ON THE STOCHASTIC EVOLUTION OF FINITE POPULATIONS

FABIO A. C. C. CHALUB* AND MAX O. SOUZA†

ABSTRACT. This work is concerned with Markov chain models used in population genetics. Namely, we start from two celebrated models, the Moran and the Wright-Fisher processes, and study them from a very general viewpoint. Our aim is threefold: to identify the algebraic structures associated to time-homogeneous processes; to study the monotonicity properties of the fixation probability, with respect to the initial condition—in particular we show that there are situations, that are by no means exceptional, where an increase in the initial presence of a type can lead to a decrease in the fixation probability of this type; to understand time-inhomogeneous processes in a more systematic way. In addition, we also discuss the traditional identification of frequency dependent fitnesses and pay-offs, extensively used in evolutionary game theory, the role of weak selection when the population is finite, and the relations between jumps in evolutionary processes and frequency dependent fitnesses.

1. Introduction

1.1. Background. The evolution of finite populations is inherently prone to stochastic effects, and these are enhanced if the population is small. Thus, the correct modelling of these effects is a key step in understanding the dynamics of such populations. In the absence of widely accepted quantitative first-principles for modelling evolutionary processes, one has to proceed phenomenologically. The Wright-Fisher process, one of the most prevalent models in mathematical population genetics, was a watershed point, and has set up much of the current paradigm in modelling finite populations [33, 81]. This process is a Markov chain with a finite number of states that are distributed multinomially, and hence not easily amenable to analysis except in some special cases [33, 79]. As a simpler process that should still capture the essential aspects of the Wright-Fisher dynamics, Moran conceived a Birth-Death process for genetics, that is contemporaneously known as the Moran model [59]. Later on, one can find a plethora of processes designed to model the evolution of genotype frequencies—see for instance [18, 22, 31, 31, 12] and references therein for many examples with different levels of rigour and generality.
Here, we are mainly concerned with Moran and Wright-Fisher processes for a population with two types. The former, as observed above, is a birth-death process, and hence one individual is replaced at each time step. On the other hand, the entire population is replaced at once in the latter. Both processes are Markov chains with two absorbing states, which are the only stationary states. They also share the same diffusion approximation—up to rescaling—for large populations and weak-selection, which suggests that, at least in this regime, they are two sides of the same coin \[56\]. However, it is not clear if these similarities extend outside this regime.

From a broader perspective, both Moran and Wright-Fisher processes — when considered in a population comprising two types, which will denote by $A$ and $B$ — belong to a class of Markov chains that are characterised by two parameters: the population size $N$, and a vector of $N + 1$ type selection probabilities, where each entry indicates how likely type $A$ is to be chosen for reproduction depending on its prevalence in the population — and hence it accounts for the natural selection effects in the model. Mutation effects are not considered and, therefore, full absence or full prevalence of a given type are absorbing states. In the former case, this type has been extinguished, while in the latter it has fixated. In addition, every state of the population is accessible from any other state in a finite number of steps. This class of models appears in so many instances, that is natural to name it: we will term it the *Kimura* class.

The axiomatization of this class will be our starting point, and this will lead to a number of questions that seem to be unnoticed in the literature:

**Fitness and type selection probabilities:** One of the most typical modelling approach is to consider that the probability of a given type will be selected for reproduction is proportional to chosen function, that depends on the state of the population. I.e., these functions, one for each type, are a modelling proxy for their reproductive success, and they are usually identified with the fitness of the corresponding type. Here, we will be somewhat more careful, and will term them reproductive fitness. In this setting, a natural question to consider is whether this notion of fitness is consistent with the classical one of Darwinian fitness — the ratio between the prevalence of a given type in two successive generations — which may be computed directly from the statistical properties of the corresponding process.

**Qualitative properties of fixation:** For models without mutation, a key quantity in understanding their evolutionary behaviour is how likely a given type will fixate as a function of its current presence in the population. For neutral processes, the fixation probability of any type is given by the current frequency of this type. In particular, the larger the fraction of individuals of a certain type, the larger is the fixation probability of this particular type. Such a monotonic behaviour of the fixation probability is usually taken for granted (e.g. \[26, 60\]) also outside the neutral regime. Here, we shall say that a process is regular, if its fixation is strictly increasing in the initial frequency of any given type. This is certainly correct in the diffusive limit \[15\], but otherwise this issue seems to be largely overlooked. Moreover, since it is possible to define neutrality by means
of fixation probabilities \cite{35,42,49}, one might ask as to what extent the fixation probability characterises an evolutionary process.

**Time-inhomogeneous evolution:** All processes described thus far are homogeneous in time. Nevertheless, real environments are not static, and the understanding of evolutionary dynamics in this changing scenario provides a new set of challenges, as evidenced by the recent results of Ashcroft et al. \cite{1}, Carja et al. \cite{14}, Cvijović et al. \cite{23}, Melbinger and Vergassola \cite{58}, Uecker and Hermisson \cite{77}. The very first question here is if the Kimura class is suitable to model this new situation. Simple examples will convince the reader that this is not the case, and at this point we will introduce the Gillespie class: a restriction of the Kimura class which is still large enough to include all previous examples, and that is also closed by products and convex combinations.

1.2. **Main results and outline.** The aim of this work is threefold: to identify natural classes of time-homogeneous evolutionary processes and to develop an algebraic theory for such processes; to identify broad qualitative features of the fixation probability—in particular the concept of regular evolution—and to characterise them; to understand time-inhomogeneous process in a more systematic way. Along these lines, we begin by introducing the Kimura class discussed above, and then proceed to obtain the following results:

**Fitness and type selection probabilities:** As observed above, in most modelling works there is a typical identification of reproductive and Darwinian fitness. We show that such identification is always consistent with the Wright-Fisher process, but it cannot be generically consistent with any birth-death process, in particular with Moran processes.

**Qualitative properties of fixation:** We point out that the explicit expression for the fixation probability for birth-death processes implies that they are regular. We then show that the Moran process is a universal regular process: given any increasing fixation vector there exists a unique choice of type selection probabilities for which the corresponding Moran process has the given fixation. We also show that the Wright-Fisher (WF) is a universal process as far as fixation of probability is concerned. Namely, given an admissible fixation vector $\mathbf{F}$, i.e. a fixation for which the only entry with a zero is the first, and the only entry with a one is the last, then there exists a WF process that has $\mathbf{F}$ as a fixation vector. Furthermore, if the fixation is increasing, this process is unique. In addition, we provide a very natural link of regular processes with stochastically ordered matrices, and with matrices that have the so-called Perron-Frobenius property. Finally, we also show that type selection probabilities implied by a smooth fixation are close to neutral, with rigorous error bounds—and hence some form of weak-selection regime implicitly holds in this case.

**Time-inhomogeneous evolution:** Starting from the Kimura class, we show that it lacks the appropriate structure to describe time-inhomogeneous processes, in the sense that the compound process may not be ergodic. Therefore, it needs to be restricted. The appropriate restriction turns out to be given by processes whose interior dynamics is given by a totally indecomposable matrix
instead of an irreducible one. In this case, it gives rise to a convex semigroup of transition matrices. Moreover, process for which the transition matrix in any given step is given by either a Moran or Wright-Fisher process belong to this class.

The outline of the work is as follows: In Section 2 we introduce the Kimura class, $\mathcal{K}$, and in particular the concept of type selection probability. We also point out that both Wright-Fisher and Moran processes belong to this general class. We then introduce the parametrisation of type selection probabilities by reproductive fitness, and show that these are consistent with Darwinian fitness if, and only if, the type selection probabilities are the expected frequency in the next generation. It turns out that this condition is satisfied by the Wright-Fisher process but not by any birth-death process. We finish this section showing a few general results on the fixation probability that while seem to be known are not conveniently available.

In Section 3, we introduce the concept of admissible fixation and of regular evolution as discussed above and three classes of progressively more general ordered matrices: strictly stochastically ordered, banded stochastically ordered and stochastically ordered. We show that every process for which the transition matrix is banded stochastically ordered is regular. We then study the Moran process and, by means of the expression for the fixation probability of a birth-death process, we conclude that it is regular. We also show that, given an increasing fixation vector, there exists a unique choice of type selection probabilities for which the corresponding Moran process has this fixation vector. In the sequel, we proceed to the study of the Wright-Fisher process, and show that being regular is equivalent to having the type selection vector increasing and also equivalent to being strictly stochastically ordered. As a by product, we show that for any admissible fixation, there exists a choice of type selection probabilities — not necessarily unique — such that this fixation is realised by the corresponding Wright-Fisher process. Uniqueness of such a choice holds, if the fixation vector is increasing. The possibility of non-regular evolution for the Wright-Fisher process is further investigated within the realm of evolutionary game theory. In this vein, we prove that type selection probabilities given by two-player game theory are always increasing, with the usual identification of fitnesses and pay-offs; hence the corresponding process is always regular. On the other hand, as soon as we move to three-player games we can find examples of non-increasing type-selection probabilities, and hence non-regular Wright-Fisher dynamics. Finally, we give a complete characterisation of regularity for matrices in the Kimura class. Indeed, we show that regularity is equivalent to a certain transformed matrix of the process having the so-called Perron-Frobenius property.

Section 4 explores what happens with type selection probabilities, if we assume that fixation is given by a smooth function, and the population is large, but still finite. In this case, for the Moran process we show that these probabilities are given by the neutral ones with a first order correction that is order $1/N$, and which can be identified with the right hand side of the associated replicator dynamics equation. For the Wright-Fisher process, we are able to show that these probabilities are given by the neutral ones together with a first order correction that is order $1/\sqrt{N}$. 
In Section 5 we deal with time-inhomogeneous processes. We introduce a new class of processes with a totally indecomposable core matrix, the Gillespie class $G$, and show that it is closed under products and convex combinations, and thus is what is known as a convex semigroup. We then study how time-inhomogeneity might affect regularity. We show that a time-inhomogeneous Wright-Fisher process that is locally regular — i.e. a process such that every transition matrix between two consecutive time steps is given by a regular Wright-Fisher process — is itself regular. On the other hand, no such a result holds for the Moran process (or birth-death processes in general). Indeed, we provide an example of two Moran matrices whose product is non-regular. This yields a deterministic version of Parrondo’s paradox.

We finish with a discussion of the results in Section 6.

2. Revisiting the basic set-up

We use italics to denote real numbers, while boldface denote either vectors in $\mathbb{R}^{N+1}$ or matrices in $\mathbb{R}^{(N+1)\times(N+1)}$. Furthermore, vectors in $\mathbb{R}^{N-1}$ and matrices in $\mathbb{R}^{(N-1)\times(N-1)}$ are denoted by bold-tilded symbols: e.g. $\tilde{M}$ is the core matrix of $M$. For the convenience of the reader, Table 1 summarises the notation used throughout the text.

### Table 1

| Name                              | Symbol | Definition |
|-----------------------------------|--------|------------|
| Kimura matrix                     | $M \in K$ | Def. 1     |
| Regular process                   | $M$    | Def. 5     |
| Gillespie matrix                  | $M \in G$ | Def. 8     |
| Core matrix                       | $\tilde{M}$ | Def. 1     |
| Associated matrix                 | $L$    | Thm. 1     |
| Fixation probability              | $F$    | Prop. 1    |
| Type Selection Probability (TSP)  | $p$ (full population), $q$ (restricted population) | Subsec 2.1 |
| (Strictly, banded) Stochastic ordered matrices | $(St^2O, BStO)StO$ | Def. 7 |
| Darwinian fitness                 | $\Psi$ (absolute) and $\Phi$ (relative) | Def. 3 |
| Reproductive fitness              | $\psi$ | Eq. (4)    |

Table 1. Table with some important notation introduced in this work.

2.1. A natural class of Markov chains. Consider a population of two types denoted by $\mathbb{A}$ and $\mathbb{B}$, respectively, with fixed size $N \in \mathbb{N}$, which evolves in the absence of mutation. We will assume that the population dynamics is described by a discrete time Markov chain, with time-homogeneous transition probabilities, and where the chain is in state $j$, if there are $j$ individuals of type $\mathbb{A}$ in the population.

As is well known [47, 74], such a chain can be completely described by specifying the transition probability from state $j$ to $i$, which we will denote by $M_{ij}$, $i, j = 0, \ldots, N$. We will further assume that $M_{ij} = M(i, j, p, N)$, where $p \in \mathbb{R}^{N+1}$ is a vector of probabilities such that $p_j \in [0, 1]$ describes the probability of a type $\mathbb{A}$ individual being selected for...
reproduction, with the chain in state $j$. We will term $\mathbf{p}$ the vector of *type selection probabilities* (TSP).

Since there is no mutation, we have that
\[
M_{i0} = \begin{cases} 
1, & i = 0 \\
0, & i = 1, \ldots, N 
\end{cases}\quad \text{and} \quad M_{iN} = \begin{cases} 
0, & i = 0, \ldots, N - 1 \\
1, & i = N 
\end{cases}.
\]

In agreement with this interpretation of $\mathbf{p}$, we will always have $p_0 = 0$, $p_N = 1$, and $0 < p_i < 1$ for $i = 1, \ldots, N - 1$, unless stated otherwise. Furthermore, we assume that all states are accessible to the dynamics, from any non-absorbing initial condition, in a sufficiently large number of steps.

