Fixation probabilities for the Moran process with three or more strategies: general and coupling results

Eliza M. Ferreira1 · Armando G. M. Neves2

Received: 24 November 2018 / Revised: 19 December 2019 / Published online: 31 May 2020
© Springer-Verlag GmbH Germany, part of Springer Nature 2020

Abstract
We study fixation probabilities for the Moran stochastic process for the evolution of a population with three or more types of individuals and frequency-dependent fitnesses. Contrary to the case of populations with two types of individuals, in which fixation probabilities may be calculated by an exact formula, here we must solve a large system of linear equations. We first show that this system always has a unique solution. Other results are upper and lower bounds for the fixation probabilities obtained by coupling the Moran process with three strategies with birth–death processes with only two strategies. We also apply our bounds to the problem of evolution of cooperation in a population with three types of individuals already studied in a deterministic setting by Núñez Rodríguez and Neves (J Math Biol 73:1665–1690, 2016). We argue that cooperators will be fixated in the population with probability arbitrarily close to 1 for a large region of initial conditions and large enough population sizes.

Keywords Markov chains · Evolutionary games · Coupling method

Mathematics Subject Classification 91A22 · 92D15 · 60J20

EMF had a Capes scholarship. AGMN was partially supported by Fundação de Amparo à Pesquisa de Minas Gerais (FAPEMIG, Brazil).

Armando G. M. Neves
aneves@mat.ufmg.br
Eliza M. Ferreira
eliza.ferreira@ufla.br

1 Departamento de Ciências Exatas, Universidade Federal de Lavras, Lavras, Brazil
2 Departamento de Matemática, Universidade Federal de Minas Gerais, Belo Horizonte, Brazil
1 Introduction

The Moran process Moran (1958) is a well-known discrete-time stochastic model for the genetic evolution of a population of fixed finite size composed of individuals of several possible types (phenotypes or genotypes) and assuming no mutations in the reproduction process. The model was created having in mind a population composed of individuals of two types, assuming that the fitnesses of the individuals depend on their types and not on the frequency of these types in the population. Later on, see Nowak et al. (2004) and Taylor et al. (2004), the process was extended to the context of Evolutionary Game Theory, see Maynard Smith and Price (1973), Hofbauer and Sigmund (1998) or Nowak (2006), in which fitnesses may depend on the frequency of the types among the population and are specified through a pay-off matrix. In Evolutionary Game Theory, the types of individuals in the population are usually referred to as the strategies adopted by the individuals. Extension of the Moran model for populations with more than two strategies was made by Wang et al. (2007) and opened up a new class of problems, as this work will show.

In this paper we will study the Moran process for populations composed of individuals adopting three or more strategies. We acknowledge previous work by Wang et al. (2007), but as their preprint has not been published, we will repeat part of their work here, adding full mathematical rigor. To be true, the whole paper is written having in mind that the number of strategies in the population is three, but the same theory can be easily extended to a larger number of strategies.

Studying populations with three or more strategies is biologically relevant, because only two strategies may not be rich enough for unveiling important phenomena. One example, studied in this paper as an application of general results, is the evolution of cooperation, in which individuals adopting Tit-for-Tat act as catalysts, fighting defectors and enhancing evolution of really cooperative strategies Nowak and Sigmund (1992). Another example Dingli et al. (2009) is a model for multiple myeloma, a type of cancer, in which we have two types of normal cells, osteoblasts and osteoclasts, and also malignant cells. According to the authors of the above paper, multiple myeloma is only an example and the paradigm of several cell types with frequency-dependent fitness might possibly be useful for understanding other types of cancer. Although we do not study this application, the methods devised here may be useful in this task.

The first difficulty in passing from two to more than two strategies in the Moran process is that we do not have anymore an exact formula such as (4) below for the fixation probabilities. A good understanding of all possible evolutionary scenarios for the Moran process with two strategies was obtained by de Souza et al. (2019) based on a classification of these scenarios provided by Taylor et al. (2004). As the cited works show, the behavior of the stochastic Moran process is naturally related to the behavior of the deterministic replicator dynamics Taylor and Jonker (1978) with the same pay-off matrix. If the number of strategies is two, we have four scenarios (without counting the trivial neutral scenario) for the replicator dynamics and eight for the Moran process Taylor et al. (2004). On the other hand, if the number of strategies is three, the replicator dynamics is much richer. Bomze (1983) showed that the number of possible phase portraits for the replicator dynamics is 47. We expect that such richness...
Fixation probabilities for the Moran process with three
strategies

2 Definitions and notations for the Moran process with three
strategies

2.1 Birth–death processes

In order to introduce the relevant concepts in this work, we start by defining birth–death processes. As will be soon clear, the Moran process for a population with only two types of individuals is a particular case of the following definition. Although there are other different definitions, the one presented here is in Nowak (2006, Sect. 6.2).

A birth–death process is a discrete-time Markov chain such that:

of behaviors in the deterministic three-strategies case is reflected also when the same population is studied as a Moran process.

In this work, after proving general results, we will also propose a method to produce upper and lower bounds for the fixation probabilities in the Moran process with three or more strategies. With these bounds we will understand e.g. the behavior of the fixation probability of a pure strategy which is a strict Nash equilibrium when it is close to fixation. On the other hand, we will see that in the opposite situation of a pure strategy which is a source in the replicator dynamics, our results are not so conclusive. Despite the progress we have made, there is still much work to be completed in the direction of providing a complete classification, as in Taylor et al. (2004), for the Moran process with three strategies. As an example, we will apply our results for a single phase-portrait among the 47 of Bomze (1983). Our choice corresponds to one of the dynamics studied in Núñez Rodríguez and Neves (2016) for the problem of evolution of cooperation.

This paper is organized as follows. In Sect. 2 we will define birth–death processes and introduce the notation for the Moran processes for two and three strategies. We will see that for three or more strategies the calculation of fixation probabilities amounts to solving a large system of linear equations. We will prove that the systems do have solutions, which are also unique.

In Sect. 3 we will present general results for obtaining upper and lower bounds for the fixation probabilities of the Moran process with three strategies by the Markov chain coupling technique. The mentioned bounds are illustrated for some concrete pay-off matrices in Sect. 4. These concrete examples will be further mentioned in the rest of the paper. In Sect. 5 we will see that the bounds of Sect. 3 can produce interesting results in some situations, e.g. when we are in the basin of attraction of a strict Nash equilibrium strategy of the replicator dynamics. Section 6 is dedicated to an example concerning the evolution of cooperation with three strategies. This example has already been studied in a deterministic setting in Núñez Rodríguez and Neves (2016). Finally, in Sect. 7 we draw some conclusions and outline open problems.

Some auxiliary results on birth–death processes necessary for proving the theorems on the Moran process with three strategies are proved in Appendix A. The proofs of most theorems stated in the main text on the Moran process with three strategies are in Appendix B together with auxiliary results.
– The set of states $S = \{0, 1, 2, \ldots, N\}$ is finite.
– States 0 and $N$ are absorbing, i.e. the transition probability from any of these states to a different state is null.
– The transition matrix of the chain is tridiagonal, i.e. from state $i \in S$ the only non-vanishing transition probabilities are to state $i$ itself and to the states $i \pm 1$ which belong to $S$. The transition probabilities of $i$ to $i \pm 1$ are both positive.

In the context of birth–death processes, a transition $i \rightarrow i + 1$ will be termed a birth and a transition $i \rightarrow i - 1$ will be called a death.

As all states in $S \setminus \{0, N\}$ are transient, we have the phenomenon of fixation: starting at any transient state, if we wait long enough, with probability 1 the state will eventually end in an absorbing state. An important question is the probability that one or the other absorbing state will be attained and how this probability depends on the initial state of the population.

The fixation probabilities can be exactly calculated by the explicit formula (4) below, known at least since Moran (1958), see also Nowak (2006). If $X_n$ denotes the state at time $n$, let $a_i$ be the birth probability

$$a_i = P(X_{n+1} = i + 1 | X_n = i)$$  \hspace{1cm} (1)

and $b_i$ be the death probability

$$b_i = P(X_{n+1} = i - 1 | X_n = i).$$  \hspace{1cm} (2)

Of course $a_0 = a_N = b_0 = b_N = 0$. We also define the birth to death ratio as

$$r_i = \frac{a_i}{b_i}.\hspace{1cm} (3)$$

If $\pi_i$ denotes the probability of fixation at state $N$ when the initial state is $i$, then, of course, the probability of fixation at 0 with the same initial state is $1 - \pi_i$. We have $\pi_0 = 0$, $\pi_N = 1$ and for $i = 1, 2, \ldots, N - 1$

$$\pi_i = \frac{1 + \sum_{j=1}^{i-1} \prod_{k=1}^{j} r_k^{-1}}{1 + \sum_{j=1}^{N-1} \prod_{k=1}^{j} r_k^{-1}},$$  \hspace{1cm} (4)

where the numerator in the right-hand side is just equal to 1 if $i = 1$.

Despite the exact result (4), proof of the difficulties in completely understanding birth–death and Moran processes for only two strategies is the number of papers on the subject since Taylor et al. (2004), e.g. Antal and Scheuring (2006), Chalub and Souza (2016), Durand and Lessard (2016) and de Souza et al. (2019).

We also define the dual birth–death process, in which births and deaths are interchanged, by

$$\overline{a}_i = b_{N-i} \quad \text{and} \quad \overline{b}_i = a_{N-i}.\hspace{1cm} (5)$$

© Springer
The fixation probability at state $N$ of the dual process can then be calculated by a formula similar to (4). The duality idea will be used in this paper in the proof of Theorem 10 in Appendix A.

### 2.2 The Moran process with two strategies

The Moran process is a stochastic model for the evolution of a population with fixed finite size equal to $N$ individuals, which are of two different types: either they adopt strategy A, or they adopt strategy B. At each time step two independent random choices are made: one individual is chosen for reproduction and one for death. We assume that the reproducing individual produces an offspring with the same type as itself, the *absence of mutations hypothesis*, and also that the offspring of the reproducing individual replaces the dying individual. The dying individual is chosen uniformly among the whole population, but the reproducing individual is chosen with probability proportional to the fitness of its type.

The Moran process is thus a birth–death process as defined above, where the state $i \in S$ is the number of A individuals in the population. When the state is $i$, $a_i$ is the probability of drawing an A individual for reproduction and a B for death, and $b_i$ is the probability of drawing a B for reproduction and an A for death.

The transition probabilities $a_i$ and $b_i$ depend on the fitnesses, which are calculated using a game-theoretic pay-off matrix. For a game with two strategies, the pay-off matrix is a $2 \times 2$ matrix $M = (m_{ij})$, where $m_{ij} > 0$ is the reward that an individual of type $i$ receives when interacting with an individual of type $j$. We will agree that individuals of types A and B will be labeled respectively as types 1 and 2 in the pay-off matrix.

The fitnesses $f_i$ of the A individuals and $g_i$ of the B individuals are defined Nowak et al. (2004) as

$$f_i = 1 - w + w \left[ \frac{m_{11} i - 1}{N - 1} + \frac{m_{12} N - i}{N - 1} \right]$$

and

$$g_i = 1 - w + w \left[ \frac{m_{21} i N - 1}{N - 1} + \frac{m_{22} N - i - 1}{N - 1} \right],$$

where the intensity of selection is $w \in [0, 1]$.

The probability of drawing an A individual for reproduction is $\frac{i f_i}{S_i}$ and the probability of choosing a B for reproduction is $\frac{(N-i) g_i}{S_i}$, where

$$S_i = i f_i + (N - i) g_i.$$  

As the death draw is defined to be uniform, then the probabilities of the transitions from $i$ to $i + 1$ and from $i$ to $i - 1$ are respectively

$$a_i = \frac{i f_i}{S_i} \frac{N - i}{N}$$

$$b_i = \frac{(N-i) g_i}{S_i} \frac{i}{N}.$$
and

\[ b_i = \frac{N - i}{S_i} \frac{i}{N}. \]  \hspace{1cm} (10)

These should be substituted in (3), giving

\[ r_i = \frac{f_i}{g_i}, \]  \hspace{1cm} (11)

which may be used in (4) for calculating the fixation probabilities for the Moran process with two strategies.

The situation in which all elements in the pay-off matrix are equal is called neutral evolution. In the neutral case, (4) yields \( \pi_i = i/N \). If \( \rho_A \) and \( \rho_B \) are respectively the fixation probability of a single A or B individual in a population of size \( N \), i.e. \( \rho_A = \pi_1 \) and \( \rho_B = 1 - \pi_{N-1} \), then we have \( \rho_A = \rho_B = 1/N \) in the neutral case.

The case in which A and B fitnesses are independent of the types frequencies is obtained by inserting \( w = 1, m_{11} = m_{12} = f \) and \( m_{21} = m_{22} = g \) in (6) and (7), so that \( f \) and \( g \) become the fitnesses of A and B individuals and \( r = f/g \) is the relative fitness of A individuals with respect to B. In this case, (4) yields

\[ \pi_i = \frac{1 - r^{-i}}{1 - r^{-N}}. \]  \hspace{1cm} (12)

The classification of the evolutionary scenarios for the Moran process from the point of view of fixation probabilities was performed by Taylor et al. (2004). Their classification scheme considers at first whether a single A in the population is more or less fit than the Bs, and also whether a single B is more or less fit than the As. This is the natural consequence of taking into account the game-theoretic dynamics for an infinite population, the so-called replicator dynamics Taylor and Jonker (1978). An important discovery by Taylor et al. (2004) is that the evolutionary scenario depends also on whether \( \rho_A \) and \( \rho_B \) are larger or smaller than their values \( 1/N \) for neutral evolution. Simple combinatorics leads to 16 scenario possibilities, but Taylor et al. (2004) prove that only eight scenarios can actually occur.

