.(11) for $4N\mu = 1$:

$$T(x_0 \to x_1) = 4N\int_{x_0}^{x_1}y^{-4N\nu}(1-y)^{-1}dy\int_0^y z^{4N\nu-1}dz ,$$
$$= \nu^{-1}\ln\frac{1-x_0}{1-x_1} .$$

As $x_1 \to 1$, this result approaches infinity. This is also illustrated by the vanishing bi-stability of the system. To ensure the validation of the expansion, we require at least that $x_1 < 1$ or $4N\mu < 1$.

III. RESULTS AND DISCUSSIONS

A. Escape time revisited

We come back to the average time to escape from $x = 0$, after which the population is not expected to return immediately. The escape time can be derived from above first passage

time by setting $x_0 \to 0, x_1 \to 1$. This statement comes from our following observations on the landscape configuration shown in Figure 1b: Near $x = a$, the landscape valley is flatly distributed and a population may wander there for some time before approaching either boundary state; and for any $a < x_1 < 1$ (x_1 not near 1), it may still come back to $x = 0$ with considerable possibilities after reaching x_1; near $x = 1$, the landscape peak grows sharply, so that a population would be caught stable there (not easily returns to $x = 0$). Thus the escape time can be approximately considered as the average time taken to reach another stable state (the substitution time in population genetics [9]). By taking $x_0 \to 0, x_1 \to 1$ in Eq.(13), we obtain Eq.(16). The convergence of Eq.(16) (mainly the series in Eq.(14)) under $4N\mu < 1$ can be verified by Raabe's test, which can be found in a typical textbook on Mathematical Analysis.

We make a further comment on Eq.(16). Under $4N\nu \ll 1$, the scale of $1/\nu$ would be much bigger than the remaining terms (of order $2N\mathcal{O}(1)$, the well-known characteristic timescale of genetic drift [22]). The interesting thing is that even though the expansion is taken on frequency x, it naturally separates the two distinct timescales, and provides a good analytical approximation for the escape time in Eq.(11). If instead $4N$ and ν^{-1} are comparable (under the constraint $4N\nu < 1$), the expansion in Eq.(16) is still valid, but there would be no clear separation between the leading term and the rest. It implies a mixed timescale of uphill and downhill movements on the landscape, illustrated by its near-flat slopes.

Assume $4N\nu, 4N\mu \ll 1$ and take the leading terms in Eq.(16), we get

$$T(0 \to 1) = \frac{1 + 2N\mu}{\nu} + 2N\mathcal{O}(1) . \qquad (17)$$

Comparisons among the numerical solution of the classical formula Eq.(11), our analytical approximation in Eq.(17), and the results in the discrete Wright-Fisher model (derived from the Master equation of Eq.(1) [23]) is given in Figure 2.

From Eq.(17), the escape time increases with N and μ; this is consistent with the biological interpretations. For increasing μ, the stronger backward mutation makes the forward transition more difficult; for $4N\mu \ll 1$, the effect is near-linear; for $4N\mu \geq 1$, it makes the transition to $x = 1$ impossible. The fixation of A_1 becomes phenomenally impossible because the attractive basin $(a, 1)$ vanishes and $x = 1$ becomes an unstable state (potential valley) then. For increasing N, it has effects from two sides: increase the number of new mutants each generation and decrease the intensity of random drift. By Eq.(17), the net effect is to increase the escape time, though in a minor order if $4N\nu \ll 1$.
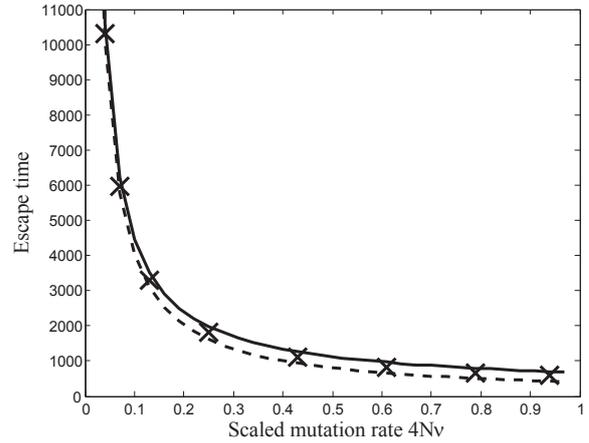

