

The probability of fixation of a single mutant in an exchangeable selection model

Sabin Lessard · Véronique Ladret

Received: 13 July 2006 / Revised: 30 November 2006 /
Published online: 25 January 2007
© Springer-Verlag 2007

Abstract The Cannings exchangeable model for a finite population in discrete time is extended to incorporate selection. The probability of fixation of a mutant type is studied under the assumption of weak selection. An exact formula for the derivative of this probability with respect to the intensity of selection is deduced, and developed in the case of a single mutant. This formula is expressed in terms of mean coalescence times under neutrality assuming that the coefficient of selection for the mutant type has a derivative with respect to the intensity of selection that takes a polynomial form with respect to the frequency of the mutant type. An approximation is obtained in the case where this derivative is a continuous function of the mutant frequency and the population size is large. This approximation is consistent with a diffusion approximation under moment conditions on the number of descendants of a single individual in one time step. Applications to evolutionary game theory in finite populations are presented.

Keywords Exchangeable model · Coalescence times · Diffusion approximation · Evolutionary game theory · Fixation probability

Mathematics Subject Classification (2000) 60J70 · 91A22 · 92D25

1 Introduction

Recently, one of the major issues in evolutionary game theory has been to incorporate a finite number of interacting individuals (e.g., [21,22,31] see also [3,14,24,26,27], for sex ratio evolution, kin selection or frequency-depen-

S. Lessard (✉) · V. Ladret
Département de mathématiques et de statistique, Université de Montréal,
Montréal, QC H3C 3J7, Canada
e-mail: lessards@dms.umontreal.ca

dent selection in structured populations). Evolutionarily stable strategies, ESS [15, 16], and related concepts based originally on fitness comparisons or stability conditions of fixation states in virtually infinite populations are currently revisited on the basis of fixation probabilities when finite populations are considered. If the population size and the mutation rate are small enough, it is reasonable to assume that mutants are introduced one at a time into previously fixed populations and that successive fixation states determine the course of long-term evolution (see, for comparison, [12] for long-term evolution in infinite populations). This raises interest for the exact conditions ensuring the replacement of one type with another as the result of the combined effects of selective forces and random genetic drift.

In [13], we have deduced a first-order approximation for the probability of fixation of a mutant type with respect to the intensity of selection for a wide variety of frequency-dependent selection models using either exact solutions for models of the Moran type for populations of any finite size N or diffusion approximations for models of the Wright–Fisher type for populations of large size N . When this probability is less than the initial frequency of the mutant type, then selection opposes the mutant type replacing the wild type, or equivalently, the wild type is favored against replacement with the mutant type. Of particular interest, in relation with the ESS concept of a resident strategy uninvadable by any mutant strategy, is finding a wild type that is favored against replacement with any mutant type at least when selection is weak enough.

In [13], we have also pointed out a major difference between viability models and fertility models when selection depends on randomly pairwise interactions between individuals and the population size is finite. The difference stems from the fact that the probability distribution of types for two interacting individuals are not the same in both models. The analysis of the Wright–Fisher model in the case of fertility selection was only heuristic however, since terms of order $1/N$ arising from excluding interactions of individuals with themselves were kept in the drift parameter of the approximating diffusion process, while this approximation should be valid only in the limit as N goes to infinity.

An interesting aspect of exchangeable models is that they allow us to consider situations more general than the Wright–Fisher model, that is, models in which the numbers of descendants left by the N individuals of the population in one time step do not necessarily follow a multinomial distribution. This is the case, for instance, in a conditional branching process given only that the population size is kept constant (see, e.g., [5, p. 103], [11], and references therein). This is also the case when the parameters of the multinomial distribution are themselves random variables. In particular, if these parameters have a constant mean but a variable variance, then it becomes possible to study the evolution of the variance in the population. In the case of haploid sex allocation, for instance, an evolutionary decrease in the variance of the individual strategy in support of sex ratio homeostasis has been predicted, since only a zero variance is favored against replacement with any mutant under weak selection [13].

The fixation probability in a Wright–Fisher model with frequency-dependent selection can be studied directly using, e.g., total-positivity arguments, and

approximated in the case of weak selection applying a first-step analysis and assuming some regularity conditions on the probability of fixation as a function of the intensity of selection [8]. The interpretation of the coefficients coming into play in such first-order approximations in terms of mean coalescence times as in [25] in a context of a haploid island model is particularly appealing, since it makes possible an extension of the results to more general settings.

Our first objective in this paper will be to ascertain the validity of the exact first-order approximation for the probability of fixation of a single mutant in terms of mean coalescence times in a finite population assuming weak frequency-dependent selection and mild regularity conditions on the transition probabilities for the frequency of the mutant. Our second objective will be to study the fixation probability in a more general setting than the Wright–Fisher model by supposing only that the numbers of descendants left by individuals of the same type in one time step are exchangeable variables. The selection model that will be considered can be seen as an extension of the neutral exchangeable model known as the Cannings model [2]. Therefore, the results obtained on the fixation probability, as well as conclusions that can be drawn from it, will prove to be robust in the sense they will apply to a wide range of models.

2 Exchangeable selection model

The Cannings exchangeable model was originally considered under neutrality [2], but it can be extended to incorporate selection. This can be done as follows. Consider a population of N genes or individuals, reproducing at discrete times $t = 0, 1, 2, \dots$. Assume two types of individuals, A and B . Given k individuals of type A labelled $1, \dots, k$ and $N - k$ individuals of type B labelled $k + 1, \dots, N$ at some current time t , the numbers of descendants left by the different individuals at the next time $t + 1$, possibly including some of the individuals at time t still present in the population at time $t + 1$, are random variables represented by $z_1(x), \dots, z_k(x)$ for the A individuals and $z_{k+1}(x), \dots, z_N(x)$ for the B individuals, where $x = k/N$ is the frequency of the A individuals at time t . The key assumption is that every marginal distribution of $(z_1(x), \dots, z_N(x))$ depends only on the numbers of variables among $z_1(x), \dots, z_k(x)$ and $z_{k+1}(x), \dots, z_N(x)$ that are present. In other words, the variables $z_1(x), \dots, z_k(x)$ are exchangeable, and the same for $z_{k+1}(x), \dots, z_N(x)$. This exchangeability assumption defines the exchangeable selection model.

In particular, the mean of $z_1(x), \dots, z_k(x)$ is assumed to be the same, suppose $\mu_A(x)$, and similarly for the mean of $z_{k+1}(x), \dots, z_N(x)$, represented by $\mu_B(x)$. Since

$$z_1(x) + \dots + z_N(x) = N, \quad (1)$$

we require that

$$x\mu_A(x) + (1 - x)\mu_B(x) = 1. \quad (2)$$

In the case of neutrality, the variables $z_1(x), \dots, z_N(x)$ do not actually depend on x and are represented by z_1, \dots, z_N . These variables are all exchangeable with common mean equal to 1. This is also the case when $x = 0$ or 1.