The above classification for the Moran process with 2 strategies was also treated in de Souza et al. (2019). That work associates to each of the 8 evolutionary scenarios a precise shape for the graphs of the fixation probabilities. Other results in the same paper are asymptotic formulae for the fixation probabilities in the limit \( N \to \infty \). As a consequence of these formulae, it can be shown that some of the evolutionary scenarios cannot happen for large enough populations.

### 2.3 The Moran process with three strategies

Consider a fixed-size population with \( N \) individuals divided into three types, say A, B and C. The state of the population at each time \( n \in \{0, 1, 2, \ldots \} \) can be specified by the number of individuals of types A and B. Obviously, if at time \( n \) we have \( i \) individuals of type A and \( j \) of type B, then we will have \( N - i - j \) type C individuals.
The state in this situation is denoted $X_n = (i, j)$. If, as before, $S = \{0, 1, 2, \ldots, N\}$, the set of all states of the Moran process with three strategies is

$$\Lambda_N = \{(i, j) : i, j \in S \text{ and } i + j \leq N\}.$$  

Consider an equilateral triangle $ABC$ of unitary side length. It will be useful to represent the set of states $\Lambda_N$ of a Moran process with three strategies as the nodes of a mesh on $ABC$, see Fig. 1. State $(i, j) \in \Lambda_N$ will be identified with the point on the mesh reached from the vertex C by the vector $\frac{i}{N} CA + \frac{j}{N} CB$ and we will sometimes speak about states and points on the mesh as synonyms. With this identification, the vertices of the triangle represent the states in which a single type is present. On the sides of the triangle, one type is absent.

For future use we define a compact set $\Lambda \subset \mathbb{R}^2$ by

$$\Lambda = \{(x, y) \in \mathbb{R}^2 : x \geq 0, y \geq 0, x + y \leq 1\}. \quad (13)$$

We identify element $(x, y) \in \Lambda$ with the point in the triangle $ABC$ reached from the vertex C by the vector $x \ CA + y \ CB$

As for two strategies, the population dynamics for the Moran process with three strategies is defined by two independent random choices at each time step: one individual is drawn to die and another is drawn to reproduce, its offspring being of the same type as itself, again an absence of mutations hypothesis. The rules are analogous to the ones already described for the Moran process with two strategies, see (8), (9), (10):

- The death choice is made with uniform distribution.
- The reproduction choice is such that the probability of an individual being drawn for reproduction is proportional to the fitness of the individual’s type. The exact specification will be given in the following paragraphs.

The fitness functions $f$, $g$ and $h$ for individuals of types respectively A, B and C may be calculated from Evolutionary Game Theory standard prescriptions. Let $M$ be a $3 \times 3$ pay-off matrix, where $m_{ij} > 0$ is the reward that an individual of type $i$ receives when interacting with an individual of type $j$. We will agree that individuals of types
A, B and C will be numbered, respectively, as types 1, 2 and 3 in the pay-off matrix. As natural extensions of (6) and (7), fitnesses are given by

\[
 f_{ij} = 1 - w + w \left[ m_{11} \frac{i - 1}{N - 1} + m_{12} \frac{j}{N - 1} + m_{13} \frac{N - i - j}{N - 1} \right]
\]

\[
 g_{ij} = 1 - w + w \left[ m_{21} \frac{i}{N - 1} + m_{22} \frac{j - 1}{N - 1} + m_{23} \frac{N - i - j}{N - 1} \right]
\]

\[
 h_{ij} = 1 - w + w \left[ m_{31} \frac{i}{N - 1} + m_{32} \frac{j}{N - 1} + m_{33} \frac{N - i - j - 1}{N - 1} \right].
\]

(14)

Parameter $w \in [0, 1]$ is again the intensity of selection. In the examples in this paper, for simplicity, we will always take $w = 1$. Consideration of other positive values for $w$ will not change any of the theoretical results.

If the state is $(i, j)$, the probability of drawing an A for reproduction is \( \frac{if_{ij}}{S_{ij}} \), where

\[
 S_{ij} = if_{ij} + jg_{ij} + (N - i - j)h_{ij}
\]

(15)

denotes the sum of the fitnesses of all individuals.

Analogously, the probabilities for drawing a B or a C for reproduction are respectively \( \frac{gg_{ij}}{S_{ij}} \) and \( \frac{(N - i - j)h_{ij}}{S_{ij}} \).

The above rules define for each $3 \times 3$ positive pay-off matrix $M$ and population size $N$ the stochastic time evolution of the population state: a discrete-time Markov chain with finite state space $\Lambda_N$. States $(0, 0), (0, N)$ and $(N, 0)$ are absorbing and all the remaining states are transient Allen (2011). Again, with probability 1, all trajectories beginning at some transient state will be absorbed and one important problem is to calculate the fixation probability of each among the three strategies as a function of the initial state.

As the number of individuals of each type may increase or decrease by 1 unit, or remain constant, the point representing the state either remains fixed, or jumps to one of the nearest 6 neighbors in the triangular mesh of Fig. 1. The Moran process for three strategies can thus be seen as a two-dimensional random walk on $\Lambda_N$.

If X and Y may stand for A, B or C, we introduce now notation $p_{ij}^{XY}$ for the probability of drawing an individual of type $X$ for reproduction and an individual of type $Y$ for death when the state is $(i, j) \in \Lambda_N$. Transition probability $p_{ij}^{AB}$, for example, may be calculated as

\[
p_{ij}^{AB} = \frac{if_{ij} \cdot j}{S_{ij} \cdot N}
\]

because the probability of drawing an A for reproduction is \( \frac{if_{ij}}{S_{ij}} \), the probability of drawing a B for death is \( \frac{j}{N} \) and the reproduction and death draws are independent.

The whole set of transition probabilities of the chain at state $(i, j)$, calculated as in the example, are
where \( x \) one time step and the corresponding transition probabilities to them from state \( \alpha \) the fixation probability of individuals of type A and B and all of them will assume the form of a term independent of \( \alpha \) and \( \beta \) respectively. Let \( \alpha_{ij} \) denote the fixation probability of A individuals if the initial state is \( \alpha_{ij} \). For the sake of future use, we observe that we may write Eqs. (14) to (17) in terms of the population fractions \( x = i/N \) and \( y = j/N \) of individuals of type A and B and all of them will assume the form of a term independent of \( N \) plus corrections that tend to 0 when \( N \to \infty \). For example, we define the deterministic fitnesses

\[
F(x, y) = 1 - w + w \left[ m_{11}x + m_{12}y + m_{13}(1 - x - y) \right],
\]

\[
G(x, y) = 1 - w + w \left[ m_{21}x + m_{22}y + m_{23}(1 - x - y) \right],
\]

\[
H(x, y) = 1 - w + w \left[ m_{31}x + m_{32}y + m_{33}(1 - x - y) \right].
\]

Then

\[
f_{ij} = 1 - w + w \frac{N}{N - 1} \left[ m_{11}x + m_{12}y + m_{13}(1 - x - y) \right] - \frac{w m_{11}}{N - 1}
\]

\[
= F(x, y) - \frac{w m_{11}}{N - 1} + w \left( \frac{N}{N - 1} - 1 \right) \left[ m_{11}x + m_{12}y + m_{13}(1 - x - y) \right]
\]

\[
= F(x, y) + O \left( \frac{1}{N} \right),
\]

where \( x = i/N, y = j/N \). Analogous expressions hold for \( g_{ij} \) and \( h_{ij} \). In several other places in this paper we will be concerned with writing quantities depending on the state \( (i, j) \) in terms of the fractions \( x = i/N \) and \( y = j/N \) and seeing a “deterministic” part independent of \( N \) and remainder terms that vanish as \( N \to \infty \).

The fixation probabilities of individuals of types A, B or C will be obtained as solutions to systems of linear equations. We proceed now to producing these equations. Let \( \alpha_{ij} \) denote the fixation probability of A individuals if the initial state is \( (i, j) \). For B and C individuals and the same initial state the fixation probabilities will be denoted respectively \( \beta_{ij} \) and \( \gamma_{ij} \). Taking into account the seven possibilities for the state after one time step and the corresponding transition probabilities to them from state \( (i, j) \), the fixation probability \( \alpha_{ij} \) may be written as

\[
\alpha_{ij} = p_{ij}^{AB} \alpha_{i+1,j-1} + p_{ij}^{BA} \alpha_{i-1,j+1} + p_{ij}^{AC} \alpha_{i-1,j} + p_{ij}^{CA} \alpha_{i-1,j+1} + p_{ij}^{BC} \alpha_{i,j-1} + p_{ij}^{const} \alpha_{ij}.
\]
Reorganizing the terms of the above expression we have

\[
\alpha_{ij} = \frac{1}{1 - p_{ij}^{\text{const}}}
\left[ p_{ij}^{AB} \alpha_{i+1,j-1} + p_{ij}^{BA} \alpha_{i-1,j+1} + p_{ij}^{AC} \alpha_{i+1,j} + p_{ij}^{CA} \alpha_{i-1,j} + p_{ij}^{BC} \alpha_{i,j+1} + p_{ij}^{CB} \alpha_{i,j-1} \right].
\]  

(20)

Since \(1 - p_{ij}^{\text{const}} = p_{ij}^{AB} + p_{ij}^{BA} + p_{ij}^{AC} + p_{ij}^{CA} + p_{ij}^{BC} + p_{ij}^{CB}\), it follows from (20) that \(\alpha_{ij}\) is a weighted average of the values of the same function in the 6 nearest neighbors of state \((i, j)\) in the triangular mesh. The same equations remain valid if \(\alpha_{ij}\) is replaced for \(\beta_{ij}\) or \(\gamma_{ij}\), but with different boundary conditions. These relations for the fixation probabilities can also be found in Wang et al. (2007).

We have one equation of the type (20) for each point of the mesh in the interior of the triangle \(ABC\). The values of \(\alpha_{ij}\) at the points of the mesh on the boundary of the triangle are known. In fact, for the points on the BC side, we have

\[
\alpha_{0,j} = 0, \quad j = 0, \ldots, N,
\]

because on this side there are no A type individuals and the absence of mutations hypothesis prohibits them to be produced by Bs or Cs. For the points on the border at sides AB and AC, we have at most two types of individuals and we may use versions of (4). On the AB side of the triangle we thus have

\[
\alpha_{i,N-i} = \frac{1 + \sum_{j=1}^{i-1} \prod_{k=1}^{j} (r_{k}^{AB})^{-1}}{1 + \sum_{j=1}^{N-1} \prod_{k=1}^{j} (r_{k}^{AB})^{-1}},
\]

(22)

for \(i = 1, 2, \ldots, N\), where \(r_{k}^{AB} = \frac{f_{k,N-k}}{f_{k,N}}\) is the relative fitness of type A individuals with respect to B individuals in the absence of type C individuals. The version of (4) for the AC side of the triangle is

\[
\alpha_{i,0} = \frac{1 + \sum_{j=1}^{i-1} \prod_{k=1}^{j} (r_{k}^{AC})^{-1}}{1 + \sum_{j=1}^{N-1} \prod_{k=1}^{j} (r_{k}^{AC})^{-1}},
\]

(23)

for \(i = 1, 2, \ldots, N\) and with \(r_{k}^{AC} = \frac{f_{k,0}}{f_{k,N}}\).

The calculation of \(\alpha_{ij}\) for the mesh points in the interior of the triangle amounts thus to solving a system of linear equations (20) with \((N - 1)(N - 2)/2\) unknowns, one for each point of the mesh in the interior of the triangle. Equations (21), (22) and (23) act as Dirichlet boundary conditions. This problem is similar to the problem of approximating the solution of a two-dimensional Dirichlet problem for the Laplace equation using finite differences, see e.g. Petrovsky (1992). In the approximation of the two-dimensional Laplace equation in a rectangular lattice, a simple arithmetic mean of 4 neighboring lattice points appears. In our problem, instead, we have a weighted average of 6 neighbors.

\(\copyright\) Springer
The problems of calculating the fixation probabilities $\beta_{ij}$ for $B$ individuals and $\gamma_{ij}$ for $C$ individuals are completely analogous. As the transition probabilities appearing in (20) are the same regardless we are calculating the fixation probability for $A$, $B$ or $C$, equations (20) are exactly the same, but boundary conditions (21), (22) and (23) must be replaced by their analogues.

We will end this section by proving that the problems of finding the fixation probabilities for the Moran process with three strategies are all well-posed in the sense that the corresponding linear systems have unique solutions. The following proofs are adaptations of the corresponding ones for the finite-differences Laplace equation, see e.g. Petrovsky (1992).

We say that $R \subset \Lambda_N$ is a connected region if for any states $(i_1, j_1)$ and $(i_2, j_2)$ in $R$ there is a path between the corresponding points in the triangle $ABC$ passing only through the links of the mesh and without passing by points representing states not in $R$. We also say that a state is in the interior of $R$ if the point representing it has 6 neighbors at distance $1/N$ in the mesh and all of them represent states in $R$. Otherwise we will say that it is a border point of $R$.