Fig. 2. Comparisons among the numerical solution of Eq.(15) (solid line), analytical approximation in Eq.(17) (dashed line), and results in the discrete model (crosses) in the range $0 < 4N\nu < 1$. Other parameter settings are: $\mu = 0.00005$, $N = 100$. The leading terms in the expansion Eq.(17) approximates the exact solutions well for $4N\nu \ll 1$; when $4N\nu \approx 1$, the two timescales are comparable and the remaining terms becomes important, so the error becomes major. The numerical solution of the discrete formula matches well with the continuous results for the whole range $0 < 4N\nu < 1$, though is found generally smaller than the latter. This may be explained by that a diffusion equation is the second order approximation of a master equation.

If we further take $\mu = 0$, Eq.(17) becomes

$$T(0 \to 1) \approx \nu^{-1} . \qquad (18)$$

The result is approximately independent of the population size N. It reminds us of the well-known formula for the rate of substitution for the neutral mutants: $2N\nu \times 1/2N = \nu$ [9]; Eq.(18) is just its inverse. We note that Eq.(18) is not simply the characteristic timescale for mutation, but coincides with it under the limit $4N\nu \ll 1$. The driving force for the process is mainly random drift ($1/2N \gg \nu$), without which the stochastic escape cannot happen. Simulation realized from single population dynamics are shown in Figure 3.

B. Compare with previous results

[11] derived the rate of escaping over a potential barrier from the diffusion equation, by calculating the probability current at the saddle point (our valley state). He assumed that the main contribution to the escape time is due to a small region near the saddle. His analytical approximation requires a finite barrier height, and that the equilibrium distribution near the potential well be approximately Gaussian. These assumptions generate an "Arrhenius exponential factor" as in Eq.(9) in the calculations. Other methods developed from the backward

$$T(0 \to 1) = 4N \int_0^1 y^{-4N\nu}(1 - y)^{-4N\mu} dy \cdot \int_0^y z^{4N\nu-1}(1 - z)^{4N\mu-1} dz , \qquad (15)$$

$$= \frac{1}{\nu} + \frac{1}{\nu}\left[2N\mu + 4N\mu \sum_{n=2}^{\infty} \prod_{k=2}^{n} \left(\frac{k - 1 + 4N\mu}{k} \right) \frac{1}{n+1} \right] + 4N(1 - 4N\mu) \sum_{n=1}^{\infty} \prod_{k=2}^{n} \left(\frac{k - 4N\mu}{k} \right) \frac{1}{(n+1)(n+4N\nu)} . \qquad (16)$$
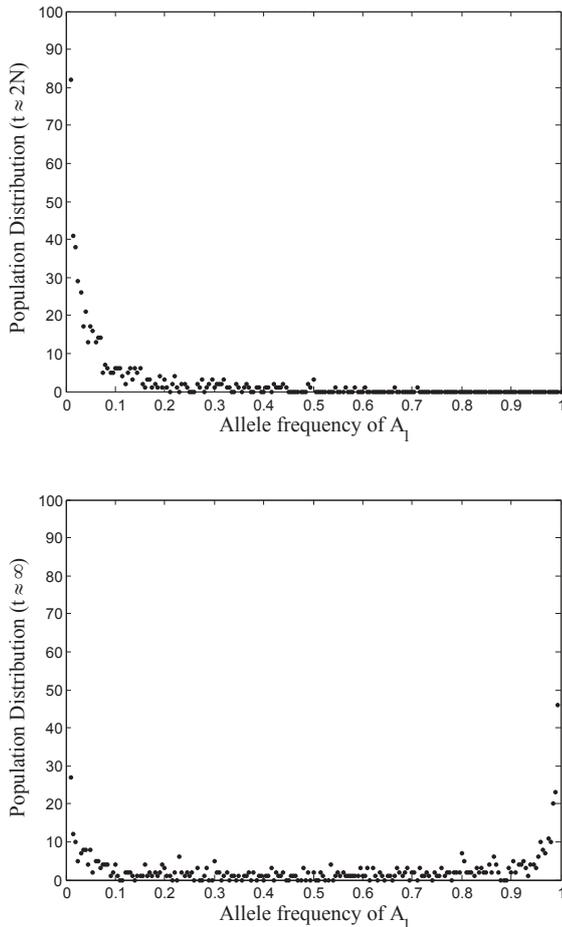
Fig. 3. Simulations realized from the stochastic differential equation (SDE) corresponding to Eq.(1). Parameter settings are: $N = 150, \mu = 0.00005, \nu = 0.00015$. The number of populations (denoted as particles in the figures) simulated is 1000, all of which are initiated at $x_0 = 0$, thus $\rho(x = x_0) = \delta(x)$. Observations of the first passage time $0 \to 1$ are made for each population. (a) Snapshot of population distribution at time $t \approx 2N$. States of populations spread around near $x = 0$, the local attractive state. (b) Snapshot of population distribution at time $t \approx \infty$ (equilibrium). Each population has at least passed the global stable state $x = 1$ once. Most populations are centered around $x = 1$ or $x = 0$. The simulated average first passage time is 6656.5 (the theoretical expectation is $1/\nu = 6666.67$).