In order to allow random drift, the numbers of descendants left by the different individuals in one time step are assumed to have some positive variance, namely $\sigma_A^2(x) > 0$ for $z_1(x), \dots, z_k(x)$ and $\sigma_B^2(x) > 0$ for $z_{k+1}(x), \dots, z_N(x)$. Denoting by $\text{Cov}_{AA}(x)$, $\text{Cov}_{BB}(x)$ and $\text{Cov}_{AB}(x)$ the covariances between two variables of the first set, two variables of the second set and one variable of the first and one of the second, respectively, we have the relationship

$$\begin{aligned} k\sigma_A^2(x) + (N - k)\sigma_B^2(x) &= -k(k - 1)\text{Cov}_{AA}(x) \\ &\quad - (N - k)(N - k - 1)\text{Cov}_{BB}(x) \\ &\quad - 2k(N - k)\text{Cov}_{AB}(x). \end{aligned} \quad (3)$$

This is a consequence of the exchangeability assumption under the constraint (1).

Weak selection is modelled by introducing a small positive parameter s , called the intensity of selection, which is such that the difference between the means $\mu_A(x)$ and $\mu_B(x)$ for $0 < x < 1$, called the coefficient of selection for A against B , can be expressed as

$$\mu_A(x) - \mu_B(x) = s(h_N(x) + O(s)), \quad (4)$$

where $O(s)$ designates a function of order s . Owing to (2), the means $\mu_A(x)$ and $\mu_B(x)$ for $0 < x < 1$ take the form

$$\mu_A(x) = 1 + s(1 - x)(h_N(x) + O(s)) \quad (5)$$

and

$$\mu_B(x) = 1 - sx(h_N(x) + O(s)), \quad (6)$$

respectively.

The function $h_N(x)$, for $0 \leq x \leq 1$, represents a rate of change in the coefficient of selection, taken with respect to the intensity of selection and evaluated at neutrality, that is, with respect to s at $s = 0$. It does not depend on the intensity of selection itself, but it may depend on the population size, N , besides being frequency-dependent. The factor $1 - x$ in (5) ensures that all terms of order smaller or equal to s vanish as x goes to 1, which corresponds to a continuity assumption on the account that $\mu_A(1) = 1$. Symmetrically, all terms of order smaller or equal to s vanish in (6) as x goes to 0. Examples of functions $h_N(x)$, most often in a polynomial form, can be found in the literature, e.g., in the case of frequency-dependent viability or fertility selection or in the case of sex allocation (see, e.g., [13], and references therein). Some are given in the next section.

Under weak selection, the variances $\sigma_A^2(x)$ and $\sigma_B^2(x)$ take the form $\sigma_N^2 + O(s)$ and the covariances $\text{Cov}_{AA}(x)$, $\text{Cov}_{BB}(x)$ and $\text{Cov}_{AB}(x)$ the form $-c_N + O(s)/N$, respectively, where

$$c_N = \frac{\sigma_N^2}{N - 1}, \tag{7}$$

with $\sigma_N^2 > 0$, owing to (3) in the case $s = 0$. The quantity c_N corresponds to the probability that two individuals randomly chosen at a given time have the same parent, possibly one of the two individuals, one step backward in time in the absence of selection. It will be assumed later on that σ_N^2 is uniformly bounded with respect to N , which is the case, e.g., for the Moran model and the Wright–Fisher model, and which ensures that c_N converges to 0 as N goes to infinity.

The frequency of A at time $t + 1$ is given by

$$\frac{z_1(x) + \dots + z_k(x)}{N}, \tag{8}$$

whose expected value is $x\mu_A(x)$ and variance is

$$\frac{k\sigma_A^2(x) + k(k - 1)\text{Cov}_{AA}(x)}{N^2}. \tag{9}$$

This variance depends on s and can be written as

$$v_s(x) = x(1 - x)c_N + O(s)/N, \tag{10}$$

where $\sigma_N^2 > 0$ is the variance in the number of descendants left by a single individual in one time step under neutrality. On the other hand, the change in the frequency of A has mean $x\mu_A(x) - x$, which depends also on s and takes the form

$$m_s(x) = sx(1 - x)(h_N(x) + O(s)). \tag{11}$$

The value of the expected change evaluated at $s = 0$ is

$$m_0(x) = 0 \tag{12}$$

while the value of its derivative with respect to s , denoted by $m'_s(x)$ and evaluated at $s = 0$, is

$$m'_0(x) = x(1 - x)h_N(x). \tag{13}$$

Some illustrative examples follow.

3 Examples

Let

$$f_A(x) = 1 + su_A(x) \quad (14)$$

and

$$f_B(x) = 1 + su_B(x) \quad (15)$$

be the fitnesses of A and B individuals, respectively, when the frequencies of A and B in the population are x and $1 - x$, respectively. Terms of order $o(s)$ are ignored in these fitnesses. The functions $u_A(x)$ and $u_B(x)$ are constant in the case of frequency-independent selection. They are linear functions of x if they result from a large number of random pairwise interactions as in linear evolutionary game theory [15, 16], but more generally linear fractional transformations of x if the interaction rates depend on the types of the interacting individuals [30]. These functions may depend on N . Moreover, we will assume in this section that they are deterministic, but they may be stochastic as it occurs in studies for sex ratio homeostasis or in conditional branching processes (see, e.g., [11, 13, 24]).

In a haploid Wright–Fisher model [6, 32], time steps correspond to non overlapping generations and the parent of every individual in the next generation is chosen independently of all others with a probability proportional to its fitness. Therefore, the vector $(z_1(x), \dots, z_k(x), z_{k+1}(x), \dots, z_N(x))$ for the numbers of descendants left in the next generation by the Nx individuals of type A and the $N(1 - x)$ individuals of type B in this order will have a multinomial distribution of parameters $(N; q_1, \dots, q_k, q_{k+1}, \dots, q_N)$ given by

$$q_1 = \dots = q_k = \frac{f_A(x)}{N\bar{f}} = \frac{1 + su_A(x)}{N(1 + s\bar{u})} \quad (16)$$

and

$$q_{k+1} = \dots = q_N = \frac{f_B(x)}{N\bar{f}} = \frac{1 + su_B(x)}{N(1 + s\bar{u})}, \quad (17)$$

where

$$\bar{f} = xf_A(x) + (1 - x)f_B(x) = 1 + s(xu_A(x) + (1 - x)u_B(x)) = 1 + s\bar{u} \quad (18)$$

represents the mean fitness. Then, the difference between the mean numbers of descendants left by single individuals of types A and B , respectively, is

$$\mu_A(x) - \mu_B(x) = \frac{f_A(x) - f_B(x)}{\bar{f}} = \frac{s(u_A(x) - u_B(x))}{1 + s\bar{u}}, \quad (19)$$

which is in the form $s(h_N(x) + O(s))$, where

$$h_N(x) = u_A(x) - u_B(x). \quad (20)$$

Moreover, the variances are

$$\sigma_A^2(x) = \frac{f_A(x)}{\bar{f}} \left(1 - \frac{f_A(x)}{N\bar{f}} \right) \quad (21)$$

and

$$\sigma_B^2(x) = \frac{f_B(x)}{\bar{f}} \left(1 - \frac{f_B(x)}{N\bar{f}} \right), \quad (22)$$

which are in the form $\sigma_N^2 + O(s)$, where

$$\sigma_N^2 = 1 - \frac{1}{N} \quad (23)$$

is the variance in the number of descendants left in the next generation by a single individual under neutrality.

