

Article

Strong Migration Limit for Games in Structured Populations: Applications to Dominance Hierarchy and Set Structure

Dhaker Kroumi and Sabin Lessard *

Département de mathématiques et de statistique, Université de Montréal, Montréal, QC H3C 3J7, Canada; E-Mail: kroumidhaker@yahoo.com

* Author to whom correspondence should be addressed; E-Mail: lessards@dms.umontreal.ca; Tel.: +1-514-343-6818.

Academic Editor: Martin A. Nowak and Christian Hilbe

Received: 2 June 2015 / Accepted: 21 August 2015 / Published: 7 September 2015

Abstract: In this paper, we deduce a condition for a strategy S_1 to be more abundant on average at equilibrium under weak selection than another strategy S_2 in a population structured into a finite number of colonies of fixed proportions as the population size tends to infinity. It is assumed that one individual reproduces at a time with some probability depending on the payoff received in pairwise interactions within colonies and between colonies and that the offspring replaces one individual chosen at random in the colony into which the offspring migrates. It is shown that an expected weighted average equilibrium frequency of S_1 under weak symmetric strategy mutation between S_1 and S_2 is increased by weak selection if an expected weighted payoff of S_1 near neutrality exceeds the corresponding expected weighted payoff of S_2 . The weights are given in terms of reproductive values of individuals in the different colonies in the neutral model. This condition for S_1 to be favoured by weak selection is obtained from a strong migration limit of the genealogical process under neutrality for a sample of individuals, which is proven using a two-time scale argument. The condition is applied to games between individuals in colonies with linear or cyclic dominance and between individuals belonging to groups represented by subsets of a given set.

Keywords: abundance in frequency; dominance hierarchy; prisoner's dilemma; reproductive value; set-structured populations; structured coalescent

1. Introduction

Population structure refers to consistent demographic differences among individuals as a function of some other attribute, such as geographic location, age, size, gender or physiological state. Over the last few years, structured population models have become a central modelling formalism in theoretical biology and game dynamics, as well as being one of the most widely used.

An important question is the effect of migration on evolution. Nagylaki [1] studied the strong migration limit in a geographically-structured population, which occurs when migration dominates all other evolutionary forces in the limit of a large population, by means of a diffusion approximation. He considers a finite number of demes represented by the integers $1, \dots, d$. Deme i is composed of $N_i = N\alpha_i$ diploid individuals considered at a single multi-allelic locus, for $i = 1, \dots, d$. Here, N and α_i denote the whole population size and the proportion of deme i in the whole population, for $i = 1, \dots, d$, with $\sum_{i=1}^d \alpha_i = 1$. Time is discrete with non-overlapping generations, and the reproduction scheme in each deme follows the Wright–Fisher model as a result of random mating. Following the production of a very large number of offspring and selection among offspring, there is migration. The probability that an individual in deme i comes from deme j is represented by m_{ij}^* , for $i, j = 1, \dots, d$. The backward migration matrix $\mathbf{M}^* = (m_{ij}^*)_{1 \leq i, j \leq d}$ is assumed to be constant and ergodic. Following migration and mutation, there is random sampling within demes to restore the deme sizes. In the limit of a large population size ($N \rightarrow \infty$), the stochastic dynamics in this structured population is described by a Wright–Fisher diffusion as in a well-mixed population with an effective population size $N_e = \alpha N$ taken as the unit of time, where:

$$\alpha = \left(\sum_{i=1}^d \frac{\pi_i^2}{\alpha_i} \right)^{-1} \quad (1)$$

Here, $(\pi_i)_{1 \leq i \leq d}$ is the stationary distribution associated with the backward migration matrix \mathbf{M}^* . Under the same assumptions, but for a haploid population and in the absence of selection, Notohara [2] showed that the genealogical process, known in this context as the structured coalescent (Herbots [3]), is described in the limit of a large population size by the standard Kingman coalescent [4–6], which is such that each pair of lineages coalesces backward in time at rate one independently of all others.

Diffusion approximations and genealogical processes are very important tools to address questions related to the effect of selection on the evolution of strategies in game dynamics. Among these questions, how cooperation can emerge and persist from interactions between individuals is of prime interest. This question has attracted increasing attention from mathematical or theoretical biologists (Axelrod and Hamilton [7], Szabò and Töke [8], Traulsen *et al.* [9], Santos and Pacheco [10], Hauert and Szabò [11], Nowak [12,13], Nowak and Sigmund [14], Ohtsuki *et al.* [15], Szabò and FÁth [16]). In order to study this question, a simple game named the prisoner’s dilemma has been considered. The simplest form of this game has payoffs in additive form with the following parameters: a donor pays a cost c to a recipient to get a benefit b , where $b > c$.

The first studies on the evolution of cooperation in structured populations assumed symmetric interactions between all members of the population. This means that the payoffs depend only on the strategies used by the interacting individuals (Hamilton [17,18], Trivers [19], Frank [20], Nowak and Sigmund [21,22], Nowak [23], Traulsen and Nowak [24], Kroumi and Lessard [25]). Even in this

case, interactions can form a complex system as in the set-structured population introduced in Tarnita *et al.* [26]: every individual of the population belongs to exactly K sets among M sets, and a cooperator cooperates only with individuals that belong to at least L of the K sets, and defects otherwise.

In living groups, repetitive interactions for access to limited resources or mating opportunities can lead to the creation of a social order. A hierarchy dominance that may depend on differences in resource holding power (Hammerstein [27], Wilson [28]) can be established so that individuals are dominant over those below them and submissive to those above them. This motivates the study of asymmetric interactions with a cost for not conforming to the established hierarchy in structured populations with linear or cyclic dominance (Tao *et al.* [29], Kroumi and Lessard [30]).

In this paper, we consider a Moran-type model for games played in a population structured into d colonies of different finite sizes. In pairwise interactions, the individuals can adopt one of two strategies, S_1 or S_2 . An individual from colony i interacts with an individual from colony j with probability q_{ij} for $i, j = 1, \dots, d$, where $\sum_{j=1}^d q_{ij} = 1$ for $i = 1, \dots, d$. The expected payoff that an individual receives determines the probability for this individual to produce an offspring. One offspring is produced at a time, and following migration, this offspring replaces one individual locally chosen at random. More precisely, if the offspring is produced in colony i and migrates to colony j with probability m_{ij} , then the offspring replaces an individual chosen at random in colony j , for $i, j = 1, \dots, d$. Finally, there is mutation of the strategy used by each individual independently of all others with probability u from one time step to the next, and when this occurs, the mutant strategy is chosen at random among S_1 and S_2 . We will find a condition for S_1 to be more abundant on average than S_2 at equilibrium in the limit of a large population under weak selection and weak mutation. This result relies on the strong migration limit of the genealogical process in the absence of selection, which is proven using a lemma for two-time scale Markov chains due to Möhle [31]. The condition will be expressed in terms of expected weighted payoffs using reproductive values as weights, which is an alternative to the use of structure coefficients (Nowak *et al.* [32]) for games in structured populations. This will allow us to give an intuitive interpretation of this condition.

The remainder of this paper is organized as follows. We present the details of the model in Section 2. We use a two-time scale convergence result that is established in Appendix A to derive the limiting genealogical process in a neutral structured population in Section 3. The equilibrium state under weak selection is studied in Section 4. In Section 5, a condition for a weighted average equilibrium frequency of S_1 to increase as the selection intensity increases from zero is deduced. This condition is applied to situations with linear or cyclic dominance hierarchy in Section 6 and to games in set-structured populations in Section 7. The results are interpreted and discussed in Section 8.

2. Model

We assume a population subdivided into d colonies represented by the integers $1, \dots, d$. Each colony is made of a finite number N_i of haploid individuals, for $i = 1, \dots, d$. Each individual in the population adopts one of two strategies, S_1 or S_2 . We assume pairwise interactions between individuals within each colony and between individuals from different colonies. More precisely, an individual from colony i interacts with an individual chosen at random from colony j with probability q_{ij} , for $j = 1, \dots, d$, with

$\sum_{j=1}^d q_{ij} = 1$, for $i = 1, \dots, d$. Then, the payoff that the individual from i receives is given by the entries of the matrix:

$$\mathbf{A}_{ij} = \begin{matrix} & \begin{matrix} S_1 & S_2 \end{matrix} \\ \begin{matrix} S_1 \\ S_2 \end{matrix} & \begin{pmatrix} a_{ij} & b_{ij} \\ c_{ij} & d_{ij} \end{pmatrix} \end{matrix} \quad (2)$$

according to the strategy adopted by the individual from i , corresponding to row S_1 or S_2 , and the strategy used by the individual from j , corresponding to column S_1 or S_2 . The expected payoffs of strategies S_1 and S_2 played by individuals in colony i are denoted by $\varphi_{1,i}$ and $\varphi_{2,i}$, respectively. It is assumed that these expected payoffs translate into fertilities, or reproductive successes, in the form:

$$f_{1,i} = 1 + s \times \varphi_{1,i} \quad (3)$$

and:

$$f_{2,i} = 1 + s \times \varphi_{2,i} \quad (4)$$

respectively, where $s > 0$ represents the intensity of selection. It is assumed throughout the paper that selection is weak, actually that the intensity of selection s is small compared to the inverse of the population size N^{-1} , where:

$$N = N_1 + \dots + N_d \quad (5)$$

The case $s = 0$ corresponds to neutrality.

Time is discrete. At each time step, an individual is chosen in the whole population with probability proportional to its fertility to produce an offspring. This offspring inherits the strategy used by its parent. If the parent is from colony i , then the offspring stays in colony i with probability m_{ii} or migrates to colony $j \neq i$ with probability m_{ij} , with $\sum_{j=1}^d m_{ij} = 1$, for $i = 1, \dots, d$. In both cases, the offspring replaces an individual chosen at random in the same colony. It is assumed throughout that the forward migration matrix $\mathbf{M} = (m_{ij})_{1 \leq i,j \leq d}$ is irreducible and aperiodic, that is ergodic. In other words, there exists some power of this migration matrix, \mathbf{M}^k for some integer $k \geq 1$, with all positive entries. Finally, strategy mutation occurs with probability u for each individual independently of all others, so that the strategy used by the individual at the next time step is chosen at random among S_1 and S_2 with probability u and remains the same with the complementary probability $1 - u$.

3. Genealogical Process in the Neutral Model

In this section, we derive the genealogical process of a sample taken from the population structured into d colonies under the neutral model in the limit of a large population size. Every individual of the population has the same fertility, which is given by one.

Consider a sample of size $n \geq 2$ at a given time step. Looking backward in time at the genealogy of this sample, the distribution of the ancestors in the d colonies at any previous time step can be described by a vector:

$$\mathbf{n} = (n_1, \dots, n_d) \quad (6)$$

where n_i denotes the number of ancestors in colony i , for $i = 1, \dots, d$. Then:

$$|\mathbf{n}| = n_1 + \dots + n_d \quad (7)$$

is the total number of ancestors.