The following result is a generalization of a familiar property of the solutions of the Laplace equation. It will be used in proving uniqueness of the solutions for the fixation probabilities:

**Proposition 1** (Maximum and minimum property) Let $R \subset \Lambda_N$ be connected. Then the maximum and the minimum of the fixation probability function $\alpha$ restricted to $R$ are on the border of $R$. The same property holds also for the maximum and the minimum of $\beta$ and $\gamma$.

**Proof** Assume that the fixation probability function $\alpha$ restricted to $R$ reaches its maximum at a state $(i_*, j_*)$ in the interior of $R$.

We know that $\alpha_{i_*, j_*}$ is the weighted average of the function $\alpha$ in the 6 nearest neighbors of $(i_*, j_*)$ on $\Lambda_N$, all of which are in $R$. Being an average, $\alpha_{i_*, j_*}$ cannot be strictly larger than any of the values of $\alpha$ at all the 6 nearest neighbors. As it is the maximum in the set with elements $(i_*, j_*)$ and its 6 nearest neighbors, then the values of $\alpha$ at all these 7 points must be the same. If we repeat the argument for the nearest neighbors of the nearest neighbors, and so on, we conclude that the occurrence of a maximum of $\alpha$ at an interior point implies that $\alpha$ is constant on $R$, so that the maximum also occurs at the border.

A similar argument holds for the minimum of $\alpha$ and also for $\beta$ and $\gamma$. \hfill \Box

We can now prove

**Theorem 1** (Uniqueness of $\alpha$, $\beta$ and $\gamma$) The linear system of Eqs. (20) for $(i, j) \in \Lambda_N$ with boundary conditions (21), (22) and (23) has a unique solution. The same holds also for the analogous systems for $\beta$ and $\gamma$.

**Proof** By ordering in some way the points of $\Lambda_N$, e.g. using lexicographic order, the linear system for $\alpha$ may be written in matrix form $EX = F$, where $E$ is a square matrix of dimension $(N-1)(N-2)/2$ and $F$ is a column matrix depending only on the boundary conditions (21), (22) and (23).
We claim that the only solution of the corresponding homogeneous linear system $EX = 0$ is the trivial $X = 0$. In fact, the set of solutions for this homogeneous system is non-empty and the maximum and minimum property holds for these solutions. As the boundary condition for $EX = 0$ is 0 at all border points of $\Lambda_N$, then the maximum and minimum of the solutions of $EX = 0$ must be 0. Thus, the only solution is the trivial one.

As a consequence, we must have $\det E \neq 0$. It follows that $E$ is invertible and $EX = F$ has a unique solution.

Uniqueness of $\beta$ and $\gamma$ follow because the same matrix $E$ appears as the coefficient matrix of the corresponding linear systems.

### 3 Coupling results

We have already seen that the fixation probabilities for the Moran process with three strategies can be calculated by solving a linear system, but we no longer have an explicit formula for the solution, as (4) in the case of two strategies. We will show however that we can provide upper and lower bounds for these probabilities. These bounds will be derived by the coupling method den Hollander (2012).

Coupling is a powerful way of comparing two or more random variables by constructing them simultaneously through the same random device. More specifically, for a given Moran process with three strategies, which we call target chain, we will construct a birth–death process as defined in Sect. 2, which we call comparison chain, in which the fixation probability is explicitly calculated. We may realize the two chains simultaneously, and, as will be seen, this will give us bounds for the fixation probability in the target chain in terms of the exactly calculated fixation probability for the comparison chain.

In what follows, we will use the notations previously introduced for the transition probabilities both for birth–death processes and Moran processes with three strategies. More concretely, $a_i$ and $b_i$ will stand respectively for birth and death probabilities at state $i$ in a birth–death process, and $p_{ij}^{XY}$ are the transition probabilities in the Moran process with three strategies introduced in (16). We also introduce

$$Z_{ij}^+ = p_{ij}^{AB} + p_{ij}^{AC} \quad \text{and} \quad Z_{ij}^- = p_{ij}^{BA} + p_{ij}^{CA}$$

for the probabilities of respectively increasing and decreasing the number of A individuals when the state is $(i, j)$.

We start with a general result:

**Theorem 2** Consider a Moran process with three strategies. Suppose that there exists a birth–death process with state space $\{0, 1, 2, \ldots, N\}$ such that for all $(i, j) \in \Lambda_N$ the birth and death probabilities satisfy

$$a_i \leq Z_{ij}^+ \quad \text{and} \quad b_i \geq Z_{ij}^-$$

and also

$$a_{i-1} \leq 1 - Z_{ij}^-.$$

□
If \( \alpha_{ij} \) denotes the fixation probability of the strategy A and \( \pi_i \) is the fixation probability at state N in the birth–death process, then

\[
\alpha_{ij} \geq \pi_i
\]  

for all \((i, j) \in \Lambda_N\).

In the above result, the target chain is the Moran process with three strategies and the comparison chain is the birth–death process. We may calculate the fixation probability \( \pi_i \) of the comparison chain by (4), thus obtaining a lower bound for the fixation probability \( \alpha_{ij} \) of the Moran process with three strategies.

The proof of the above result, based on coupling the target and the comparison chains, can be seen at B.1. Here we expose an idea of the argument.

Let \((i_n, j_n)\) denote the state of the target chain \(n\) time steps after the time 0 when both chains start. At the same time, let \(i'_n\) denote the state of the comparison chain. Suppose that \(i_0 = i'_0\). The first condition in (25) states that at time 0 and also when it happens that \(i_n = i'_n\), if the chains are realized using the same random device, whenever there is a birth in the comparison chain, the number of type A individuals in the target chain will increase. In a similar way, if \(i_n = i'_n\), the second condition in (25) ensures that whenever the number of type A individuals decreases, there will be a death in the comparison chain. With both conditions (25) and \(i_0 = i'_0\), we guarantee that \(i_1 \geq i'_1\).

In order to guarantee that we will always have \(i_n \geq i'_n\), we must be careful that, in case \(i'_n = i_n - 1\), there is no possibility of decreasing the number of A individuals in the target chain and increasing the state of the comparison chain. Condition (26) is used for that. With \(i_0 = i'_0\), (25) and (26) we must have \(i_n \geq i'_n\) for \(n = 1, 2, \ldots\). Hence, whenever there is fixation in state N of the comparison chain, there will be fixation of strategy A in the target chain. Thus the fixation probability \(\pi_i\) of the comparison chain is a lower bound to \(\alpha_{ij}\).

An analogous result can be used to find an upper bound for the fixation probability in a Moran process with three strategies.

Theorem 3 Consider a Moran process with three strategies and the same notation introduced before Theorem 2. If there exists a birth–death process with state space \(\{0, 1, 2, \ldots, N\}\) such that for all \((i, j) \in \Lambda_N\) the birth and death probabilities satisfy

\[
a_i \geq Z_{ij}^+ \quad \text{and} \quad b_i \leq Z_{ij}^-
\]  

and also

\[
b_{i+1} \leq 1 - Z_{ij}^+,
\]  

then

\[
\alpha_{i,j} \leq \pi_i
\]  

for all \((i, j) \in \Lambda_N\).
At this point the reader may wonder if we can in fact find comparison chains satisfying the hypotheses in Theorems 2 and 3, so that upper or lower bounds for Moran processes with three strategies are produced. We show now that such comparison chains do exist in the important case of a Moran process with three strategies and frequency independent fitnesses. Although frequency independent fitnesses may be thought of as too much trivial, we know no other bounds for this particular case. Moreover, it will be seen that the following result will suggest how to obtain comparison chains for the general case.

**Theorem 4** Consider a Moran process for three types of individuals A, B and C and population size $N$, with frequency-independent fitnesses respectively given by $f > 0$, $g > 0$ and $h > 0$. Suppose without loss of generality that $f > g > h$. Then the following bounds hold for all $(i, j) \in \Lambda_N$:

\[
1 - \left(\frac{f}{g}\right)^{-i} \leq \alpha_{ij} \leq \frac{1 - \left(\frac{f}{h}\right)^{-i}}{1 - \left(\frac{f}{h}\right)^{-N}} \tag{31}
\]

\[
1 - \left(\frac{g}{f}\right)^{-j} \leq \beta_{ij} \leq \frac{1 - \left(\frac{g}{h}\right)^{-j}}{1 - \left(\frac{g}{h}\right)^{-N}} \tag{32}
\]

\[
1 - \left(\frac{h}{f}\right)^{-(N-i-j)} \leq \gamma_{ij} \leq \frac{1 - \left(\frac{h}{g}\right)^{-(N-i-j)}}{1 - \left(\frac{h}{g}\right)^{-N}}. \tag{33}
\]

Inequalities $f > g > h$ state that A individuals are fitter than Bs, which in turn are fitter than Cs. Focusing now on A individuals, the intuition behind the bounds above is that it becomes easier for As to fixate if we replace all Bs by Cs. The upper bound for $\alpha_{ij}$ in (31) is just the fixation probability for A in a population with $i$ A individuals and $N - i$ Cs calculated by (12). Similarly, it is harder for As to fixate if we replace Cs by Bs, and the lower bound in (31) is just the fixation probability for a population of $i$ As and $N - i$ Bs. The upper and lower bounds in (32) and (33) are analogous. The rigorous proof for this intuition can be found in B.2. It uses Theorems 2 and 3 with comparison chains obtained replacing individuals of one type by individuals of the other two types.

An interesting consequence of Theorem 4 concerns the behavior of the fixation probabilities $\alpha, \beta$ and $\gamma$ when population size $N$ tends to infinity. For large populations we expect that the randomness inherent in the Moran process becomes less important and, if the Moran process is compatible with natural selection, only the fittest individuals should survive. In order to prove this compatibility, it is necessary that the fractions $x, y$ and $1 - x - y$ of A, B and C individuals are fixed, whereas the respective numbers thereof tend to infinity. This idea is precisely defined if we define for $(x, y) \in \Lambda$
where \([z]\) denotes the integer closest to \(z\) and the set \(A\) was defined in (13).

**Corollary 1** Consider the same hypotheses as in Theorem 4. Then, for any \((x, y) \in A\) with \(x \neq 0\), \(A_N(x, y) \xrightarrow{N \to \infty} 1\), \(B_N(x, y) \xrightarrow{N \to \infty} 0\) and \(\Gamma_N(x, y) \xrightarrow{N \to \infty} 0\).

**Proof** By Theorem 4,

\[
A_N(x, y) \geq 1 - \left( \frac{g}{f} \right)^{\lfloor N x \rfloor}.
\]

The result for \(A_N\) is proven by taking the limit \(N \to \infty\) and using \(f > g > 0\). The results for \(B_N\) and \(\Gamma_N\) follow because they are both probabilities and their sum with \(A_N\) equals 1. \(\square\)

The following theorem is a generalization of the ideas presented in Theorem 4 and works as a general “recipe” for constructing comparison chains satisfying hypotheses of Theorems 2 and 3. Its proof can be seen at B.3. Contrary to Theorem 4, the proof for the next result requires that the population size \(N\) is large enough. We are not sure whether this is a necessary condition.

**Theorem 5** Consider a Moran process with three strategies and the notation introduced in Theorem 2. Define a comparison birth–death process with birth and death probabilities respectively given by

\[
a_i^{\text{low}} = \min_{0 \leq j \leq N-i} Z_{ij}^+ \quad \text{and} \quad b_i^{\text{low}} = \max_{0 \leq j \leq N-i} Z_{ij}^-.
\]

Let \(\pi_i^{\text{low}}\) denote the fixation probability in state \(N\) of the comparison birth–death process and \(\alpha_{ij}\) denote the Moran process fixation probability for type A in the initial state \((i, j)\). Then, for large enough \(N\),

\[
\alpha_{ij} \geq \pi_i^{\text{low}}.
\]

Similarly, if we define another birth–death process with fixation probability \(\pi_i^{\text{up}}\) by taking

\[
a_i^{\text{up}} = \max_{0 \leq j \leq N-i} Z_{ij}^+ \quad \text{and} \quad b_i^{\text{up}} = \min_{0 \leq j \leq N-i} Z_{ij}^-,
\]

then, for large enough \(N\),

\[
\alpha_{ij} \leq \pi_i^{\text{up}}.
\]
4 Some numerical examples of lower and upper bounds derived from Theorem 5

Besides being very general, because it works for every possible pay-off matrix, Theorem 5 is also easy to implement numerically. A drawback is that the proof given in this paper requires that the population size \( N \) is large enough. In the following examples we explicitly calculate the lower and upper bounds of Theorem 5 for three different choices of pay-off matrix. In some cases we also compare the lower and upper bounds with fixation probabilities numerically evaluated by solving linear system (20) with the appropriate boundary conditions. It results that, for all pay-off matrices and all values of \( N \) used, the numerically calculated fixation probabilities lie between the corresponding lower and upper bounds, showing that the hypothesis of large enough \( N \) in Theorem 5 seems to be not too much stringent, if at all necessary. The main point here is not arguing whether the hypothesis is necessary, but illustrating that the upper and lower bounds of Theorem 5 may give quite good estimates in some examples, and also that they may be not very informative in other examples. We will also return to some of the given examples in the rest of this paper.