In a Moran model [20], generations are overlapping and at each time step an individual chosen with a probability proportional to its fitness produces an offspring and this offspring replaces an individual chosen at random among all individuals but the parent. Then, the numbers of descendants left by the N individuals at the next time step form a vector $(z_1(x), \dots, z_k(x), z_{k+1}(x), \dots, z_N(x))$ which is a permutation of $(2, 0, 1, \dots, 1)$ and whose distribution is determined by the type of the individual leaving two descendants in one time step, including the individual itself, and the type of the individual being replaced and leaving none. Since an individual of type A leaves two descendants in one time step with probability

$$P(z_1(x) = 2) = \frac{f_A(x)}{N\bar{f}}, \quad (24)$$

one with probability

$$P(z_1(x) = 1) = \left(1 - \frac{f_A(x)}{N\bar{f}} \right) \left(1 - \frac{1}{N-1} \right), \quad (25)$$

and zero otherwise, we find that

$$\mu_A(x) = 1 - \frac{1}{N-1} + \frac{f_A(x)}{(N-1)\bar{f}}, \quad (26)$$

and similarly for $\mu_B(x)$, so that

$$\mu_A(x) - \mu_B(x) = \frac{f_A(x) - f_B(x)}{(N-1)\bar{f}}, \quad (27)$$

from which

$$h_N(x) = \frac{u_A(x) - u_B(x)}{N-1}. \quad (28)$$

Moreover, ignoring terms of order s , the variances are

$$\sigma_N^2 = \frac{2}{N}. \quad (29)$$

Notice that the Moran model can be extended by allowing the individual chosen with a probability proportional to its fitness to produce either $N\psi - 1$ offspring with probability $N^{-\alpha}$ or 1 with probability $1 - N^{-\alpha}$, which replace the same number of individuals chosen at random among all individuals but the parent [4]. The parameter α is positive, while the parameter ψ is a proportion such that $N\psi$ is a positive integer. Then, the vector $(z_1(x), \dots, z_k(x), z_{k+1}(x), \dots, z_N(x))$ is a permutation of $(2, 0, 1, \dots, 1)$ with probability $N^{-\alpha}$ and a permutation of $(N\psi, 0, \dots, 0, 1, \dots, 1)$ with 0 appearing $N\psi - 1$ times and 1 appearing $N(1 - \psi)$ times. In such a case, an individual of type A leaves $N\psi$ descendants in one time step with probability

$$P(z_1(x) = N\psi) = \frac{f_A(x)}{N^{1+\alpha}\bar{f}}, \quad (30)$$

two with probability

$$P(z_1(x) = 2) = \left(1 - \frac{1}{N^\alpha}\right) \frac{f_A(x)}{N\bar{f}}, \quad (31)$$

one with probability

$$P(z_1(x) = 1) = \left(1 - \frac{f_A(x)}{N\bar{f}}\right) \left(\frac{N^{1-\alpha}(1 - \psi) + (N-2)(1 - N^{-\alpha})}{N-1}\right) \quad (32)$$

and zero otherwise. This leads to

$$\mu_A(x) = 1 - \left(\frac{1}{N-1} - \frac{f_A(x)}{(N-1)\bar{f}}\right) \left(1 + \frac{N\psi - 2}{N^\alpha}\right) \quad (33)$$

and similarly for $\mu_B(x)$, from which

$$h_N(x) = \left(\frac{u_A(x) - u_B(x)}{N - 1} \right) \left(1 + \frac{N\psi - 2}{N^\alpha} \right). \tag{34}$$

Moreover, we have

$$\sigma_N^2 = \frac{2}{N} + \frac{N\psi(N\psi - 1) - 2}{N^{1+\alpha}} \tag{35}$$

for the variance in the number of descendants left in one time step by a single individual in the absence of selection. Notice that this variance is not bounded as N goes to infinity if $\alpha < 1$.

4 Fixation probability for a mutant type under weak selection

In this section, we consider the fate of a mutant type A in the long run. Suppose that A is represented k times at $t = 0$ and let X_t be the frequency of A at times $t = 0, 1, 2, \dots$. We have a Markov chain on the states i/N for $i = 0, 1, \dots, N$, with initial distribution $X_0 = k/N$ and fixation states $x = 0$ and $x = 1$, while all other states are transient. The probability of transition from i/N to j/N in one time step will depend on the intensity of selection and will be represented by $P_{ij}(s)$, for $i, j = 0, 1, \dots, N$. It will be assumed throughout that the transition probabilities and their derivatives with respect to s , which are supposed to exist, are continuous at $s = 0$. The one-step transition matrix $P = (P_{ij}(s))$ satisfies $P^T = P^{(T)}$, where $P^{(T)} = (P_{ij}^{(T)}(s))$ represents the T -step transition matrix.

From its initial state k/N , the chain will converge in probability to a random variable X_∞ which takes the value 1 with some probability $p_N(s)$ and 0 with the complementary probability $1 - p_N(s)$, where $p_N(s)$ is the probability of fixation of A as a function of the intensity of selection. Notice that $p_N(0) = k/N$, since one of the initial individuals will be the ancestor of all the population in the long run and it will be one chosen at random by symmetry if no selection takes place.