Let $\mathbf{n}(\tau)$ be the distribution of the ancestors $\tau \geq 0$ time steps back. Given an initial sample of n individuals, this is a discrete-time Markov chain with state space:

$$E_n = \{\mathbf{n} = (n_1, \dots, n_d) : n_1 + \dots + n_d \leq n\} \tag{8}$$

Let $\Pi_{\mathbf{n},\mathbf{n}'}$ be the transition probability from state \mathbf{n} to state \mathbf{n}' . A possible transition is from \mathbf{n} to $\mathbf{n} - \mathbf{e}_i$, for $i \in \{1, \dots, d\}$, such that $n_i \geq 1$. Here, \mathbf{e}_i is a d -dimensional unit vector with the i -th component equal to one and all other components equal to zero. This transition is obtained if one of the n_i ancestors in colony i produced an offspring who stayed in colony i and is one of the $n_i - 1$ other ancestors in this colony, or if one of the n_j ancestors in colony $j \neq i$ produced an offspring who migrated to colony i and is one of the n_i ancestors in this colony. All this occurs with probability:

$$\Pi_{\mathbf{n},\mathbf{n}-\mathbf{e}_i} = \frac{n_i}{N} m_{ii} \frac{(n_i - 1)}{N_i} + \sum_{j \neq i} \frac{n_j}{N} m_{ji} \frac{n_i}{N_i} = \left(\frac{n_i(n_i - 1)}{\alpha_i} m_{ii} + \sum_{j \neq i} \frac{n_i n_j}{\alpha_i} m_{ji} \right) \times \frac{1}{N^2} \tag{9}$$

Here, $\alpha_i = N_i/N$ is the proportion of colony i in the whole population, for $i = 1, \dots, d$. Another possible transition is from \mathbf{n} to $\mathbf{n} - \mathbf{e}_i + \mathbf{e}_j$, for $i, j \in \{1, \dots, d\}$, such that $n_i \geq 1$ and $i \neq j$. This occurs if an individual in colony j other than the n_j ancestors in this colony produced an offspring who migrated to colony i and is one of the n_i ancestors in this colony. The probability of this event is:

$$\Pi_{\mathbf{n},\mathbf{n}-\mathbf{e}_i+\mathbf{e}_j} = \frac{(N_j - n_j)}{N} m_{ji} \frac{n_i}{N_i} = \frac{n_i \alpha_j}{\alpha_i} m_{ji} \times \frac{1}{N} - \frac{n_i n_j}{\alpha_i} m_{ji} \times \frac{1}{N^2} \tag{10}$$

The last transition with positive probability is to stay in the same state, for which we have:

$$\begin{aligned} \Pi_{\mathbf{n},\mathbf{n}} &= 1 - \sum_{i:n_i \geq 1} \Pi_{\mathbf{n},\mathbf{n}-\mathbf{e}_i} - \sum_{\substack{i \neq j \\ n_i \geq 1}} \Pi_{\mathbf{n},\mathbf{n}-\mathbf{e}_i+\mathbf{e}_j} \\ &= 1 - \left(\sum_{\substack{i \neq j \\ n_i \geq 1}} \frac{n_i \alpha_j}{\alpha_i} m_{ji} \right) \times \frac{1}{N} - \left(\sum_{i:n_i \geq 1} \frac{n_i(n_i - 1)}{\alpha_i} m_{ii} \right) \times \frac{1}{N^2} \end{aligned} \tag{11}$$

From Equations (9)–(11), the transition matrix $\Pi_N = (\Pi_{\mathbf{n},\mathbf{n}'})_{\mathbf{n},\mathbf{n}' \in E_n}$ can be decomposed into the form:

$$\Pi_N = \mathbf{I} + \frac{\mathbf{A}}{N} + \frac{\mathbf{B}}{N^2} \tag{12}$$

Here, \mathbf{I} is an identity matrix of a size given by the number of elements in the state space E_n , and $\mathbf{A} = (a_{\mathbf{n},\mathbf{n}'})_{\mathbf{n},\mathbf{n}' \in E_n}$ is an infinitesimal generator whose non-null entries are given by:

$$a_{\mathbf{n},\mathbf{n}'} = \begin{cases} \frac{n_i \alpha_j}{\alpha_i} m_{ji} & \text{if } \mathbf{n}' = \mathbf{n} - \mathbf{e}_i + \mathbf{e}_j, \text{ for } i \neq j \text{ and } n_i \geq 1 \\ - \sum_{\substack{i \neq j \\ n_i \geq 1}} \frac{n_i \alpha_j}{\alpha_i} m_{ji} & \text{if } \mathbf{n}' = \mathbf{n} \end{cases} \tag{13}$$

Moreover, the non-null entries of $\mathbf{B} = (b_{\mathbf{n},\mathbf{n}'})_{\mathbf{n},\mathbf{n}' \in E_n}$ are given by:

$$b_{\mathbf{n},\mathbf{n}'} = \begin{cases} \frac{n_i(n_i-1)}{\alpha_i} m_{ii} + \sum_{j \neq i} \frac{n_i n_j}{\alpha_i} m_{ji} & \text{if } \mathbf{n}' = \mathbf{n} - \mathbf{e}_i, \text{ for } i \text{ such that } n_i \geq 1 \\ - \frac{n_i n_j}{\alpha_i} m_{ji} & \text{if } \mathbf{n}' = \mathbf{n} - \mathbf{e}_i + \mathbf{e}_j, \text{ for } i \neq j \text{ and } n_i \geq 1 \\ - \sum_{i:n_i \geq 1} \frac{n_i(n_i-1)}{\alpha_i} m_{ii} & \text{if } \mathbf{n}' = \mathbf{n} \end{cases} \tag{14}$$

Now, let Γ_k be the subset of all possible states with k ancestors, namely:

$$\Gamma_k = \{\mathbf{n} = (n_1, \dots, n_d) : |\mathbf{n}| = n_1 + \dots + n_d = k\} \tag{15}$$

for $k = 1, \dots, n$. Note that the set E_n is the disjoint union of the subsets $\Gamma_1, \Gamma_2, \dots, \Gamma_n$. With respect to these subsets in this order, the matrices \mathbf{A} and \mathbf{B} whose non-null entries $a_{\mathbf{n},\mathbf{n}'}$ and $b_{\mathbf{n},\mathbf{n}'}$ are given by Equations (13) and (14), respectively, can be expressed in the block forms:

$$\mathbf{A} = \begin{pmatrix} \mathbf{A}_1 & \mathbf{0} & \dots & \mathbf{0} \\ \mathbf{0} & \mathbf{A}_2 & \dots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \dots & \mathbf{A}_n \end{pmatrix} \tag{16}$$

and:

$$\mathbf{B} = \begin{pmatrix} \mathbf{0} & \mathbf{0} & \mathbf{0} & \dots & \mathbf{0} & \mathbf{0} \\ \mathbf{B}_{2,1} & \mathbf{B}_{2,2} & \mathbf{0} & \dots & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{B}_{3,2} & \mathbf{B}_{3,3} & \dots & \mathbf{0} & \mathbf{0} \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \dots & \mathbf{B}_{n,n-1} & \mathbf{B}_{n,n} \end{pmatrix} \tag{17}$$

Here, $\mathbf{0}$ denotes a zero matrix of any dimension.

The exponential matrix $\exp\{m\mathbf{A}\}$ for every integer $m \geq 1$ takes the block form:

$$\exp\{m\mathbf{A}\} = \begin{pmatrix} \exp\{m\mathbf{A}_1\} & \mathbf{0} & \dots & \mathbf{0} \\ \mathbf{0} & \exp\{m\mathbf{A}_2\} & \dots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \dots & \exp\{m\mathbf{A}_n\} \end{pmatrix} \tag{18}$$

Note that \mathbf{A}_k is the infinitesimal generator of an irreducible Markov chain on a finite state space, which is Γ_k , for $k = 1, \dots, n$. Consequently, the limit matrix:

$$\mathbf{P}_k := \lim_{m \rightarrow \infty} \exp\{m\mathbf{A}_k\} \tag{19}$$

exists for $k = 1, \dots, n$, and so does:

$$\mathbf{P} := \lim_{m \rightarrow \infty} \exp\{m\mathbf{A}\} = \begin{pmatrix} \mathbf{P}_1 & \mathbf{0} & \dots & \mathbf{0} \\ \mathbf{0} & \mathbf{P}_2 & \dots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \dots & \mathbf{P}_n \end{pmatrix} \tag{20}$$

Note also that \mathbf{P}_k is actually a rank one matrix with the stationary distribution associated with \mathbf{A}_k in every row, for $k = 1, \dots, n$. It remains to find this distribution.

Let (π_1, \dots, π_d) be the stationary distribution associated with \mathbf{A}_1 . By definition, this distribution satisfies the equation:

$$(\pi_1, \dots, \pi_d)\mathbf{A}_1 = (0, \dots, 0) \tag{21}$$

with:

$$\pi_1 + \dots + \pi_d = 1 \tag{22}$$

Note that $\mathbf{A}_1 = N(\mathbf{M}^* - \mathbf{I})$, where $\mathbf{M}^* = (m_{ij}^*)_{1 \leq i, j \leq d}$ is the backward migration matrix, whose entries are given by:

$$m_{ij}^* = \begin{cases} \frac{m_{ji}N_j}{NN_i} & \text{if } i \neq j \\ 1 - \sum_{t \neq i} \frac{m_{ti}N_t}{NN_i} & \text{if } i = j \end{cases} \tag{23}$$

This is obtained from Equation (13) in the case where $\mathbf{n} = \mathbf{e}_i$ for $i = 1, \dots, d$. The entry m_{ij}^* is the probability that an individual chosen at random in colony i comes from colony j one time step back. Moreover,

$$(\pi_1, \dots, \pi_d)\mathbf{M}^* = (\pi_1, \dots, \pi_d) \tag{24}$$

which means that (π_1, \dots, π_d) is the stationary distribution associated with the backward migration matrix \mathbf{M}^* . The stationary distribution associated with \mathbf{A}_k , for $k = 2, \dots, n$, can be expressed with respect to this distribution. For $\mathbf{n} = (n_1, \dots, n_d) \in \Gamma_k$, we have:

$$\pi_{\mathbf{n}} = \binom{|\mathbf{n}|}{n_1, n_2, \dots, n_d} \pi_1^{n_1} \times \dots \times \pi_d^{n_d} \tag{25}$$

where:

$$\binom{|\mathbf{n}|}{n_1, n_2, \dots, n_d} = \frac{|\mathbf{n}|!}{n_1! \times n_2! \times \dots \times n_d!} \tag{26}$$

The proof of Equation (25) is given in Appendix B.

By definition of the stationary distribution $(\pi_{\mathbf{n}})_{\mathbf{n} \in \Gamma_k}$, we have:

$$\mathbf{P}_k \mathbf{A}_k = \mathbf{0} \tag{27}$$

for $k = 1, \dots, n$, from which:

$$\mathbf{P}\mathbf{A} = \mathbf{A}\mathbf{P} = \begin{pmatrix} \mathbf{P}_1 \mathbf{A}_1 & \mathbf{0} & \dots & \mathbf{0} \\ \mathbf{0} & \mathbf{P}_2 \mathbf{A}_2 & \dots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \dots & \mathbf{P}_n \mathbf{A}_n \end{pmatrix} = \mathbf{0} \tag{28}$$

Owing to Proposition 1 in Appendix A, we conclude that:

$$\lim_{N \rightarrow \infty} (\mathbf{\Pi}_N)^{\lfloor tN^2 \rfloor} = \lim_{N \rightarrow \infty} \left(\mathbf{I} + \frac{\mathbf{A}}{N} + \frac{\mathbf{B}}{N^2} \right)^{\lfloor tN^2 \rfloor} = \mathbf{P} \exp\{t\mathbf{G}\} \tag{29}$$

where $\mathbf{G} = \mathbf{P}\mathbf{B}\mathbf{P}$ and $\lfloor tN^2 \rfloor$ denotes the floor value of tN^2 , which is defined as the greatest integer less or equal to the real number tN^2 . Using the block forms of \mathbf{B} and \mathbf{P} given in Equations (17) and (20), respectively, we have:

$$\mathbf{G} = \begin{pmatrix} \mathbf{0} & \mathbf{0} & \mathbf{0} & \dots & \mathbf{0} & \mathbf{0} \\ \mathbf{P}_2 \mathbf{B}_{2,1} \mathbf{P}_1 & \mathbf{P}_2 \mathbf{B}_{2,2} \mathbf{P}_2 & \mathbf{0} & \dots & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{P}_3 \mathbf{B}_{3,2} \mathbf{P}_2 & \mathbf{P}_3 \mathbf{B}_{3,3} \mathbf{P}_3 & \dots & \mathbf{0} & \mathbf{0} \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \dots & \mathbf{P}_n \mathbf{B}_{n,n-1} \mathbf{P}_{n-1} & \mathbf{P}_n \mathbf{B}_{n,n} \mathbf{P}_n \end{pmatrix} \tag{30}$$