All these assumptions combined lead to the following definition:

**Definition 1** (The Kimura class of matrices). *Let $\mathbf{M}$ be a $(N+1) \times (N+1)$ stochastic matrix. We say that $\mathbf{M}$ is *Kimura*, if*

\[
\mathbf{M} = \begin{pmatrix} 1 & \mathbf{a}^\top & 0 \\ 0 & \mathbf{M} & 0 \\ 0 & \mathbf{b}^\top & 1 \end{pmatrix},
\]

*where $\mathbf{M}$ is a $(N-1) \times (N-1)$ sub-stochastic irreducible matrix, $\mathbf{0}$ is the zero vector in $\mathbb{R}^{N-1}$, and with $\mathbf{a}$ and $\mathbf{b}$ non-zero, non-negative vectors in $\mathbb{R}^{N-1}$. We will also say that $\mathbf{M}$ is the *core matrix* of $\mathbf{M}$, and that $\mathbf{a}$ is the vector of one-step extinction rate, and $\mathbf{b}$ is the vector one-step fixation rate. The class of Kimura matrices is denoted by $\mathcal{K}$.*

Notice that since the sum of an irreducible matrix with a non-negative matrix is itself irreducible, we immediately have that the Kimura class is convex.

**Remark 1.** *We termed the matrices in definition 1 after the celebrated sentence from the geneticist Motoo Kimura that states that “A mutant gene which appeared in a finite population will eventually either be lost from the population or fixed (established) in it” [51]. The Kimura theory, however, is also know from the importance given to the neutral and quasi-neutral evolution (see [50]). Here, we depart as much as possible from the neutral theory.*

This class can be loosely described as time homogeneous Markov chains with two absorbing states, and ergodic transition states. It is implicitly the standard class used in many applications (see, e.g., the discussion on the Wright-Fisher model without mutations at Charlesworth and Charlesworth [18] or, more generally, Markov models with absorbing states at Taylor and Karlin [74]). When considering time-inhomogeneous processes, this class will have to be restricted; see section 5.

Finally, we note that in this work, all $(N+1) \times (N+1)$ transition matrices, except if otherwise stated, will be Kimura.

2.2. **Two archetypical evolutionary processes.** Two important evolutionary processes that fit the framework given above are the generalised Moran and generalised Wright-Fisher processes, that are defined as follows:
**Moran process (M):** The Moran process is a death-birth process — hence with overlapping generations. Given a population with \( N \) individuals at state \( i \), one individual is selected to die with probability \( 1/N \), and he/she is replaced by an individual of type \( A \) with a probability \( p_i \) and a type \( B \) individual with probability \( 1 - p_i \). Therefore, the transition matrix is given by

\[
M_{ij} = \begin{cases} 
\frac{j}{N}(1-p_j), & i = j - 1, \\
\frac{N-p_j}{N}(1-(1-p_j)), & i = j, \\
\frac{j}{N}p_j, & i = j + 1, \\
0, & |i - j| > 1.
\end{cases}
\]

**Wright-Fisher process (WF):** The Wright-Fisher model is an evolutionary process with non-overlapping generations. Given a population of constant size \( N \) at state \( j \), it is replaced by a new population, where the probability that the new generation is at state \( i \) is given by

\[
M_{ij} = \binom{N}{i} p_j^i (1-p_j)^{N-i}.
\]

Let \( X_k \) be the number of type \( A \) individuals in the population, under the corresponding evolutionary process. Then, the expected number of type \( A \) individuals, given the state \( j \), in the next step is given by

\[
\mathbb{E} [X_{k+1} | X_k = j] = \sum_{i=0}^{N} i M_{ij}.
\]

**Definition 2.** We say that an evolutionary process is neutral if the corresponding population process, \( X_k \), satisfies \( \mathbb{E} [X_{k+1} | X_k = j] = j, \) for all \( j \). Indeed, neutrality as defined by means of an exchangeable model is equivalent, for a two types populations, to the property that the population process of a given type is a martingale—cf. [12, 13].

Notice that some references define neutrality from properties of the type selection probability vector \( p \) [11, 26, 31] or from the fixation probability always being equal to the initial presence [33, 42, 49]. Here we opt for the definition from the Darwinian fitness, i.e, that their conditional expected presence is the same as the observed one [38, 62]. For a population with two types, these definitions are equivalent, as we will see below.

For the processes above we have the following result:

**Lemma.** Let \( p_j \) be the probability that an individual of type \( A \) is selected for reproduction, when the population is at state \( j \), for either the Moran or the Wright-Fisher process. Let \( X_k^M \) and \( X_k^WF \) be the populations processes associated to the Moran and to the Wright-Fisher models, respectively. Then, the expected number of type \( A \) individuals in the next time-step \( (k+1) \) when the population is at state \( j \) at time-step \( k \) is given by:

\[
\begin{align*}
\mathbb{E} [X^M_{k+1} | X^M_k = j] &= j + p_j - j/N; \\
\mathbb{E} [X^WF_{k+1} | X^WF_k = j] &= Np_j; \\
\end{align*}
\]

Note that \( \mathbb{E} [X^WF_{k+1} | X^WF_k = j] = \mathbb{E} [X^M_{k+1} | X^M_k = j] \) for \( j = 0, \ldots, N \) if, and only if, \( p_j = j/N \). In this case, both conditional expectations are equal to \( j \), and hence both processes are neutral.
Proof. For the Moran process, we have
\[
(j - 1)M_{j-1,j} + jM_{jj} + (j + 1)M_{j+1,j} = \frac{j(j - 1)}{N} + \frac{j^2 + (j + 1)(N - j) - j(N - j) - (j - 1)p_j}{N}
\]
\[
= j + p_j - \frac{j}{N},
\]
and (2) follows.

For the WF process, we have that
\[
0 = \frac{\partial p_j}{\partial p_j} \left[ \sum_{i=0}^{N} \binom{N}{i} p_j^i (1 - p_j)^{N-i} \right]
\]
\[
= \sum_{i=0}^{N} \binom{N}{i} p_j^{i-1} (1 - p_j)^{N-i-1} (i(1 - p_j) - p_j(N - i))
\]
\[
= \sum_{i=0}^{N} \binom{N}{i} p_j^{i-1} (1 - p_j)^{N-i-1} (i - Np_j),
\]
from which (3) follows.

Notice that Equation (3) further implies, for the WF process, that
\[ p_j \]
is the expected frequency of type A in the population in the next generation, given that in the present generation such frequency is \( j/N \).

2.3. Type selection probabilities and Darwinian fitness. As observed by Orr [64], biologists broadly agree on the essence of the idea of fitness, although they give many different definitions. Fundamentally, fitness measures the ability of individuals to survive and reproduce in some environment. This can be measured through the expected number of type A individuals in the next generation, from a given state—this is the Darwinian fitness [42]. On the other hand, in many different modelling approaches — as in evolutionary game theory, where fitness is identified with the game pay-off — fitness is used as a proxy to the probability to select a given type, when determining the next reproduction event; see, for example [45]. Here, we call this function reproductive fitness.

Let us denote the reproductive fitnesses by \( \varphi^{(A,B)} : \{0, 1, \ldots, N\} \rightarrow \mathbb{R}_+ \). Then, these fitness functions are typically related to the type selection probabilities by
\[
(4) \quad p_j = \frac{j \varphi^{(A)}(j)}{j \varphi^{(A)}(j) + (N - j) \varphi^{(B)}(j)}.
\]
This assumes that the population is well-mixed, and it is in agreement with the intuitive idea that \( \varphi^{(A)}/\varphi^{(B)} \) indicates the probability to be selected for reproduction.

Equation (4) is widely used in many works dealing either with the Wright-Fisher process (see, e.g., [2, 17, 45, 75]) or specifically with the Moran process [15, 37, 62, 63, 75]. Nevertheless, we will now see that the identification between Darwinian and
reproductive fitnesses is consistent only with the Wright-Fisher process for all possible cases. We follow [42] in the definition below. See also [22, 55]

**Definition 3.** Let $X_k^*$ be a stochastic population process, with finite population $N$. Let us define the (discrete time) Darwinian fitness as

$$
\Psi_k^A(j) := \frac{\mathbb{E}[X_{k+1}^* | X_k^* = j]}{j}, \quad \Psi_k^B(j) := \frac{(N - \mathbb{E}[X_{k+1}^* | X_k^* = j])}{(N - j)}
$$

The relative Darwinian fitness is

$$
\Phi_k(j) := \frac{\Psi_k^A(j)}{\Psi_k^B(j)}.
$$

For time independent Markov chains, we assume that both $\Psi_k^A, \Psi_k^B$ (and, therefore, $\Phi_k$) do not depend on $k$, and henceforth we will drop the subscript $k$.

**Lemma.** An evolutionary stochastic process with finite population is consistent both with reproductive fitness and with Darwinian fitness if, and only if, the corresponding population process satisfies:

$$
\mathbb{E}[X_{k+1} | X_k = j] = N p_j.
$$

**Proof.** From the Darwinian fitness definition, we have

$$
\Phi(j) = \frac{N - j}{j} \frac{\mathbb{E}[X_{k+1}^{WF} | X_k^{WF} = j]}{N - \mathbb{E}[X_{k+1}^{WF} | X_k^{WF} = j]}
$$

while from the reproductive fitness parametrisation, we have that

$$
\Phi(j) = \frac{N - j}{j} \frac{p_j}{1 - p_j},
$$

and the result follows. □

**Remark 2.** The above result together with formulas (2) and (3) implies that the Moran process is consistent with both reproductive fitness and with Darwinian fitness if, and only if, it is neutral. More precisely, when using the Moran process to describe the evolution of a population, we may not identify the Darwinian fitness and the reproductive fitness. In this vein, there are two possibilities: alternatively

1. One can use the reproductive fitness for modelling, but should not expect that the type selection probabilities determine the expected value of type $A$ in the next generation.

2. One can calibrate a Moran model by the Darwinian fitness. However, in this case the type selection probabilities are given by

$$
p_j = \frac{j}{N} + \frac{j(N - j)(\Phi(j) - 1)}{N - j + j\Phi(j)}.
$$
2.4. **Fixation probabilities.** The fixation vector is a left eigenvector of the transition matrix associated to the eigenvalue one, with the first entry being zero, and normalised such that the last entry is one \[47, 74\]. It can be also characterised in a more algebraically fashion—cf. \[40, 47\]—as follows

**Proposition 1.** Let \( M \in \mathcal{K} \). Then, there exists a unique vector \( \tilde{F} \in \mathbb{R}^{N-1} \), with \( 0 < \tilde{F}_i < 1 \), such that 
\[
\tilde{F}^\dagger = \begin{pmatrix} 0 & \tilde{F}^\dagger \tilde{M} + b^\dagger & 1 \end{pmatrix} = \begin{pmatrix} 0 & \tilde{F}^\dagger & 1 \end{pmatrix}.
\]

This will be satisfied if, and only if, we have
\[
(6) \quad \tilde{F}^\dagger (I - \tilde{M}) = b^\dagger.
\]

Since \( \tilde{M} \) is sub-stochastic and irreducible, \( I - \tilde{M} \) is invertible and \( (I - \tilde{M})^{-1} \) is positive—cf. \[3\]. Hence \( \tilde{F} \) is uniquely defined, and positive. A similar calculation with \( 1 - F \) instead of \( F \) yields
\[
(7) \quad (1 - \tilde{F})^\dagger (I - \tilde{M}) = a^\dagger.
\]

Hence, using the same argument as above, we conclude that \( 1 - \tilde{F} \) is positive, and hence that we have \( \tilde{F}_i < 1 \).

The vector \( F \) and, by extension, the vector \( \tilde{F} \) are called the **fixation vectors** associated to the process \( M \).

**Remark 3.** It is possible for a process not in the Kimura class to fixate. Indeed, let \( \mathcal{K}_0 \) the set of matrices with the form given by Eq. \(7\) with \( \rho(\tilde{M}) < 1 \). Then it is easy to see that a process without mutation fixates if, and only if, its transition matrix belongs to \( \mathcal{K}_0 \), although the corresponding fixation may not be admissible. On the other hand, if we write \( \mathcal{K}_1 \subset \mathcal{K}_0 \) with \( a, b > 0 \), then every matrix in \( \mathcal{K}_1 \) has an admissible fixation. It can be also shown that \( \mathcal{K}_1 \) is closed through convex combinations and multiplications. Although \( \mathcal{K}_0 \) and \( \mathcal{K}_1 \) are not our primary interest in this work, most of the results presented here could be easily extended to \( \mathcal{K}_1 \), and some even to \( \mathcal{K}_0 \). In subsection \(3.6\), we study examples that belong to these classes.

3. **Regular Fixation**

3.1. **Preliminary definitions and results.**

**Definition 4** (Admissible fixation vector). A fixation vector \( F \) satisfying \( 0 < F_i < 1, \ i = 1, \ldots, N-1 \), is termed admissible. Proposition\(7\) then states that any fixation vector from a process whose transition matrix is Kimura is admissible.
We say that \( u \in \mathbb{R}^N \) is increasing (non decreasing) if for all \( i > j \), we have \( u_i > u_j \) \((u_i \geq u_j)\).