diffusion equation obtained similar results [12]. In population genetics, however, strong random genetic drift would generate non-Gaussian (even divergent) equilibrium distributions. The adaptive landscape is typically flat at the middle states and divergent at the boundaries (Figure 1b). This constrains the application of the classical escape formulae in certain cases: The peak-shifting rates are often numerically approximated or can be analytically discussed only under very special settings in population genetics [24][5]. Our expansion Eq.(13) taken on the escape formula Eq.(11) near the divergent state makes it possible to analytically approximate the results in such cases. It is proved to be valid for all the conditions where a bi-stable adaptive landscape is maintained. Our results show that the finiteness of the escape time does not require the finiteness of

the potential peak (or barrier height), but corresponds to the finiteness of the partial integral of equilibrium distribution in that attractive basin (see Eq.(12)).

As mentioned previously, our result in Eqs.(17)(18) can be compared with that derived as the rate of substitution in population genetics [3][4]. First, the estimation of neutral substitution rate ν [9] only gives a characteristic rate for the fixation of A_1 to take place. We cannot have a detailed knowledge about how the transition time depends on other factors of the system. An example is the effect of population size N, though it may be of minor impacts under $\nu \ll 1/4N$. If instead ν and $1/4N$ become comparable, the simple estimation would not work any more. We've shown that our analytical approximation is valid for the whole range $0 < 4N\nu < 1$; when $4N\nu$ is not very small, the timescales of the uphill and downhill movements become mixed and the substitution time should be expressed by a more complex form. Second, the evolutionary transition time calculated in the present work can be more generally applied in population genetics than the substitution time, because all evolutionary factors are here treated in a unified and consistent manner. For example, it allows the existence of considerable backward mutation ($0 < \mu < 1/4N$), which may make the fixation probability of A_1 incalculable.

[6] proposed to calculate the leading (non-zero) eigenvalue of Eq.(1), which gives the flux between different equilibria. The authors claimed that their general expression of transition rate agrees with the two approaches (backward diffusion equation and rate of substitution) in the limits of very high and very low mutation, respectively. Our first comment is that, in this sense, our expansion also bridges the two approaches in the two extreme conditions — we make the first method applicable in the second situation. Second, their method failed under very weak selection $s < 4\mu$. It requires the existence of two peaks on the "deterministic equilibrium", which does not necessarily capture the correct dynamical features of the system. For example, even in the neutral case ($s = 0, 4N\mu, 4N\nu < 1$), the equilibrium distribution (and our adaptive landscape) has two local maxima, implying the long-term bi-stability. We've showed that our result can be applied under strong genetic drift. Third, part of their solution is numerical and can only be analytically expressed under limiting conditions ($4N\mu \ll 1$ or $4N\mu \gg 1$). Our method, in contrast, can be used for analytical results not only under extreme conditions, but in the cases with intermediate parameter values (when ν and $1/4N$ are comparable). It naturally separates two distinct timescales from the complex form of Eq.(11), and gives a good analytical approximation for the escape time. Our result provides a complete answer for the present bi-stable models.