The non-null entries of $\mathbf{G} = (G_{\mathbf{n},\mathbf{n}'})_{\mathbf{n},\mathbf{n}' \in E_n}$ are given by:

$$G_{\mathbf{n},\mathbf{n}'} = \begin{cases} \pi_{\mathbf{n}'} \sum_{\mathbf{x},\mathbf{y} \in \Gamma_k} \pi_{\mathbf{x}} b_{\mathbf{x},\mathbf{y}} & \text{if } \mathbf{n}' \in \Gamma_k \\ \pi_{\mathbf{n}'} \sum_{\mathbf{x} \in \Gamma_k, \mathbf{y} \in \Gamma_{k-1}} \pi_{\mathbf{x}} b_{\mathbf{x},\mathbf{y}} & \text{if } \mathbf{n}' \in \Gamma_{k-1} \end{cases} \tag{31}$$

for $\mathbf{n} \in \Gamma_k$ for $k = 2, \dots, n$. We have:

$$\begin{aligned} \sum_{\mathbf{x} \in \Gamma_k, \mathbf{y} \in \Gamma_{k-1}} \pi_{\mathbf{x}} b_{\mathbf{x},\mathbf{y}} &= \sum_{x_1 + \dots + x_d = k} \sum_{i=1}^d \binom{k}{x_1, \dots, x_d} \pi_1^{x_1} \dots \pi_d^{x_d} \times \frac{x_i(x_i - 1)m_{ii}}{\alpha_i} \\ &+ \sum_{x_1 + \dots + x_d = k} \sum_{i \neq j} \binom{k}{x_1, \dots, x_d} \pi_1^{x_1} \dots \pi_d^{x_d} \times \frac{x_i x_j m_{ji}}{\alpha_i} \end{aligned} \tag{32}$$

The first term on the right-hand side of Equation (32) is:

$$\begin{aligned} \sum_{i=1}^d \frac{k(k-1)m_{ii}}{\alpha_i} \pi_i^2 \sum_{\substack{x_1 + \dots + x_d = k \\ x_i \geq 2}} \binom{k-2}{x_1, \dots, x_i - 2, \dots, x_d} \pi_1^{x_1} \dots \pi_i^{x_i - 2} \dots \pi_d^{x_d} \\ = \sum_{i=1}^d \frac{k(k-1)m_{ii}}{\alpha_i} \pi_i^2 (\pi_1 + \dots + \pi_d)^{k-2} \\ = \sum_{i=1}^d \frac{k(k-1)m_{ii}}{\alpha_i} \pi_i^2 \end{aligned} \tag{33}$$

while the second term is:

$$\begin{aligned} \sum_{i \neq j} \frac{k(k-1)m_{ji}}{\alpha_i} \pi_i \pi_j \sum_{\substack{x_1 + \dots + x_d = k \\ x_i \geq 1, x_j \geq 1}} \binom{k-2}{x_1, \dots, x_i - 1, \dots, x_j - 1, \dots, x_d} \pi_1^{x_1} \dots \pi_i^{x_i - 1} \dots \pi_j^{x_j - 1} \dots \pi_d^{x_d} \\ = \sum_{i \neq j} \frac{k(k-1)m_{ji}}{\alpha_i} \pi_i \pi_j (\pi_1 + \dots + \pi_d)^{k-2} \\ = \sum_{i \neq j} \frac{k(k-1)m_{ji}}{\alpha_i} \pi_i \pi_j \end{aligned} \tag{34}$$

These expressions into Equation (32) give:

$$\sum_{\mathbf{x} \in \Gamma_k, \mathbf{y} \in \Gamma_{k-1}} \pi_{\mathbf{x}} b_{\mathbf{x},\mathbf{y}} = k(k-1)\lambda \tag{35}$$

where:

$$\lambda = \sum_{i=1}^d \frac{m_{ii}}{\alpha_i} \pi_i^2 + \sum_{i \neq j} \frac{m_{ji}}{\alpha_i} \pi_i \pi_j \tag{36}$$

Finally, using the fact that:

$$\sum_{\mathbf{y} \in \Gamma_{k-1}} b_{\mathbf{x},\mathbf{y}} + \sum_{\mathbf{y} \in \Gamma_k} b_{\mathbf{x},\mathbf{y}} = 0 \tag{37}$$

the non-null entries of \mathbf{G} given in Equation (32) take the form:

$$G_{\mathbf{n},\mathbf{n}'} = \begin{cases} -k(k-1)\pi_{\mathbf{n}'}\lambda & \text{if } \mathbf{n}' \in \Gamma_k \\ k(k-1)\pi_{\mathbf{n}'}\lambda & \text{if } \mathbf{n}' \in \Gamma_{k-1} \end{cases} \tag{38}$$

for $\mathbf{n} \in \Gamma_k$ for $k = 2, \dots, n$. Note that $G_{\mathbf{n},\mathbf{n}'} = 0$ for $\mathbf{n} \in \Gamma_1$ and $\mathbf{n}' \in E_n$. In summary, we have the following result in the limit of a large population size.

Theorem 1. *The strong migration limit of the genealogical process in a structured population, taking $N^2/(2\lambda)$ time steps as the unit of time as the population size $N \rightarrow \infty$, is given by:*

$$\lim_{N \rightarrow \infty} (\Pi_N)^{\lfloor (2\lambda)^{-1}tN^2 \rfloor} = \mathbf{P} \exp\{t\mathbf{H}\} := \Pi(t) \tag{39}$$

for $t > 0$, where $\mathbf{H} = (2\lambda)^{-1}\mathbf{G} = (h_{\mathbf{n},\mathbf{n}'})_{\mathbf{n},\mathbf{n}' \in E_n}$ has non-null entries given by:

$$h_{\mathbf{n},\mathbf{n}'} = \begin{cases} -\frac{k(k-1)}{2}\pi_{\mathbf{n}'} & \text{if } \mathbf{n}' \in \Gamma_k \\ \frac{k(k-1)}{2}\pi_{\mathbf{n}'} & \text{if } \mathbf{n}' \in \Gamma_{k-1} \end{cases} \tag{40}$$

for $\mathbf{n} \in \Gamma_k$, for $k = 2, \dots, n$.

Remark 1. Equation (39) means that, in the limit as $N \rightarrow \infty$ for $k = 2, \dots, n$, the ancestors are in state $\mathbf{n} \in \Gamma_k$ with probability $\pi_{\mathbf{n}}$ as long as their number is k , while this number decreases by one at rate $k(k-1)/2$ per $N^2/(2\lambda)$ time steps, for $k = 2, \dots, n$. In other words, after a scaled time of exponential distribution with parameter $k(k-1)/2$, the number of ancestors jumps from k to $k-1$, and these ancestors are found in state \mathbf{n}' , with probability $\pi_{\mathbf{n}'}$ for $\mathbf{n}' \in \Gamma_{k-1}$. Note that the number of ancestors is described by the standard Kingman coalescent in a well-mixed population (see Kingman [4–6]).

Remark 2. The limiting process for the number of ancestors in the structured population of size N corresponds to the limiting process for the Moran model (see Moran [33,34]) in a well-mixed population of size $N_e = N/\sqrt{\lambda}$ with $\frac{N_e^2}{2} = \frac{N^2}{2\lambda}$ time steps as the unit of time as $N \rightarrow \infty$, where λ is given in Equation (36). The parameter λ is a measure of mixing, and N_e is an effective population size that takes into account the population structure.

4. Equilibrium State

Suppose without loss of generality that the individuals in the population occupy ordered sites, such that the sites of colony 1 come first, then the sites of colony 2 come second, and so on, up to the sites of colony d . The state of the population at a given time step is represented by the N -dimensional vector $\delta = (\delta_1, \dots, \delta_N)$, where:

$$\delta_l = \begin{cases} 1 & \text{if } S_1 \text{ at site } l \\ 0 & \text{if } S_2 \text{ at site } l \end{cases} \tag{41}$$

for $l = 1, \dots, N$. With q_{ij} being the probability for an individual from colony i to interact with an individual chosen at random from colony j , for $i, j = 1, \dots, d$, and assuming that an individual can interact with itself, the expected payoffs of strategies S_1 and S_2 in colony i for $i = 1, \dots, d$ are given by:

$$\varphi_{1,i} = \sum_{j=1}^d q_{ij}(a_{ij}x_j + b_{ij}(1-x_j)) \tag{42}$$

and:

$$\varphi_{2,i} = \sum_{j=1}^d q_{ij}(c_{ij}x_j + d_{ij}(1 - x_j)) \tag{43}$$

respectively, where:

$$x_j = \frac{k_j}{N_j} = \frac{1}{N_j} \sum_{l=N_1+\dots+N_{j-1}+1}^{N_1+\dots+N_j} \delta_l \tag{44}$$

is the frequency of S_1 in colony j , for $j = 1, \dots, d$ (with the convention that $N_0 = -N_1$).

An offspring is produced according to the corresponding fertilities $f_{1,i}$ and $f_{2,i}$ given in Equations (3) and (4). This offspring migrates from colony i to colony j with probability m_{ij} and replaces an individual chosen at random in colony j , for $j = 1, \dots, d$. Moreover, the strategy of each individual mutates into a strategy chosen at random among S_1 and S_2 with probability u , and this occurs independently for all individuals. Then, the conditional expected value of the new state of the population takes the form:

$$\mathbb{E}[\delta'|\delta] = (1 - u)\mathbf{R}(s)\delta + \frac{u}{2}\mathbf{u} \tag{45}$$

where \mathbf{u} is an all-ones N -dimensional vector, while:

$$\mathbf{R}(s) = (r_{ij}(s)\mathbf{U}_{ij})_{1 \leq i,j \leq d} + \mathbf{I} - \mathbf{D}(\mathbf{r}(s)) \tag{46}$$

where $\mathbf{D}(\mathbf{r}(s))$ is a diagonal matrix with the vector:

$$\mathbf{r}(s) = (r_i(s)\mathbf{u}_i)_{1 \leq i \leq d} \tag{47}$$

on the main diagonal. Here, \mathbf{U}_{ij} denotes an all-ones $N_i \times N_j$ -matrix and \mathbf{u}_i an all-ones N_i -dimensional vector. Moreover,

$$r_{ij}(s) = \frac{f_{1,j}m_{ji}}{N\bar{f}N_i} \tag{48}$$

is the probability for a given individual in colony i to be replaced by an offspring produced by an individual playing S_1 in colony j , if there is any, while:

$$r_i(s) = \sum_{j=1}^d \frac{N_j\bar{f}_j m_{ji}}{N\bar{f}N_i} \tag{49}$$

is the total probability for a given individual in colony i to be replaced by an offspring.

Note that:

$$\bar{f}_j = x_j f_{1,j} + (1 - x_j) f_{2,j} = 1 + s \times \bar{\varphi}_j \tag{50}$$

and:

$$\bar{f} = \sum_{j=1}^d \alpha_j (x_j f_{1,j} + (1 - x_j) f_{2,j}) = 1 + s \times \bar{\varphi} \tag{51}$$

are the average fertilities in colony j and in the whole population, respectively, where:

$$\bar{\varphi}_j = x_j \varphi_{1,j} + (1 - x_j) \varphi_{2,j} \tag{52}$$

and:

$$\bar{\varphi} = \sum_{j=1}^d \alpha_j (x_j \varphi_{1,j} + (1 - x_j) \varphi_{2,j}) \tag{53}$$

are the corresponding average expected payoffs. Therefore,

$$r_{ij}(s) = r_{ij}(0) + sr_{ij}(0)(\varphi_{1,j} - \bar{\varphi}) + o(s) \tag{54}$$

and:

$$r_i(s) = r_i(0) + s \sum_{j=1}^d \frac{N_j m_{ji}}{N N_i} (\bar{\varphi}_j - \bar{\varphi}) + o(s) \tag{55}$$

with:

$$r_{ij}(0) = \frac{m_{ji}}{N N_i} \tag{56}$$

and:

$$r_i(0) = \sum_{j=1}^d \frac{N_j m_{ji}}{N N_i} \tag{57}$$

Note that the stochastic matrix:

$$\mathbf{R}(0) = (r_{ij}(0) \mathbf{U}_{ij})_{1 \leq i, j \leq d} + \mathbf{I} - \mathbf{D}(\mathbf{r}(0)) \tag{58}$$

has a stationary distribution given by:

$$\mathbf{v}(0)^T = \left(\frac{\pi_1}{N_1} \mathbf{u}_1^T, \dots, \frac{\pi_d}{N_d} \mathbf{u}_d^T \right) \tag{59}$$

with T for transpose, where (π_1, \dots, π_d) is the stationary distribution of the backward migration matrix \mathbf{M}^* defined in Equation (23). Moreover,

$$\mathbf{R}(s) = \mathbf{R}(0) + s \frac{d\mathbf{R}}{ds}(0) + o(s) \tag{60}$$

and:

$$\mathbf{r}(s) = \mathbf{r}(0) + s \frac{d\mathbf{r}}{ds}(0) + o(s) \tag{61}$$

where $o(s)$ stands here for a matrix or a vector whose entries or components are functions little- o with respect to s as $s \rightarrow 0$.