**Definition 5** (Regular process). We say that an evolution process such that the transition matrix belongs to the Kimura class is regular, if the associated fixation vector is increasing. By extension, we shall say that a fixation vector that is increasing is a regular fixation. Notice that any regular fixation is necessarily admissible.

**Remark 4.** A regular evolutionary process conforms to the intuitive idea that the probability of fixation of a given type increases when the number of individuals of this type increases—cf. [26, 52, 72].

We begin by giving a sufficient condition for a process to be regular. In order to do so, we will need the concept of stochastic ordering of probability vectors:

**Definition 6** (Vector stochastic ordering). We say that two vectors \( u, v \in \Delta^N := \{ x \in \mathbb{R}^{N+1} | x_i \geq 0, \sum_i x_i = 1 \} \) are stochastically ordered, \( u \succ v \), if for all \( n = 1, \ldots, N \), we have that \( \sum_{i=n}^{N} u_i \geq \sum_{i=n}^{N} v_i \). If all inequalities are strict, then we say \( u \succ \succ v \).

Matrices whose columns are stochastically ordered will also be relevant:

**Definition 7** (Ordered matrices). Consider a \( N \times N \) matrix \( A \). We say that \( A \) is stochastically ordered (SO, \( A \in \text{StO}_N \)) if all column vectors are stochastically ordered, i.e., if for all \( i > j \), we have that \( A_{\cdot,i} \succ A_{\cdot,j} \). We say that \( A \) is strictly stochastically ordered (SSO, \( A \in \text{St}^2\text{O}_N \)) if for all \( i > j \), we have that \( A_{\cdot,i} \succ \succ A_{\cdot,j} \). Finally, we say that \( A \) is a banded stochastically ordered matrix (BSO, \( A \in \text{BStO}_N \)), if it is stochastically ordered, and if \( \sum_{i=1}^{N} A_{i,j+1} \geq \sum_{i=1}^{N} A_{i,j} \) for \( n \in \{ j + 1 - k_1, j + 1 + k_2 \} \cap \{ 1, N \} \) for \( k_1, k_2, \geq 1 \). We say that the matrix \( A \) has upper (lower) band of size \( k_1 \) \((k_2)\), resp. If \( k_1 = k_2 \), we say that \( A \) has a band of size \( k_1 \).

We are now ready to state our first result:

**Theorem 1.** Let \( M \) be a \((N+1) \times (N+1)\) Kimura transition matrix as in Definition 7 that is also banded stochastically ordered. Then \( M \) is regular.

*Proof.* Following Keilson and Kester [18], we let \( T \) be a \((N+1) \times (N+1)\) matrix such that \( T_{ij} = 1 \) for \( j \geq i \) and 0 otherwise. Therefore \( T_{i,i-1}^{-1} = -1 \), for \( i = 1, \ldots, N \), \( T_{ii} = 1 \) for \( i = 0, \ldots, N \) and 0 otherwise. A vector \( v \) is increasing if, and only if, \( v^\dagger T^{-1} \) has positive entries, except possibly for the first. Also, a matrix \( M \) is stochastically ordered if, and only if, \( \text{TMT}^{-1} \) is positive except possibly for the Oth column and row. Indeed, if either \( i = 0 \) or \( j = 0 \), we have \((\text{TMT}^{-1})_{ij} = \delta_{ij} \), and for \( i, j > 0 \), we have that

\[
(\text{TMT}^{-1})_{ij} = \sum_{k,l} T_{ik} M_{kl} T_{lj}^{-1} = \sum_k T_{ik} A_{kj} - \sum_k T_{ik} A_{k,j-1} = \sum_{k=i}^{N} M_{kj} - \sum_{k=i}^{N} M_{k,j-1}.
\]

Therefore

\[
\text{TMT}^{-1} = \begin{pmatrix} 1 & \mathbf{0}^\dagger \\ \mathbf{0} & \mathbf{L} \end{pmatrix},
\]

and we conclude that \( M \) is strictly stochastically ordered if, and only if, \( \mathbf{L} \) is a positive matrix. In general, if \( M \) is banded stochastically ordered, then the corresponding \( \mathbf{L} \) is
non-negative, and at least the diagonal, main superdiagonal, and lower superdiagonal are positive, and hence \( L \) is primitive. We say that \( L \) is the associated matrix to \( M \).

Notice that \( \rho(TMT^{-1}) = \rho(L) = 1 \), and that it is an eigenvalue of multiplicity two—both algebraically and geometrically—of \( TMT^{-1} \) and a simple eigenvalue of \( L \). In addition, it is easily checked that \( e_0 \) is both a left and right eigenvector of \( TMT^{-1} \) associated to the eigenvalue one. Let \( w_0 = (0 \ w^\dagger) \) be a left eigenvector of \( TMT^{-1} \) associated to one, that is linearly independent from \( e_0 \). Then, \( w \) is a left eigenvector of \( L \) with associated eigenvalue equals to 1, and the Perron-Frobenius theorem implies that \( w \) can be taken to be positive.

Let \( F \) be the fixation vector of \( M \). We will now show that \( F^\dagger = \frac{w_0^\dagger T}{\sum_{i=1}^N w_i} \) and hence that \( F \) is increasing.

Indeed, the first entry of the right hand side is zero, and last entry is one. Also, we have
\[
F^\dagger M = \frac{w_0^\dagger TMT^{-1}T}{\sum_{i=1}^N w_i} = \frac{w_0^\dagger T}{\sum_{i=1}^N w_i} = F^\dagger.
\]

\[ \square \]

Remark 5. It turns out that we can characterise the regular matrices \( M \) as exactly as those matrices for which the corresponding \( L \) matrix has a positive left eigenvector associated to its spectral radius, which is more general than to require \( L \) to be primitive. This will be discussed in Subsection 3.5.

3.2. Moran process is regularly universal. The Moran process is a special case of a more general class known as birth-death (BD) processes. A general BD process is characterised by a transition matrix that satisfy \( M_{i,j} = 0 \), if \(|i - j| > 1\). Let \( X_k \) denote the corresponding population process. Then it is easily checked that
\[
E[X_{k+1} | X_k = j] = j + M_{j+1,j} - M_{j-1,j}.
\]

If \( M \) is a Kimura tri-diagonal matrix, then the fixation vector \( F \) is given by Grinstead and Snell [40], Karlin and Taylor [47]:
\[
F_i = c^{-1} \sum_{l=1}^i \prod_{k=1}^{l-1} \frac{M_{k,k-1}}{M_{k,k+1}}, \quad c = \sum_{l=1}^N \prod_{k=1}^{l-1} \frac{M_{k,k-1}}{M_{k,k+1}}.
\]

In particular, every tri-diagonal process is regular. We now turn to a more detailed study of Moran processes:

Theorem 2. Let \( F \) be an admissible fixation vector. Then \( F \) is the fixation vector of some Moran process if, and only if, \( F \) is increasing. Moreover, in the latter case, the type fixation probabilities of the Moran process that realises such a vector are given by
\[
p_j = \frac{j(F_j - F_{j-1})}{j(F_j - F_{j-1}) + (N-j)(F_{j+1} - F_j)} \in (0, 1), \quad j = 1, \ldots, N - 1.
\]

Proof. Assume \( F \) is the fixation vector of a Moran process. Then it is increasing as a special case of Equation (8).
Conversely, assume that \( F \) is increasing, and recall that \( F \) is a fixation vector if, and only if, it satisfies the recursion

\[
F_{j-1} \frac{j}{N}(1 - p_j) + F_j \left( \frac{j}{N} p_j + \frac{N - j}{N} (1 - p_j) \right) + F_{j+1} \frac{N - j}{N} p_j = F_j.
\]

This can be rewritten as

\[
(N - j) p_j G_j - j(1 - p_j) G_{j-1} = 0, \quad j = 1, \ldots, N - 1,
\]

where \( G_j \) is the marginal gain in fixation:

\[
G_j := F_{j+1} - F_j.
\]

Equation (10) is usually solved for \( G_j \), which leads to a special case of Equation (8). However, it can be also uniquely solved for \( p_j \) yielding equation (9), which satisfies \( 0 < p_j < 1 \).

**Remark 6.** Equation (9) indicates that the marginal increasing in the fixation probability—i.e. the increase in fixation probability, when the presence of type \( X \) increase by one—can be understood as a reproductive fitness, cf. Equation (4). More precisely, we let

\[
\varphi(A)(j) := F_j - F_{j-1} > 0,
\]

\[
\varphi(B)(j) := (1 - F_j) - (1 - F_{j+1}) = F_{j+1} - F_j > 0.
\]

Although a BD process do not need to be stochastically ordered to be regular, the class of stochastically ordered matrices will be of interest when discussing time-inhomogeneous processes, and thus we will digress about this point. First, we observe that a BD process is banded stochastically ordered if, and only if, we have

\[
M_{j+1,j} + M_{j,j+1} < 1, \quad j = 1, \ldots, N - 2.
\]

In particular, the Moran process is banded stochastically ordered if, and only if,

\[
\left( \frac{1 - j}{N} \right) p_j + \frac{j + 1}{N} (1 - p_{j+1}) < 1, \quad j = 1, \ldots, N - 2.
\]

This immediately yields the following result

**Lemma.** Let \( M \) be the Moran matrix associated to type selection probability \( p \).

1. If \( p \) is increasing, we have that \( M \) is banded stochastically ordered;
2. If \( p \) is such that \( \frac{1}{N+1} < p_j < 1 - \frac{1}{N+1} \), then \( M \) is banded stochastically ordered.

3.3. Regularity and irregularity in the Wright Fisher process. The discussion of the Wright-Fisher process requires considerably more work. We begin with a Lemma that will turn out to be a very important tool

**Lemma.** Let \( F \) be a vector of fixation probabilities. Define

\[
\Upsilon_F(p) = \sum_{i=0}^{N} F_i \binom{N}{i} p^i (1 - p)^{N-i}.
\]

Then \( \Upsilon_F \) is continuous, with \( \Upsilon_F(0) = 0 \) and \( \Upsilon_F(1) = 1 \). Furthermore, if \( F \) is increasing, i.e. \( 0 = F_0 < F_1 < \cdots < F_N = 1 \), then \( \Upsilon_F(p) \) is an increasing function in \([0, 1] \).
Proof. Continuity is immediate. Also, a direct computation yields
\[ \Upsilon_F(0) = F_0 = 0 \quad \text{and} \quad \Upsilon_F(1) = \left( \begin{array}{c} N \\ N \end{array} \right) F_N = 1. \]
Let \( p \in (0, 1) \). Then
\[
\varphi'_F(p) = \frac{1}{p(1-p)} \sum_{i=0}^{N} F_i \left( \begin{array}{c} N \\ i \end{array} \right) p^i (1-p)^{N-i} (i - Np)
\]
\[
= \frac{1}{p(1-p)} \left[ \sum_{i=0}^{\lfloor Np \rfloor} F_i \left( \begin{array}{c} N \\ i \end{array} \right) p^i (1-p)^{N-i} (i - Np) + \sum_{i=\lfloor Np \rfloor + 1}^{N} F_i \left( \begin{array}{c} N \\ i \end{array} \right) p^i (1-p)^{N-i} (i - Np) \right]
\]
\[
> \frac{F_{\lfloor Np \rfloor}}{p(1-p)} \sum_{i=0}^{\lfloor Np \rfloor} \left( \begin{array}{c} N \\ i \end{array} \right) p^i (1-p)^{N-i} (i - Np) + \sum_{i=\lfloor Np \rfloor + 1}^{N} \left( \begin{array}{c} N \\ i \end{array} \right) p^i (1-p)^{N-i} (i - Np)
\]
\[
= \frac{F_{\lfloor Np \rfloor}}{p(1-p)} \sum_{i=0}^{N} \left( \begin{array}{c} N \\ i \end{array} \right) p^i (1-p)^{N-i} (i - Np) = 0.
\]

We are now in a position to characterise the regular WF processes:

**Theorem 3.** Let \( M \) be the transition matrix of the Wright Fisher process associated to the type selection probability vector \( p \). The three conditions below are equivalent.

1. The process \( M \) is regular.
2. The matrix \( M \) is strictly stochastically ordered.
3. The vector \( p \) is increasing.

Proof. \( \blacklozenge \Rightarrow \blacklozenge \) Define \( h_n(p) = \sum_{i=n}^{N} \left( \begin{array}{c} N \\ i \end{array} \right) p^i (1-p)^{N-i} \). It is clear that \( h_0(p) = 1 \) and \( h'_0(p) = 0 \) for all \( p \). Note that for \( p \in (0, 1) \) and \( n \geq 1 \)
\[ h'_n(p) = \frac{1}{p(1-p)} \sum_{i=n}^{N} \left( \begin{array}{c} N \\ i \end{array} \right) p^i (1-p)^{N-i} (i - pN). \]
We define
\[ \gamma_n := \frac{\sum_{i=n}^{N} \left( \begin{array}{c} N \\ i \end{array} \right) i p^i (1-p)^{N-i}}{\sum_{i=n}^{N} \left( \begin{array}{c} N \\ i \end{array} \right) p^i (1-p)^{N-i}}, \]
and \( \gamma_0 = p \). Furthermore, \( \gamma_n \) is the mean of \( i/N \), from \( i = n \) to \( i = N \) with probability distribution given by \( \left( \begin{array}{c} N \\ i \end{array} \right) p^i (1-p)^{N-i} \), and therefore \( \gamma_N > \gamma_{N-1} > \cdots > \gamma_1 > \gamma_0 = p \).
In particular \( h'_n(p) > 0 \) for \( n > 0 \). From the fact that \( \sum_{i=n}^{N} M_{ij} = h_n(p_j) > h_n(p_{j-1}) \), for all \( n > 0 \) (with equality for \( n = 0 \)), we conclude that \( M \) is strictly stochastically ordered.