One feature of Theorem 5, explaining why results may be not very good in some cases, is that the lower and upper bounds \( \pi^\text{low} \) and \( \pi^\text{up} \) were obtained minimizing \( Z_{ij}^+ \) (resp. maximizing \( Z_{ij}^- \)) over all \( j \). We believe that better bounds may be produced using Theorems 2 and 3 with different choices of comparison chains more adapted to the particular pay-off matrices of the cases to be studied.

As a first example, consider the pay-off matrix

\[
M = \begin{pmatrix}
  3.00 & 0.67 & 2.61 \\
  2.33 & 1.00 & 1.40 \\
  2.97 & 0.90 & 2.25
\end{pmatrix},
\]

which appears in the context of a model for the evolution of cooperation Núñez Rodríguez and Neves (2016). In Sect. 6 we will describe the strategies A, B and C and explain where the matrix comes from. For now, we show in Table 1 the numerically calculated fixation probabilities for all three strategies at point \((i, j) = N(2/5, 3/10)\) and various values of \( N \), and the lower and upper bounds for them obtained by using Theorem 5 and its analogues for B and C fixation probabilities. The chosen point, shown in Fig. 5, is very far from the red region in that picture in which it is easier, see Theorem 6, to prove that the A fixation probability tends to 1 as \( N \to \infty \). Despite that, results in Table 1 for \( \alpha, \beta \) and \( \gamma \) strongly support that if the fractions of A, B and C individuals are fixed respectively at values 2/5, 3/10 and 3/10, the A fixation probability tends to 1 as \( N \to \infty \).

We can also conjecture that the B fixation probability at the chosen point tends to 0 as \( N \to \infty \). This is a consequence of Theorem 10.

In our second example, the pay-off matrix is

\[
M = \begin{pmatrix}
  2 & 2.1 & 2.15 \\
  2.2 & 1.75 & 1.83 \\
  2.1 & 1.8 & 1.82
\end{pmatrix}.
\]
At each column the diagonal element is the smallest pay-off. This fact implies that all three equilibria of the replicator dynamics in which the population is composed of individuals following a single strategy are sources.

In Fig. 2 we show for this pay-off matrix and two values of \( N \) the lower and upper bounds of Theorem 5 for the A fixation probability \( \alpha_{ij} \) as functions of the fraction \( i/N \) of A individuals. We also plot in the left panel of the same figure the numerically calculated values of \( \alpha_{ij} \) for each value of \( i \) and a random choice of \( j \) between 1 and \( N - i - 1 \). Notice that for all values of \( i \) and already for \( N = 40 \) the fixation probability lies between the lower and upper bounds. In the right panel we show only the lower and upper bounds for the same pay-off matrix and \( N = 100 \). Notice that in this case the upper bounds are not very interesting, because they are very close to 1. But the lower bounds are also very close to 1 even for small values of \( i \). This fact suggests that, despite the whole population adopting strategy A being a repulsive equilibrium in the deterministic dynamics, for fixed \((x, y)\) in the interior of \( \Lambda \) we have \( A_N(x, y) \) tending to 1 as \( N \to \infty \).

We do not have a proof of the above claim, but pay-off matrix (38) is just a perturbation of (43), for which we will indeed prove that \( A_N(x, y) \xrightarrow{N \to \infty} 1 \) if \( d \) is close enough to 0. The argument for that is given after Theorem 9.
Fig. 3 Upper and lower bounds for the A fixation probability as functions of the fraction of A individuals. The pay-off matrix is (39). Population sizes are $N = 40$ in the left panel and $N = 200$ in the right panel.

Our last example here is pay-off matrix

$$M = \begin{pmatrix} 3 & 1 & 5 \\ 4 & 3 & 1 \\ 1 & 4 & 3 \end{pmatrix},$$

(39)

an example of the well-known rock-scissors-paper game Hofbauer and Sigmund (1998). In those games all equilibria in which the whole population is of a single type correspond to saddle points of the replicator equation. As shown in Fig. 3, for $N = 40$ the lower and upper bounds for the A fixation probability are not very informative, as the former are close to 0, and the latter are close to 1. Increasing population size to $N = 200$, right panel in the same figure, makes things worse. That behavior is typical of the closeness to a pure-strategy saddle point.

5 Strict Nash equilibria and related results

In general, consequences of Theorem 5 depend on knowing for each $i$ the location of the maximum or of the minimum among the values $Z_{ij}^+$, $j = 1, 2, \ldots, N - i$, but explicitly finding these points may be a complicated task. The results in this section refer to important situations in which Theorem 5 may be used without the need of locating the maximum or the minimum of the $Z_{ij}^+$.

One situation is the $N \to \infty$ limit of the fixation probability of a strategy when this strategy is a strict Nash equilibrium. We remind that strategy A is a strict Nash equilibrium (see e.g. Hofbauer and Sigmund (1998) or Nowak (2006)) if $m_{11} > m_{i1}$ for $i = 2$ and $i = 3$. It can be shown that if A is a strict Nash equilibrium, then the point $(1, 0) \in \Lambda$ corresponding to the whole population being of type A is an asymptotically stable equilibrium for the replicator dynamics. There exists then a basin of attraction $B$ for this equilibrium, such that every orbit of the replicator dynamics starting at a point in $B$ will end in $(1, 0)$ in the infinite time limit. As an important similarity of the Moran process with deterministic dynamics we will show that if strategy A is a strict Nash equilibrium, then $A_N(x, y) \nrightarrow N \to \infty 1$ for all $(x, y)$ in $B$. 
We start with a less strong result, showing that $A_N(x, y)$ tends to 1 as $N \to \infty$ in a neighborhood of the point $(1, 0)$:

**Theorem 6** Consider a Moran process with three strategies such that strategy A is a strict Nash equilibrium. Then there exists $x^* \in [0, 1)$ such that $\lim_{N \to \infty} A_N(x, y) = 1$ for all $(x, y) \in \Lambda$ with $x > x^*$.

The proof of the preceding result, along with some other results necessary for it, may be found in B.4.

In order to extend the applicability of Theorem 6 to the whole basin of attraction $B$ of the strict Nash equilibrium in the replicator dynamics, we take some time here to introduce some different ideas and a strong result borrowed from Benaim and Weibull (2003).

For every $(x, y) \in \Lambda$ we define the Moran vector field as the two-dimensional vector $V(x, y) = (V_1(x, y), V_2(x, y))$ with components given by

$$V_1(x, y) = P_{AB}(x, y) + P_{AC}(x, y) - P_{BA}(x, y) - P_{CA}(x, y)$$

and

$$V_2(x, y) = P_{BA}(x, y) + P_{BC}(x, y) - P_{AB}(x, y) - P_{CB}(x, y).$$

If $L$ and $M$ stand for A, B or C, functions $p_{LM}^{ij}(x, y)$ are defined as the deterministic limits of the $p_{ij}^{LM}$ transition probabilities appearing in (16). The relationship between $p_{LM}^{ij}(x, y)$ and $p_{ij}^{LM}$ is just like the one between $F(x, y)$ and $f_{ij}$ exemplified in (19). As an example,

$$p_{ij}^{AB} = \frac{ijf_{ij}}{NS_{ij}} = P_{AB}(x, y) + O(\frac{1}{N}),$$

where $x = i/N$, $y = j/N$ and

$$P_{AB}(x, y) = \frac{xyF(x, y)}{xF(x, y) + yG(x, y) + (1 - x - y)H(x, y)}.$$

The system of differential equations

$$\begin{cases}
\dot{x} = V_1(x, y) \\
\dot{y} = V_2(x, y)
\end{cases}$$

will be referred to as the Moran replicator equations.

The above differential equations are not the standard replicator equations Taylor and Jonker (1978) restricted to the $xy$ plane, but the reader may easily show that at each $(x, y) \in \Lambda$, $V(x, y)$ is the corresponding vector field in the standard replicator equations multiplied by the positive scalar $[xF(x, y) + yG(x, y) + (1 - x - y)H(x, y)]^{-1}$. As a consequence, although the solutions to both systems are not the same, the orbits in $\Lambda$ of these solutions are exactly the same curves, but traversed with different speeds. Of course the basins of attraction of equilibria are the same both for the standard and Moran replicator equations.
Lemma 1 in Benaim and Weibull (2003) relates the solutions of mean-field differential equations over finite time intervals with the related Markov chains. In our concrete case, that lemma shows that solutions of the deterministic Moran replicator equations (40) uniformly approximate the stochastic Moran processes with the same pay-off matrix over finite time intervals. Although the lemma is applicable to more general situations, we state it here without proof and adapted to the situation and notation of this work. We start with some notations.

Let \( \xi(t, x_0, y_0) = (x(t), y(t)) \) be the solution at time \( t \) of Eq. (40) with initial condition \( (x(0), y(0)) = (x_0, y_0) \in \Lambda \). Let \( N \) be the population size and \( \hat{X}^N(t) = \frac{1}{N}(i_{[Nt]}, j_{[Nt]}) \) be the rescaled state at time \( [Nt] \) (notice the multiplication by \( N \)) of the Moran process with three strategies and initial state \( (i, j) = ([Nx_0], [Ny_0]) \in \Lambda_N \). The distance at time \( t \) between \( \xi(t, x_0, y_0) \) and \( \hat{X}^N(t) \) may be defined as

\[
||\hat{X}^N(t) - \xi(t, x_0, y_0)||_\infty = \max\{|x(t) - \frac{i_{[Nt]}}{N}|, |y(t) - \frac{j_{[Nt]}}{N}|\}.
\]  

Another distance is the uniform convergence norm

\[
D^N(T, x_0, y_0) = \max_{0 \leq t \leq T} ||\hat{X}^N(t) - \xi(t, x_0, y_0)||_\infty.
\]  

Lemma 1 in Benaim and Weibull (2003) may then be stated as

**Theorem 7** (Benaim and Weibull) There exists a scalar \( c > 0 \) such that, for any \( \epsilon > 0 \), \( T > 0 \), and any \( N \) large enough

\[
\text{Prob}\{D^N(T, x_0, y_0) \geq \epsilon\} \leq 4e^{-\epsilon^2 cN}.
\]

In other words, Theorem 7 states that for the same initial conditions, finite time \( T \), any given \( \epsilon > 0 \) and large enough population \( N \), the probability that at any time \( t \in [0, T] \) the Moran process state deviates more than \( \epsilon \) of the solution \( (x(t), y(t)) \) of (40) tends to 0 as \( N \to \infty \).

An important difference between Theorem 7 and results about fixation probabilities such as Theorem 6 is that fixation is an infinite time phenomenon in the deterministic dynamics (40). Thus we cannot directly apply Theorem 7 to conclude anything about fixation probabilities, but we may use it together with some other theorem. The next theorem is a general result connecting Theorem 7 with fixation probabilities.

**Theorem 8** Let \( p \in [0, 1] \) and \( U \subset \Lambda \) be an open set such that \( A_N(x, y) \to p \) as \( N \to \infty \) for all \((x, y) \in U\). If, for some finite \( T \), \( \xi(T, x_0, y_0) \in U \), then \( A_N(x_0, y_0) \to p \) as \( N \to \infty \).

**Proof** Let \( \epsilon > 0 \) be such that the ball \( B_\epsilon \) of radius \( \epsilon \) centered in \( \xi(T, x_0, y_0) \) is contained in \( U \). \( A_N(x_0, y_0) \) may be written as the sum of two terms: the fixation probability of strategy A starting at \((x_0, y_0)\) and conditioned to passing at time \([NT]\) through \( B_\epsilon \) and the probability of starting at \((x_0, y_0)\) and passing elsewhere at time \([NT]\). Theorem 7 implies that the latter tends to 0 as \( N \to \infty \). The former may be
written as the product of the probability of the Moran process taking from \((x_0, y_0)\) to \(B_\epsilon\) in time \([NT]\) by the probability of fixation of A starting in \(B_\epsilon\). The latter tends to \(p\) by hypothesis and the former tends to 1 by Theorem 7.

The following result extends the result of Theorem 6 to the whole basin of attraction of a strict Nash equilibrium in the deterministic dynamics:

**Corollary 2** Consider a Moran process with three strategies such that strategy A is a strict Nash equilibrium. Let \(B\) be the set of points \((x_0, y_0)\) such that \(\lim_{t \to \infty} \xi(t, x_0, y_0) = (1, 0)\), i.e. the deterministic basin of attraction of strategy A. Then, for all \((x, y) \in B\), \(A_N(x, y) \rightarrow 1\).

**Proof** Let \(x^*\) be as in Theorem 6 and let \(U\) be the interior of the set \(\{(x, y) \in \Lambda; x > x^*\}\). Theorem 6 guarantees that the hypothesis of Theorem 8 for \(U\) is fulfilled with \(p = 1\). We also know that any orbit of the Moran replicator equation starting in \(B\) must enter \(U\) in some finite time. So the second hypothesis in Theorem 8 holds and the result is thus a consequence of it.

Theorem 9, our next result, deals with a situation which is in some sense inverse to that in Theorem 6. We suppose now that \(F(1, 0) < G(1, 0)\) and \(F(1, 0) < H(1, 0)\). This, by continuity, implies that in a neighborhood of the point \((1, 0) \in \Lambda\) strategy A is the least fit. In the replicator dynamics, this hypothesis implies that vertex A is a source, i.e. all orbits close to \((1, 0)\) must move away from it.