At equilibrium, we have:

$$\mathbb{E}[\boldsymbol{\delta}'] = \mathbb{E}[\mathbb{E}[\boldsymbol{\delta}' | \boldsymbol{\delta}]] = (1 - u) \mathbb{E}[\mathbf{R}(s) \boldsymbol{\delta}] + \frac{u}{2} \mathbf{u} = \mathbb{E}[\boldsymbol{\delta}] \tag{62}$$

The scalar product with $\mathbf{v}(0)$ yields:

$$\langle \mathbf{v}(0), \mathbb{E}[\boldsymbol{\delta}] \rangle = (1 - u) \langle \mathbf{v}(0), \mathbb{E}[\mathbf{R}(s) \boldsymbol{\delta}] \rangle + \frac{u}{2} \tag{63}$$

Note that this equilibrium equation in the neutral case ($s = 0$ with \mathbb{E}_0 denoting expectation under this condition) gives:

$$\begin{aligned} \langle \mathbf{v}(0), \mathbb{E}_0[\boldsymbol{\delta}] \rangle &= (1 - u) \langle \mathbf{v}(0), \mathbf{R}(0) \mathbb{E}_0[\boldsymbol{\delta}] \rangle + \frac{u}{2} \\ &= (1 - u) \langle \mathbf{v}(0) \mathbf{R}(0), \mathbb{E}_0[\boldsymbol{\delta}] \rangle + \frac{u}{2} \\ &= (1 - u) \langle \mathbf{v}(0), \mathbb{E}_0[\boldsymbol{\delta}] \rangle + \frac{u}{2} \end{aligned} \tag{64}$$

from which:

$$\langle \mathbf{v}(0), \mathbb{E}_0[\boldsymbol{\delta}] \rangle = \frac{1}{2} \tag{65}$$

Under weak selection, it follows from Equations (60) and (63) that:

$$\begin{aligned} \langle \mathbf{v}(0), \mathbb{E}[\boldsymbol{\delta}] \rangle &= (1 - u) \left(\langle \mathbf{v}(0), \mathbf{R}(0) \mathbb{E}[\boldsymbol{\delta}] \rangle + s \langle \mathbf{v}(0), \mathbb{E} \left[\frac{d\mathbf{R}}{ds}(0) \boldsymbol{\delta} \right] \rangle \right) + \frac{u}{2} + o(s) \\ &= (1 - u) \left(\langle \mathbf{v}(0) \mathbf{R}(0), \mathbb{E}[\boldsymbol{\delta}] \rangle + s \langle \mathbf{v}(0), \mathbb{E} \left[\frac{d\mathbf{R}}{ds}(0) \boldsymbol{\delta} \right] \rangle \right) + \frac{u}{2} + o(s) \\ &= (1 - u) \left(\langle \mathbf{v}(0), \mathbb{E}[\boldsymbol{\delta}] \rangle + s \langle \mathbf{v}(0), \mathbb{E} \left[\frac{d\mathbf{R}}{ds}(0) \boldsymbol{\delta} \right] \rangle \right) + \frac{u}{2} + o(s) \end{aligned} \tag{66}$$

This equation gives an approximation in the case of weak selection, that is for $s > 0$ small enough.

5. Condition for Weak Selection to Favour a Strategy over Another

In this section, we will prove the main result below.

Theorem 2. *Strategy S_1 is favoured by weak selection in the strong migration limit of a structured population with payoff matrices given by (2) for strategies S_1 and S_2 under weak mutation, in the sense that S_1 is more abundant than S_2 in expected weighted average equilibrium frequency for a weak enough intensity of selection, if:*

$$\sum_{i,j,k=1}^d \pi_i \frac{\alpha_j}{\alpha_i} m_{ji} q_{jk} (a_{jk} + b_{jk} - c_{jk} - d_{jk}) > 0 \tag{67}$$

with α_i being the proportion of individuals in colony i , m_{ji} the probability that an offspring migrates from colony j to colony i , π_i the stationary proportion of ancestors of an individual that are in colony i in the neutral model and q_{jk} the probability for an individual in colony j to interact with an individual in colony k .

Proof. The equilibrium Equation (66) implies that:

$$\begin{aligned} \langle \mathbf{v}(0), \mathbb{E}[\boldsymbol{\delta}] \rangle &= \frac{1}{2} + \frac{s(1 - u)}{u} \langle \mathbf{v}(0), \mathbb{E} \left[\frac{d\mathbf{R}}{ds}(0) \boldsymbol{\delta} \right] \rangle + \frac{o(s)}{u} \\ &= \frac{1}{2} + \frac{s(1 - u)}{u} \langle \mathbf{v}(0), \mathbb{E}_0 \left[\frac{d\mathbf{R}}{ds}(0) \boldsymbol{\delta} \right] \rangle + \frac{o(s)}{u} \end{aligned} \tag{68}$$

Note that:

$$\langle \mathbf{v}(0), \mathbb{E}[\boldsymbol{\delta}] \rangle = \sum_{i=1}^d \pi_i \mathbb{E}[x_i] \tag{69}$$

is a weighted expected frequency of strategy S_1 at equilibrium. If the intensity of selection $s > 0$ is small enough, then:

$$\langle \mathbf{v}(0), \mathbb{E}[\boldsymbol{\delta}] \rangle > \langle \mathbf{v}(0), \mathbb{E}_0[\boldsymbol{\delta}] \rangle = \frac{1}{2} \tag{70}$$

if:

$$\langle \mathbf{v}(0), \mathbb{E}_0\left[\frac{d\mathbf{R}}{ds}(0)\boldsymbol{\delta}\right] \rangle > 0 \tag{71}$$

This is a condition for weak selection to favour S_1 in the sense that strategy S_1 is more abundant in weighted average frequency at equilibrium than strategy S_2 . If the inequality is reversed, then weak selection favours S_2 in the same sense.

With the assumptions of the model, the vector $\mathbf{v}(0)$ takes the expression given in Equation (59) and:

$$\frac{d\mathbf{R}}{ds}(0) = \left(\frac{m_{ji}}{NN_i} (\varphi_{1,j} - \bar{\varphi}) \mathbf{U}_{ij} \right)_{1 \leq i, j \leq d} - \mathbf{I} \left(\mathbf{u}_i \sum_{j=1}^d \frac{N_j m_{ji}}{NN_i} (\bar{\varphi}_j - \bar{\varphi}) \right)_{1 \leq i \leq d} \tag{72}$$

where $\bar{\varphi}_j$ and $\bar{\varphi}$ are given in Equations (52) and (53), respectively. Therefore, we have:

$$\begin{aligned} \langle \mathbf{v}(0), \mathbb{E}_0\left[\frac{d\mathbf{R}}{ds}(0)\boldsymbol{\delta}\right] \rangle &= \sum_{i,j=1}^d \pi_i \frac{N_j m_{ji}}{NN_i} \mathbb{E}_0 [\varphi_{1,j} x_j (1 - x_i) - \varphi_{2,j} x_i (1 - x_j)] \\ &+ \sum_{i,j=1}^d \pi_i \frac{N_j m_{ji}}{NN_i} \mathbb{E}_0 [x_i \bar{\varphi}] - \sum_{i,j=1}^d \pi_i \frac{N_j m_{ji}}{NN_i} \mathbb{E}_0 [x_j \bar{\varphi}] \end{aligned} \tag{73}$$

Moreover, Equation (24) entails that:

$$\begin{aligned} \sum_{i,j=1}^d \pi_i \frac{N_j m_{ji}}{NN_i} \mathbb{E}_0 [x_j \bar{\varphi}] &= \sum_{j=1}^d \left(\sum_{i=1}^d \pi_i \frac{N_j m_{ji}}{NN_i} \right) \mathbb{E}_0 [x_j \bar{\varphi}] \\ &= \sum_{j=1}^d \left(\sum_{i=1}^d \pi_j \frac{N_i m_{ij}}{NN_j} \right) \mathbb{E}_0 [x_j \bar{\varphi}] \\ &= \sum_{i,j=1}^d \pi_i \frac{N_j m_{ji}}{NN_i} \mathbb{E}_0 [x_i \bar{\varphi}] \end{aligned} \tag{74}$$

with the last equality obtained by permuting the indices i and j . Therefore, the condition for weak selection to favour S_1 becomes:

$$\sum_{i,j=1}^d \pi_i \frac{N_j m_{ji}}{NN_i} \mathbb{E}_0 [\varphi_{1,j} x_j (1 - x_i) - \varphi_{2,j} x_i (1 - x_j)] > 0 \tag{75}$$

Using the expressions given in Equations (42) and (43) for the expected payoffs of S_1 and S_2 in colony j , we find that:

$$\mathbb{E}_0 [\varphi_{1,j} x_j (1 - x_i) - \varphi_{2,j} (1 - x_j) x_i] = \sum_{k=1}^d q_{jk} \times \tag{76}$$

$$(a_{j,k} \mathbb{E}_0 [x_k x_j (1 - x_i)] + b_{j,k} \mathbb{E}_0 [(1 - x_k) x_j (1 - x_i)] - c_{j,k} \mathbb{E}_0 [x_k (1 - x_j) x_i] - d_{j,k} \mathbb{E}_0 [(1 - x_k) (1 - x_j) x_i])$$

In the neutral model, a permutation of strategies S_1 and S_2 does not change the expected value of a product of their equilibrium frequencies. Consequently, we have:

$$\mathbb{E}_0[x_k x_j (1 - x_i)] = \mathbb{E}_0[(1 - x_k)(1 - x_j)x_i] \quad (77)$$

and:

$$\mathbb{E}_0[(1 - x_k)x_j(1 - x_i)] = \mathbb{E}_0[x_k(1 - x_j)x_i] \quad (78)$$

Moreover, in the strong migration limit with $N^2/(2\lambda)$ time steps as the unit of time as $N \rightarrow \infty$ and under weak mutation, so that $\theta = uN^2/(2\lambda)$ remains constant, the above expected values are all equal to the probability that exactly two given individuals out of three, irrespective of the colonies that they are in, use the same strategy (see Appendix C). Ignoring common positive factors and writing N_j/N_i as α_j/α_i , we get the condition given in Equation (67) for weak selection to favour S_1 , which completes the proof. \square

6. Application to Cooperation with Dominance Hierarchy

Dominance hierarchy in animals is a form of social structure in which a linear or nearly linear ranking exists, with each animal dominant over those below it and submissive to those above it in the hierarchy. In this section, we study the evolution of cooperation in two cases of dominance hierarchy. The first case is linear dominance, which involves global interactions. Colonies form a complete graph. An individual from any colony i can interact with an individual from any colony j ($q_{ij} > 0$ for $i, j = 1, \dots, d$). The second case is cyclic dominance, which assumes only local interactions. Colonies are distributed over a one-dimensional cycle with each colony connected to the two nearest-neighbour colonies. An individual from colony i can interact only with an individual from the same colony or from the two nearest-neighbour colonies ($q_{ij} > 0$ only if $i = j, j - 1$ or $j + 1$ for $i, j = 1, \dots, d$ with the convention that 0 stands for d and $d + 1$ for 1). In both cases, the condition for weak selection to favour cooperation is deduced from Theorem 2 for any irreducible and aperiodic migration matrix. This condition is developed further under the assumptions of random interactions and uniform migration in the case of linear dominance and under the assumptions of random interactions and local symmetric migration in the case of cyclic dominance.