\( \blacklozenge \Rightarrow \blacklozenge \) It follows immediately from Theorem \( \blacklozenge \).

\( \blacklozenge \Rightarrow \blacklozenge \) Since \( M \) is regular, we have that the fixation vector \( F \) is increasing. From lemma \( \blacklozenge \), \( \Upsilon_F \) is increasing, with \( \Upsilon_F(0) = 0 \) and \( \Upsilon_F(1) = 1 \). Furthermore, \( \Upsilon_F(p_i) = \)
$F_i$. We conclude that $p_0 = 1 - p_N = 0$ and that the vector $p = \Upsilon_F^{-1}(F)$ is strictly increasing.

The result above shows that not every choice of $p$ yields a regular WF process. This naturally leads to the question of what is the class of non-regular fixation probabilities that the WF process can realise. The next result shows, perhaps surprisingly, that any admissible fixation can be realised, although not necessarily uniquely.

**Theorem 4** (Universality of the Wright-Fisher process). Let $F$ be an admissible fixation vector. Then there exists at least one WF matrix that has $F$ as a fixation vector. In addition, if $F$ is increasing, than such WF matrix is unique.

**Proof.** Notice that a given $F$ is the fixation vector of the Wright-Fisher process defined by $p$ if, and only if, we have

$$\Upsilon_F(p_j) = F_j, \quad j = 0, \ldots, N.$$ 

On the other hand, from Lemma 3.3 we have that $\Upsilon_F(0) = 0$, and $\Upsilon_F(1) = 1$. In addition, we always have, for $p \in (0, 1)$,

$$0 < \Upsilon_F(p) < \sum_{i=0}^{N} \binom{N}{i} p^i(1-p)^{N-i} = (p + (1 - p))^N = 1.$$ 

Since $\Upsilon_F$ is continuous, the intermediate value theorem implies that $\Upsilon_F$ is onto $[0, 1]$. Thus, given any admissible fixation vector $F$, there exists at least one type selection probability vector $p$ such that $\Upsilon_F(p_j) = F_j$. Furthermore, if $F$ is increasing, then $\Upsilon_F$ is also increasing and therefore $p$ is uniquely defined. \hfill $\Box$

### 3.4. Evolutionary game theory and regularity of WF processes.

Most of the cases of non-constant fitnesses functions studied in the mathematical literature considers fitnesses obtained from evolutionary game theory, where pay-off are computed using two-player games. This corresponds to affine fitnesses functions, which are the simplest class of non constant functions. In the next result, we show that in its simplest and traditional setting, 2-player games, evolutionary game theory leads to increasing type selection probabilities, and hence to regular WF processes. At the end of this section, however, we show that non-regular fixation patterns are present in the WF process at the next level of generalization, i.e., in three-players game theory (quadratic fitnesses functions).

**Proposition 2.** If fitnesses functions are positive affine, then the type selection probability vector $p$ is increasing.

**Proof.** Let $\Phi^{(a)}(i) = ai + b(N-i) + \alpha$ and $\Phi^{(b)}(i) = ci + d(N-i) + \beta$, with $a, b, c, d, \alpha, \beta > 0$. We will show that the sequence

$$p_i := \frac{i \Phi^{(a)}(i)}{i \Phi^{(a)}(i) + (N-i) \Phi^{(b)}(i)} = \frac{i(ai + b(N-i) + \alpha)}{i(ai + b(N-i) + \alpha) + (N-i)(ci + d(N-i) + \beta)}$$

is increasing. Initially, let us show that

$$f(x) := \frac{x}{1 - cx + d(1-x) + \beta},$$
with \( \alpha' = \alpha/N \) and \( \beta' = \beta/N \), is increasing in the interval \([0,1]\). Let \( x = i/n \) and write

\[
p_i = \left(1 + \frac{(N - i)(ci + d(N - i) + \beta)}{iai + b(N - i) + \alpha}\right)^{-1} = \left(1 + \frac{1 - x cx + d(1 - x) + \beta'}{x ax + b(1 - x) + \alpha'}\right)^{-1}
\]

\[
= (1 + f(x)^{-1})^{-1}.
\]

Differentiating \( f \) (and dropping primes), we find

\[
f'(x) = \frac{g(x)}{(1 - x)^2(cx + d(1 - x) + \beta)^2},
\]

where \( g(x) = (d + \beta)(b - a)(x - 1)^2 + (c - d)(a + \alpha)x^2 + (d + \beta)(a + \alpha) \). \( f \) is increasing if and only if \( g \) is positive. Function \( g(x) \) is quadratic with \( g(0) = (b + \alpha)(d + \beta) > 0 \) and \( g(1) = (a + \alpha)(c + \beta) > 0 \). Furthermore, we define

\[
x_0 := \left(1 + \frac{(a + \alpha)(c - d)}{(d + \beta)(b - a)}\right)^{-1},
\]

such that \( g'(x_0) = 0 \). If \( x_0 \notin [0,1] \), then \( g \) is monotone in \([0,1]\) and from the fact that it is positive on the borders, it will be positive everywhere. Now, assume that \( x_0 \in [0,1] \).

We have that

\[
g(x_0) = \frac{(a + \alpha)((d - c)\alpha + (a - b)\beta + ad - bc)}{a - b}x_0,
\]

and \( \frac{(a+\alpha)(c-d)}{(d+\beta)(b-a)} > 0 \). Therefore

(1) If \( c - d > 0 \) and \( b - a > 0 \). Then, \( bc > ad \) and consequently \( g(x_0) > 0 \).

(2) If \( c - d < 0 \) and \( b - a < 0 \). Then, \( bc < ad \) and consequently \( g(x_0) > 0 \).

We conclude that \( g(x) > 0 \) for all \( x \in [0,1] \). We have that \( f \) is increasing in \([0,1]\) and we conclude \( p_i = (1 + f(i/n)^{-1})^{-1} \) is increasing. \( \square \)

**Corollary 1.** If the fitnesses functions are positive affine, then the Wright-Fisher matrix is regular.

A direct calculation, however, shows that if we depart from the realm of 2-player games then we can find very different results:

**Remark 7** (A non-regular three-player game). Let \( \varphi^{(A)}(x) = 15 - 24x + 10x^2 \) and \( \varphi^{(B)}(x) = 1 + 14x^2 \), which are strictly positive in the interval \([0,1]\), then \( p_i \) given from (4) is not increasing. These functions can be obtained from 3-player game theory, with \( a_0 = 15, a_1 = 3, a_2 = 1, b_0 = 1, b_1 = 1, b_2 = 15 \), where \( a_k \) (\( b_k \)) is the pay-off of a type \( A \) (\( B \), respectively) player against \( k \) other players. Note that the relative fitness \( \Psi^{(A)}/\Psi^{(B)} = \varphi^{(A)}/\varphi^{(B)} \) is decreasing and is associated to coexistence games (i.e., \( \Psi^{(A)}/\Psi^{(B)} > 1 \) for \( x \) near zero, and \( \Psi^{(A)}/\Psi^{(B)} < 1 \) for \( x \) near one).

3.5. **Characterisation of regular processes.** It turns out that we can characterise the regular matrices in terms of what is termed the Perron-Frobenius property. Following Johnson and Tarazaga [46], we denote by \( \text{PF}_N \) the set of \( N \times N \) matrices \( A \) that have the Perron-Frobenius property, i.e.,

(1) \( \rho(A) \) is a simple eigenvalue;
(2) there exist positive right and left eigenvectors of $A$ associated to $\rho(A)$.

**Theorem 5.** Let $M$ be a Kimura transition matrix, and let

$$TMT^{-1} = \begin{pmatrix} 1 & 0^T \\ 0 & L \end{pmatrix},$$

where $T$ is the matrix defined in Theorem 1. Then $M$ is regular if, and only if, $L \in PF_N$.

**Proof.** Let $F$ be the fixation vector associated to $M$, and let

$$\begin{pmatrix} F \end{pmatrix}^T (1 \ 0) = (0 \ G_1 \ G_2 \ \cdots \ G_N), \quad G_j = F_j - F_{j-1}.$$

Then $M$ is regular if, and only if, $\tilde{G}$ is positive.

Moreover, from the structure of $TMT^{-1}$, we immediately have that $\rho(L) = 1$, and that it is a simple eigenvalue of $L$. Indeed, $e_0$ is a right and left eigenvector for $TMT^{-1}$, and $(0 \ 1^T)$ is a right eigenvector of $TMT^{-1}$, and thus $1$ is a right eigenvector of $L$.

On the other hand, $\tilde{G}$ is a real left-eigenvector of $L$. Since we always have $G_1 > 0$, we have that $L \in PF_N$ if, and only if, $\tilde{G}$ is positive.

**Remark 8.** Notice that for $i = 1, \ldots, N$, we always have

$$\sum_{j=1}^{N} (TMT^{-1})_{ij} = \sum_{k=i}^{N} \sum_{j=1}^{N} (M_{kj} - M_{k,j-1}) = \sum_{k=i}^{N} (M_{kN} - M_{k0}) = 1.$$

In particular, this means that $(1 \ 1)^T$ is a right-eigenvector of $M$ associated to the eigenvalue one, and hence that $1$ is a right eigenvalue of $L$ associated to the same eigenvalue. Notice also that the column-wise sum of $L$ yields information on the difference of the marginal increase in the expected presence of the first type after one step:

$$\sum_{i=1}^{N} (TMT^{-1})_{ij} = \sum_{i=1}^{N} \sum_{k,l=0}^{N} T_{ik} M_{kl} T_{lj}^{-1} = \sum_{k,l=0}^{N} k M_{kl} T_{lj}^{-1}$$

$$= \sum_{k=0}^{N} E[X_{k+1} \mid X_k = l] T_{lj}^{-1} = E[X_{k+1} \mid X_k = j] - E[X_{k+1} \mid X_k = j - 1], j > 0.$$

In particular, this implies that $\lim_{k \to \infty} L^k$ is a matrix that is constant by columns with each column $j$ being $F_j - F_{j-1}$.

As observed in Noutsos [61], Tarazaga et al. [73], there are matrices with negative entries in $PF_N$. However, $A \in PF_N$ if, and only if, there exists $k_0 \in \mathbb{N}$ such that $A^k$ is positive for $k \geq k_0$—cf. [46]. Such matrices are termed eventually positive.

Since $M$ is strictly stochastically ordered if, and only if, $L$ is positive, we have that for a regular matrix, there exists $k_0 \in \mathbb{N}$ such that $M_k$ is strictly stochastically ordered, if $k \geq k_0$. Such matrices are called eventually strictly stochastically ordered—cf. [48].

**Remark 9.** We can use Theorem 5 to prove that BD processes are regular even if they are not stochastically ordered, and without using the explicit expression for the fixation.
probability. Indeed, for such processes we have that $L$ is a tridiagonal matrix with the non-zero entries given by

$$
L_{j,j+1} = M_{j,j+1} \\
L_{j+1,j} = M_{j+1,j} \\
L_{j,j} = 1 - M_{j,j+1} - M_{j+1,j}
$$

with $j = 0, \ldots, N - 2$ in the first two equations, and $j = 0, \ldots, N - 1$ in the last equation.

Thus $L$ is an irreducible matrix with non-negative off-diagonal elements, and hence is an irreducible Metzler matrix. Such a matrix has the Perron-Frobenius property $[3, 9]$, and hence BD processes are regular.

3.6. Alternative processes. We finish this section with some comments about other models presented in the literature. We begin by discussing two models introduced as alternative dynamics that belong to the Kimura class and for which all the theory developed so far applies directly. In the sequel, we discuss two models that are likely to be unrealistic for most biological populations, and hence should be largely taken as pedagogical examples. The first one belongs to class $K_1$ and, as noted in Remark 3, all general theorems apply. The second belongs to the class $K_0$, and it can be seen as process in the boundary of the Kimura class.