Before enunciating Theorem 9, we make a definition for Moran processes with three strategies analogous to another definition made in Appendix A for birth–death processes.

Let \(i > i^*\). We define \(\alpha_{i,j \mid i^*}\) as the probability with initial condition \((i, j)\) that strategy A fixates without ever returning to any among the states \((i^*, k)\) \(k = 0, 1, \ldots, N - i^*\). In other words, in \(\alpha_{i,j \mid i^*}\) we take into account only events in which strategy A fixes and the number of A individuals is always larger than \(i^*\). Similarly, for \(x > x^*\) and \(N\) large enough so that \([Nx] > [Nx^*]\), we define \(A_{N \backslash x^*}(x, y) = \alpha_{[Nx],[Ny],[Nx^*]}\).

With this definition, we have:

**Theorem 9** Consider a Moran process with three strategies. If \(F(1, 0) < G(1, 0)\) and \(F(1, 0) < H(1, 0)\), then there exists \(x^* \in [0, 1)\) such that

\[
\lim_{N \to \infty} A_{N \backslash x^*}(x, y) = 0
\]

for all \((x, y) \in \Lambda\) with \(x > x^*\).

The proof of Theorem 9 is just a repetition of the ideas in the proof of Theorem 6, using Theorem 14 stated and proved in B.4.

One might naively believe that in the situation of Theorem 9 it could be true that \(A_N(x, y) \to 0\) as \(N \to \infty\), but the above result is much weaker. In fact, the stronger result described here is false. This is shown by the following example, in which the hypotheses of Theorem 9 are fulfilled, but we have for all \((x, y) \in \Lambda\) with \(x > 0\) that \(A_N(x, y) \to 1\).
Consider the pay-off matrix

\[ M = \begin{pmatrix} a & b & b \\ c & d & d \\ c & d & d \end{pmatrix} \]  

with \( a, b, c \) and \( d \) all positive, \( a < c \) and \( d < b \). The peculiar form of \( M \) implies that individuals adopting strategies B and C have the same fitness for any population composition. As a result, for the sake of calculating the fixation probabilities it is as if we had only two strategies. Thus \( \alpha_{ij} \) is independent of \( j \) and can be calculated by (4). Inequality \( a < c \) guarantees that \( F(1, 0) < G(1, 0) = H(1, 0) \). The other inequality \( d < b \) implies that strategies B and C are not Nash equilibria. This is a hypothesis necessary for using Theorem 5 in de Souza et al. (2019). Using the notation of the above cited work, see also (45) ahead in this paper, by taking \( d \) close enough to 0 we get \( L(1) < 0 \). According to Theorem 5 in de Souza et al. (2019), we will have

\[ \lim_{N \to \infty} A_N(x, y) = 1 \]  

for all \( (x, y) \in \Lambda \) with \( x > 0 \).

As a further illustration of the same fact, see again Fig. 2. The pay-off matrix (38) used there is just a perturbation of (43). In the numerical results presented we can see that the fixation probability of the repulsive strategy A seems to tend to 1 as \( N \to \infty \). As (38) does not have the exact form of (43), this is only a conjecture.

The example after Theorem 9 shows that if strategy A is the least fit and it is close to fixation, then for large \( N \) its fixation probability may even be close to 1. This will not happen if A is the least fit strategy in a region away from its fixation. This is the content of our last result in this section.

**Theorem 10** Consider a Moran process with three strategies. If there exists \( x^* \in (0, 1) \) such that \( F(x, y) < G(x, y) \) and \( F(x, y) < H(x, y) \) \( \forall (x, y) \in \Lambda \) with \( x < x^* \), then \( \lim_{N \to \infty} A_N(x, y) = 0 \) for all \( (x, y) \in \Lambda \) with \( x < x^* \).

It is important to remember that, with the exception of Theorem 4, all the results presented since the beginning of Sect. 3 refer to the fixation of the A strategy. By making appropriate adjustments in the hypotheses, analogous results are also valid for the fixation of strategies B and C.

**6 The evolution of cooperation with three strategies from a stochastic point of view**

In Núñez Rodríguez and Neves (2016) some results were proved for the replicator dynamics in a model for the evolution of cooperation in a population with three strategies. In this section, after a brief description of the problem, we illustrate the use of the results proved so far for the Moran process with three strategies in gaining some understanding of the stochastic version of the results obtained in that paper.

Nowak and Sigmund (1992) considered a population with 100 types of individuals following different reactive strategies for the infinitely repeated prisoner’s dilemma (IRPD). In the prisoner’s dilemma, individuals may at each interaction either cooperate or defect. Reactive strategies are characterized by two parameters: loyalty and forgiveness. The loyalty of a reactive strategy is the probability that the player adopting
this strategy cooperates after receiving cooperation in the previous interaction. The forgiveness is the probability that the player cooperates after receiving a defection in the previous interaction. Although reactive strategies are characterized by probabilities, the pay-offs in the IRPD are deterministically calculated as an average of infinitely many interactions Nowak (2006). Nowak and Sigmund numerically solved the replicator dynamics for this population, a system of 99 ordinary differential equations. Their numerical solution suggested that among the 100 strategies, only three play a prominent role in their numerical experiment.

The first prominent strategy is ALLD: individuals which always defect. The second important strategy is ATFT (almost tit-for-tat): individuals with loyalty close to 1 and small positive forgiveness. The loyalty and forgiveness values used by Nowak and Sigmund (1992) were respectively 0.99 and 0.01, but the results in Núñez Rodríguez and Neves (2016) hold as long as loyalty is close enough to 1 and forgiveness is close enough to 0. The third strategy, which we will denote as G (generous), has loyalty equal to 1, and positive forgiveness \( q \) larger than the forgiveness of the ATFT.

In the numerical experiment by Nowak and Sigmund (1992), initially the population frequency of most strategies is strongly depleted, with the exception of strategies closest to ALLD. After this initial period, the frequency of strategies close to ATFT increased, almost attaining the whole population. But the ultimate winner in their simulations was a surprisingly cooperative strategy which they called GTFT (generous tit-for-that), i.e. a strategy of the G kind with an optimum value of \( q \) which allows the followers of this strategy not to be too much exploited by defectors. This optimum value of \( q \) turned out to be \( 1/3 \) for the parameter values used in their experiment.

Núñez Rodríguez and Neves (2016) studied a simplified version of the extremely complicated population in Nowak and Sigmund (1992). They considered only three kinds of individuals adopting the aforementioned prominent reactive strategies: ALLD, ATFT and G. The forgiveness parameter \( q \) of the G individuals may be varied. The results of Núñez Rodríguez and Neves (2016) show that, according to the value of \( q \), there are several different scenarios for which strategies survive in the infinite time limit and, consequently, different types of evolution of cooperation, or non-evolution of cooperation, may occur. In particular, they prove that there exists a threshold value for \( q \) under which the result of the numerical experiment by Nowak and Sigmund holds. More precisely, in a population with the three types of individuals above mentioned, existence is proved of a threshold \( q_{GTFT} \) such that for \( q < q_{GTFT} \) there will be a region with positive area such that for initial conditions in this region only G individuals will survive. In a complementary region, also with positive area, only ALLD individuals will survive. When \( q < q_{GTFT} \), the scenario is termed one in which full evolution of cooperation holds.

We will consider in this section only the full evolution of cooperation scenario (FECS). Matrix (37), considered before in Sect. 4, is an example of pay-off matrix in this scenario. Its calculation relies on the stationary distribution of a Markov chain with 4 states and the pay-offs are averages of the parameters \( T, R, P \) and \( S \) of the prisoner’s dilemma weighted by the stationary distribution. The whole calculation is explained in Núñez Rodríguez and Neves (2016) and uses formulas (8)-(11) in that paper. The parameter values for obtaining (37), as well as some other useful information, are stated in the caption of Fig. 4. Strategy A (numbered 1 in the matrix) is G, strategy B
Fig. 4 Some orbits for the replicator dynamics with pay-off matrix in (37). Strategy A is G (generous), strategy B is ALLD (always defect), and strategy C is ATFT (almost tit-for-tat). The parameters for the prisoner’s dilemma (see Núñez Rodríguez and Neves (2016) or Nowak and Sigmund (1992)) are $T = 5$, $R = 3$, $P = 1$, $S = 0$. The G strategy has a forgiveness parameter $q = 1/3$ and the ATFT has loyalty equal to 0.9 and forgiveness 0.1. Besides the vertices of the triangle, the only equilibria of the dynamics in the full evolution of cooperation scenario are points $P_{12}$ and $P_{23}$. $P_{12}$ is a source and $P_{23}$ is a saddle point. Strategies ALLD and G are strict Nash equilibria. The heteroclinic orbit joining $P_{12}$ and $P_{23}$ separates the basins of attraction of both strict Nash equilibria. If $F$, $G$ and $H$ denote as elsewhere in this paper the deterministic fitnesses (18), the red line is the set of points in which $F = G$. The green line is the set of points in which $F = H$. The blue line is the set of points in which $G = H$ (color figure online)

(numbered 2) is ALLD, and strategy C (numbered 3) is ATFT. The reader consulting Núñez Rodríguez and Neves (2016) should be aware that the numbering of strategies here differs with respect to that paper. Fig. 4 shows some orbits of the replicator dynamics for the pay-off matrix (37) and the clear partition of $Λ$ in regions of positive area in which either only strategy A or only strategy B survive.

Because $m_{11}$ is the only maximal coefficient in the first column of (37), we know that strategy A is a strict Nash equilibrium. Similarly, $m_{22}$ is the only maximal coefficient in the second column and B is also a strict Nash equilibrium. But $m_{33}$ is not a maximal coefficient in the third column and strategy C is not a Nash equilibrium. These conclusions on which strategies are or are not Nash equilibria are generally valid for the FECS. We also know that for the FECS there is no interior equilibrium for the dynamics and no equilibrium on side $AC$. On the side $AC$ A is always fitter than C. On the side $AB$ we have an equilibrium, depicted as $P_{12}$ in Fig. 4, such that above $P_{12}$ on that side, B is fitter than A, but A is fitter than B below $P_{12}$. On the side $BC$ we have an equilibrium $P_{23}$ such that B is fitter than C above it and C is fitter than B below it.

We also show in Fig. 4 the lines in which the deterministic fitnesses $F$, $G$ and $H$ defined in (18) are pairwise equal. These isoclines divide the $ABC$ triangle in regions where the fitness ranking is fixed. It can be seen that above the $G = H$ isocline, the blue line in Fig. 4, we have $G > H > F$, between the $G = H$ and $F = G$ (red in Fig. 4) isoclines we have $H > G > F$, between the $F = G$ and $F = H$ (green in Fig. 4) isoclines we have $H > F > G$ and, finally, below the $F = H$ isocline we have $F > H > G$.

Let $x_1^*$ denote the fraction of A individuals at the point in which the $F = H$ isocline intercepts the $AB$ side. The fitness ranking given above shows that if $x > x_1^*$, then $F > H > G$. We can then readily apply Theorem 6 to conclude that for initial conditions
Fig. 5 The red, green and blue lines and points \( P_{12} \) and \( P_{23} \) are the same as defined in the caption of Fig. 4. The red region close to vertex A is the one in which we can use Theorem 6 to prove that \( A_N(x, y) \) tends to 1 when \( N \to \infty \). By Corollary 2, this region can be extended to the much larger region below the heteroclinic orbit of the replicator equations joining \( P_{12} \) and \( P_{23} \), also shown in the figure. The blue region close to B is, by an analogue of the same theorem, the one in which we can prove that \( B_N(x, y) \) tends to 1 when \( N \to \infty \). Again, it can be extended by Corollary 2, to the entire region above the heteroclinic orbit. The point close to the center of the triangle corresponds to the state \((x, y) = (2/5, 3/10)\). See Table 1 for values of the fixation probabilities and respective lower and upper bounds at this point (color figure online).

\[(x, y)\) in the red region close to vertex A in Fig. 5 we have \( \lim_{N \to \infty} A_N(x, y) = 1 \). Similarly, if \( y^*_1 \) denotes the fraction of B individuals at the point \( P_{23} \), then, for \( y > y^*_1 \) we have \( G > H > F \). By an analogue of the same theorem, \( \lim_{N \to \infty} B_N(x, y) = 1 \) if \((x, y)\) is in the blue region in Fig. 5.

Another of our results we can use is Theorem 10. If we denote \( y^*_2 \) the fraction of type B individuals at point \( P_{12} \), then the fitness ranking already exhibited shows that B is the least fit strategy for all \((x, y) \in \Lambda \) with \( y < y^*_2 \). By an analogue of Theorem 10, \( B_N(x, y) \xrightarrow{N \to \infty} 0 \) for such points.

Although \( B_N(x, y) \xrightarrow{N \to \infty} 0 \) does not necessarily mean that the fixation probability for strategy A should be large, we may use Corollary 2 to extend the red region in Fig. 5 in which we know that \( A_N(x, y) \xrightarrow{N \to \infty} 1 \) to the much larger region below the heteroclinic orbit joining \( P_{12} \) and \( P_{23} \), the basin of attraction of point A in the replicator dynamics. Similarly, the region in which \( B_N(x, y) \xrightarrow{N \to \infty} 1 \) can be extended to the basin of attraction of point B in the replicator equation, i.e. the region above the heteroclinic orbit joining \( P_{12} \) and \( P_{23} \).