Being uniformly bounded, the chain will also converge in mean. Therefore, we have

$$E_s(X_\infty - X_0) = \lim_{T \rightarrow \infty} E_s(X_T - X_0), \tag{36}$$

where E_s designates expectation when the intensity of selection is s . Writing

$$X_T - X_0 = \sum_{t=0}^{T-1} (X_{t+1} - X_t), \tag{37}$$

the above equation takes the form

$$E_s(X_\infty - X_0) = \sum_{t=0}^{\infty} E_s(X_{t+1} - X_t). \quad (38)$$

This equation becomes

$$p_N(s) - p_N(0) = \sum_{t=0}^{\infty} E_s(m_s(X_t)), \quad (39)$$

using the identity

$$E_s(X_{t+1} - X_t) = E_s(E_s(X_{t+1} - X_t | X_t)) \quad (40)$$

and the fact that the conditional expectation of $X_{t+1} - X_t$ given X_t is $m_s(X_t)$. Notice that

$$E_s(m_s(X_t)) = \sum_{j=0}^N m_s(j/N) P_{kj}^{(t)}(s). \quad (41)$$

Then, differentiating with respect to s in (39) yields (see the second paragraph below Result 1)

$$p'_N(s) = \sum_{t=0}^{\infty} \frac{d}{ds} E_s(m_s(X_t)), \quad (42)$$

where the derivative on the right-hand side can be expressed in the form

$$E'_s(m_s(X_t)) + E_s(m'_s(X_t)), \quad (43)$$

with

$$E'_s(m_s(X_t)) = \sum_{j=0}^N m_s(j/N) \left[\frac{d}{ds} P_{kj}^{(t)}(s) \right] \quad (44)$$

and

$$E_s(m'_s(X_t)) = \sum_{j=0}^N \left[\frac{d}{ds} m_s(j/N) \right] P_{kj}^{(t)}(s). \quad (45)$$

Finally, evaluating at $s = 0$ and using (11), we find that

$$p'_N(0) = E_0(m'_0(X_t)), \quad (46)$$

where E_0 denotes expectation when $s = 0$. This derivative represents the rate of change in the probability of fixation of the mutant type A , taken with respect to the intensity of selection and evaluated at neutrality. It gives the first-order effect of the intensity of selection on the fixation probability. Appealing to (13), we have the following approximation under weak selection.

Result 1 *In the exchangeable selection model for a population of fixed size N with a small intensity of selection s and a first-order coefficient of selection for a mutant type A in the form $sh_N(x)$, the probability of fixation of the mutant type A is approximated by*

$$p_N(s) = p_N(0) + sp'_N(0) + o(s), \quad (47)$$

where $p_N(0)$ is the initial frequency of A and

$$p'_N(0) = \sum_{t=0}^{\infty} E_0(X_t(1 - X_t)h_N(X_t)), \quad (48)$$

with E_0 denoting expectation under neutrality.

Result 1 holds for any population size and in this sense it is complementary to a diffusion approximation for a large population size. Moreover, we will see later on (Result 4 and Discussion) that it is consistent with a diffusion approximation when the population size is large and selection very weak under moment conditions on the number of descendants left in one time step by a single individual.

The approach and some of the arguments leading to Result 1 can be found in Rousset [25], but (42) which claims that the derivative of the series in (39) is the series of the derivatives has to be justified. This will be the case if the series in (42), or equivalently the derivative of $E_s(X_T)$ in (36), converges uniformly for s in a neighborhood of 0. The states being fixed and their number finite, it suffices to show that this is the case for the derivative of the T -step transition matrix $P^{(T)}$. A formal proof of this statement under our assumptions is provided in the Appendix.

5 Fixation probability for a single mutant

In this section, we focus on a mutant type whose initial frequency is $1/N$ and we assume that the rate of change in its selection coefficient is in a polynomial form, namely,

$$h_N(x) = a_0 + a_1x + \cdots + a_nx^n, \quad (49)$$

for some $n \geq 0$ and some constants a_0, \dots, a_n . Then, the rate of change in its fixation probability becomes

$$p'_N(0) = \sum_{i=0}^n a_i d_i, \tag{50}$$

where

$$d_i = \sum_{t=0}^{\infty} E_0 \left(X_t^{i+1} - X_t^{i+2} \right). \tag{51}$$

Defining $\xi_l = 1$ if individual l at time t is of type A and 0 otherwise for $l = 1, \dots, N$, the frequency of A at time t can be expressed as

$$X_t = \frac{1}{N} \sum_{l=1}^N \xi_l \tag{52}$$

and its i th power as

$$X_t^i = \frac{1}{N^i} \sum_{\substack{i_1, \dots, i_N \geq 0 \\ i_1 + \dots + i_N = i}} \binom{i}{i_1, \dots, i_N} \xi_1^{i_1} \dots \xi_N^{i_N}. \tag{53}$$

The variables ξ_l being exchangeable when $s = 0$ and satisfying $\xi_l^{i_l} = \xi_l$ if $i_l \geq 1$ and 1 if $i_l = 0$, the expected value of X_t^i under neutrality will be

$$E_0(X_t^i) = \frac{1}{N^i} \sum_{l=1}^{\min(i, N)} N_{[l]} S_{il} E_0(\xi_1 \dots \xi_l), \tag{54}$$

where $N_{[l]} = N(N - 1) \dots (N - l + 1)$ and

$$S_{il} = \frac{1}{l!} \sum_{\substack{r_1, \dots, r_l \geq 1 \\ r_1 + \dots + r_l = i}} \binom{i}{r_1, \dots, r_l}. \tag{55}$$

This is a Stirling number of the second kind (see, e.g., [1]), that is, the number of ways of partitioning a set of i elements into l nonempty subsets, which can also be expressed as

$$S_{il} = \sum_{k=1}^l (-1)^{l-k} \frac{k^i}{k!(l-k)!} \tag{56}$$

using an inclusion–exclusion argument. Moreover, we have

$$E_0(\xi_1 \dots \xi_l) = \frac{1}{N} P_0(T_l \leq t), \tag{57}$$

where P_0 represents probability in the neutral model and T_l stands for the coalescence time of l lineages, that is, the number of steps backward in time starting from l individuals at time t and ending with their most recent common ancestor. If $T_l \leq t$, then this common ancestor and its l descendants at time t are all of type A with probability $1/N$, the frequency of type A at time 0. Otherwise, this probability is 0. Writing $P_0(T_l \leq t) = 1 - P_0(T_l > t)$, we find

$$E_0(X_t^i) = \frac{1}{N} - \frac{1}{N^{i+1}} \sum_{l=1}^{\min(i,N)} N_{[l]} S_{il} P_0(T_l > t). \tag{58}$$

Since $\sum_{t=0}^{\infty} P_0(T_l > t) = E_0(T_l)$ and $E_0(T_1) = 0$, we get

$$d_i = \sum_{l=2}^{\min(i+2,N)} N_{[l]} \left[\frac{S_{i+2,l}}{N^{i+3}} - \frac{S_{i+1,l}}{N^{i+2}} \right] E_0(T_l), \tag{59}$$

with the convention that $S_{i+1,i+2} = 0$. Thus, we have proved the following.

Result 2 *In the exchangeable selection model for a population of fixed size N , if the rate of change in the coefficient of selection for a mutant allele against a wild type is in the form*

$$h_N(x) = \sum_{i=0}^n a_i x^i, \tag{60}$$

then the rate of change in the probability of fixation of this allele when represented once in the population takes the exact form

$$p'_N(0) = \sum_{i=0}^n a_i d_i, \tag{61}$$

where d_i is given in (59).