Pairwise Comparison. Another Birth-Death process that has also been used as a model of evolutionary dynamics is the so-called pairwise-comparison process (PC) $[76]$, whose transition matrix is given by:

$$
M_{ij} = \begin{cases} 
\frac{j(N-j)(1-q_j)}{N^2} , & i = j - 1 , \\
\frac{j(N-j)}{N^2} q_j , & i = j + 1 , \\
0 , & |i-j| > 1 .
\end{cases}
$$

where $q_j (1 - q_j)$ is the probability that $A$ replaces $B$ ($B$ replaces $A$, resp.) in a pair contest. This process satisfies

$$
\mathbb{E}[X_{k+1}^{PC}|X_k^{PC} = j] = j + \frac{j(N-j)}{N^2}(2q_j - 1)
$$

Hence a PC process is neutral when $q_j = \frac{1}{2}$. We point out that this fact was implicitly stated in Altrock and Traulsen $[1]$, Hilbe $[44]$. As a matter of fact, when the intensity of selection converges to zero, the replacement probability converges to $\frac{1}{2}$. Moreover, it is easily verified that every PC process is banded stochastically ordered, and therefore regular, inasmuch as we have that

$$
M_{j+1,j} + M_{j,j+1} = \frac{j(N-j)}{N^2} q_j + \frac{(j+1)(N-j-1)}{N^2}(1-q_{j-1}) < \frac{1}{4} + \frac{1}{4} < 1 .
$$

Furthermore, given any increasing fixation vector $F$ we have it is realised by a PC process upon choosing

$$q_j = \frac{G_{j-1}}{G_j + G_{j-1}} \in (0, 1),$$
with $G_j$ given by Equation (11). Finally, we observe that while $q$ is also a type selection probability, it is based on a sample over pairs instead of a sample over the whole population.

Eldon-Wakeley. This model was introduced in [27]. It is an intermediate model between the Moran and the Wright-Fisher process, in which at each time step, one individual is selected to reproduce, according to a TSP vector $\mathbf{p}$, and begets $U - 1$ new individuals, $U \in \{1, \ldots, N\}$. The parent persists, while its offspring replace $U - 1$ individuals who are selected with equal probability to die among the remaining individuals. The original work studied the neutral case, i.e., $p_i = i/N$, when it can be easily checked that $F_i = i/N$. Using the notation of the current work, the transition matrix is given by

$$M_{ij} = p_j M_{ij}^{(1)} + (1 - p_j) M_{ij}^{(2)},$$

where

$$M_{ij}^{(1)} = \binom{N - j}{i - j} \binom{j - 1}{U - 1 - i + j} \binom{N - 1}{U - 1}^{-1},$$

$$M_{ij}^{(2)} = \binom{j}{j - i} \binom{N - 1 - j}{U - 1 + i - j} \binom{N - 1}{U - 1}^{-1}.$$

We use that $\binom{n}{b} = 0$ whenever $b < 0$ or $b > a$, and therefore $M^{(1)}(M^{(2)})$ is lower (upper, respect.) triangular matrix. After some simplifications, we find that for $N + 1 \geq n \geq 1$, $0 \leq j \leq N$,

$$\sum_{i=n}^{N} (M_{i,j+1}^{(1)} - M_{ij}^{(1)}) = \frac{(n - U)(N - U)(U - 1)}{(N - j)(j - 1)} \begin{cases} > 0, & U + j \geq n \geq \max\{j + 1, U + 1\} \\ = 0, & \text{otherwise} \end{cases},$$

$$\sum_{i=n}^{N} (M_{i,j+1}^{(2)} - M_{ij}^{(2)}) = \frac{(N - j - 2)(j - 1)}{(N - U)} \begin{cases} > 0, & \min\{j + 1, N - U\} \geq n \geq j + 2 - U \\ = 0, & \text{otherwise} \end{cases}.$$
replaces the entire population. Here, we show how a small variation of our approach would apply in this case. Given a TSP vector $p$, we define the generalised $\Lambda_1$-process by the matrix

$$M_{ij} = \begin{cases} 
\frac{1-p_j}{N}, & i = 0, j = 0, \ldots, N; \\
1 - \frac{1}{N}, & i = j, j = 1, \ldots, N - 1; \\
\frac{p_j}{N}, & i = N, j = 0, \ldots, N.
\end{cases}$$

A direct calculation shows that $F_j = p_j$, and hence that this is model is regular if, and only if, the corresponding TSP is increasing. Also, another direct calculation shows that $M \in St^{2\mathbb{O}}$ if, and only if, $p$ is increasing. Hence, Theorem 3 also holds for the $\Lambda_1$ process.

We stress that, since this model belongs to the $K_1$ class, all the generic results holds for the $\Lambda_1$ process, including the discussion on time-inhomogeneous processes in Section 5. Nevertheless, the reasons as why this model is considered by their authors as “unrealistic” are possible the same as it is not included in the Kimura class, i.e., that mixed states are not necessarily accessible, even considering long time intervals, from any mixed initial condition.

Lethal mutation. This process, i.e., such that $p_i = 0$ for $i \in \{1, \ldots, N - 1\}$, is non-Kimura. Assume, however a family of increasing TSPs $p^{(\epsilon)}$, $\epsilon \geq 0$ such that $p^{(\epsilon)} \to p^{(0)} = (0, 0, \ldots, 1)$. For each value of $\epsilon > 0$ we define the WF transition matrix $M^{(\epsilon)}$ and it is clear that $M_i^{(\epsilon)} \to 0$ for $i < N$ and $M_{N,i}^{(\epsilon)} \to 1$. The continuity of $F$ with respect to $M$ — that follows from Proposition 1 — implies that, as $\epsilon \to 0$, $F^{(\epsilon)} \to (0, 0, \ldots, 1)$, which is the fixation probability of the matrix $M^{(0)}$. On the other hand, if we assume the Moran process, then $M_i^{(\epsilon)} \to i/N$, $M_{N,i}^{(\epsilon)} \to (N-i)/N$ and $M_{i+1,i}^{(\epsilon)} \to 0$ for $i < N$. The fixation vector is exactly the same as the one for the WF process and the limit of the fixation vector when $\epsilon \to 0$ is also the fixation vector of the limit matrix. We conclude that the limit models of both M and WF models, in the case of lethal mutation, can be understood as limits of Kimura regular models, and therefore they belong to the boundary of the set of regular matrices. It is clear from the previous discussion that the fixation vector $F = (0, 0, \ldots, 0, 1)$ is to be expected, independently of the precise way this model is built. The limit model belongs to the class $K_0$ discussed in Remark 5, and therefore the fixation vector exists but is not admissible. Finally, notice that if type A represents a lethal mutation, then the assumption of constant population is artificial when applied to state $i = N$, and this explains the discontinuity of fixation in the limit model.

4. Regular and smooth fixation for large populations

Given an increasing fixation vector $F$, Theorems 2 and 4 show that there are unique vectors $p_M$ and $p_{WF}$ that realises this fixation vector for the Moran process and for the Wright-Fisher process, respectively. We now want to study the behaviour of type selection vectors, if the following conditions are met:

1. $F_j = \phi(j/N)$, with $\phi : [0, 1] \to [0, 1]$ being sufficiently smooth;
2. $N$ is large, but still finite.
It will turn out that such assumptions imply on some sort of weak-selection regime. We begin with Moran process:

**Proposition 3.** Let \( \phi : [0,1] \rightarrow [0,1] \) be a \( C^2 \) function, and assume that \( F_j = \phi(j/N) \). Assume also that \( N \) is sufficiently large. Then

\[
p_j^{(N)} = \frac{j}{N} - \delta_N \frac{j}{N} \left( 1 - \frac{j}{N} \right) \frac{\phi''(j/N)}{\phi'(j/N)} + O(1/N^2),
\]

with \( 0 < \delta_N < 2/N \), is such that the fixation vector associated with a Moran process given by the type selection probability \( p \) is \( F \).

**Proof.** From Theorem 2 and on using the mean value theorem, we have

\[
p_j^{(N)} = \frac{(j/N)\phi'(x_1)}{(1 - j/N) \phi'(x_2) + (j/N)\phi'(x_1)} x_1 \in \left( (j-1)/N, j/N \right), x_2 \in \left( j/N, (j+1)/N \right)
\]

\[
\frac{j}{N} - \delta_N \frac{j}{N} \left( 1 - \frac{j}{N} \right) \frac{\phi''(j/N)}{\phi'(j/N)} + O(1/N^2),
\]

with \( 0 < \delta_N < 2/N \). \( \square \)

**Remark 10.** Notice that if we use the fixation probability yielded by Replicator-Diffusion equation, then \(-\phi''(x)/\phi'(x)\) is the gradient of selection, i.e., the difference between fitnesses of types \( \mathbb{A} \) and \( \mathbb{B} \), in the weak selection regime [15, 17].

In order to deal with this question for the WF process, we need a result from approximation theory:

**Lemma (See 29, Section 3.6).** Let \( f : [0,1] \rightarrow \mathbb{R} \) be a Lipschitz continuous function, with Lipschitz constant \( K \), and let

\[
\Upsilon_{f}^{N}(x) = \sum_{i=0}^{N} f(i/N)B_{i,N}(x), \quad B_{i,N}(x) = \binom{N}{i} x^i (1-x)^{N-i}.
\]

Then

\[
\| f - \Upsilon_{f}^{N} \|_{\infty} \leq \frac{9K}{4N^{1/2}}.
\]

This result implies in the following important result

**Theorem 6 (Continuity of fixation and weak-selection).** Assume that the fixation probability is described by an increasing continuously differentiable function \( \phi : [0,1] \rightarrow \mathbb{R} \), such that there exists constants \( K_-, K_+ > 0 \), with \( K_- \leq \phi'(x) \leq K_+ \), \( x \in [0,1] \). Consider the WF process with a population of size \( N \), and let \( p^{(N)} \) be a vector of type selection probabilities such that the corresponding fixation vector \( F^{(N)} \) satisfies \( F_j^{(N)} = \phi(j/N) \). Then

\[
\left| \frac{j}{N} - p_j^{(N)} \right| \leq \frac{9K_+}{4K_-N^{1/2}}, \quad j = 1, \ldots, N-1.
\]
Proof. On one hand, we have by the mean value theorem that
\[ \left| \phi\left(\frac{j}{N}\right) - \phi(p_j^{(N)}) \right| = \left| \phi'(\bar{z}) \right| \left| \frac{j}{N} - p_j^{(N)} \right| \geq K \left| \frac{j}{N} - p_j^{(N)} \right|, \]
for some \( \bar{z} \) in the open interval delimited by \( j/N \) and \( p_j^{(N)} \). On the other hand, we also have by Theorem 4 that
\[ \left| \Upsilon_N^{\phi}(p_j^{(N)}) - \phi(p_j^{(N)}) \right| \leq 9K \frac{1}{4N^{1/2}}. \]
Since \( \Upsilon_N^{\phi}(p_j^{(N)}) = \phi(j/N) \), the result follows. \qed

5. Ergodicity and regularity in time-inhomogeneous processes

5.1. An algebra of evolutionary matrices. Understanding evolution in fluctuating environments has always been an important issue — cf. the earlier review in [32] or the entire chapter devoted to this topic in [38] — and it has been gaining importance recently cf. [5, 53]. This question leads naturally to the study of products of transition matrices, and how it might impact on the ergodicity of the non-absorbing states, and on the regularity of such processes.

We are thus led to consider whether the class discussed in Section 2.1 is closed under products. Since the product of irreducible or primitive matrices is not necessarily irreducible nor primitive, such a closure is not to be expected. Indeed, given two tridiagonal matrices such that their core have null diagonal, they are irreducible, but their product — a pentadiagonal matrix with super- and sub-diagonal identically zero in the kernel — is not. Incidentally, tridiagonal matrices with null diagonal is a well-known device used for computation biologists to speed up simulations for fixation. However, it will turn out from the discussion below that this device will usually not work for time-inhomogeneous processes. In order to deal with these difficulties, we begin by restricting the Kimura class as follows:

Definition 8 (The Gillespie class of matrices). We say that a matrix \( A \) is totally indecomposable if there are no permutation matrices \( P \) and \( Q \) such that \( PAQ = \begin{pmatrix} B & 0 \\ 0 & D \end{pmatrix} \), with \( B, D \) non-trivial square matrices and \( 0 \) the null matrix. We say that a Kimura transition matrix \( M \) is a Gillespie matrix if \( \tilde{M} \) is totally indecomposable.

Remark 11. We termed the matrices in definition 8 after John H. Gillespie who, as far as we know, was the first to systematically study time-inhomogeneous evolutionary processes for finite populations using stochastic techniques, cf. [38].

Totally indecomposable matrices can be also characterised as irreducible matrices that have positive diagonal [52]. Such a characterisation immediately implies that the product and convex combinations of totally indecomposable matrix are again totally indecomposable, and this leads to the following result:

Proposition 4. The class of Gillespie matrices is a convex set and it is closed by multiplication.

Remark 12.
(1) The idea behind the use of totally indecomposable products is the fact that there is no two subset of non-homogeneous states, i.e., \( I_1, I_2 \subset \{1, \ldots, N-1\} \), \( I_1, I_2 \neq \emptyset \) such that \( I_2 \) cannot be reached from any state in \( I_1 \).

(2) Any totally indecomposable matrix is primitive, and in particular irreducible. Thus, the Gillespie class is contained in the Kimura class, and all previous results apply. On the other hand, all examples of Kimura matrices studied so far are also Gillespie matrices, therefore the Gillespie class includes all ergodic Birth-death, Wright-Fisher, pairwise comparison and Eldon-Wakely processes.

(3) The class of Gillespie matrices is, thus, what is known as a convex semi-group, i.e., it is a semi-group that is also a convex set, with the group operation being compatible with the convex structure [8, 69].

5.2. Periodic environments. One feature of periodic varying environments is that, in general, the fixation probability depends not only on the initial presence, but also on the current time state of the environment. Indeed, consider a periodic environment of period \( l \), and let the corresponding transition matrices be \( M_0, \ldots, M_{l-1} \). We extend the indices of the matrices for all integers, such that if \( k \equiv k' \mod l \), then \( M_k = M_{k'} \).