6.1. Linear Dominance

Assume that individuals from colony i are better competitors than individuals from colony j , for $i, j = 1, \dots, d$, such that $i < j$. Under this assumption, there are two types of interactions: symmetric interactions, which occur between individuals in the same colony, and asymmetric interactions, which occur between individuals in different colonies (see, e.g., Krebs and Davies [35] and Walters and Seyfarth [36]). In the context of the repeated prisoner's dilemma, two strategies are available: tit-for-tat (*TFT*) and always defect (*AllD*) (see, e.g., Axelrod [37] and Axelrod and Hamilton [7]). Let $R \geq 1$ be the number of rounds of the game, $b > 0$ be the benefit of an individual if its opponent cooperates and $c > 0$ be the cost that a cooperator incurs, with $b > c$.

The payoff matrix for symmetric interactions within colony i is:

$$\mathbf{A}_{ii} = \begin{matrix} & \begin{matrix} TFT & AllD \end{matrix} \\ \begin{matrix} TFT \\ AllD \end{matrix} & \begin{pmatrix} R(b-c) & -c \\ b & 0 \end{pmatrix} \end{matrix} \tag{79}$$

for $i = 1, \dots, d$. In asymmetric interactions, a defector from colony i incurs a defection cost β when in interaction with a defector from colony $j < i$ (Tao *et al.* [29]; Kroumi and Lessard [30]). Therefore, the payoff matrix is $\mathbf{A}_{ij} = \mathbf{A}_{ii}$ for $i < j$, but:

$$\mathbf{A}_{ij} = \begin{matrix} & \begin{matrix} TFT & AllD \end{matrix} \\ \begin{matrix} TFT \\ AllD \end{matrix} & \begin{pmatrix} R(b-c) & -c - (R-1)\beta \\ b - (R-1)\beta & -R\beta \end{pmatrix} \end{matrix} \tag{80}$$

for $j < i$. Note that the case $R = 1$ corresponds to the standard additive prisoner’s dilemma for symmetric, as well as asymmetric interactions. The payoffs as defined in Equation (2) satisfy:

$$a_{jk} + b_{jk} - c_{jk} - d_{jk} = \begin{cases} R(b-c) - b - c & \text{if } 1 \leq j \leq k \leq d \\ R(b + \beta - c) - b - c & \text{if } 1 \leq k < j \leq d \end{cases} \tag{81}$$

Then, the condition given in Equation (67) for *TFT* to be favoured under weak selection and mutation becomes:

$$(R(b-c) - b - c) \sum_{i,j=1}^d \pi_i \frac{\alpha_j}{\alpha_i} m_{ji} + \frac{R\beta}{d} \sum_{i,j=1}^d \sum_{k=1}^{j-1} \pi_i \frac{\alpha_j}{\alpha_i} m_{ji} q_{jk} > 0 \tag{82}$$

This condition can be written into the equivalent form:

$$R > \frac{b+c}{b-c+a_L\beta} \tag{83}$$

where:

$$a_L = \frac{\sum_{i,j=1}^d \sum_{k=1}^{j-1} \pi_i \frac{\alpha_j}{\alpha_i} m_{ji} q_{jk}}{\sum_{i,j=1}^d \pi_i \frac{\alpha_j}{\alpha_i} m_{ji}} \tag{84}$$

In the case of random interactions ($q_{jk} = 1/d$) and colonies of the same size ($\alpha_i = 1/d$) with uniform migration ($m_{ji} = 1 - m$ if $i = j$ and $m/(d - 1)$ otherwise), so that $\pi_i = 1/d$ for $i = 1, \dots, d$, we have $a_L = (d - 1)/(2d)$. Figure 1 shows the migration and interaction graphs in this case. Moreover, the above condition takes the form:

$$R > \frac{d(b+c)}{d(b-c) + \frac{d-1}{2}\beta} \tag{85}$$

This extends a result found in Kroumi and Lessard [30] in the particular case where $d = 3$.

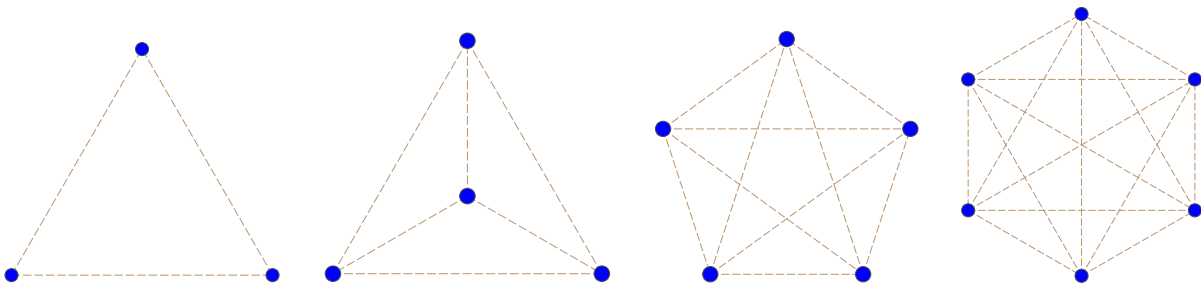


Figure 1. Examples with linear dominance. The individuals of the population are distributed over d colonies. From left to right: $d = 3$; $d = 4$; $d = 5$; $d = 6$. Colonies are represented by full circles. Interactions and migration are possible between colonies connected by dashed edges. The interaction and migration graphs are both complete graphs.

6.2. Cyclic Dominance

Cyclic dominance is inspired from the rock-paper-scissors game (Linhart [38]). Rock is wrapped by paper; paper is cut by scissors; and scissors are crushed by rock. This simple game is popular among children and adults to decide on trivial disputes that have no obvious winner. In the case of cyclic dominance with $d \geq 3$ colonies, we assume interactions only between individuals in adjacent colonies contrary to the case of linear dominance where there were interactions between individuals in all colonies. More precisely, we assume that an individual from colony i can interact only with an individual from colony $i - 1$, i or $i + 1$, for $i = 2, \dots, d - 1$. An individual from colony 1 can interact only with an individual from the same colony or from colony 2 or d , while an individual from colony d can interact only with an individual from colony 1, $d - 1$ or d . We assume also that individuals from colony i dominate individuals from colony $i + 1$, for $i = 1, \dots, d - 1$, and individuals from colony d dominate individuals from colony 1. As in the case of linear dominance, there are symmetric interactions and asymmetric interactions. Symmetric interactions within colony $i = 1, \dots, d$ result in the payoff matrix:

$$\mathbf{A}_{ii} = \begin{matrix} & \begin{matrix} TFT & ALLD \end{matrix} \\ \begin{matrix} TFT \\ ALLD \end{matrix} & \begin{pmatrix} R(b - c) & -c \\ b & 0 \end{pmatrix} \end{matrix} \tag{86}$$

The payoffs of individuals in colony i in asymmetric interaction with dominated individuals from colony $i + 1$ (or 1 if $i = d$) are given by the same matrix as for symmetric interactions, that is:

$$\mathbf{A}_{i,i+1} = \mathbf{A}_{d,1} = \mathbf{A}_{i,i} \tag{87}$$

for $i = 1, \dots, d - 1$. However, if an individual in colony i is in asymmetric interaction with a dominant individual in colony $i - 1$ (or d of $i = 1$), then the payoff matrix is:

$$\mathbf{A}_{i,i-1} = \mathbf{A}_{1,d} = \begin{matrix} & \begin{matrix} TFT & ALLD \end{matrix} \\ \begin{matrix} TFT \\ ALLD \end{matrix} & \begin{pmatrix} R(b - c) & -c - (R - 1)\beta \\ b - (R - 1)\beta & -R\beta \end{pmatrix} \end{matrix} \tag{88}$$

for $i = 2, \dots, d$.

In this case, the condition given in Equation (67) for TFT to be favoured under weak selection and mutation takes the form:

$$(R(b - c) - b - c) \sum_{i,j=1}^d \pi_i \frac{\alpha_j}{\alpha_i} m_{ji} + R\beta \sum_{i,j=1}^d \pi_i \frac{\alpha_j}{\alpha_i} m_{ji} q_{j,j-1} > 0 \tag{89}$$

This can be written as:

$$R > \frac{b + c}{b - c + a_C \beta} \tag{90}$$

where:

$$a_C = \frac{\sum_{i,j=1}^d \pi_i \frac{\alpha_j}{\alpha_i} m_{ji} q_{j,j-1}}{\sum_{i,j=1}^d \pi_i \frac{\alpha_j}{\alpha_i} m_{ji}} \tag{91}$$

In the case of random interactions between individuals belonging to the same colony and the two adjacent colonies ($q_{j,j-1} = q_{j,j} = q_{j,j+1} = 1/3$ for $j = 1, \dots, d$ with the convention that 0 stands for d and $d + 1$ for 1) with colonies of the same size ($\alpha_i = 1/d$) and symmetric local migration ($m_{ji} = 1 - m$ if $i = j$, $m/2$ if $i = j - 1$ or $j + 1$, and zero otherwise), so that $\pi_i = 1/d$ for $i = 1, \dots, d$, we have $a_C = 1/3$. The migration and interaction graphs in this case are illustrated in Figure 2. In this case, the above condition takes the form:

$$R > \frac{b + c}{b - c + \frac{\beta}{3}} \tag{92}$$

This condition in the case of cyclic dominance is the same as the condition given in Equation (85) in the case of linear dominance for $d = 3$, in agreement with Kroumi and Lessard [30]. This result conforms nicely with intuition, because for $d = 3$, the regular cycle is identical to the complete graph. However, it is more stringent than the condition given in Equation (85) as soon as $d > 3$ (see Figure 3).

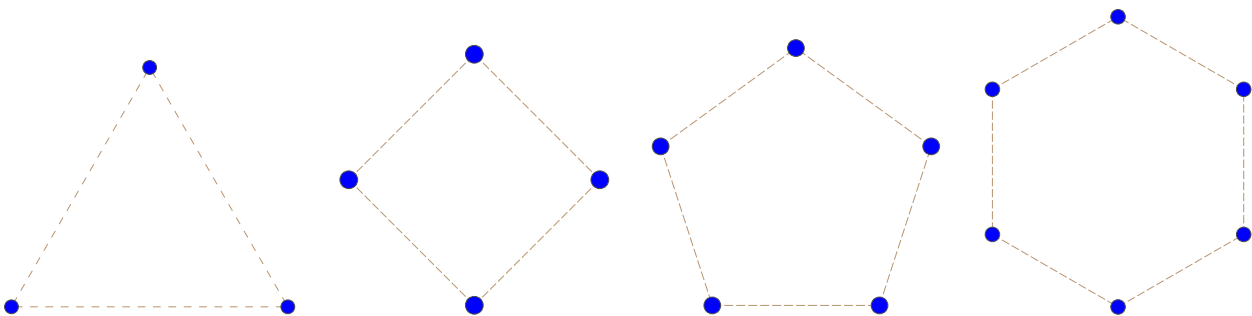


Figure 2. Examples with cyclic dominance. The individuals of the population are distributed over d colonies on a circle. From left to right: $d = 3$; $d = 4$; $d = 5$; $d = 6$. Colonies are represented by full circles. Interactions and migration are possible only between colonies connected by dashed edges. The interaction and migration graphs are regular cycles.