Notice that the expected coalescence times in (59) satisfy the formula

$$E_0(T_i) = \frac{1}{1 - p_{ii}} + \sum_{l=2}^{i-1} \frac{p_{il}}{1 - p_{ii}} E_0(T_l), \tag{62}$$

where

$$p_{il} = \frac{\binom{N}{l}}{\binom{N}{i}} \sum_{\substack{r_1, \dots, r_l \geq 1 \\ r_1 + \dots + r_l = i}} E_0 \left(\binom{z_1}{r_1} \cdots \binom{z_l}{r_l} \right) \tag{63}$$

is the probability for i individuals to have l parents one time step back (see, e.g., [18]). In particular, we have

$$E_0(T_2) = \frac{1}{p_{21}} = \frac{N - 1}{E_0(z_1(z_1 - 1))} = \frac{N - 1}{\sigma_N^2} = c_N^{-1}. \tag{64}$$

The next expected values can be found recursively.

6 Approximation in the case of a large population size

In order to get an approximation in the case of a large population size, we will make the following assumption.

Assumption A $\sup_N E_0(z_1^k) < \infty$ for all $k \geq 2$.

Assumption A means that the number of descendants of a single individual in one time step under neutrality exhibits moments of every order uniformly bounded with respect to N . In particular, the variance σ_N^2 is uniformly bounded with respect to N . Assumption A holds, e.g., for the Moran model and the Wright–Fisher model, but not necessarily for models with a highly skewed distribution for the number of descendants among individuals (see, e.g., [4]).

Under assumption A, the ancestral process starting from l individuals with c_N^{-1} time steps as the unit of time converges under neutrality to Kingman’s [10] coalescent as N goes to infinity (see, e.g., [17,29], for more details). As a consequence, the time to the most recent common ancestor in number of c_N^{-1} time steps, that is, the scaled coalescence time $\tau_l = c_N T_l$, has a mean under neutrality $E_0(\tau_l) = c_N E_0(T_l)$ which converges to $2(1 - 1/l)$ as N goes to infinity. Therefore, we have

$$\lim_{N \rightarrow \infty} \sigma_N^2 d_i = \frac{2}{(i + 1)(i + 2)}. \tag{65}$$

This yields the following approximation.

Result 3 *If the population size N is large and Assumption A holds, then the rate of change in the probability of fixation of a single mutant in Result 2 reduces to the approximation*

$$p'_N(0) \approx \sum_{i=0}^n \frac{2a_i\sigma_N^{-2}}{(i+1)(i+2)}, \tag{66}$$

where σ_N^2 is the variance in the number of descendants of a single individual in one time step under neutrality.

Now, let us consider any continuous function $h_N(x)$, for $0 \leq x \leq 1$, as rate of change in the selection coefficient for the mutant allele. Owing to Weierstrass theorem, there exists a sequence of polynomials that converges uniformly to $h_N(x)$, that is, given any $\epsilon > 0$, there exists a polynomial

$$g_N(x) = \sum_{i=0}^n a_i x^i \tag{67}$$

such that

$$|h_N(x) - g_N(x)| < \epsilon, \tag{68}$$

for $0 \leq x \leq 1$. Then, we have

$$\left| \sum_{t=0}^{\infty} E_0(X_t(1 - X_t)(h_N(X_t) - g_N(X_t))) \right| \leq \epsilon d_0, \tag{69}$$

where d_0 is defined in (51) and is such that $\sigma_N^2 d_0$ tends to 1 as N tends to infinity owing to (65). On the other hand, Result 3 ensures that

$$\sum_{t=0}^{\infty} E_0(X_t(1 - X_t)g_N(X_t)) \approx \sum_{i=0}^n \frac{2a_i\sigma_N^{-2}}{(i+1)(i+2)}, \tag{70}$$

if N is large enough and Assumption A holds, which can be written in the form

$$\sum_{t=0}^{\infty} E_0(X_t(1 - X_t)g_N(X_t)) \approx 2\sigma_N^{-2} \int_0^1 \int_0^y g_N(x) dx dy. \tag{71}$$

Finally, using the fact that

$$\left| \int_0^1 \int_0^y h_N(x) dx dy - \int_0^1 \int_0^y g_N(x) dx dy \right| \leq \int_0^1 \int_0^y |h_N(x) - g_N(x)| dx dy \leq \frac{\epsilon}{2}, \tag{72}$$

we find the approximation

$$\left| p'_N(0) - 2\sigma_N^{-2} \int_0^1 \int_0^y h_N(x) dx dy \right| \approx 2\epsilon\sigma_N^{-2}. \quad (73)$$

Therefore, we have established the following.

Result 4 *If the rate of change in the coefficient of selection for a mutant allele in the exchangeable selection model for a population of large size N under Assumption A, represented by $h_N(x)$, is continuous for $0 \leq x \leq 1$, then we have the approximation*

$$p'_N(0) \approx 2\sigma_N^{-2} \int_0^1 \int_0^y h_N(x) dx dy \quad (74)$$

for the rate of change in the probability of fixation of this allele when represented once in the population.

7 Discussion

Introducing selection into the Cannings exchangeable model, we have studied the rate of change in the fixation probability for a mutant type, more precisely, the derivative of the probability for the mutant type to become fixed in the population, the derivative being taken with respect to the intensity of selection, represented by s , and evaluated at $s = 0$. This rate has been obtained rigorously assuming mild regularity conditions on the transition probabilities for the type frequencies under weak selection. Moreover, it has been expressed exactly in terms of mean coalescence times under neutrality in the case of a single mutant in a population of fixed size N whose coefficient of selection has a rate of change in a polynomial form with respect to the type frequencies. This rate, denoted by $h_N(x)$ where x is the current frequency of the mutant type, is defined as the derivative of the difference between the mean number of descendants from one mutant and the mean number from one non mutant in one time step, again with respect to s and evaluated at $s = 0$.

In a haploid Wright–Fisher model with numbers of descendants left by the individuals of the population in one time step given by a multinomial vector $(z_1(x), \dots, z_N(x))$ whose parameters are determined by fertility differences resulting from pairwise interactions between parents having additive effects with intensity in the form $s = \gamma/(N - 1)$, we found [13]

$$h_N(x) = a_0 + a_1x, \quad (75)$$

where

$$a_0 = (\mathbf{p}_A - \mathbf{p}_B)^T M \mathbf{p}_B - \frac{\mathbf{p}_A^T M \mathbf{p}_A}{N} + \frac{\mathbf{p}_B^T M \mathbf{p}_B}{N} \tag{76}$$

and

$$a_1 = (\mathbf{p}_A - \mathbf{p}_B)^T M (\mathbf{p}_A - \mathbf{p}_B). \tag{77}$$

Here, M is a game matrix, while \mathbf{p}_A and \mathbf{p}_B are frequency vectors that represent mixed strategies used by the individuals of the mutant type A and the individuals of the wild type B , respectively. On the other hand, the mean coalescence times in a exact neutral Wright–Fisher model can be calculated directly. We find $E_0(T_2) = N$ and

$$E_0(T_3) = \frac{N(4N - 3)}{(3N - 2)}, \tag{78}$$

from which (55) and (59) yield

$$d_0 = \frac{N - 1}{N} \tag{79}$$

and

$$d_1 = \frac{N - 1}{3N - 2}. \tag{80}$$

Finally, Result 2 gives

$$p_N \left(\frac{\gamma}{N - 1} \right) = \frac{1}{N} + \frac{\gamma}{N} \left(a_0 + \frac{N}{(3N - 2)} a_1 \right) + o \left(\frac{\gamma}{N} \right) \tag{81}$$

as approximation for the probability of fixation of a single mutant A under weak selection. This is in agreement with the heuristic diffusion approximation in [13] if N is large and the direct first-step analysis in [8] if we set $w = \gamma/N$.