Let
\[
P_k = M_{k+l-1} M_{k+l-2} \cdots M_{k+1} M_k,
\]
i.e., \( P_k \) identifies the products of transition matrices after \( l \) steps (one period) when the process starts with \( M_k \). We also define \( P_\infty := \lim_{n \to \infty} P_0^n \).

In the following result, we prove that the joint fixation probability will not depend on the initial condition, if and only if the fixation probabilities associated to all instantaneous evolutions are the same.

**Lemma.** We have that
\[
\lim_{n \to \infty} P_k^n = P_\infty, \quad k = 1, \ldots, l-1,
\]
if, and only if, the fixation vectors of \( M_k \), \( k = 0, \ldots, l-1 \) are the same.

**Proof.** First, we recall that if \( A \) and \( B \) are two Gillespie matrices, then we have that \( \lim_{k \to \infty} A^k = \lim_{k \to \infty} B^k \) if, and only if, they have the same fixation probability. Note also that due to the semi-group property, the fact that all \( M_k \) matrices are Gillespie implies that all \( P_k \) are Gillespie.

\( \Leftarrow \) Let \( F \) be such that \( F M_k = F \) for all \( k \), with \( F_0 = 1 - F_N = 0 \). It is immediate that \( F P_k = F \) for all \( k \), and hence Equation (12) holds.

\( \Rightarrow \) First, we notice that \( M_{k+l} P_k = P_{k+1} M_k \), and therefore \( M_k P^n_{k+1} = M_{k+l} P^n_k = P^n_{k+1} M_k \). Taking \( n \to \infty \) and using Equation (12) yields
\[
M_k P_\infty = P_\infty M_k, \quad k = 0, \ldots, l-1.
\]

Let
\[
M_k = \begin{pmatrix}
1 & a_k^\dagger & 0 \\
0 & M_k & 0 \\
0 & b_k^\dagger & 1
\end{pmatrix}
\quad \text{and} \quad
P_\infty = \begin{pmatrix}
1 & (1 - \bar{F})^\dagger & 0 \\
0 & \bar{F} & 0 \\
0 & \bar{F}^\dagger & 1
\end{pmatrix},
\]
with $F^\dagger = (0 \ F^\dagger \ 1)$ the fixation vector of all $P_k$. Together with the identity in Equation (13) we find
\[
\tilde{F}^\dagger M_k + b_k^\dagger = \tilde{F}^\dagger
\]
for all $k$, and hence that $F$ is the fixation vector of $M_k$. \hfill \Box

There is still one caveat before finishing this subsection. The fact that $G$ is a convex semi-group implies that for any finite sequence $M_0, \ldots, M_{l-1} \in G$, we have that $P := M_{l-1} \cdots M_0 \in G$, and that $P^n \to P^\infty$, when $n \to \infty$. Nothing is said about the convergence of the product
\[
\prod_{k=0}^{\infty} M_k.
\]
In order to investigate such convergence, first we observe that all the partial products are of the form
\[
\cdots, P_0^n, M_0P_0^n, M_1M_0P_0^n, \ldots, M_{l-2} \cdots M_0P_0^n, P_0^{n+1}, \ldots.
\]
A necessary and sufficient condition for convergence of the product is that all terms in the above equation converge to the same limit, i.e. the infinite product converges if, and only if, the following equalities are satisfied
\[
P_\infty = M_0P_\infty = M_1M_0P_\infty = \cdots = M_{l-2} \cdots M_0P_\infty.
\]
But since $M_kP_\infty = P_\infty$, these equalities are all satisfied, and convergence follows.

Notice, however, that this argument does not work for non-periodic products of the matrices $M_k$ — in particular, it does not apply for random products. This will be discussed in Subsection 5.4.

5.3. Regular and non-regular evolution. We now address the following question: assume that a time inhomogeneous process is stepwise regular. Is such a process itself regular? In order to answer this question, we begin with the following result

**Lemma.** Let $\mathcal{M}_N$ be one of the sets $\text{StO}_N$, $\text{St}^2O_N$ or $\text{BStO}_N$. Then if $A, B \in \mathcal{M}_N$, then $AB \in \mathcal{M}_N$ and $\mu A + (1 - \mu)B \in \mathcal{M}_N$ for all $\mu \in [0, 1]$. Furthermore, $\text{St}^2O_N \subset \text{BStO}_N \subset \text{StO}_N$.

**Proof.** Let the matrix $T$ be as defined in Theorem 1. It is clear that $TAT^{-1}$ is positive if and only if $A \in \text{St}^2O_N$, $TAT^{-1}$ is non-negative with the three principal diagonals positive if and only if $A \in \text{BStO}_N$ and $TAT^{-1}$ is non-negative if and only if $A \in \text{StO}_N$. The result follows noting that $TX_1X_2T^{-1} = TX_1T^{-1}TX_2T^{-1}$ and $T(\mu X_1 + (1 - \mu)X_2)T^{-1} = \mu TX_1T^{-1} + (1-\mu)TX_2T^{-1}$. \hfill \Box

**Theorem 7.** The intersection of the set of banded stochastically ordered matrices with the set of regular Gillespie matrices is a convex semigroup.

**Proof.** The result follows from Lemma 5.3, Proposition 4 and Theorem 1. \hfill \Box

**Corollary 2.** Let $\mathcal{R}$ be one of the following set of matrices:

1. $\text{WF}$ matrices with increasing $p$ (or, equivalently, regular $\text{WF}$ matrices).
2. $M$ matrices with increasing $p$. 
(3) $M$ matrices with $p \in (\epsilon_N, 1 - \epsilon_N)$, $\epsilon_N = 1/(N+1)$.
(4) The union of any two of the previous sets or of all three.

Then the set generated by convex combinations and finite products of elements of $\mathcal{R}$ is a convex sub-semigroup of regular Gillespie matrices.

**Remark 13.** Notice that Theorem 3 and Corollary 2 show that a non-homogeneous WF process, i.e., a process instantaneously given by a WF matrix, that is regular at every instant is itself regular. In particular, time inhomogeneity alone cannot induce non-regular fixation patterns. See [38].

For the Moran, and for tri-diagonal processes in general, multiplication outside of the class identified in Corollary 2 can produce non-regular matrices. Indeed, let $p_1 = (0, \frac{1}{7}, \frac{6}{7}, 1)$ and $p_2 = (0, \frac{6}{7}, \frac{1}{7}, 1)$.

Then the corresponding Moran matrices are given by

$$M_1 = \begin{pmatrix} 1 & \frac{2}{7} & 0 & 0 \\ 0 & \frac{13}{21} & \frac{2}{7} & 0 \\ 0 & \frac{2}{7} & \frac{21}{2} & 0 \\ 0 & 0 & 0 & \frac{7}{1} \end{pmatrix} \quad \text{and} \quad M_2 = \begin{pmatrix} 1 & \frac{1}{7} & 0 & 0 \\ 0 & \frac{2}{7} & \frac{8}{13} & 0 \\ 0 & \frac{2}{7} & \frac{21}{2} & 0 \\ 0 & 0 & 0 & \frac{7}{1} \end{pmatrix}. $$

Let

$$M_3 = M_1M_2 = \begin{pmatrix} 1 & \frac{23}{147} & \frac{8}{147} & 0 \\ 0 & \frac{44}{485} & \frac{44}{485} & 0 \\ 0 & \frac{44}{485} & \frac{44}{485} & 0 \\ 0 & \frac{44}{485} & \frac{44}{485} & 1 \end{pmatrix},$$

and let $F_i$, $i = 1, 2, 3$, be the corresponding fixation vectors. It is easy, though tedious, to check that

$$F_1 = (0, \frac{1}{5}, \frac{4}{5}, 1)^\dagger, \quad F_2 = (0, \frac{22}{25}, \frac{13}{25}, 1)^\dagger, \quad F_3 = (0, \frac{244}{485}, \frac{241}{485}, 1)^\dagger.$$

Hence $M_3$ is not regular despite the fact that $M_1$ and $M_2$ are regular.

**Remark 14.** This example can be generalised straightforwardly any even dimensions (and also adapted to odd ones). Indeed, assume $N+1$ is even let $p_1 = (0, \frac{1}{k}, \frac{k-1}{k}, \ldots, \frac{1}{k}, \frac{k-1}{k}, 1)$ and $p_2 = (0, \frac{1}{k}, \frac{1}{k}, \ldots, \frac{1}{k}, \frac{1}{k}, 1)$. Then for sufficient large $k$ the product of $M_1$ and $M_2$ is non-regular.

The above example can be considered a deterministic version in evolutionary dynamics of Parrondo’s paradox in economy [41, 67]. In this sense, if evolution is described by $M_1$ in summer, and by $M_2$ in winter, then we obtain the following conclusions:

(1) In an environment that is always summer or winter, the fixation probability of type A with two individuals is larger than with just one individual. However, in the switching case the reverse holds. See [65] for a similar finding in biochemistry.
The probability of fixation when there is only one individual in the switching case is larger than in every state of the non-switching case, while when there are two individuals it is smaller. This is the same “surprising effect” presented in \[4\], however with deterministic time evolution, i.e., no stochasticity was assumed in the time evolution of the model. See also \[58, 82\].

**Remark 15.** Let us define a matrix $N$ such that $N_{ii} = 0$ and $N_{ij} = \frac{M_{ij}}{\sum_{k \neq i} M_{ki}}$ for $i \neq j$ and let $F$ be the fixation probability associated to $M$. It is immediate to prove that $FN = F$. In effect

$$\sum_i F_i N_{ij} = \frac{\sum_{i \neq j} F_i M_{ij}}{\sum_{k \neq i} M_{ki}} = \frac{F_j - F_j M_{jj}}{1 - M_{jj}} = F_j .$$

This observation has been used to speed up numerical computations of the vector $F$. See, e.g. \[7\].

On the other hand, given two Moran matrices $M_1$ and $M_2$, with fixation vectors $F_1$ and $F_2$, one obtains matrices $N_1$ and $N_2$, using the procedure described above, that have the same fixation vectors. Nevertheless, the fixation vectors of $M_1 M_2$ and $N_1 N_2$ will be different, unless in very special cases.

### 5.4. Aperiodic evolution.

As observed in Subsection 5.2, the convergence of non-periodic products is not guaranteed by the semi-group property, despite the fact that all finite products belong to $G$. However, we will now show that the results in Daubechies and Lagarias \[24\] (see also \[10\]) can be applied in our present context. Following Daubechies and Lagarias \[24\], we say that a set $S = \{M_0, \ldots, M_{l-1}\}$ is a Left Convergent Product (LCP) set if for any sequence of integers $\{d_i\}_{i=1}^{\infty}$ with $0 \leq d_i \leq l - 1$ we have that the left product

$$\prod_{i=1}^{\infty} M_{d_i} = \cdots M_{d_n} \cdots M_{d_2} M_{d_1}$$

is well defined. Let us also write

$$TM_k T^{-1} = \begin{pmatrix} 1 & 0 \\ 0 & L_k \end{pmatrix} ,$$

and therefore the set $S$ may be defined by the $(N + 1) \times (N + 1)$ matrices $M_k$ or by the $N \times N$ matrices $L_k$.

We now show the following result:

**Lemma.** Let $S$ be a finite set with $S \subset StO_{N+1}$. Assume that there exists an integer $m$, such that for all $k \geq m$ we have that all products $L_{d_k} \cdots L_{d_1}$ have a positive column. Then $S$ is an LCP set.

**Proof.** Under the assumptions, the matrices $L_k$ are row stochastic matrices with a single eigenvalue $\lambda = 1$, cf. Remark 8. The existence of the integer $m$ implies (for the transpose) in condition (C4) of Theorem 6.1 of \[24\], from which the result follows. $\Box$

As a special case of the Lemma, we observe that if $S \subset BStO$ then every product $\prod_{i=1}^{n} M_{d_i}$ is positive, for $n > N$, and hence the set is LCP. Thus, the same condition
that guarantees regularity also ensures that arbitrary products drawn from a finite set exists.

From a more general perspective, if the set $S$ is infinite, and if the factors are drawn following an ergodic stationary stochastic process then, since all matrices are stochastic, their product converges almost surely—cf. [43]. Notice that the conditions required for LCP here are more restrictive, however the conclusions are stronger in the sense that every infinite product is definite and not just almost all. Notice also, that under the hypothesis we have that the limit function, as defined in Daubechies and Lagarias [24], is continuous.