C. Our adaptive landscape

Our construction of the adaptive landscape in Eq.(6) comes principally from the novel decomposition of system's dynamical components [17][18]. It is closely associated with the equilibrium distribution. The adaptive landscape corresponds to the potential energy in a Boltzmann-Gibbs distribution

as shown in Eq.(7), if the normalization constant in Eq.(3) is finite. It defines a measure of probabilistic potential, and all dynamical factors are treated equivalently. It gives a full picture of the system's middle-and-long-term behaviors. This can be compared to the classical fitness landscape, which only presents the effect of selection. Other effects, such as mutation and drift, are not integrated into the expression and require separated attentions. Wright's shifting balance theory, for example, tried to describe the evolution under the interaction of selection, mutation, and random drift on the fitness landscape. It has been under controversy whether this "artificial" mechanism is truly relevant.

An extension to the classical fitness landscape is the so-called "deterministic equilibrium" [6], which integrates all the "deterministic" factors (mutation, selection, etc.) of the system and defines a landscape. Compared to the "stochastic equilibrium" (system's equilibrium distribution), it does not include genetic drift. Such definition does not match exactly to the long-term behaviors of the system, just like the fitness landscape. One example is the case of weak selection, where there is only one peak on their deterministic equilibrium, and their method fails. We've shown that strong genetic drift would lead to bi-stable distribution in the long-term evolution and at equilibrium. Our treatment for the non-Gaussian distribution can be applied there.

In analogy with thermodynamics, [25] defined three scalar functions in biological evolution: a (negative) energy for selection, a potential function for mutation, and an entropy for random drift. Our present work, instead, integrates all above factors into an evolutionary adaptive landscape, and treat them with unified viewpoint and approaches. Other factors (e.g. frequency-dependent selection) can be readily added into our formulas. [26] used the analogy to thermodynamics to study the evolution of macroscopic quantitative traits, in a way that is independent of the microscopic details. However, there are cases (e.g. when mutation rate is low) that their maximum entropy approximation would not work. Also, their method requires a normalizable equilibrium distribution. They used the normalization constant as the generating function for macroscopic variables, which plays a "major role" in their calculations. Our method does not have such restrictions and the equilibrium distribution can be unnormalizable. As $\nu, \mu \to 0$, the equilibrium distribution in Figure 1a would develop Dirac delta functions at the boundary [27]; our adaptive landscape, instead, changes continuously under this limit (see Eq.(8)).

IV. Conclusion

We have calculated the expected time to escape from an infinite adaptive peak. We've shown that such an infinite adaptive peak does not necessarily mean biological fixation, and the peak-shifting time can be analytically approximated. Our results extend the application of Kramers' classical formula to the non-Gaussian distribution cases, and break through the limitations of previous efforts. Our adaptive landscape has served as a both qualitative and quantitative tool for describing the system's middle-and-long-term behaviors and guiding our calculations. Our work has provided a complete description for the bi-stabilities in the present model.