The corresponding Moran model with the vector $(z_1(x), \dots, z_N(x))$ being a permutation of $(2, 0, 1, \dots, 1)$ has $h_N(x)/(N - 1)$ with $h_N(x)$ given in (75) as rate of change in the selection coefficient and $E_0(T_l) = N(N - 1)(1 - 1/l)$ for $l \geq 2$ as mean coalescence times. We find

$$d_0 = \frac{(N - 1)^2}{2N} \tag{82}$$

and

$$d_1 = \frac{(N - 1)^2(N + 1)}{6N^2}, \tag{83}$$

from which we deduce the approximation

$$p_N \left(\frac{\gamma}{N-1} \right) = \frac{1}{N} + \frac{\gamma}{2N} \left(a_0 + \frac{(N+1)}{3N} a_1 \right) + o \left(\frac{\gamma}{N} \right) \tag{84}$$

for the probability of fixation of a single mutant A under weak selection, which is close to (81) if N is large and γ is replaced with 2γ . This factor 2 differentiating the Wright–Fisher model and the Moran model is familiar (see, e.g., [5, p. 121]).

In a general exchangeable model, $E_0(T_2) = c_N^{-1}$ and the mean coalescence times $E_0(T_l)$ for $l \geq 3$ can be found recursively. If the rate of change in the selection coefficient is constant and given by a_0 , then the rate of change in the fixation probability is $a_0 d_0$, where

$$d_0 = \frac{(N-1)^2}{N^2 \sigma_N^2}. \tag{85}$$

A model of this kind for the evolution of the variance of the sex ratio with σ_N^2 converging to 1 as N goes to infinity was studied by a diffusion approximation in [13].

If the rate of change in the selection coefficient is proportional to $a_0 + a_1 x$ with a_0 and a_1 given by (76) and (77), respectively, in the context of a linear game, then the rate of change in the fixation probability will be proportional to

$$a_0 d_0 + a_1 d_1 = d_0 (\mathbf{p}_A - \mathbf{p}_B)^T \tilde{M} \mathbf{p}_B + \left(\frac{N d_1 - d_0}{N - 2} \right) (\mathbf{p}_A - \mathbf{p}_B)^T \tilde{M} (\mathbf{p}_A - \mathbf{p}_B), \tag{86}$$

where

$$\tilde{M} = M - \frac{1}{N} (M + M^T). \tag{87}$$

If $a_0 d_0 + a_1 d_1 < 0$, then weak selection opposes A replacing B , in the sense that the probability of fixation of a single mutant A , expressed as $p_N(s)$, is less than the initial frequency of A , given by $1/N$, when the selection intensity s is small enough. If $N d_1 > d_0$, this is the case for every $\mathbf{p}_A \neq \mathbf{p}_B$ close enough to \mathbf{p}_B if and only if for every $\mathbf{p}_A \neq \mathbf{p}_B$ we have

$$(\mathbf{p}_A - \mathbf{p}_B)^T \tilde{M} \mathbf{p}_B \leq 0 \tag{88}$$

and, in case of equality,

$$(\mathbf{p}_A - \mathbf{p}_B)^T \tilde{M} (\mathbf{p}_A - \mathbf{p}_B) < 0. \tag{89}$$

This means that, ignoring degeneracies, a necessary and sufficient condition for weak selection opposing \mathbf{p}_A replacing \mathbf{p}_B for every $\mathbf{p}_A \neq \mathbf{p}_B$ close enough to \mathbf{p}_B is that \mathbf{p}_B is an ESS for the game matrix \tilde{M} [15, 16]. Notice that this does not

preclude weak selection favoring \mathbf{p}_A replacing an ESS \mathbf{p}_B if \mathbf{p}_A is far enough from \mathbf{p}_B .

In the case of two pure strategies $\mathbf{p}_A = (1, 0)$ and $\mathbf{p}_B = (0, 1)$ for a 2×2 game matrix

$$M = \begin{pmatrix} a & b \\ c & d \end{pmatrix}, \tag{90}$$

for instance, with $a > c$ and $d > b$, which means that \mathbf{p}_A and \mathbf{p}_B are the best replies to themselves, weak selection favors A replacing B if $a_0d_0 + a_1d_1 > 0$. This condition is equivalent to

$$\frac{d - b + \frac{a+d}{N}}{a - b - c + d} < \frac{d_1}{d_0}. \tag{91}$$

The left-hand member of this inequality as N goes to infinity corresponds to the unstable equilibrium frequency of the pure strategy \mathbf{p}_A for the replicator dynamics in an infinite population (see, e.g., [7, p. 147]). The right-hand member reduces to $1/3$ in the limit as long as one remains in the domain of application of Kingman’s [10] coalescent. This generalizes the $1/3$ -law originally proposed by Nowak et al. [21] for the Moran model to a larger class of models. (See also [8], for the Wright–Fisher model.) In general however, the ratio d_1/d_0 may converge to a different limit.

The approximation for the rate of change in the probability of fixation of a single mutant in a large population given in Result 3 strongly depends on Assumption A that the number of descendants left by a single individual in one time step under neutrality has bounded moments as N goes to infinity. If Assumption A does not hold, then the ancestral process with c_N^{-1} time steps as the unit of time may still converge as N goes to infinity, but to a more general coalescent process which admits multiple mergers as the Λ -coalescent [19, 23, 28]. This is the case for the extended Moran model presented in the example section if $\alpha \leq 2$ [4]. In such a case, the limit of $\sigma_N^2 d_i$ may still exist, but it will generally lead to a different approximation.