**Remark 16 (Limit function).** As observed above, under the assumptions of Lemma 5.4, if $|S| = l$, and if we write $S_l$ to denote the set of sequences $d = \{d_k\}_{k=1}^\infty$, with $d_k \in \{0, 1, \ldots, l-1\}$, endowed with the metric $D(d, d') := l^{-r}$, where $r$ is the first index such that $d_r \neq d'_r$. Then there exists a continuous matrix function $M_\infty : S_l \rightarrow M_{N+1}(\mathbb{R})$ describing all the possible products with elements from $S$. Namely, we have that

$$
\prod_{k=1}^{\infty} M_{d_k} = M_\infty(d) = \begin{pmatrix} 1 & 1 - F_1(d) & \cdots & 1 - F_{N-1}(d) & 0 \\
0 & 0 & \cdots & 0 & 0 \\
0 & F_1(d) & \cdots & F_{N-1}(d) & 1 \end{pmatrix}
$$

In the present context, this implies that the functions $F_i : S_l \rightarrow [0, 1]$ are uniformly continuous in the topology induced in the space of sequences by the metric $D$, i.e., given $\epsilon > 0$ there exists $\delta > 0$ such that $|F_i(d) - F_i(d')| < \epsilon$, whenever $D(d, d') < \delta$. This representation implies two results: (i) that the fixation probability converge uniformly along the product; (ii) after a sufficient large, but finite, number of steps, the fixation probability is known with large precision. In particular, one can compute the fixation probability of an arbitrary infinite product (random or not) of matrices drawn from an LCP set using a finite and potentially small sub-product.

**Remark 17 (Mixtures vs random products).** Given a finite set $S = \{M_1, \ldots, M_l\} \subset \mathcal{G}$, the convex combination of elements of $S$ can be interpreted as a mixture of matrices in $S$. While mixtures enjoy a long tradition in statistics and other areas, their applicability in evolutionary dynamics seems to be first pointed out by Der et al. [22]. Notice, however, that the stochastic processes associated to the transition matrices are somewhat more limited under the mixture approach than under the random matrix product approach. As an example, take $l = 2$, and let $X_k, k \in \mathbb{N}$ be i.i.d, with values $M_1$ or $M_2$, and binomial distribution of parameter $p$. Then

$$
\mathbb{E}[X_k] = \mathbb{E}[X_1] = pM_1 + (1 - p)M_2.
$$

Also, let $X^n = \prod_{k=1}^{n} X_k$. Then, because of independence, we have

$$
\mathbb{E}[X^n] = \mathbb{E}[X_1]^n
$$

Thus a mixture replaces the stochastic process for the transition matrices by its expected value under the assumption of independence (or stationary ergodicity), while a random product allows more generic behaviour.
6. Discussion

We believe that the contribution of this work goes along three main lines: (i) an axiomatisation of the algebraic properties of evolutionary processes; (ii) a qualitative study of fixation in finite populations, including the identification and characterisation of regularity; (iii) the study of how to compose basic processes in order to model environmental modifications, i.e., the rigorous construction of time-inhomogeneous evolutionary processes. In this final section, we give an unified view of this work, showing how these three components may be seen as parts of the same general framework.

Indeed, revisiting the typical basic setup allowed us to introduce the parametrisation through the so-called type selection probabilities. They have the advantage of being directly accessible, and hence are particularly helpful for understanding basic principles of the models — as some of the results obtained here already indicate.

Along the second line, we completely linked regular Wright-Fisher (WF) process and the newly introduced vector of type selection probabilities. Furthermore, we also showed the existence of non-regular WF processes. In general, WF process are associated to micro-evolution, i.e., step-by-step evolution, while large jumps are possible, but rare \cite{18}. However, for certain choices of the frequency-dependent fitnesses functions, if a population has evolved to a monomorphic configuration (i.e., type \(A\), say, has fixed) it will be more likely that such a fixation occurred through a large jump from an intermediate step, than after a long and continuous process, where \(x\), the fraction of type \(A\) individuals in the population, on average, steadily increases towards 1 from smaller values. As an example, let us consider the case where environmental conditions have not changed (i.e., fitnesses functions are the same) since \(t = 0\), when the population was mixed, and at a later time we find a population in the state \(x = 1\). What was the most probable state of the population at time \(t = 0\)? On the absence of further intermediate measurements, and on using a maximum likelihood estimator, we conclude that the most probable state is given by \(i^* = \arg \max_{i \in \{1, \ldots, N-1\}} F_i\). Assuming that all mixed states are equally probable at \(t = 0\), an application of Bayes’ Theorem will give the same answer. Indeed, if \(F_i\) is initially increasing, then it drops close to zero for larger values of \(i\), and it eventually increases until \(F_N = 1\) only near \(x = 1\), then it is clear that fixation of type \(A\) is possible, but it is more likely to happen, if it avoids larger values of \(i\) — or, in simpler words, if it jumps from intermediate values of \(i\), when both types are present in comparable amounts, straight into fixation, i.e., with \(i = N\).

Notice that, according to Theorem \ref{thm:main}, large jumps in the WF process will be possible only if fitnesses functions are not affine in \(x\). This shows that, within the WF framework, truly multiplayer games might have much more complex dynamics. As a matter of fact, we will show elsewhere that any fixation pattern, and therefore any relative fitness, can be well approximated by pay-offs from \(d\)-player game theory, provided \(d\) is large enough. Affine functions correspond to 2-player games; therefore, discontinuities in evolution (jumps) are associated to interactions in the population involving necessarily more than 2 individuals at the same time and that cannot be reduced to a series of pairwise interactions. One possible example is the evolution of the language \cite{6}. As human evolution is regulated by complex social interaction \cite{54}, we may expect frequent
discontinuities in evolutionary traits, specially if more types (i.e., pure strategies in a game) are allowed \[39\].

The topic of small versus large changes in evolution (or, in other words, the compatibility between micro- and macro-evolution) is an import one. This is the traditional dichotomy between the gradualist view and the punctuational view of evolution. See, for example, the discussion in Charlesworth et al. \[19\], Frazzetta \[36\] and references therein. As explained in Frazzetta \[36\] “large steps in evolution are more infrequent than small ones (…). But that fact alone cannot be used to dismiss large-step change.” More precisely 22% of substitutional changes at the DNA level can be attributed to punctuational evolution \[66\]. In view of the discussion in section 3 we find perfect compatibility between models used primarily for the study microevolution (the \textit{Fisher’s microscope} \[36, 78\]) and jumps in the evolutionary process. It is important to stress that here we discuss discontinuous evolution from mixed populations to homogeneous one, without intermediate mixed states, but the word \textit{macroevolution} has many different meanings \[28\]; our approach may describe discontinuities in the fossil record \[36\], not speciation \[28\].

As a consequence of the discussion in section 4, if the population is large and the fixation probability is the restriction of an increasing smooth function (the same for all sufficiently large $N$), then we are are forced to be in the quasi-neutral (or weak-selection) regime. This conforms to the idea that an allele conferring great advantage will typically have a large effect (see \[36\] and references therein), alternatively, if the force of selection is small, the process is regular and the fixation probability is continuous: no jumps are allowed.

In this work, we also offer the construction of an algebraic framework to study theoretical population genetics. Indeed, this idea is not new and can be traced back at least to the Ph.D thesis of Cotterman and Shannon \[20, 71\] — cf. \[21\] for an interesting historical perspective on these thesis. In this work, we formulate a general theory for evolutionary process in finite populations of haploid type, constant size, without mutations but considering very general effects from natural selection. Most of the modelling in population genetics assume a constant transition matrix between all possible states in a population. In particular, they assume a choice of a certain stochastic process (like Moran, Wright-Fisher, pairwise comparison among many others). In this work, we were able to formulate stochastic processes where these, say, classical processes are combined. We consider two different possibilities:

1. Given processes $M_1$ and $M_2$, we consider a new stochastic process given by the convex combination with parameter $\mu \in (0, 1)$. This is what is known as a mixture of processes, and as observed in Remark 17 replaces the random product of these matrices by the corresponding expected value. This yields a time homogeneous mean-field approximation of this stochastic behaviour.

2. Given processes $M_1$ and $M_2$, we consider a new stochastic process given by the product of the transition matrices. This represents a time inhomogeneous evolution, with the inhomogeneity being either deterministic—as in the case of periodical variations—or random.
Mixtures, or convex combinations, have been already used in modelling evolution in Eldon and Wakeley [27], but its importance in evolutionary processes seems to be first pointed out by Der et al. [26]. Inhomogeneous Markov chains have been considered in evolutionary models previously—a recent example is [4]. On the other hand, we are not aware of a unified treatment and the identification of the underlying algebraic structure of transition matrices usually employed in modelling evolution dynamics—convex semi-group of evolutionary matrices.

In this work, we paid particular attention to the behaviour of the fixation probability, in particular to the study of its monotonicity with respect to the initial presence of a given type. We have also paid special care to grantee that the evolution processes are always ergodic. This led us to the introduction of the Gillespie class, as a set of matrices that is closed under multiplication (representing time inhomogeneity in evolution) and convex combinations (representing mixtures) that includes as particular cases the Moran process and the Wright-Fisher process. Furthermore, we define a subset where regularity is preserved under the same operations. Some of the qualitative results obtained in this class of matrices will not depend on the details of the modelling assumption, which is important because neither model is a first-principle model. We have also identified subclasses within the Gillespie class, which preserves regularity under composition.

We showed that it is possible to build an evolutionary dynamics version of the Parrondo’s paradox (i.e., type A has a larger fixation probability than type B, given a certain initial condition, in two different environments but a smaller one in the switching environment) for the Moran process, but that this is not possible in the Wright-Fisher case.

We also hope that this manuscript sheds some light on the role of the neutral evolution. More specifically: assume a non-neutral Gillespie matrix $M$ and a neutral one $N$. The associated fixation vectors are $F_M$ and $F_N$, respectively, where $F_N = (0, N^{-1}, 2N^{-1}, \ldots, 1)$. Stochastic processes given by $\alpha M + (1 - \alpha)N$, $MN$ and $NM$ have, in general, fixation vectors that are different from $F_M$ (and, clearly, from $F_N$). In this sense, it shows that neutrality is a property that strongly depends on the environment and the interactions within the population. For instance, let us say that the evolution is given by the deterministic environment $MN$, and therefore we would like to say that in the odd steps, evolution is instantaneously neutral. However, as the effect of the neutral evolution is context-dependent, we may not extend it to the entire process and say that it is neutral part of the time, or that it is neutral with a certain probability. This makes the definition of neutrality used here closer to the concept of iso-neutrality, and not to the stronger concept of euneutrality, cf. [68]. In [26], the conditional expectation defines neutrality, but the definition of a stronger concept, called “pure drift (...) process” requires that the variance is also comparable with of the neutral Wright-Fisher process (the same as in the case of the Moran process, up to multiplicative constants). Therefore, neutral matrices can induce distinctive behaviours in a a stochastic process depending, possibly, in the associated higher-order moments.
In previous works of the authors [15, 16], diffusive limits of Moran and Wright-Fisher process were calculated. A general Kimura equation, called in these works “replicator-diffusion equation” was obtained as an approximated PDE for the stochastic process. Considering what was discussed above, an important follow up of the current work would be the derivation of a diffusion approximation for time dependent fitness. Some derivations of time dependent Kimura equation appear in the literature, but they are obtained from semi-heuristic considerations, and not as large population limits of basic stochastic processes (see [23, 77]). In this vein, a particular question to be tackled would be how reversed dominance (i.e., type A dominates B in summer and the reverse holds in winter) would be able to generate metastable (quasi-stationary) intermediate states. This would suggest that the existence of metastable states might be a natural consequence of a changing environment, and is not necessarily (as sometimes claimed) a strategy that species develop to deal with uncertain future environment [14]. This will clearly depend on the ratio between characteristic intergeneration and oscillatory timescales, and the strength of stochastic effects determining the environment conditions.

Acknowledgements. FACCC was partially supported by FCT/Portugal Strategic Project UID/MAT/00297/2013 (Centro de Matemática e Aplicações, Universidade Nova de Lisboa) and by a “Investigador FCT” grant. FACCC is also in debt for useful discussions in preliminary ideas of this work with Alexandre Baraviera (Universidade Federal do Rio Grande do Sul, Brazil) and Charles Johnson (College of William and Mary, USA). MOS was partially supported by CNPq under grant # 308113/2012-8. MOS also thanks the hospitality of the Universidade Nova de Lisboa and the partial support under grant UID/MAT/00297/2013. MOS further thanks preliminary discussions of some the ideas in this work with the working group in evolutionary game theory at Universidade Federal Fluminense.

References

[1] Altrock, P. M. and Traulsen, A. (2009). Fixation times in evolutionary games under weak selection. New J. Phys., 11(1):013012.
[2] Antal, T. and Scheuring, I. (2006). Fixation of strategies for an evolutionary game in finite populations. B. Math. Biol., 68(8):1923–1944.
[3] Arrow, K. J. (1989). A “dynamic” proof of the Frobenius-Perron theorem for Metzler matrices. Probability, statistics, and mathematics, Pap. in Honor of Samuel Karlin, 17-26 (1989).
[4] Ashcroft, P., Altrock, P. M., and Galla, T. (2014a). Fixation in finite populations evolving in fluctuating environments. J. R. Soc. Interface, 11(100).
[5] Ashcroft, P., Altrock, P. M., and Galla, T. (2014b). Fixation in finite populations evolving in fluctuating environments. J. R. Soc. Interface, 11(100):20140663.
[6] Atkinson, Q. D., Meade, A., Venditti, C., Greenhill, S. J., and Pagel, M. (2008). Languages evolve in punctuational bursts. Science, 319(5863):588.
[7] Barbosa, V. C., Donangelo, R., and Souza, S. R. (2010). Early appraisal of the fixation probability in directed networks. Phys. Rev. E, 82:046114.
[8] Berg, C. (1990). Positive definite and related functions on semigroups. The analytical and topological theory of semigroups, Conf., Oberwolfach/Ger. 1989, De Gruyter Expo. Math. 1, 253-278 (1990).
[9] Berman, A. and Plemmons, R. J. (1979). Nonnegative matrices. *The Mathematical Sciences, Classics in Applied Mathematics*, 9.