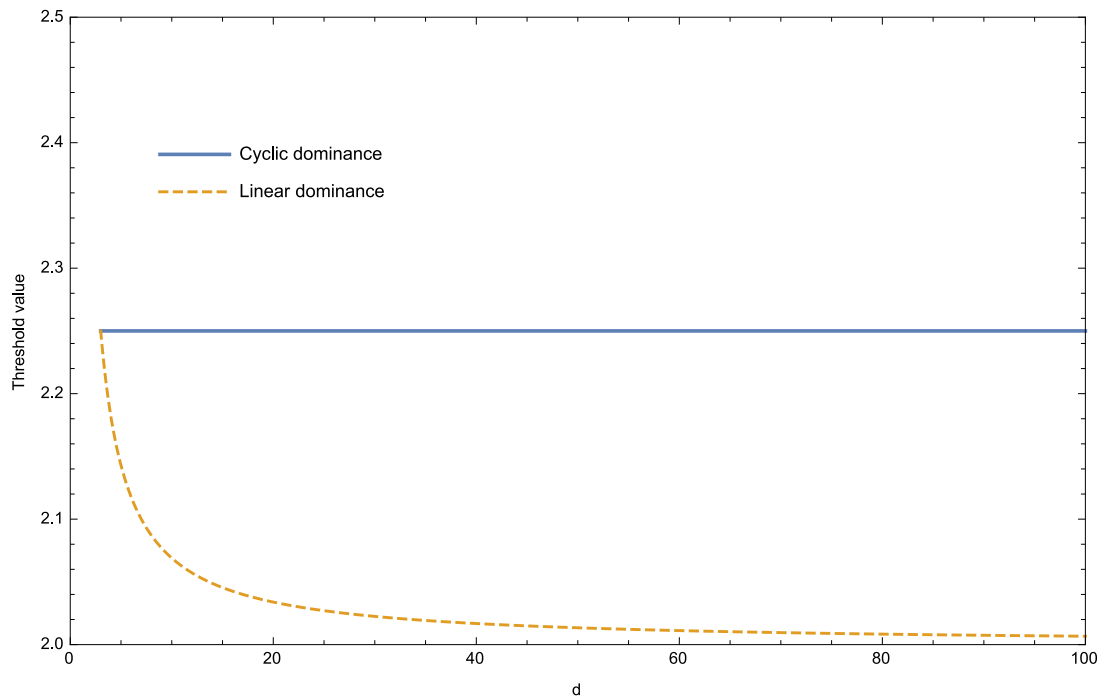


Figure 3. Exact threshold value that R must exceed for cooperation to be favoured by weak selection under linear dominance and cyclic dominance in the $N \rightarrow \infty$ limit as a function of the number d of colonies, for $b = 2$, $c = 1$ and $\beta = 1$. The threshold value is the same for $d = 3$, since the complete graph coincides with the cyclic graph in this case. However, while this value remains constant in the case of cyclic dominance as d increases, it decreases in the case of linear dominance, making cooperation more favourable.

7. Application to Games in Set-Structured Populations

Tarnita *et al.* [26] consider a population composed of N individuals and M sets. They assume that each individual belongs exactly to K sets, which corresponds to a phenotype. Interactions occur only between individuals within the same sets (individuals interact as many times as the number of sets to which they both belong). With M sets numbered $1, \dots, M$, every individual l is represented by a M -dimensional vector $\mathbf{x}_l = (x_{l,i})_{1 \leq i \leq M}$. Here, $x_{l,i} = 1$ if individual l belongs to set i and zero otherwise, for $i = 1, \dots, M$, with exactly K components equal to one and $M - K$ components equal to zero. The individuals represented by the same vector \mathbf{x} belong to the same colony represented by this vector. Here, there are $d = \binom{M}{K}$ colonies. We assume that an offspring inherits the K sets of his parent represented by \mathbf{x} with probability $m_{\mathbf{x},\mathbf{x}}$ and chooses K sets represented by \mathbf{y} with probability $m_{\mathbf{x},\mathbf{y}}$. These are phenotype mutation probabilities. The number of interactions between two individuals represented by $\mathbf{y} = (y_i)_{1 \leq i \leq M}$ and $\mathbf{z} = (z_i)_{1 \leq i \leq M}$, respectively, is given by:

$$\langle \mathbf{y}, \mathbf{z} \rangle = \sum_{i=1}^M y_i z_i \tag{93}$$

Therefore,

$$q_{\mathbf{y},\mathbf{z}} = \frac{\langle \mathbf{y}, \mathbf{z} \rangle \alpha_{\mathbf{z}}}{\sum_{\mathbf{x}} \langle \mathbf{y}, \mathbf{x} \rangle \alpha_{\mathbf{x}}} \tag{94}$$

where α_x is the proportion of colony x . Two strategies are in use, represented by S_1 and S_2 . The payoff matrix for i in y against j in z is denoted by $A_{y,z}$.

Suppose that phenotype mutation occurs at random, so that $m_{x,y} = 1/d$ for every couple of phenotypes (x, y) . Then, the backward matrix takes the form:

$$m_{x,y}^* = \begin{cases} \frac{\alpha_y}{dN\alpha_x} & \text{if } x \neq y \\ 1 - \frac{1-\alpha_x}{dN\alpha_x} & \text{if } x = y \end{cases} \tag{95}$$

The stationary distribution of this matrix is given by:

$$\pi_x = \frac{\alpha_x^2}{\sum_y \alpha_y^2} \tag{96}$$

Therefore, the condition given in Equation (67) for S_1 to be favoured by weak selection becomes:

$$\sum_{y,z} \alpha_y q_{y,z} (a_{y,z} + b_{y,z} - c_{y,z} - d_{y,z}) > 0 \tag{97}$$

Written in the form:

$$\sum_{y,z} \alpha_y q_{y,z} (a_{y,z} + b_{y,z}) > \sum_{y,z} \alpha_y q_{y,z} (c_{y,z} + d_{y,z}) \tag{98}$$

the condition means that the expected payoff of S_1 exceeds the expected payoff of S_2 when the expected frequencies of S_1 and S_2 are equal among the individuals of the same phenotype for every phenotype.

Suppose now that an S_1 -individual uses strategy S_1 with an opponent only if the two individuals belong to at least L common sets, where $1 \leq L \leq K$ is a fixed constant, and uses S_2 otherwise. On the other hand, an S_2 -individual always uses strategy S_2 . In this case, the payoff matrix for an individual in colony x against an individual in colony y takes the form:

$$A_{x,y} = \begin{matrix} & \begin{matrix} S_1 & S_2 \end{matrix} \\ \begin{matrix} S_1 \\ S_2 \end{matrix} & \begin{pmatrix} R & S \\ T & P \end{pmatrix} \end{matrix} \tag{99}$$

if $\langle x, y \rangle \geq L$, and:

$$A_{x,y} = \begin{matrix} & \begin{matrix} S_1 & S_2 \end{matrix} \\ \begin{matrix} S_1 \\ S_2 \end{matrix} & \begin{pmatrix} P & P \\ P & P \end{pmatrix} \end{matrix} \tag{100}$$

if $1 \leq \langle x, y \rangle \leq L - 1$ (note that the meaning of R is different here from the previous section). Moreover, we have:

$$a_{x,y} + b_{x,y} - c_{x,y} - d_{x,y} = \begin{cases} R + S - T - P & \text{if } \langle x, y \rangle \geq L \\ 0 & \text{otherwise} \end{cases} \tag{101}$$

In this case, the condition given in Equation (67) for S_1 to be favoured by weak selection reduces to:

$$(R + S - T - P) \sum_{x,y} \sum_{z: \langle y,z \rangle \geq L} \pi_x \frac{\alpha_y}{\alpha_x} m_{y,x} q_{y,z} > 0 \tag{102}$$

which is the same as:

$$R + S > T + P \tag{103}$$

This condition, known as risk dominance in a coordination game (Harsanyi and Selten [39]), does not depend on the population structure.

8. Discussion

Our main result (Theorem 2) for S_1 to be favoured by weak selection over S_2 in a structured population under weak strategy mutation in the limit of a large population with payoffs in pairwise interactions depending on the locations of the players is given by Equation (67), which can be written in the form:

$$\sum_{i,j,k=1}^d \pi_i \frac{\alpha_j}{\alpha_i} m_{ji} q_{jk} \left(\frac{a_{j,k} + b_{j,k}}{2} \right) > \sum_{i,j,k=1}^d \pi_i \frac{\alpha_j}{\alpha_i} m_{ji} q_{jk} \left(\frac{c_{j,k} + d_{j,k}}{2} \right) \tag{104}$$

Here, π_i is the limiting proportion of time back that a single lineage spends in colony i in the absence of selection, for $i = 1, \dots, d$. It represents the expected contribution of colony i to the whole population in the long run forward in time, called its reproductive value, under the neutral model. With α_i representing the proportion of colony i and m_{ij} the probability for an offspring from colony i to migrate to colony j for $i, j = 1, \dots, d$, the quantity $\pi_i m_{ij} / \alpha_i$ represents an expected relative reproductive value of an offspring produced by an individual in colony i . On the other hand, every individual interacts with an S_1 -individual and with an S_2 -individual with the same probability $1/2$ in a neutral population at equilibrium, since then, the probability that any given individual uses strategy S_1 or S_2 is equal to the probability that the most recent mutant ancestor of this individual used strategy S_1 or S_2 , which is $1/2$ in each case from the assumptions on strategy mutation. We have the same approximate probability $1/2$ in an equilibrium population under weak selection. With α_j being the probability that an individual chosen at random in the whole population belongs to colony j and q_{jk} , the probability for an individual from colony j to interact with an individual from colony k , for $i, j, k = 1, \dots, d$, the left-hand side of Equation (104) can be interpreted as the expected payoff of S_1 weighted by relative reproductive values of offspring in an equilibrium population near neutrality. The right-hand side of Equation (104) has a similar interpretation for S_2 , and the inequality guarantees that S_1 is more abundant on average than S_2 at equilibrium near neutrality if individuals are weighted by their relative reproductive values.

This interpretation is very intuitive. It is an alternative to the use of structure coefficients (Nowak *et al.* [32]) for games in structured populations. Moreover, this interpretation suggests an effective payoff matrix (Lessard [40]):

$$\tilde{\mathbf{A}} = \sum_{i,j,k=1}^d \pi_i \frac{\alpha_j m_{ji}}{\alpha_i} q_{jk} \mathbf{A}_{jk} \tag{105}$$

where \mathbf{A}_{jk} is the payoff matrix for an individual from colony j in interaction with an individual from colony k . This means that the game in the structured population is equivalent to a game in a well-mixed population with this matrix as the payoff matrix.

This result has been obtained for a structured population reproducing according to a Moran model with one individual replaced at a time in the strong migration limit as the population size N tends to infinity. We have shown (Theorem 1) that the genealogical process in the neutral model, which is

described by the transition matrix in Equation (12) from one time step to the previous one, where \mathbf{A} and \mathbf{B} are given in Section 3, tends to the standard Kingman coalescent (Kingman [6]) for the number of ancestors if $N^2/(2\lambda)$ time steps are taken as the unit of time where λ is given by Equation (36), while the ancestors are distributed independently among the colonies according to the stationary distribution $(\pi_i)_{1 \leq i \leq d}$. The proof (Appendix A) relies on a two-time scale argument and uses a lemma due to Möhle [31]. A similar result was proven in a different way in Notohara [2] in the case of a subdivided population that reproduces according to a Wright–Fisher model.

Our main result has been applied to the situation of dominance hierarchy with d colonies in decreasing order of dominance in the case of linear dominance and in counter-clockwise order of dominance in the case of cyclic dominance. Individuals in a given colony can interact with individuals in all other colonies in the case of linear dominance, but only with individuals in the same colony or in the two adjacent colonies in the case of cyclic dominance. Considering the strategies *TFT* and *AllD* in a repeated additive prisoner’s dilemma and a cost for defection against a dominant defector, it has been shown that linear dominance is more favourable than cyclic dominance for increasing the expected frequency of *TFT* at equilibrium as soon as $d > 3$. This has been obtained under the assumptions of colonies of the same size with uniform or symmetric migration and random interactions.