For a large population under Assumption A and for any continuous rate of change in the coefficient of selection $h_N(x)$, we have deduced the approximation

$$p_N(s) \approx \frac{1}{N} + 2s\sigma_N^{-2} \int_0^1 \int_0^y h_N(x) dx dy \tag{92}$$

for the probability of fixation of a single mutant under weak selection. This approximation neglects terms of order s^2 , uniformly with respect to N , and terms of the form $so(1)$ where $o(1)$ tends to 0 as N goes to infinity. It makes sense if s is small compared to $1/N$ when N is large. This agrees with a diffusion approximation if we take c_N^{-1} time steps as the unit of time and assume $\gamma = s(N - 1)$

small enough for N large enough. Then, the drift and diffusion parameters of the diffusion process, $a(x)$ and $b(x)$, are approximated by $\gamma x(1-x)\sigma_N^{-2}h_N(x)$ and $x(1-x)$, respectively, and the fixation probability by (see, e.g., [5, p. 140], and references therein)

$$p_N(s) = \frac{\int_0^{1/N} \psi(y) dy}{\int_0^1 \psi(y) dy}, \quad (93)$$

where

$$\psi(y) = \exp \left\{ -2 \int_0^y \frac{a(x)}{b(x)} dx \right\}. \quad (94)$$

Using the approximation

$$\psi(y) \approx 1 - 2\gamma\sigma_N^{-2} \int_0^y h_N(x) dx, \quad (95)$$

we find

$$p_N(s) \approx \frac{1}{N} + 2 \left(\frac{N-1}{N} \right) s \sigma_N^{-2} \int_0^1 \int_0^y h_N(x) dx dy \quad (96)$$

as approximation for the fixation probability of a single mutant, in agreement with (92) for N large enough. Notice that this approximation makes sense only if s is much smaller than $1/N$, the value of $p_N(0)$, and this is consistent with γ small. Moreover, the diffusion approximation requires that $\sigma_N^{-2}h_N(x)$ converges uniformly as N goes to infinity besides other conditions on higher-order moments for the change of the mutant frequency (see, e.g., [9, Chap. 15]). These conditions are satisfied, e.g., for the Moran model and the Wright–Fisher model. If Assumption A does not hold however, a diffusion approximation may not be available.

Acknowledgments This research was supported in part by the Natural Sciences and Engineering Research Council of Canada. We thank Reinhard Bürger for comments and suggestions.

Appendix

We want to prove that the derivative of the T -step transition matrix, that is,

$$\frac{dP^{(T)}}{ds} = \left(\frac{d}{ds} P_{ij}^{(T)}(s) \right), \quad (97)$$

converges uniformly for s in a neighborhood of 0 as $T \rightarrow \infty$. Since $P^{(T)} = P^T$, where P represents the one-step transition matrix, we have

$$\frac{dP^{(T)}}{ds} = \sum_{t=0}^{T-1} P^t \frac{dP}{ds} P^{T-t-1}. \tag{98}$$

Since 0 and 1 are absorbing states, it follows that the transition matrix P can be written in the form

$$P = \begin{bmatrix} 1 & \mathbf{0} & 0 \\ \mathbf{x} & A & \mathbf{y} \\ 0 & \mathbf{0} & 1 \end{bmatrix}, \tag{99}$$

where $\mathbf{0}$ denotes a zero row vector of dimension $N - 1$, while \mathbf{x} and \mathbf{y} are two column vectors of dimension $N - 1$ and A is a square matrix of order $N - 1$. Thus, a straightforward calculation leads to the expression

$$\begin{bmatrix} 0 & \mathbf{0} & 0 \\ \left(S_1^{(T)} \frac{d\mathbf{x}}{ds} + S_3^{(T)} \mathbf{x} \right) & S_2^{(T)} & \left(S_1^{(T)} \frac{d\mathbf{y}}{ds} + S_3^{(T)} \mathbf{y} \right) \\ 0 & \mathbf{0} & 0 \end{bmatrix} \tag{100}$$

for the derivative (98), where

$$S_1^{(T)} = \sum_{t=0}^{T-1} A^t, \tag{101}$$

$$S_2^{(T)} = \sum_{t=0}^{T-1} A^t \frac{dA}{ds} A^{T-t-1}, \tag{102}$$

$$S_3^{(T)} = \sum_{t=0}^{T-1} A^t \frac{dA}{ds} \left(\sum_{k=0}^{T-t-1} A^k \right). \tag{103}$$

These are all square matrices of order $N - 1$. Using the infinite norm

$$\|\mathbf{x}\| = \max_{i=1, \dots, N-1} |x_i|, \tag{104}$$

for a vector $\mathbf{x} = (x_1, \dots, x_{N-1})$, the associated norm for a square matrix $A = (A_{ij})$ of order $N - 1$ is

$$\|A\| = \sup_{\mathbf{x} \neq \mathbf{0}} \frac{\|A\mathbf{x}\|}{\|\mathbf{x}\|} = \max_{i=1, \dots, N-1} \sum_{j=1}^{N-1} |A_{ij}|. \tag{105}$$

This norm has the property that

$$\|AB\| \leq \|A\|\|B\|, \tag{106}$$

for two square matrices A and B of order $N - 1$.

Since $P(0)$ is the transition matrix for a standard Cannings exchangeable model, its eigenvalues in decreasing order, denoted by $\lambda_0 \geq \lambda_1 \geq \lambda_2 \geq \lambda_3 \geq \dots \geq \lambda_N$, satisfy $1 = \lambda_0 = \lambda_1 > \lambda_2 = 1 - c_N \geq \lambda_3 \dots \geq \lambda_N \geq 0$ ([2], see, e.g., [5]). Hence, the eigenvalues of $A(0)$, namely $\lambda_2 \geq \lambda_3 \geq \dots \geq \lambda_N$, are all strictly less than 1 in absolute value. Consequently, the matrix $I - A(0)$ is invertible and $\|A(0)^T\|$ converges to 0 as T goes to infinity. Therefore, we have that $\|(I - A(0))^{-1}\|$ is finite and

$$\|A^{T_0}(0)\| = \max_{i=1, \dots, N-1} \sum_{j=1}^{N-1} P_{ij}^{(T_0)}(0) < 1, \tag{107}$$

for some $T_0 \geq 1$. Assuming that $A(s)$ and $dA(s)/ds$ are continuous at $s = 0$, we can find some neighbourhood V of $s = 0$ in which $\|(I - A(s))^{-1}\|$ and $\|dA(s)/ds\|$ are uniformly bounded and

$$\|A^{T_0}(s)\| = \phi < 1. \tag{108}$$

Moreover, we have always $\|A(s)\| \leq 1$. Writing $t \geq 1$ in the form $t = \lambda T_0 + k$, where $\lambda \geq 0$ and $0 \leq k \leq T_0 - 1$, it follows from (106) and (108) that

$$\|A^t(s)\| \leq \|A^{T_0}(s)\|^\lambda \|A(s)\|^k \leq \phi^\lambda \tag{109}$$

and

$$\sum_{t=0}^\infty \|A^t(s)\| \leq T_0 \sum_{\lambda=0}^\infty \phi^\lambda = \frac{T_0}{1 - \phi}, \tag{110}$$

for all s in V . On the other hand, we have

$$S_1^{(T)} - (I - A)^{-1} = (I - A)^{-1}(-A^T), \tag{111}$$

from which

$$\|S_1^{(T)} - (I - A)^{-1}\| \leq \|(I - A)^{-1}\| \|A^T\|, \tag{112}$$

for all $T \geq 1$. Moreover, if $T \geq 2t_0 + 1$, then

$$\|S_2^{(T)}\| \leq 2 \|A^{t_0}\| \left\| \frac{dA}{ds} \right\| \sum_{t \geq 0} \|A^t\|, \tag{113}$$

which is obtained from the expression

$$S_2^{(T)} = \sum_{t=0}^{t_0-1} A^t \frac{dA}{ds} A^{T-t-1} + \sum_{t=t_0}^{T-1} A^t \frac{dA}{ds} A^{T-t-1}. \tag{114}$$