[10] Bru, R., Elsner, L., and Neumann, M. (1994). Convergence of infinite products of matrices and inner-outer iteration schemes. *Electron. T. Numer. Anal.*, 2(3):183–193.

[11] Bürger, R. (2000). *The mathematical theory of selection, recombination and mutation*. Chichester: Wiley.

[12] Cannings, C. (1974). The latent roots of certain Markov chains arising in genetics: a new approach, I. Haploid models. *Adv. Appl. Probab.*, pages 260–290.

[13] Cannings, C. (1975). The latent roots of certain Markov chains arising in genetics: a new approach, II. Further haploid models. *Adv. Appl. Probab.*, pages 264–282.

[14] Carja, O., Liberman, U., and Feldman, M. W. (2014). Evolution in changing environments: Modifiers of mutation, recombination, and migration. *Proc. Nat. Acad. Sci. USA*, 111(50):17935–17940.

[15] Chalub, F. A. C. C. and Souza, M. O. (2009). From discrete to continuous evolution models: a unifying approach to drift-diffusion and replicator dynamics. *Theor. Pop. Biol.*, 76(4):268–277. Also available as an Arxiv preprint: 0811.0203.

[16] Chalub, F. A. C. C. and Souza, M. O. (2014). The frequency-dependent Wright-Fisher model: diffusive and non-diffusive approximations. *J. Math. Biol.*, 68(5):1089–1133. Available also as an Arxiv preprint.

[17] Chalub, F. A. C. C. and Souza, M. O. (2016). Fixation in large populations: a continuous view of a discrete problem. *J. Math. Biol.*, 72(1-2):283–330.

[18] Charlesworth, B. and Charlesworth, D. (2010). *Elements of Evolutionary Genetics*. Roberts and Company Publishers, Greenwood Village, Colorado.

[19] Charlesworth, B., Lande, R., and Slatkin, M. (1982). A neo-darwinian commentary on macroevolution. *Evolution*, 36(3):474–498.

[20] Cotterman, C. W. (1940). *A calculus for statistico-genetics*. PhD thesis, The Ohio State University.

[21] Crow, J. F. (2001). Shannon’s brief foray into genetics. *Genetics*, 159(3):915–917.

[22] Crow, J. F. and Kimura, M. (1970). *An introduction to population genetics theory*. Harper International Edition., New York.

[23] Cvijović, I., Good, B. H., Jerison, E. R., and Desai, M. M. (2015). Fate of a mutation in a fluctuating environment. *Proc. Nat. Acad. Sci. USA*, 112(36):E5021–E5028.

[24] Daubechies, I. and Lagarias, J. C. (1992). Sets of matrices all infinite products of which converge. *Linear Algebra Appl.*, 161:227–263.

[25] Der, R., Epstein, C., and Plotkin, J. B. (2012). Dynamics of neutral and selected alleles when the offspring distribution is skewed. *Genetics*, 191(4):1331–1344.

[26] Der, R., Epstein, C. L., and Plotkin, J. B. (2011). Generalized population models and the nature of genetic drift. *Theor. Popul. Biol.*, 80(2):80–99.

[27] Eldon, B. and Wakeley, J. (2006). Coalescent processes when the distribution of offspring number among individuals is highly skewed. *Genetics*, 172(4):2621–2633.

[28] Erwin, D. H. (2000). Macroevolution is more than repeated rounds of microevolution. *Evol. Dev.*, 2(2):78–84.
[29] Estep, D. (2002). *Practical Analysis in One Variable*. Undergraduate Texts in Mathematics. Springer-Verlag New York.

[30] Ethier, S. N. and Kurtz, T. G. (1986). *Markov processes*. Wiley Series in Probability and Mathematical Statistics: Probability and Mathematical Statistics. John Wiley & Sons Inc., New York. Characterization and convergence.

[31] Ewens, W. J. (2004). *Mathematical Population Genetics. I: Theoretical Introduction*. 2nd ed. Interdisciplinary Mathematics 27. New York, NY: Springer.

[32] Felsenstein, J. (1976). *The theoretical population genetics of variable selection and migration*. *Annu. Rev. Genet.*, 10(1):253–280.

[33] Fisher, R. A. (1922). On the dominance ratio. *Proc. Royal Soc. Edinburgh*, 42:321–341.

[34] Fisher, R. A. (1930). The distribution of gene ratios for rare mutations. *Proc. Royal Soc. Edinburgh*, 50:214–219.

[35] Fontdevila, A. (2011). *The Dynamic Genome: A Darwinian Approach*. Oxford University Press.

[36] Frazzetta, T. H. (2012). Flatfishes, Turtles, and Bolyerine Snakes: Evolution by Small Steps or Large, or Both? *Evol. Biol.*, 39(1):30–60.

[37] Fudenberg, D., Nowak, M. A., Taylor, C., and Imhof, L. A. (2006). Evolutionary game dynamics in finite populations with strong selection and weak mutation. *Theor. Popul. Biol.*, 70(3):352–363.

[38] Gillespie, J. H. (1991). *The causes of molecular evolution*. Oxford University Press.

[39] Gokhale, C. S. and Traulsen, A. (2010). Evolutionary games in the multiverse. *Proc. Nat. Acad. Sci. USA*, 107(12):5500–5504.

[40] Grinstead, C. and Snell, J. (1997). Introduction to probability. *American Mathematical Society*.

[41] Harmer, G. P., Abbott, D., Taylor, P. G., and Parrondo, J. M. R. (2000). Parrondo’s paradoxical games and the discrete Brownian ratchet. In Abbott, D and Kish, LB, editor, *Unsolved problems of noise and fluctuations*, volume 511 of *AIP Conference Proceedings*, pages 189–200. 2nd International Conference on Unsolved Problems of Noise and Fluctuations (UPoN 99), Adelaide, Australia, Jul 12-15, 1999.

[42] Hartle, D. L. and Clark, A. G. (2007). *Principles of Population Genetics*. Sinauer, Massachusetts.

[43] Hennion, H. (1997). Limit theorems for products of positive random matrices. *Ann. Probab.*, 25(4):1545–1587.

[44] Hilbe, C. (2011). Local replicator dynamics: A simple link between deterministic and stochastic models of evolutionary game theory. *B. Math. Biol.*, 73(9):2068–2087.

[45] Imhof, L. A. and Nowak, M. A. (2006). Evolutionary game dynamics in a Wright-Fisher process. *J. Math. Biol.*, 52(5):667–681.

[46] Johnson, C. R. and Tarazaga, P. (2004). On matrices with Perron-Frobenius properties and some negative entries. *Positivity*, 8(4):327–338.

[47] Karlin, S. and Taylor, H. M. (1975). *A first course in stochastic processes*. Academic Press [A subsidiary of Harcourt Brace Jovanovich, Publishers], New York-London, Second edition.
[48] Keilson, J. and Kester, A. (1977). Monotone matrices and monotone Markov processes. *Stoch. Proc. Appl.*, 5(3):231–241.
[49] Kimura, M. (1962). On the probability of fixation of mutant genes in a population. *Genetics*, 47:713–719.
[50] Kimura, M. (1983). *The Neutral Theory of Molecular Evolution*. Cambridge: University Press., UK.
[51] Kimura, M. and Ohta, T. (1969). Average number of generations until extinction of an individual mutant gene in a finite population. *Genetics*, 63(3):701–712.
[52] Lewin, M. (1971). On nonnegative matrices. *Pacific J. Math.*, 36(3):753–759.
[53] Lorenzi, T., Chisholm, R. H., Desvillettes, L., and Hughes, B. D. (2015). Dissecting the dynamics of epigenetic changes in phenotype-structured populations exposed to fluctuating environments. *J. Theor. Biol.*
[54] Mathew, S. and Perreault, C. (2015). Behavioural variation in 172 small-scale societies indicates that social learning is the main mode of human adaptation. *P. Roy. Soc. B-Biol. Sci.*, 282(1810).
[55] Maynard Smith, J. (1998). *Evolutionary genetics*. Oxford University Press, Oxford.
[56] McCandlish, D. M., Epstein, C. L., and Plotkin, J. B. (2015a). Formal properties of the probability of fixation: Identities, inequalities and approximations. *Theor. Popul. Biol.*, 99:98–113.
[57] McCandlish, D. M., Epstein, C. L., and Plotkin, J. B. (2015b). Formal properties of the probability of fixation: identities, inequalities and approximations. *Theor. Popul. Biol.*, 99:98–113.
[58] Melbinger, A. and Vergassola, M. (2015). The impact of environmental fluctuations on evolutionary fitness functions. *Scientific reports*, 5.
[59] Moran, P. A. P. (1962). *The Statistical Process of Evolutionary Theory*. Clarendon Press, Oxford.
[60] Nassar, R. F. and Cook, R. D. (1974). Ultimate probability of fixation and time to fixation or loss of a gene under a variable fitness model. *Theor. Appl. Genet.*, 44(6):247–254.
[61] Noutsos, D. (2006). On Perron-Frobenius property of matrices having some negative entries. *Linear Algebra Appl.*, 412(2-3):132–153.
[62] Nowak, M. A. (2006). *Evolutionary Dynamics: Exploring the Equations of Life*. The Belknap Press of Harvard University Press, Cambridge, MA.
[63] Nowak, M. A., Sasaki, A., Taylor, C., and Fudenberg, D. (2004). Emergence of cooperation and evolutionary stability in finite populations. *Nature*, 428(6983):646–650.
[64] Orr, H. A. (2009). Fitness and its role in evolutionary genetics. *Nat. Rev. Genet.*, 10(8):531–539.
[65] Osipovitch, D. C., Barratt, C., and Schwartz, P. M. (2009). Systems chemistry and Parrondo’s paradox: computational models of thermal cycling. *New J. Chem.*, 33(10):2022–2027.
[66] Pagel, M., Venditti, C., and Meade, A. (2006). Large punctuational contribution of speciation to evolutionary divergence at the molecular level. *Science*, 314(5796):119–121.
[67] Parrondo, J. M. R., Harmer, G. P., and Abbott, D. (2000). New paradoxical games based on Brownian ratchets. Phys. Rev. Lett., 85(24):5226–5229.
[68] Proulx, S. R. and Adler, F. R. (2010). The standard of neutrality: still flapping in the breeze? J. Evolution. Biol., 23(7):1339–1350.
[69] Ressel, P. (1987). Integral representations on convex semigroups. Math. Scand., 61:93–111.
[70] Schuster, P. (2011). The Mathematics of Darwin’s Theory of Evolution: 1859 and 150 Years Later. In Chalub, F. A. C. C. and Rodrigues, J. F., editor, Mathematics Of Darwin’S Legacy, Mathematics and Biosciences in Interaction, pages 27–66. Conference on Mathematics of Darwin’s Legacy, Univ Lisbon, Lisbon, PORTUGAL, NOV 23-24, 2009.
[71] Shannon, C. E. (1940). An algebra for theoretical genetics. PhD thesis, Massachusetts Institute of Technology, Cambridge, MA. Ph.D. thesis in Mathematics.
[72] Tan, S., Lü, L., Yu, X., and Hill, D. (2012). Monotonicity of fixation probability of evolutionary dynamics on complex networks. In IECON 2012-38th Annual Conference on IEEE Industrial Electronics Society, pages 2337–2341. IEEE.
[73] Tarazaga, P., Raydan, M., and Hurman, A. (2001). Perron-Frobenius theorem for matrices with some negative entries. Linear Algebra Appl., 328(1-3):57–68.
[74] Taylor, H. M. and Karlin, S. (1998). An introduction to stochastic modeling. Academic Press Inc., San Diego, CA, Third edition.
[75] Traulsen, A., Pacheco, J. M., and Imhof, L. A. (2006). Stochasticity and evolutionary stability. Phys. Rev. E, 74:021905.
[76] Traulsen, A., Pacheco, J. M., and Nowak, M. A. (2007). Pairwise comparison and selection temperature in evolutionary game dynamics. J. Theor. Biol., 246(3):522–529.
[77] Uecker, H. and Hermisson, J. (2011). On the Fixation Process of a Beneficial Mutation in a Variable Environment. Genetics, 188(4):915–U281.
[78] Waxman, D. and Welch, J. (2005). Fisher’s microscope and Haldane’s ellipse. Am. Nat., 166(4):447–457.
[79] Wright, S. (1931). Evolution in mendelian populations. Genetics, 16(2):0097–0159.
[80] Wright, S. (1937). The distribution of gene frequencies in populations. Proc. Nat. Acad. Sci. USA, 23:307–320.
[81] Wright, S. (1938). The distribution of gene frequencies under irreversible mutations. Proc. Nat. Acad. Sci. USA, 24:253–259.
[82] Yakushina, T., Saakian, D. B., Bratus, A., and Hu, C.-K. (2015). Evolutionary Games with Randomly Changing Payoff Matrices. J. Phys. Soc. Jpn., 84(6).