Another application concerns the set-structured population as introduced in Tarnita *et al.* [26], but with colonies of fixed relative sizes and reproduction according to a Moran model instead of a Wright–Fisher model. With uniform mutation from one subset of sets to another of the same size, which defines the phenotype of an individual, the condition for S_1 to be favoured by weak selection is that, for an individual chosen at random in the whole population, the expected payoff of S_1 exceeds the expected payoff of S_2 near neutrality. With strategy S_1 actually used only if the number of common sets to which the two players belong exceeds some threshold, it has been shown that the condition for S_1 to be favoured by weak selection reduces to a condition known as risk dominance (Harsanyi and Selten [39]) as in a well-mixed population. Note that the same result was obtained in Tarnita *et al.* [26] in the case of a high rate of phenotype mutation, which corresponds to strong migration from one phenotype to another.

9. Conclusions

Conditions for a strategy to be more abundant on average at equilibrium than another strategy in a structured population under weak selection and weak strategy mutation but strong migration can be expressed in terms of reproductive values in the neutral model on the basis of a two-time scale argument. This can be applied, for instance, to two-player games in populations with set structure or dominance hierarchy. In the case of dominance hierarchy, it has been shown that the condition for *TFT* to be more abundant on average than *AllD* in a repeated additive prisoner’s dilemma with a given cost for defection against a dominant defector is generally less stringent with linear dominance than with cyclic dominance. In the case of set structure with strong migration from one set to another, it has been shown that the condition for a strategy to be more abundant on average than another corresponds to risk dominance.

Author Contributions

The authors have contributed equally to this paper.

Conflicts of Interest

The authors declare no conflict of interest.

Appendix

A. Convergence Result

For every square $n \times n$ matrix $\mathbf{A} = (a_{ij})_{1 \leq i, j \leq n}$, define:

$$\|\mathbf{A}\| := \max_{1 \leq i \leq n} \left(\sum_{j=1}^n |a_{ij}| \right) \tag{A1}$$

and:

$$\exp \{ \mathbf{A} \} := \sum_{k=0}^{\infty} \frac{\mathbf{A}^k}{k!} \tag{A2}$$

Note that, for any square $n \times n$ matrices \mathbf{A} and \mathbf{B} ,

$$\|\mathbf{AB}\| \leq \|\mathbf{A}\| \|\mathbf{B}\| \tag{A3}$$

Moreover,

$$\exp \{ \mathbf{A} + \mathbf{B} \} = \exp \{ \mathbf{A} \} \exp \{ \mathbf{B} \} \tag{A4}$$

if $\mathbf{AB} = \mathbf{BA}$. Finally, \mathbf{I} and $\mathbf{0}$ denote the identity matrix and the zero matrix, respectively.

Proposition 1. *Let $(c_N)_{N \geq 1}$ be a sequence of positive real numbers with $\lim_{N \rightarrow \infty} c_N = 0$. Let \mathbf{A} and \mathbf{B} be square matrices of size n , such that $\|\exp\{\mathbf{A}\}\| = 1$ and $\mathbf{P} := \lim_{m \rightarrow \infty} \exp\{m\mathbf{A}\}$ exists with $\mathbf{AP} = \mathbf{PA} = \mathbf{0}$. Then, we have:*

$$\lim_{N \rightarrow \infty} (\mathbf{I} + c_N \mathbf{A} + c_N^2 \mathbf{B})^{\lfloor tc_N^{-2} \rfloor} = \mathbf{P} \exp\{t\mathbf{G}\} \tag{A5}$$

for every real number $t > 0$, where $\mathbf{G} = \mathbf{PBP}$.

Proof. Since $c_N \rightarrow 0$ as $N \rightarrow \infty$, there exists an integer $N_0 \geq 1$, such that $c_N < 1$ and:

$$c_N \|\mathbf{A} + c_N \mathbf{B}\| < 1 \tag{A6}$$

for every integer $N \geq N_0$. Then:

$$(\mathbf{I} + c_N \mathbf{A} + c_N^2 \mathbf{B})^{\lfloor tc_N^{-2} \rfloor} = \exp \left\{ \lfloor tc_N^{-2} \rfloor \log (\mathbf{I} + c_N \mathbf{A} + c_N^2 \mathbf{B}) \right\} \tag{A7}$$

where:

$$\log (\mathbf{I} + c_N \mathbf{A} + c_N^2 \mathbf{B}) = \sum_{l=1}^{\infty} \frac{(-1)^{l+1} c_N^l (\mathbf{A} + c_N \mathbf{B})^l}{l} \tag{A8}$$

For more details on the logarithm of a matrix, see, e.g., Culver [41] or Higham [42]. Therefore,

$$\begin{aligned}
 (\mathbf{I} + c_N \mathbf{A} + c_N^2 \mathbf{B})^{\lfloor tc_N^{-2} \rfloor} &= \exp \{ c_N \lfloor tc_N^{-2} \rfloor (\mathbf{A} + c_N \mathbf{B}) \} \times \exp \left\{ -\frac{c_N^2 \lfloor tc_N^{-2} \rfloor (\mathbf{A} + c_N \mathbf{B})^2}{2} \right\} \\
 &\times \exp \left\{ c_N^2 \lfloor tc_N^{-2} \rfloor \mathbf{E}_N \right\} \tag{A9}
 \end{aligned}$$

where:

$$\mathbf{E}_N = \sum_{l=3}^{\infty} \frac{(-1)^{l+1} c_N^{l-2} (\mathbf{A} + c_N \mathbf{B})^l}{l} \tag{A10}$$

For the third term on the right-hand side of Equation (A9), we have:

$$\begin{aligned}
 \|\mathbf{E}_N\| &\leq \sum_{l=3}^{\infty} \frac{c_N^{l-2} \|\mathbf{A} + c_N \mathbf{B}\|^l}{l} \leq \sum_{l=1}^{\infty} c_N^l \|\mathbf{A} + c_N \mathbf{B}\|^{l+2} \\
 &\leq \frac{c_N \|\mathbf{A} + c_N \mathbf{B}\|^3}{1 - c_N \|\mathbf{A} + c_N \mathbf{B}\|} \rightarrow 0 \tag{A11}
 \end{aligned}$$

as $N \rightarrow \infty$. Consequently,

$$\lim_{N \rightarrow \infty} \mathbf{E}_N = \mathbf{0} \tag{A12}$$

From this and the fact that $c_N^2 \lfloor tc_N^{-2} \rfloor \rightarrow t$ as $N \rightarrow \infty$, we have:

$$\lim_{N \rightarrow \infty} \exp \{ c_N^2 \lfloor tc_N^{-2} \rfloor \mathbf{E}_N \} = \mathbf{I} \tag{A13}$$

On the other hand,

$$\lim_{N \rightarrow \infty} \exp \left\{ -\frac{c_N^2 \lfloor tc_N^{-2} \rfloor}{2} (\mathbf{A} + c_N \mathbf{B})^2 \right\} = \exp \left\{ -\frac{t}{2} \mathbf{A}^2 \right\} \tag{A14}$$

Finally, note that:

$$\exp \{ \mathbf{A} + c_N \mathbf{B} \} = \exp \{ \mathbf{A} \} + c_N \mathbf{D}_N \tag{A15}$$

where:

$$\mathbf{D}_N = \mathbf{D} + c_N \mathbf{R}_N \tag{A16}$$

with:

$$\mathbf{D} = \sum_{l=1}^{\infty} \frac{1}{l!} \sum_{m_1+m_2=l-1} \mathbf{A}^{m_1} \mathbf{B} \mathbf{A}^{m_2} \tag{A17}$$

and:

$$\mathbf{R}_N = \sum_{l=2}^{\infty} \frac{1}{l!} \sum_{k=2}^l c_N^{k-2} \sum_{m_1+\dots+m_{k+1}=l-k} \mathbf{A}^{m_1} \mathbf{B} \mathbf{A}^{m_2} \mathbf{B} \dots \mathbf{A}^{m_k} \mathbf{B} \mathbf{A}^{m_{k+1}} \tag{A18}$$

Here, the summation is over all non-negative integers m_1, \dots, m_{k+1} , which sum up to $l - k$. Moreover,

$$\begin{aligned} \|\mathbf{R}_N\| &\leq \sum_{l=2}^{\infty} \frac{1}{l!} \sum_{k=2}^l c_N^{k-2} \sum_{m_1+\dots+m_{k+1}=l-k} \|\mathbf{A}^{m_1} \mathbf{B} \mathbf{A}^{m_2} \mathbf{B} \dots \mathbf{A}^{m_k} \mathbf{B} \mathbf{A}^{m_{k+1}}\| \\ &\leq \sum_{l=2}^{\infty} \frac{1}{l!} \sum_{k=2}^l c_N^{k-2} \sum_{m_1+\dots+m_{k+1}=l-k} \|\mathbf{A}\|^{m_1+\dots+m_{k+1}} \|\mathbf{B}\|^k \\ &\leq \sum_{l=2}^{\infty} \frac{1}{l!} \sum_{k=2}^l c_N^{k-2} \binom{l}{k} \|\mathbf{A}\|^{l-k} \|\mathbf{B}\|^k \\ &\leq \sum_{l=2}^{\infty} \frac{(\|\mathbf{A}\| + \|\mathbf{B}\|)^l}{l!} \\ &\leq \exp\{\|\mathbf{A}\| + \|\mathbf{B}\|\} < \infty \end{aligned} \tag{A19}$$

Note that $\binom{l}{k}$ represents the number of ways $l - k$ indistinguishable balls can be put into $k + 1$ distinguishable cells. Furthermore, a lemma proved in Möhle [31] guarantees that:

$$\lim_{N \rightarrow \infty} (\exp\{\mathbf{A}\} + c_N \mathbf{D}_N)^{\lfloor c_N^{-1} \rfloor} = \mathbf{P} \exp\{\mathbf{Q}\} \tag{A20}$$

where:

$$\mathbf{P} = \lim_{N \rightarrow \infty} (\exp\{\mathbf{A}\})^m = \lim_{N \rightarrow \infty} \exp\{m\mathbf{A}\} \tag{A21}$$

and:

$$\mathbf{Q} = \lim_{N \rightarrow \infty} \mathbf{P} \mathbf{D}_N \mathbf{P} = \mathbf{P} \mathbf{D} \mathbf{P} \tag{A22}$$

Therefore,

$$\begin{aligned} \lim_{N \rightarrow \infty} \exp\{c_N \lfloor t c_N^{-2} \rfloor (\mathbf{A} + c_N \mathbf{B})\} &= \lim_{N \rightarrow \infty} (\exp\{\mathbf{A} + c_N \mathbf{B}\})^{c_N \lfloor t c_N^{-2} \rfloor} \\ &= \lim_{N \rightarrow \infty} \left((\exp\{\mathbf{A}\} + c_N \mathbf{D}_N)^{\lfloor c_N^{-1} \rfloor} \right)^{\lfloor c_N^{-1} \rfloor^{-1} c_N \lfloor t c_N^{-2} \rfloor} \\ &= (\mathbf{P} \exp\{\mathbf{Q}\})^t \\ &= \mathbf{P} \exp\{t\mathbf{Q}\} \end{aligned} \tag{A23}$$

since $\lfloor c_N^{-1} \rfloor^{-1} c_N \lfloor t c_N^{-2} \rfloor \rightarrow t$ as $N \rightarrow \infty$. The last equality comes from the fact that:

$$\mathbf{P} \exp\{\mathbf{Q}\} = \mathbf{P} \left(\sum_{k=0}^{\infty} \frac{\mathbf{Q}^k}{k!} \right) = \mathbf{P} + \sum_{k=1}^{\infty} \frac{\mathbf{P} \mathbf{D}^k \mathbf{P}}{k!} = \left(\sum_{k=0}^{\infty} \frac{\mathbf{Q}^k}{k!} \right) \mathbf{P} = \exp\{\mathbf{Q}\} \mathbf{P} \tag{A24}$$