Finally, one can write

$$S_3^{(T)} = S_1^{(T)} \frac{dA}{ds} (I - A)^{-1} - S_2^{(T)} A (I - A)^{-1}. \tag{115}$$

Combining (112), (113) and (115) with (109) and (110), we conclude that the derivative (98) converges uniformly in V as T goes to infinity and

$$\lim_{T \rightarrow \infty} \frac{dP^{(T)}}{ds} = \begin{bmatrix} 0 & \mathbf{0} & 0 \\ \mathbf{u} & O & \mathbf{v} \\ 0 & \mathbf{0} & 0 \end{bmatrix}, \tag{116}$$

where O denotes a zero square matrix of order $N - 1$, while

$$\mathbf{u} = (I - A)^{-1} \frac{d\mathbf{x}}{ds} + (I - A)^{-1} \frac{dA}{ds} (I - A)^{-1} \mathbf{x}$$

and

$$\mathbf{v} = (I - A)^{-1} \frac{d\mathbf{y}}{ds} + (I - A)^{-1} \frac{dA}{ds} (I - A)^{-1} \mathbf{y}.$$

References

1. Abramowitz, M., Stegun, I.A.: Handbook of Mathematical Functions. Dover, New York (1965)
2. Cannings, C.: The latent roots of certain Markov chains arising in genetics: a new approach. I. Haploid models. *Adv. Appl. Probab.* **6**, 260–290 (1974)
3. Cherry, J.L.: Selection in a subdivided population with dominance or local frequency dependence. *Genetics* **163**, 1511–1518 (2003)
4. Eldon, B., Wakeley, J.: Coalescent processes when the distribution of offspring number among individuals is highly skewed. *Genetics* **172**, 2621–2633 (2006)
5. Ewens, W.J.: *Mathematical Population Genetics 1. Theoretical Introduction*, 2nd edn. Springer, New York (2004)
6. Fisher, R.A.: *The Genetical Theory of Natural Selection*. Clarendon, Oxford (1930)
7. Hofbauer, J., Sigmund, K.: *The Theory of Evolution and Dynamical Systems*. Cambridge University Press, Cambridge (1988)
8. Imhof, L.A., Nowak, M.A.: Evolutionary game dynamics in a Wright–Fisher process. *J. Math. Biol.* **52**, 667–681 (2006)
9. Karlin, S., Taylor, H.M.: *A Second Course in Stochastic Processes*. Academic, New York (1981)
10. Kingman, J.F.C.: Exchangeability and the evolution of large populations. In: Koch, G., Spizzichino, F. (eds.) *Exchangeability in Probability and Statistics*, pp. 97–112. North-Holland, Amsterdam (1982)
11. Lambert, A.: Probability of fixation under weak selection: a branching process unifying approach. *Theoret. Popul. Biol.* **69**, 419–441 (2006)

12. Lessard, S.: Evolutionary dynamics in frequency-dependent two-phenotype models. *Theoret. Popul. Biol.* **25**, 210–234 (1984)
13. Lessard, S.: Long-term stability from fixation probabilities in finite populations: New perspectives for ESS theory. *Theoret. Popul. Biol.* **68**, 19–27 (2005)
14. Leturque, H., Rousset, F.: Dispersal, kin selection, and the ideal free distribution in a spatially heterogeneous population. *Theoret. Popul. Biol.* **62**, 169–180 (2002)
15. Maynard Smith, J.: The theory of games and the evolution of animal conflicts. *J. Theoret. Biol.* **47**, 209–221 (1974)
16. Maynard Smith, J., Price, G.R.: The logic of animal conflict. *Nature* **246**, 15–18 (1973)
17. Möhle, M.: Robustness results for the coalescent. *J. Appl. Probab.* **35**, 438–447 (1998)
18. Möhle, M.: The time back to the most recent common ancestor in exchangeable population models. *Adv. Appl. Probab.* **36**, 78–97 (2004)
19. Möhle, M., Sagitov, S.: A classification of coalescent processes for haploid exchangeable population models. *Ann. Probab.* **29**, 1547–1562 (2001)
20. Moran, P.A.P.: Random processes in genetics. *Proc. Camb. Philos. Soc.* **54**, 60–71 (1958)
21. Nowak, M.A., Sasaki, A., Taylor, C., Fudenberg, D.: Emergence of cooperation and evolutionary stability in finite populations. *Nature* **428**, 646–650 (2004)
22. Orzack, S.H., Hines, W.G.S.: The evolution of strategy variation: will an ESS evolve?. *Evolution* **59**, 1183–1193 (2005)
23. Pitman, J.: Coalescents with multiple collisions. *Ann. Probab.* **27**, 1870–1902 (1999)
24. Proulx, S.R.: The ESS under spatial variation with applications to sex allocation. *Theoret. Popul. Biol.* **58**, 33–47 (2000)
25. Rousset, F.: A minimal derivation of convergence stability measures. *J. Theoret. Biol.* **221**, 665–668 (2003)
26. Rousset, F., Billiard, S.: A theoretical basis for measures of kin selection in subdivided populations: Finite populations and localized dispersal. *J. Evol. Biol.* **13**, 814–825 (2000)
27. Roze, D., Rousset, F.: The robustness of Hamilton’s rule with inbreeding and dominance: Kin selection and fixation probabilities under partial sib mating. *Am. Nat.* **164**, 214–231 (2004)
28. Sagitov, S.: The general coalescent with asynchronous mergers of ancestral lines. *J. Appl. Probab.* **36**, 1116–1125 (1999)
29. Tavaré, S., Zeitouni, F.: Lectures on Probability Theory and Statistics, Ecole d’Été de Probabilités de Saint-Flour XXXI - 2001, Lecture Notes in Mathematics, Vol. 1837, Springer, New York (2004)
30. Taylor, C., Nowak, M.A.: Evolutionary game dynamics with non-uniform interaction rates. *Theoret. Popul. Biol.* **69**, 243–252 (2006)
31. Wild, G., Taylor, P.D.: Fitness and evolutionary stability in game theoretic models of finite populations. *Proc. R. Soc. B* **271**, 2345–2349 (2004)
32. Wright, S.: Evolution in Mendelian populations. *Genetics* **16**, 97–159 (1931)