7 Conclusions

When we started working on the subject of Moran process with three or more strategies, we had as an optimistic goal to provide a complete classification of all possible behaviors, either as Taylor et al. (2004) did for the Moran process with two strategies, or as Bomze (1983) did for the replicator dynamics with three strategies. We see we are still very far from achieving this goal, but we believe that this paper may be a good starting point for further work.
As a first important achievement, we introduced coupling of stochastic processes as a tool for obtaining results for the Moran process with three or more strategies. More specifically, we used coupling to obtain upper and lower bounds of general validity for the fixation probabilities.

We do not claim that our general upper and lower bounds in Theorem 5 are optimal. In fact, although Theorems 2 and 3 allow more flexibility, the general recipe of Theorem 5 with its maxima and minima taken over all $j$ may produce too small lower bounds or too large upper bounds. Examples in which these bounds are either very informative or very poor are provided in Sect. 4. Nonetheless, the general recipe is powerful enough for proving in some cases, see results in Sect. 5, that some strategies may fixate or be extinct with large probability if the population size $N$ is large.

In Sect. 6 we applied these results to a concrete problem in which two among the strategies are strict Nash equilibria and the third strategy is not a Nash equilibrium. In this case we were able to partition the whole region $\Lambda$ into regions in which either population related to Nash equilibria fixates with probability tending to 1 as the population size $N$ tends to $\infty$. As a result, we provided a stochastic analog of the full evolution of cooperation scenario in Núñez Rodríguez and Neves (2016).

We hope that our results may prove useful in other applications, or else, that work in other applications may suggest some better bounds for fixation probabilities. We believe that an interesting and challenging problem for the future is studying the behavior of fixation probabilities in cases in which there exist attractors of the replicator dynamics not at the vertices of $\Lambda$. We provided just after Theorem 9 an example of such a problem. We were able to study it, because it amounted to a Moran process with only two strategies, but its generalization given by pay-off matrix (38) is already beyond our present results. Still more challenging are the situations in which the replicator dynamics has an equilibrium in the interior of $\Lambda$, such as for pay-off matrix (39). When this interior equilibrium is asymptotically stable, we are still missing a quantity playing the same role for the Moran process with three or more strategies as $L(1)$ in (45) does for the process with two strategies, see de Souza et al. (2019). Another open problem is the study of fixation probabilities for small or intermediate values of $N$ and the possible transitions in their behaviors until the $N \to \infty$ regime is reached, see again de Souza et al. (2019) for results of this type for the two-strategy Moran process.

Acknowledgements This study was financed in part by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Finance Code 001. AGMN thanks Fábio Chalub for many fruitful discussions. The authors acknowledge the anonymous referees and the Associate Editor for many corrections improving the clarity of this work and for suggestions, including combining Lemma 1 in Benaim and Weibull (2003), Theorem 7 with our Theorem 6, strengthening a lot the results we had in a previous version of this paper.

A Some results on birth–death processes

This appendix collects some results which we did not want to insert in the main text of the paper, because they have to do only with birth–death processes. Despite that, these results were used in the proof of the theorems in Sect. 5, all of them referring to Moran
processes with three strategies. The more important results here are Theorems 11 and 12, which are cited in the proofs of Appendix B. The propositions which precede them are necessary for their proofs.

The first result here deals with comparing fixation probabilities for two birth–death processes in which the birth to death ratio is larger in one process than in the other. We observe that this result might be proved by a coupling argument similar to the one shown in Theorem 2. We opt here for a direct proof using the exact expressions (4) for the fixation probabilities.

**Proposition 2** Consider two birth–death processes with the same set of states \( S = \{0, 1, 2, \ldots, N\} \). Let \( r_i \equiv a_i/b_i \) be the birth to death ratio in the first process and \( s_i \equiv a'_i/b'_i \) be the ratio in the second process. Let also \( \pi_i \) and \( \pi'_i \) denote the respective fixation probabilities in state \( N \). If \( r_i > s_i \) for \( i = 1, 2, \ldots, N - 1 \), then \( \pi_i > \pi'_i \) for all \( i \in S \setminus \{0, N\} \).

**Proof** We start by rewriting expression (4) for \( \pi_i \):

\[
\pi_i = \frac{1 + \sum_{j=1}^{i-1} \prod_{k=1}^{j} r_k^{-1}}{1 + \sum_{j=1}^{N-1} \prod_{k=1}^{j} r_k^{-1}} = \frac{1 + \sum_{j=1}^{i-1} \prod_{k=1}^{j} r_k^{-1}}{1 + \sum_{j=1}^{N-1} \prod_{k=1}^{j} r_k^{-1}}.
\]

A similar expression may be written for \( \pi'_i \) just by writing \( s_k \) in place of \( r_k \). Let

\[
d_i = \frac{\sum_{j=i}^{N-1} \prod_{k=1}^{j} r_k^{-1}}{1 + \sum_{j=1}^{i-1} \prod_{k=1}^{j} r_k^{-1}}
\]

be the denominator minus 1 in the last expression and \( d'_i \) be the same expression with \( r_k \) replaced by \( s_k \). We will prove that \( d'_i - d_i > 0 \), which of course implies \( \pi_i > \pi'_i \).

\[
d'_i - d_i = \left( \sum_{j=i}^{N-1} \prod_{k=1}^{j} s_k^{-1} \right) \left( 1 + \sum_{j=1}^{i-1} \prod_{k=1}^{j} r_k^{-1} \right) \left( 1 + \sum_{j=1}^{i-1} \prod_{k=1}^{j} s_k^{-1} \right) - \left( \sum_{j=i}^{N-1} \prod_{k=1}^{j} r_k^{-1} \right) \left( 1 + \sum_{j=1}^{i-1} \prod_{k=1}^{j} s_k^{-1} \right) \left( 1 + \sum_{j=1}^{i-1} \prod_{k=1}^{j} s_k^{-1} \right)
\]

\[
= \frac{\sum_{j=i}^{N-1} \prod_{k=1}^{j} s_k^{-1} - \prod_{k=1}^{i} r_k^{-1}}{\left( 1 + \sum_{j=1}^{i-1} \prod_{k=1}^{j} r_k^{-1} \right) \left( 1 + \sum_{j=1}^{i-1} \prod_{k=1}^{j} s_k^{-1} \right)}
\]

\[
+ \frac{\left( \sum_{j=i}^{N-1} \prod_{k=1}^{j} s_k^{-1} \right) \left( \sum_{j=i}^{N-1} \prod_{k=1}^{j} r_k^{-1} \right) - \left( \sum_{j=i}^{N-1} \prod_{k=1}^{j} r_k^{-1} \right) \left( \sum_{j=i}^{N-1} \prod_{k=1}^{j} s_k^{-1} \right)}{\left( 1 + \sum_{j=1}^{i-1} \prod_{k=1}^{j} r_k^{-1} \right) \left( 1 + \sum_{j=1}^{i-1} \prod_{k=1}^{j} s_k^{-1} \right)}.
\]

Using the fact that \( s_k^{-1} > r_k^{-1} > 0 \) for all \( k \), both numerator and denominator in the first term in the last expression are clearly positive. To see that the second term is
positive, too, notice that its denominator is the same as the first term, and its numerator, with some patience, may be rewritten as

\[
\sum_{j=1}^{i-1} \left( \prod_{k=1}^{j} r_k^{-1} s_k^{-1} \right) \sum_{\ell=1}^{N-i} \left( \prod_{m=j+1}^{N-\ell} s_m^{-1} - \prod_{m=j+1}^{N-\ell} r_m^{-1} \right),
\]

now manifestly positive. \[\square\]

**Proposition 3** Consider a birth–death process with state space \( S = \{0, 1, 2, \ldots, N\} \). If \( r_i \) is the ratio of birth to death probabilities and \( i^* \in \{1, 2, \ldots, N-1\} \) is some fixed state, then the probability \( \pi_{i \setminus i^*} \) that the process starts at \( i > i^* \) and fixates at state \( N \) without ever passing by state \( i^* \) is

\[
\pi_{i \setminus i^*} = \frac{1 + \sum_{\ell=1}^{i-i^*-1} \prod_{k=1}^{\ell} r_k^{i^*+k}}{1 + \sum_{\ell=1}^{N-i-i^*-1} \prod_{k=1}^{\ell} r_k^{i^*+k}}.
\]

**Proof** Our result (44) may be obtained from (4), noticing that the boundary condition \( \pi_0 = 0 \) is replaced by \( \pi_{i^* \setminus i^*} = 0 \) and the set of states \( \{0, 1, 2, \ldots, N\} \) is replaced by \( \{i^*, i^* + 1, \ldots, N\} \). \[\square\]

The following result, Proposition 4, is just a straightforward adaptation of the result in Proposition 2 to the fixation probability \( \pi_{i \setminus i^*} \) defined in Proposition 3. As the proof is a mere repetition, we do not write it here.

**Proposition 4** Consider two birth–death processes with the same set of states \( S = \{0, 1, 2, \ldots, N\} \). Let \( r_i \equiv a_i/b_i \) and \( s_i \equiv a'_i/b'_i \) be the respective birth to death ratios. Suppose that there exists \( i^* \) such that \( r_i > s_i \) for \( i = i^* + 1, i^* + 2, \ldots, N - 1 \). If \( i > i^* \) and \( \pi_{i \setminus i^*} \) and \( \pi'_{i \setminus i^*} \) denote the fixation probabilities in state \( N \) with the additional condition that the process never passes by state \( i^* \), then \( \pi_{i \setminus i^*} > \pi'_{i \setminus i^*} \) for all \( i \in \{i^* + 1, i^* + 2, \ldots, N - 1\} \).

The next result is the key ingredient in the proof of the important Theorem 13, which generalizes Theorem 6 in the main text.

**Theorem 11** Suppose that for large enough values of \( N \) we have a family of birth–death processes with birth to death ratios \( r_i^{(N)} \) and fixation probabilities \( \pi_i^{(N)} \), \( i = 1, 2, \ldots, N-1 \). Let \( x \in (0, 1) \) and \( \Pi_N(x) \equiv \pi_{i_N(x)}^{(N)} \) if there exists \( s > 1 \) and \( x^* \in (0, 1) \) such that for \( N \) large enough and \( i > N x^* \) we have \( r_i^{(N)} > s \), then

\[
\lim_{N \to \infty} \Pi_N(x) = 1
\]

for \( x > x^* \).
Proof Let $\pi_i'$ be the fixation probability of a birth–death process with frequency independent fitness $s_i = s$ and let $\pi_i^\prime\backslash i^\ast$ be as in Proposition 4. Summing the geometric progressions appearing in (44) when $r_i$ is replaced by $s$, we get

$$\pi_i^\prime\backslash i^\ast = \frac{1 - s^{-(i-i^\ast)}}{1 - s^{-(N-i^\ast)}}.$$ 

Suppose $x > x^\ast$ and $N$ large enough so that $[Nx] > [Nx^\ast]$. Of course, $\pi^{(N)}_{[Nx]} \geq \pi^{(N)}_{[Nx]\backslash [Nx^\ast]}$. As, by Proposition 4, we have $\pi^{(N)}_{[Nx]\backslash [Nx^\ast]} > \pi^{(N)}_{[Nx^\ast]\backslash [Nx^\ast]}$, then

$$\pi^{(N)}_{[Nx]} > \frac{1 - s^{-([Nx]-[Nx^\ast])}}{1 - s^{-(N-[Nx^\ast])}}.$$ 

Our conclusion follows because, if $s > 1$, the last expression tends to 1 when $N \to \infty$.  

The next result here is quite analogous to Theorem 11 in its proof, but it comes with a surprise: although we will be able to prove that, under the stated hypotheses, $\pi^{(N)}_{[Nx]\backslash [Nx^\ast]}$ tends to 0 as $N \to \infty$, we cannot conclude that $\pi^{(N)}_{[Nx]}$ tends to 0, too.

**Theorem 12** Suppose that for large enough values of $N$ we have a family of birth–death processes with birth to death ratios $r_i^{(N)}$, and fixation probabilities $\pi^{(N)}_{i|N}$, $i = 1, 2, \ldots, N - 1$. Suppose also that there exist $0 < s < 1$ and $x^\ast \in (0, 1)$ such that for all $N$ and $i > Nx^\ast$ we have $r_i^{(N)} < s$. If $x > x^\ast$ and $\Pi^{N\backslash x^\ast}(x) \equiv \pi^{(N)}_{[Nx]\backslash [Nx^\ast]}$ then

$$\lim_{N \to \infty} \Pi^{N\backslash x^\ast}(x) = 0.$$ 

The proof of the above result is analogous to the proof of Theorem 11 and is left to the interested reader. We comment instead on why we cannot arrive at a result completely analogous to Theorem 11.

The first reason is that inequality $\pi^{(N)}_{[Nx]} \geq \pi^{(N)}_{[Nx]\backslash [Nx^\ast]}$ used in proving Theorem 11 is still valid and we cannot in general conclude that a quantity larger than or equal to something tending to 0 tends to 0, too.