References

[1] R. A. Fisher, *The Genetical Theory of Natural Selection.* Clarendon, Oxford, 1930.
[2] S. Wright, "Evolution in mendelian populations," *Bull. Math. Biol.,* vol. 52, pp. 241–295, 1931.
[3] M. Kimura, "Some problems of stochastic processes in genetics," *Ann. Math. Stat.,* vol. 28, pp. 882–901, 1957.
[4] ——, "On the probability of fixation of mutant genes in a population," *Genetics,* vol. 47, pp. 713–719, 1962.
[5] R. Lande, "Expected time for random genetic drift of a population between stable phenotypic states," *Pr,* vol. 82, pp. 7641–7645, 1985.
[6] N. H. Barton and S. Rouhani, "The frequency of shifts between alternative equilibria," *J. Theor. Biol.,* vol. 125, pp. 397–418, 1987.
[7] S. Gavrilets, "Evolution and speciation on holey adaptive landscape," *Trends Ecol. Evol.,* vol. 12, pp. 307–312, 1997.
[8] P. Ao, "Global view of bionetwork dynamics: adaptive landscape," *J. Genet. Genomics,* vol. 36, pp. 63–73, 2009.
[9] J. H. Gillespie, *Population Genetics: A Concise Guide.* The Johns Hopkins University press, Baltimore, 1998.
[10] D. Zhou and H. Qian, "Fixation, transient landscape, and diffusion dilemma in stochastic evolutionary game dynamics," *Phys. Rev. E,* vol. 84, pp. 031907–1, 2011.
[11] H. A. Kramers, "Brownian motion in a field of force and the diffusion model of chemical reactions," *Physica,* vol. 7, pp. 284–304, 1940.
[12] C. W. Gardiner, *Handbook of Stochastic Methods.* Springer-Verlag, Berlin, 1985.
[13] D. Waxman and S. Gavrilets, "20 questions on adaptive dynamics," *J. Evol. Biol.,* vol. 18, pp. 1139–1154, 2005.
[14] S. Wright, "The roles of mutation, inbreeding, crossbreeding and selection in evolution," in *Proc. Sixth Int. Cong. Genetics,* 1932.
[15] W. Ewens, *Mathematical Population Genetics.* Springer-Verlag, New York, 2004.
[16] N. G. van Kampen, *Stochastic Processes in Physics and Chemistry.* Elsevier, Amsterdam, 2007.
[17] P. Ao, "Potential in stochastic differential equations: novel construction," *J. Phys. A: Math. Gen.,* vol. 37, pp. L25–L30, 2004.
[18] P. Ao, C. Kwon, and H. Qian, "On the existence of potential landscape in the evolution of complex systems," *Complexity,* vol. 12, pp. 19–27, 2007.
[19] X. M. Zhu, L. Yin, L. Hood, and P. Ao, "Caltulating biolgoical behaviors of epigentic states in the phage lambda life cycle," *Funct. Integr. Genomics,* vol. 4, pp. 188–195, 2004.
[20] P. Ao, "Laws in darwinian evolutionary theory," *Phys. Life Rev.,* vol. 2, pp. 117–156, 2005.
[21] P. Ao, D. Galas, L. Hood, and X. M. Zhu, "Cancer as robust intrinsic state of endogenous molecular-cellular network shaped by evolution," *Med. Hypotheses,* vol. 70, pp. 678–684, 2008.
[22] M. Kimura, "Diffusion models in population genetics," *J. Appl. Probab.,* vol. 1, pp. 177–232, 1964.
[23] R. A. Blythe and A. J. McKane, "Stochastic models of evolution in genetics, ecology and linguistics," *J. Stat. Mech. Theor. Exp.,* vol. 07, p. P07018, 2007.
[24] M. Kimura and T. Ohta, "The average number of generations until fixation of a mutant gene in a finite population," *Genetics,* vol. 61, pp. 763–771, 1969.
[25] N. H. Barton and J. B. Coe, "On the application of statistical physics to evolutionary biology," *J. Theor. Biol.,* vol. 259, pp. 317–324, 2009.
[26] N. H. Barton and H. P. de Vladar, "Statistical mechanics and the evolution of polygenic quantitative traits," *Genetics,* vol. 181, pp. 997–1011, 2009.
[27] A. J. McKane and D. Waxman, "Singular solutions of the diffusion equation of population genetics," *J. Theor. Biol.,* vol. 247, pp. 849–858, 2007.