More than that, we know that in a birth–death process for two strategies we can fulfill the hypotheses of Theorem 12 and still have $\lim_{N \to \infty} \pi^{(N)}_{[Nx]} = 1$. This is proved for Moran processes with two strategies in de Souza et al. (2019), Theorem 5, if certain conditions on the pay-off matrix are valid. The conditions are $m_{11} < m_{21}, m_{12} > m_{22}$, i.e. neither of the two strategies is a Nash equilibrium, and

$$L(1) \equiv -\int_0^1 \log \frac{1 - w + w[m_{11}t + m_{12}(1 - t)]}{1 - w + w[m_{21}t + m_{22}(1 - t)]} dt < 0.$$  

(45)
In the interesting situation in which the hypotheses of Theorem 12 are fulfilled and we also have \( \lim_{N \to \infty} \pi_{[N,x]}^{(N)} = 1 \), we have for large \( N \) and \( x > x^\ast \) both \( \Pi_N(x) \) close to 1 and \( \Pi_{N \setminus x^\ast}(x) \) close to 0. This means that although fixation at state \( N \) is very probable, most probably the chain will pass at least once (thus, it will probably pass many times) by \( x^\ast \) before fixation occurs. An application of the above phenomenon, in which a source strategy in the replicator dynamics fixates with high probability in the Moran process, is given by McLoone et al. (2018).

### B Proofs of the theorems in the main text

#### B.1 Proof of Theorem 2

For each \((i, j) \in \Lambda_N\), consider a first partition of the interval \([0, 1]\) illustrated at the left-hand part of Fig. 6 and defined as follows. Let \( q_k(i, j), k = 1, 2, \ldots, 7 \) be the accumulated sums of the transition probabilities at \((i, j)\) ordered as \( p_{ij}^{AB}, p_{ij}^{AC}, p_{ij}^{BC}, p_{ij}^{const}, p_{ij}^{CB}, p_{ij}^{CA} \) and \( p_{ij}^{BA} \). For definiteness, \( q_1(i, j) = p_{ij}^{AB}, q_2(i, j) = q_1(i, j) + p_{ij}^{AC}, q_3(i, j) = q_2(i, j) + p_{ij}^{BC}, \) and so on. Of course, \( q_7(i, j) = 1 \).

For \( k = 1, 2, \ldots, 7 \), let \( I_k(i, j) \) be disjoint intervals with lengths \( q_k(i, j) \):

\[
I_1(i, j) = [0, q_1(i, j)), \\
I_k(i, j) = [q_{k-1}(i, j), q_k(i, j)), \quad k = 2, 3, \ldots, 6
\]
and

\[ I_7(i, j) = [q_6(i, j), 1]. \]

Notice that the sum of the lengths of \( I_1(i, j) \) and \( I_2(i, j) \) is \( Z_{ij}^+ \) and that the sum of the lengths of \( I_6(i, j) \) and \( I_7(i, j) \) is equal to \( Z_{ij}^- \).

For each \( i \in S \), we construct a second partition of \([0, 1]\) illustrated at the right-hand side of Fig. 6, similar to the first partition and related to the comparison chain:

\[ J_1(i) = [0, a_i), \quad J_2(i) = [a_i, 1 - b_i), \]

and

\[ J_3(i) = [1 - b_i, 1]. \]

Conditions (25) imply

\[ J_1(i) \subset I_1(i, j) \cup I_2(i, j) \quad \text{and} \quad J_3(i) \supset I_6(i, j) \cup I_7(i, j). \quad (46) \]

The following vectors are the possible displacements in the state of the target chain:

\[ d_1 = (1, -1), \quad d_2 = (1, 0), \quad d_3 = (0, 1), \quad d_4 = (0, 0), \quad d_5 = (0, -1), \quad d_6 = (-1, 0) \]

and

\[ d_7 = (-1, 1). \]

Suppose that at time 0 the states of the target and comparison chains are respectively \( X_0 = (i_0, j_0) \) and \( Y_0 = i_0 \). The coupling of the target and comparison chains is accomplished by a sequence of independent uniformly distributed random variables \( U_1, U_2, U_3, \ldots \in [0, 1] \) which will determine the state of both chains at all subsequent times.

The state of both chains at time 1 will be obtained by displacements calculated as functions of \( U_1 \), then at time 2 by displacements calculated as functions of \( U_2 \), and so on. The way these displacements are calculated is as follows.

We declare that if the state of the target chain at time \( \ell - 1 \) is \((i_{\ell - 1}, j_{\ell - 1})\), then the \( \ell \)-th displacement of the target chain will be \( d_k \) if \( U_\ell \in I_k(i_{\ell - 1}, j_{\ell - 1}), k = 1, 2, \ldots, 7, \) \( \ell = 1, 2, \ldots \). For the comparison chain, we declare that if its state at time \( \ell - 1 \) is \( i'_{\ell - 1} \), \( \ell = 1, 2, \ldots \), then the displacement of the state of the comparison chain will be 1, 0, or -1, respectively, if \( U_\ell \) is in \( J_1(i'_{\ell - 1}), J_2(i'_{\ell - 1}) \) or \( J_3(i'_{\ell - 1}) \). Notice that the construction up to now is such that the probabilities of the possible displacements of both chains are all correctly distributed according to the chains’ transition probabilities.

A fundamental observation is that, due to (46), whenever \( U_1 \) is such that there is a birth in the comparison chain, then the number of A individuals in the target chain will increase. And also, whenever \( U_1 \) is such that the number of A individuals decreases in the target chain, then there is a death in the comparison chain. As \( i'_0 = i_0 \), it follows that \( i'_1 \geq i'_k \).

We will prove by induction that \( i_k \geq i'_k \forall k \in \mathbb{N} \). Suppose that \( i_k \geq i'_k \) for a certain \( k \in \mathbb{N} \). By the same reasoning used in proving that \( i_1 \geq i'_1 \), we see that if \( i_k = i'_k \), then \( i_{k+1} \geq i'_{k+1} \). If \( i_k \geq i'_k + 2 \), then the conclusion \( i_{k+1} \geq i'_{k+1} \) also holds, because \( i_{k+1} \geq i_k - 1 \) and \( i'_{k+1} \leq i'_k + 1 \).
The only case in which it remains to prove that \( i_{k+1} + 1 \geq i'_{k+1} \) is when \( i_k = i'_{k} + 1 \). In this case, we use condition (26), which we had not used, yet. This condition proves that if \( U_k \) is such that a birth occurs in the comparison chain, then the number of A individuals in the target chain will not decrease and we will still have \( i_{k+1} + 1 \geq i'_{k+1} \).

We have thus realized simultaneously the target and comparison chains according to their respective transition matrices in a way that the initial states are respectively \((i_0, j_0)\) and \(i_0\) and whenever there is fixation at state \(N\) for the comparison chain, then there will be fixation of strategy A in the target chain. Thus \( \alpha_{i_0, j_0} \geq \pi_{i_0} \). As \( i_0 \) is arbitrary, the theorem is proved. \( \square \)

### B.2 Proof of Theorem 4

Before proving the theorem, we state and prove the following result, used in the proof of Theorem 4 and other results ahead.

**Proposition 5** Let \( Z_{ij}^\pm \) be defined as in (24). Then, for each fixed value of \( i \), the minimum of \( Z_{ij}^+ \) and the maximum of \( Z_{ij}^- \) for \( j \in \{0, 1, \ldots, N - i\} \) are attained at the same value of \( j \). Also, the maximum of \( Z_{ij}^+ \) and the minimum of \( Z_{ij}^- \) for \( j \in \{0, 1, \ldots, N - i\} \) are attained at the same value of \( j \).

**Proof** Just notice that \( Z_{ij}^+ = i f_{ij} \frac{N - i}{N} \) and \( Z_{ij}^- = j g_{ij} + (N - i - j) h_{ij} \frac{i}{N} \) may be rewritten as \((1 - i f_{ij}) \frac{i}{N}\). For fixed \( i \) the value of \( j \) minimizing \( i f_{ij} \) will maximize \( 1 - i f_{ij} \). \( \square \)

We can now prove Theorem 4.

Let \( a_i \) be the probability of increasing the number of A individuals from \( i \) to \( i + 1 \) in a population with only A and B individuals. We also define \( b_i \) as the probability of decreasing the number of A individuals from \( i \) to \( i - 1 \) in a population with only A and B individuals. As in (24), let \( Z_{ij}^+ \) and \( Z_{ij}^- \) be respectively the probabilities of increasing and decreasing the number of A individuals from \( i \) to \( i \pm 1 \) in a population with A, B and C individuals and frequency independent fitnesses. We have

\[
\begin{align*}
a_i &= \frac{if}{if + (N - i)g} \frac{N - i}{N}, \\
b_i &= \frac{(N - i)g}{if + (N - i)g} \frac{i}{N}, \\
Z_{ij}^+ &= \frac{if}{if + jg + (N - i - j)h} \frac{N - i}{N}, \\
Z_{ij}^- &= \frac{jg + (N - i - j)h}{if + jg + (N - i - j)h} \frac{i}{N}.
\end{align*}
\]

As \( g > h \), then \( if + (N - i)g = if + jg + (N - i - j)g \geq if + jg + (N - i - j)h \) for all \((i, j) \in \Lambda_N\). It follows that

\[
a_i = Z_{i,N-i}^+ \leq Z_{ij}^+
\]  

(47)

for all \((i, j) \in \Lambda_N\).

By Proposition 5, \( b_i = Z_{i,N-i}^- \geq Z_{ij}^- \) for all \((i, j) \in \Lambda_N\). We have thus proved that conditions (25) in Theorem 2 are fulfilled. The lower bound in (31) will result if we prove that (26) is true.
In fact, it can be seen, after some tedious manipulations, that

\[ 1 - Z_{ij}^- - a_{i-1} \]

\[ = \frac{1}{N S_{ij} S_{i-1,N-i+1}} \left[ i(i-1)^2 f^2 + (N - i + 1)(N - i)g S_{ij} \right. \]

\[ + \left. i f[(N - i - j)(g - h) + [(i - 1)(N - i) + i]g] + [jg + (N - i - j)h]f \right] \].

As all terms around the curly brackets in the above expression are obviously non-negative, as well as the denominator \( N S_{ij} S_{i-1,N-i+1} \), then condition (26) is satisfied and the lower bound in (31) proved.

All the remaining bounds can be proved in an analogous way, either using Theorem 2 or Theorem 3.

\[ \square \]

### B.3 Proof of Theorem 5

We will show that the birth–death process (35) satisfies the hypotheses of Theorem 2 for \( N \) large enough. The proof that the process defined by (36) satisfies the hypotheses of Theorem 3 for large enough \( N \) is analogous.

By (35), we automatically have for each \( i \) that \( a_{low}^i \leq Z_{ij}^+ \) and \( b_{low}^i \geq Z_{ij}^- \) for all \( j \) such that \( (i, j) \in \Lambda_N \). To complete the proof, we need to show that \( a_{low}^i \leq 1 - Z_{ij}^- \) if \( N \) is large enough.

To see that, we write

\[ 1 - Z_{ij}^- - Z_{i-1,j}^+ = (1 - Z_{ij}^- - Z_{ij}^+) + (Z_{ij}^+ - Z_{i-1,j}^+). \]  

(48)

The first term \( 1 - Z_{ij}^- - Z_{ij}^+ \) is the probability at state \( (i, j) \) that the number of A individuals remains constant. It can be written as the sum \( p_{ij}^{AA} + (p_{ij}^{BC} + p_{ij}^{CB} + p_{ij}^{BB} + p_{ij}^{CC}) \), in which \( p_{ij}^{AA} \) vanishes only if \( i = 0 \) and the sum of the remaining four terms vanishes only if \( i = N \).

Writing \( x = i / N \) and \( y = j / N \) and using a reasoning similar to the one exemplified in (19), we get

\[ 1 - Z_{ij}^- - Z_{ij}^+ = C_1(x, y) + C_2(x, y) + O\left(\frac{1}{N}\right), \]  

(49)

where

\[ C_1(x, y) = \frac{x^2 F(x, y)}{xF(x, y) + yG(x, y) + (1 - x - y)H(x, y)} \]

comes from \( p_{ij}^{AA} \) and

\[ C_2(x, y) = \frac{(1 - x)(yG(x, y) + (1 - x - y)H(x, y))}{xF(x, y) + yG(x, y) + (1 - x - y)H(x, y)} \]
comes from the sum \( p_{ij}^{BC} + p_{ij}^{CB} + p_{ij}^{BB} + p_{ij}^{CC} \).

Observe that both \( C_1 \) and \( C_2 \) are continuous functions with values in \([0, 1]\) in the compact triangle \( \Lambda \) defined in (13). Moreover \( C_1(1, 0) = 1 \) and \( C_2(0, y) = 1 \), so that there exist \( x_1, x_2 \in [0, 1], x_1 < x_2 \) such that \( C_2(x, y) \geq 1/2 \) if \((x, y) \in \Lambda \) with \( x \leq x_1 \) and \( C_1(x, y) \geq 1/2 \) if \((x, y) \in \Lambda \) with \( x \geq x_2 \). As neither \( C_1 \) nor \( C_2 \) vanishes for the points \((x, y) \in \Lambda \) with \( x \in [x_1, x_2] \), then their sum has a positive minimum value \( \mu \) in this set. Of course the minimum value of \( C_1 + C_2 \) in \( \Lambda \) cannot be smaller than the smallest between \( \mu \) and \( 1/2 \), being then positive and independent of \( N \). This proves that \( 1 - Z_{ij}^- - Z_{ij}^+ \) is bounded away from 0 for large enough \( N \).

Using the same ideas,

\[
Z_{ij}^+ = D(x, y) + O\left(\frac{1}{N}\right),
\]

with

\[
D(x, y) = \frac{x(1 - x)F(x, y)}{xF(x, y) + yG(x, y) + (1 - x - y)H(x, y)}.
\]

Then the second summand in the right-hand side of (48) becomes

\[
Z_{ij}^+ - Z_{i-1,j}^+ = D(x, y) - D\left(x - \frac{1}{N}, y\right) + O\left(\frac{1}{N}\right) = -\frac{1}{N} \frac{\partial D}{\partial x}(x, y) + O(\frac{1}{N}).
\]

We have thus shown that one of the terms in the right-hand side of (48) is positive and \( O(1) \), and the other is \( O(\frac{1}{N}) \). This proves that \( 1 - Z_{ij}^- - Z_{ij}^+ > 0 \) for all \((i, j) \in \Lambda_N \) for large enough \( N \) and the proof is completed.

\[\square\]

**B.4 Proof of Theorem 6 and related results**

The reader will see that the proof of Theorem 6 is a consequence of this more general result:

**Theorem 13** Consider a Moran process with three strategies. Suppose there exist \( s > 1, N^* \in \mathbb{N} \) and \( x^* \in [0, 1) \) such that if \( N \geq N^* \) and \( \frac{i}{N} > x^* \), then

\[
\frac{Z_{ij}^+}{Z_{ij}^-} \geq s
\]

holds \( \forall j \in \{0, 1, \ldots, N - i\} \). Then

\[
\lim_{N \to \infty} A_N(x, y) = 1
\]

for all \((x, y) \in \Lambda \) with \( x > x^* \).
Proof Let \( x \in (0, 1) \) and consider the lower bound comparison birth–death process defined in Theorem 5 by (35). By Proposition 5, we know that the maximum over \( j \) of \( Z_{[N_x], j}^+ \) and the minimum over \( j \) of \( Z_{[N_x], j}^- \) occur at the same value \( \bar{j}(x) \in [0, 1, \ldots, [N_x]] \). In other words,

\[
R_{[N_x]}^{low} = \frac{a_{[N_x]}^{low}}{b_{[N_x]}^{low}} = \frac{Z_{[N_x], \bar{j}(x)}^+}{Z_{[N_x], \bar{j}(x)}^-}.
\]

Suppose now that \( x > x^* \) and take \( N \geq N^* \) and also large enough so that \([N_x]/N > x^*\). Then \( R_{[N_x]}^{low} \) is strictly greater than \( s \) for all \( x > x^* \). By Theorem 11 in Appendix A, we conclude that \( \lim_{N \to \infty} \pi_{[N_x]}^{low} = 1 \) for all \( x > x^* \). As, by Theorem 5, \( \pi_{[N_x]}^{low} \leq \alpha_{[N_x],[Ny]} \), the theorem is proved. \( \square \)

We can now proceed with proving Theorem 6.

Proof We will show that there exist \( x^*, s \) and \( N^* \) as in the hypotheses of Theorem 13. The result will then follow as a consequence of that theorem.

In fact, if strategy A is a strict Nash equilibrium, then \( F(1, 0) > G(1, 0) \) and \( F(1, 0) > H(1, 0) \), see (18), and, by continuity, we have a neighborhood of \((1, 0)\) in \( \Lambda \) in which the deterministic fitness \( F \) is strictly larger than both \( G \) and \( H \).

Let \( x_1 \) be the greatest lower bound of the values \( x \in [0, 1] \) such that \( F(x, y) > G(x, y) \) and \( F(x, y) > H(x, y) \) hold simultaneously for all \( y \) such that \( (x, y) \in \Lambda \).

In analogy with what we did in (49), we may rewrite \( Z_{ij}^+/Z_{ij}^- \) as an asymptotic term \( R(x, y) \), where \( x = i/N \) and \( y = j/N \), plus corrections that tend to 0 as \( N \to \infty \). We obtain

\[
R(x, y) = \frac{(1 - x)F(x, y)}{yG(x, y) + (1 - x - y)H(x, y)},
\]

which is continuous in \( \Lambda \setminus (1, 0) \).

Choose \( x^* \in (x_1, 1) \) and define \( \Lambda^* = \{(x, y) \in \Lambda ; x^* \leq x < 1\} \). If we define \( R^* = \inf_{(x,y)\in \Lambda^*} R(x, y) \), we claim that \( R^* > 1 \).

In fact, although \( R \) is not defined at \((1,0)\), both \( F/G \) and \( F/H \) are continuous at this point. So, we define \( S(x, y) = \min \left\{ \frac{F(x,y)}{G(x,y)}, \frac{F(x,y)}{H(x,y)} \right\} \), which is continuous in the compact set \( \bar{\Lambda}^* = \Lambda^* \cup \{(1,0)\} \). Let \( s^* \) be the minimum value of \( S \) on \( \bar{\Lambda}^* \). As \( F(x, y) > G(x, y) \) and \( F(x, y) > H(x, y) \) in \( \bar{\Lambda}^* \), then \( s^* > 1 \). Moreover, \( R(x, y) \geq S(x, y) \) for \((x,y) \in \Lambda^* \). Thus \( R^* \geq s^* \), proving our claim that \( R^* > 1 \).

We will now estimate the difference between \( Z_{ij}^+/Z_{ij}^- \) and \( R(\frac{i}{N}, \frac{j}{N}) \). Using (24), we have

\[
\frac{Z_{ij}^+}{Z_{ij}^-} = \frac{(1 - \frac{i}{N})f_{ij}}{\frac{j}{N}g_{ij} + (1 - \frac{i}{N} - \frac{j}{N})h_{ij}}.
\]

Using also the definition (50) of \( R \), we get

\[
\frac{Z_{ij}^+}{Z_{ij}^-} - R(\frac{i}{N}, \frac{j}{N})
\]

(51)

\( \square \) Springer
there exists a constant 

\[ \lim_{N \to \infty} \frac{1}{N} \sum_{i=1}^{N} G \left( \frac{i}{N}, \frac{j}{N} \right) - \frac{1}{N} \sum_{i=1}^{N} G \left( \frac{i}{N}, \frac{j}{N} \right) \]

In (52) is bounded below by 

\[ \left( \frac{1}{N} \right)^2 M_2^2 \]

By the continuity in \( \Lambda \) of \( G \) and \( H \) and by formulas analogous to (19) for \( G \) and \( H \), we know that \( M_1 \) is finite and positive.

We can also find an upper bound for the \( \phi_{i,j,N} \) in the numerator. Let \( M_2 = \max_{(x,y) \in \Lambda} \{ F(x, y), G(x, y), H(x, y) \} \). By (19) and analogous expressions, there also exists a constant \( c > 0 \) such that for all \( N \in \mathbb{N} \),

\[ \max_{(i,j) \in \Lambda_N} \{ |F(i/N, j/N) - f_{ij}|, |G(i/N, j/N) - g_{ij}|, |H(i/N, j/N) - h_{ij}| \} < \frac{c}{N} \]

Using these bounds, we get \( |\phi_{i,j,N}| < (1 - \frac{i}{N})^2 M_2 \frac{c}{N} \).

Putting together the bounds for numerator and denominator in (52), we can see that there exists a constant \( K \) such that

\[ \left| \frac{Z_{ij}^+}{Z_{ij}^-} - R \left( \frac{i}{N}, \frac{j}{N} \right) \right| < \frac{K}{N} \]

Let now \( s = \frac{1}{2} (R^* + 1) \) and \( N^* \) be the smallest integer not smaller than \( \frac{K}{R^* - s} \).

Then, for \( N > N^* \) and \( i/N > x^* \) we have \( Z_{ij}^+ / Z_{ij}^- \geq s \) for all \( j \in \{0, 1, \ldots, N - i\} \). By Theorem 13 we conclude that \( \lim_{N \to \infty} A_N(x, y) = 1 \) for all \( (x, y) \in \Lambda \) with \( x > x^* \).

For proving Theorem 9, we start by stating a preparatory result analogous to Theorem 13:

**Theorem 14** Consider a Moran process with three strategies. Suppose there exist \( s \in (0, 1) \), \( N^* \in \mathbb{N} \) and \( x^* \in [0, 1) \) such that if \( N \geq N^* \) and \( i/N > x^* \), then

\[ \frac{Z_{ij}^+}{Z_{ij}^-} \leq s \]
Fixation probabilities for the Moran process with three… 313

holds \( \forall j \in \{0, 1, \ldots, N - i\} \). Then

\[
\lim_{N \to \infty} A_{N \setminus x^*}(x, y) = 0
\]

for all \((x, y) \in \Lambda\) with \(x > x^*\).

We do not write a complete proof of Theorem 14, because it is analogous to the proof of Theorem 13, but we explain the important differences. First of all, instead of using a lower bound comparison birth–death process, we take an upper bound (36). The conclusion is a consequence of Theorem 12 in Appendix A.

For Theorem 10, we only sketch the proof, because it is again similar to preceding ones.

**Proof** We use the upper bound (36) for \(\alpha_{ij}\) in Theorem 5. In order to prove this theorem, we should show that for fixed \(x < x^*\) we have \(\pi_{\lfloor N x \rfloor}^{up} \to 0\) when \(N \to \infty\). Hypotheses \(F(x, y) < G(x, y)\) and \(F(x, y) < H(x, y)\) if \((x, y) \in \Lambda\) and \(x < x^*\) make sure that \(r_{ij}^{up} \equiv a_{ij}^{up} b_{ij} \leq s < 1\) if \(N\) is large enough and \(i < N x^*\).

The result that, for \(x < x^*, \pi_{\lfloor N x \rfloor}^{up} \to 0\) when \(N \to \infty\) may be attained in two equivalent ways. One is to develop for \(i < i^*\) a result analogous to Proposition 3 for the probability of fixation at state 0 of the comparison chain without attaining state \(i^*\). The other way is proving by Theorem 11 that the fixation probability at state \(N\) of the dual process, see (5), tends to 1 as \(N \to \infty\).

\[\square\]

**References**

Allen LJS (2011) An introduction to stochastic processes with applications to biology. Chapman & Hall/CRC, Boca Raton
Antal T, Scheuring I (2006) Fixation of strategies for an evolutionary game in finite populations. Bull Math Biol 68(8):1923–1944
Benaim M, Weibull JW (2003) Deterministic approximation of stochastic evolution in games. Econometrica 71:873–903
Bomze IM (1983) Lotka–Volterra and replicator dynamics: a two dimensional classification. Biol Cybern 48:201–211
Chalub FACC, Souza MO (2016) Fixation in large populations: a continuous view of a discrete problem. J Math Biol 72(1):283–330
Dingli D, Chalub FA, Santos FC, Van Segbroeck S, Pacheco JM (2009) Cancer phenotype as the outcome of an evolutionary game between normal and malignant cells. Br J Cancer 101(7):1130–6
Durand G, Lessard S (2016) Fixation probability in a two-locus intersexual selection model. Theor Popul Biol 109:75–87
den Hollander F (2012) Probability theory: the coupling method. http://websites.math.leidenuniv.nl/probability/lecturenotes/CouplingLectures.pdf . Accessed 8 June 2018
de Souza EP, Ferreira EM, Neves AGM (2019) Fixation probabilities for the Moran process in evolutionary games with two strategies: graph shapes and large population asymptotics. J Math Biol 78:1033–1065
Hofbauer J, Sigmund K (1998) Evolutionary games and population dynamics. Cambridge University Press, Cambridge
Maynard Smith J, Price G (1973) The logic of animal conflicts. Nature 246:15–18
McLoone B, Fan WTL, Pham A, Smead R, Loewe L (2018) Stochasticity, selection, and the evolution of cooperation in a two-level Moran model of the snowdrift game. Complexity. https://doi.org/10.1155/2018/9836150
Moran PAP (1958) Random processes in genetics. Proc Camb Philos Soc 54(1):60–71
Nowak M (2006) Evolutionary dynamics, 1st edn. The Belknap Press of Harvard University Press, Cambridge
Nowak MA, Sigmund K (1992) Tit for tat in heterogeneous populations. Nature 355:255–253
Nowak MA, Sasaki A, Taylor C, Fudenberg D (2004) Emergence of cooperation and evolutionary stability in finite populations. Nature 428(6983):646–650. https://doi.org/10.1038/nature02414
Núñez Rodríguez I, Neves AGM (2016) Evolution of cooperation in a particular case of the infinitely repeated prisoner’s dilemma with three strategies. J Math Biol 73(6):1665–1690
Petrovsky IG (1992) Lectures on partial differential equations. Dover Publications, Mineola
Taylor C, Fudenberg D, Sasaki A, Nowak MA (2004) Evolutionary game dynamics in finite populations. Bull Math Biol 66(6):1621–1644
Taylor PD, Jonker LB (1978) Evolutionary stable strategies and game dynamics. Math Biosci 40:145–156
Wang J, Fu F, Wang L, Xie G (2007) Evolutionary game dynamics with three strategies in finite populations. arXiv preprint arXiv:physics/0701315

Publisher’s Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.