This ascertains the equality for every $t > 0$ that is an integer, then a rational number and, finally, a real number by continuity. Using Equations (A13), (A14) and (A23), the limit in Equation (A9) as $N \rightarrow \infty$ gives:

$$\lim_{N \rightarrow \infty} (\mathbf{I} + c_N \mathbf{A} + c_N^2 \mathbf{B})^{\lfloor t c_N^{-2} \rfloor} = \mathbf{P} \exp\{t\mathbf{Q}\} \exp\left\{-\frac{t}{2} \mathbf{A}^2\right\} \tag{A25}$$

On the other hand, the condition $\mathbf{PA} = \mathbf{AP} = \mathbf{0}$ ensures the equalities:

$$\begin{aligned} \mathbf{Q} &= \mathbf{PDP} = \sum_{l=1}^{\infty} \frac{1}{l!} \sum_{m_1+m_2=l-1} \mathbf{PA}^{m_1} \mathbf{BA}^{m_2} \mathbf{P} \\ &= \mathbf{PBP} + \sum_{l=2}^{\infty} \frac{1}{l!} \sum_{m_1+m_2=l-1} \mathbf{PA}^{m_1} \mathbf{BA}^{m_2} \mathbf{P} \\ &= \mathbf{PBP} = \mathbf{G} \end{aligned} \tag{A26}$$

and:

$$\mathbf{P} \exp \left\{ -\frac{t\mathbf{A}^2}{2} \right\} = \mathbf{P} \left(\sum_{k=0}^{\infty} \frac{(-1)^k \mathbf{A}^{2k}}{2^k k!} \right) = \mathbf{P} + \sum_{k=1}^{\infty} \frac{(-1)^k \mathbf{PA}^{2k}}{2^k k!} = \mathbf{P} \tag{A27}$$

Using these equalities in Equation (A25) yields:

$$\begin{aligned} \lim_{N \rightarrow \infty} (\mathbf{I} + c_N \mathbf{A} + c_N^2 \mathbf{B})^{\lfloor tc_N^{-2} \rfloor} &= \mathbf{P} \exp\{t\mathbf{G}\} \exp \left\{ -\frac{t\mathbf{A}^2}{2} \right\} \\ &= \exp\{t\mathbf{G}\} \mathbf{P} \exp \left\{ -\frac{t\mathbf{A}^2}{2} \right\} \\ &= \exp\{t\mathbf{G}\} \mathbf{P} \\ &= \mathbf{P} \exp\{t\mathbf{G}\} \end{aligned} \tag{A28}$$

This completes the proof. \square

B. Stationary Distribution Associated with \mathbf{A}_k

In order to show that:

$$\pi_{\mathbf{n}} = \binom{|\mathbf{n}|}{n_1, n_2, \dots, n_d} \pi_1^{n_1} \times \dots \times \pi_d^{n_d} \tag{B1}$$

for $\mathbf{n} \in \Gamma_k$, is the stationary distribution associated with $\mathbf{A}_k = (a_{\mathbf{n}, \mathbf{n}'})_{\mathbf{n}, \mathbf{n}' \in \Gamma_k}$ for $k \geq 2$, we have to check that:

$$\sum_{\mathbf{n} \in \Gamma_k} \pi_{\mathbf{n}} a_{\mathbf{n}, \mathbf{n}'} = 0 \quad \text{with} \quad \sum_{\mathbf{n} \in \Gamma_k} \pi_{\mathbf{n}} = 1 \tag{B2}$$

for $\mathbf{n}' \in \Gamma_k$. First note that, owing to the expression in Equation (13) for the non-null entries of \mathbf{A}_k , we have $a_{\mathbf{n}, \mathbf{n}'} \neq 0$ if $\mathbf{n}' = \mathbf{n} - \mathbf{e}_i + \mathbf{e}_j$ for $i \neq j$ and $n_i \geq 1$. In this case, we have:

$$\begin{aligned} \pi_{\mathbf{n}'} &= \binom{|\mathbf{n}|}{n_1, \dots, n_i - 1, \dots, n_j + 1, \dots, n_d} \pi_1^{n_1} \times \dots \times \pi_i^{n_i - 1} \times \dots \times \pi_j^{n_j + 1} \times \dots \times \pi_d^{n_d} \\ &= \binom{|\mathbf{n}|}{n_1, \dots, n_d} \times \frac{n_i}{n_j + 1} \pi_1^{n_1} \times \dots \times \pi_d^{n_d} \times \frac{\pi_j}{\pi_i} \\ &= \pi_{\mathbf{n}} \times \frac{n_i \pi_j}{(n_j + 1) \pi_i} \end{aligned} \tag{B3}$$

Therefore, for $\mathbf{n}' \in \Gamma_k$, we obtain:

$$\begin{aligned} \sum_{\mathbf{n} \in \Gamma_k} \pi_{\mathbf{n}} a_{\mathbf{n}, \mathbf{n}'} &= \sum_{\mathbf{n} \neq \mathbf{n}'} \pi_{\mathbf{n}} a_{\mathbf{n}, \mathbf{n}'} + \pi_{\mathbf{n}'} a_{\mathbf{n}', \mathbf{n}'} \\ &= \sum_{\substack{j: n'_j \geq 1 \\ i \neq j}} \frac{(n'_i + 1) \alpha_j}{\alpha_i} m_{ji} \times \pi_{\mathbf{n}'} \times \frac{n'_j \pi_i}{(n'_i + 1) \pi_j} + \pi_{\mathbf{n}'} a_{\mathbf{n}', \mathbf{n}'} \\ &= \sum_{j: n'_j \geq 1} \frac{n'_j \pi_{\mathbf{n}'}}{\pi_j} \sum_{i \neq j} \frac{\alpha_j}{\alpha_i} m_{ji} \pi_i + \pi_{\mathbf{n}'} a_{\mathbf{n}', \mathbf{n}'} \end{aligned} \tag{B4}$$

On the other hand, by definition of the stationary distribution (π_1, \dots, π_d) associated with \mathbf{A}_1 , we have:

$$\sum_{i=1}^d \pi_i a_{\mathbf{e}_i, \mathbf{e}_j} = 0 \tag{B5}$$

for $j = 1, \dots, d$. This entails that:

$$\sum_{i \neq j} \pi_i \frac{\alpha_j m_{ji}}{\alpha_i} = \sum_{i \neq j} \pi_j \frac{\alpha_i m_{ij}}{\alpha_j} \tag{B6}$$

for $j = 1, \dots, d$. Plugging this into Equation (B4) yields:

$$\begin{aligned} \sum_{\mathbf{n} \in \Gamma_k} \pi_{\mathbf{n}} a_{\mathbf{n}, \mathbf{n}'} &= \left(\sum_{j: n'_j \geq 1} \sum_{i \neq j} \frac{n'_j \alpha_i m_{ij}}{\alpha_j} \right) \pi_{\mathbf{n}'} + \pi_{\mathbf{n}'} a_{\mathbf{n}', \mathbf{n}'} \\ &= -a_{\mathbf{n}', \mathbf{n}'} \pi_{\mathbf{n}'} + \pi_{\mathbf{n}'} a_{\mathbf{n}', \mathbf{n}'} = 0 \end{aligned} \tag{B7}$$

This establishes the first equation in Equation (B2). It suffices to use the generalized binomial theorem to get:

$$\begin{aligned} \sum_{\mathbf{n} \in \Gamma_k} \pi_{\mathbf{n}} &= \sum_{\mathbf{n}: |\mathbf{n}|=k} \binom{|\mathbf{n}|}{n_1, \dots, n_d} \pi_1^{n_1} \times \dots \times \pi_d^{n_d} \\ &= (\pi_1 + \dots + \pi_d)^k = 1 \end{aligned} \tag{B8}$$

which is the second equation in Equation (B2).

C. Probability of Identical Strategies

With $N^2/(2\lambda)$ time steps as the unit of time, let the scaled mutation rate defined by:

$$\theta = \frac{uN^2}{2\lambda} \tag{C1}$$

be constant as $N \rightarrow \infty$. We consider two probabilities:

$$\phi_2 = \mathbb{P}_0 \{I \text{ and } J \text{ have the same strategy}\} \tag{C2}$$

and

$$\phi_3 = \mathbb{P}_0 \{I, J \text{ and } K \text{ have the same strategy}\} \tag{C3}$$

where I, J and K are three individuals in the population chosen at random at the same time step and \mathbb{P}_0 is the probability at equilibrium under neutrality ($s = 0$). Let $t_N^{(2)}(I, J)$ be the number of time steps back before the coalescence of the lineages of two individuals, I and J , chosen at random at the same time step. Similarly, let $t_N^{(3)}(I, J, K)$ be the number of time steps back before the first coalescence of two lineages among the lineages of three individuals, I, J and K , chosen at random at the same time step, and $t_N^{(2)}(I, J, K)$ be the supplementary number of time steps before the coalescence of the two remaining lineages. From the limiting genealogical process described in Theorem 1, the scaled coalescence time $\frac{2\lambda t_N^{(3)}(I, J, K)}{N^2}$ and the vector of rescaled coalescence times $\left(\frac{2\lambda t_N^{(3)}(I, J, K)}{N^2}, \frac{2\lambda t_N^{(2)}(I, J, K)}{N^2}\right)$ converge in distribution to a continuous random variable $\tau(k, j)$ and a continuous random vector $(\tau^{(3)}(I, J, K), \tau^{(2)}(I, J, K))$, respectively, as $N \rightarrow \infty$, whose densities are given by:

$$f_2(\tau) = e^{-\tau} \tag{C4}$$

and:

$$f_3(\tau_3, \tau_2) = 3e^{-(3\tau_3 + \tau_2)} \tag{C5}$$

respectively, for $\tau, \tau_2, \tau_3 > 0$. As in Kroumi and Lessard [25], by conditioning on these coalescence times, we find:

$$\phi_2 = \int_0^\infty \left(\frac{1 + e^{-\theta\tau}}{2}\right) e^{-\tau} d\tau = \frac{\theta + 1}{2\theta + 1} \tag{C6}$$

and:

$$\phi_3 = \int_0^\infty \int_0^\infty \left(\frac{1 + e^{-2\theta\tau_3} + 2e^{-2\theta(\tau_2 + \tau_3)}}{4}\right) 3e^{-(\tau_2 + 3\tau_3)} d\tau_2 d\tau_3 = \frac{\theta + 2}{2(2\theta + 1)} \tag{C7}$$

Then, by symmetry, we have:

$$\mathbb{E}_0[x_k x_j x_i] = \phi_3/2, \quad \mathbb{E}_0[x_k x_j] = \phi_2/2, \quad \mathbb{E}_0[x_k] = 1/2 \tag{C8}$$

for $i, j, k = 1, \dots, d$. Therefore, we obtain:

$$\mathbb{E}_0[x_k x_j (1 - x_i)] = \mathbb{E}_0[x_k (1 - x_j) x_i] = \frac{\phi_2 - \phi_3}{2} = \frac{\theta}{4(2\theta + 1)} \tag{C9}$$

and:

$$\mathbb{E}_0[(1 - x_k) x_j (1 - x_i)] = \mathbb{E}_0[(1 - x_k) (1 - x_j) x_i] = \frac{1 - 2\phi_2 + \phi_3}{2} = \frac{\theta}{4(2\theta + 1)} \tag{C10}$$

for $i, j, k = 1, \dots, d$.

References

1. Nagylaki, T. The strong-migration limit in geographically structured populations. *J. Math. Biol.* **1980**, *9*, 101–114.
2. Notohara, M. The strong-migration limit for the genealogical process in geographically structured populations. *J. Math. Biol.* **1993**, *31*, 115–122.

3. Herbots, H.M. The structured coalescent. In *Progress in Population Genetics and Human Evolution*; Donnelly, P., Tavaré, S., Eds.; Springer: New York, NY, USA, 1997; pp. 231–255.
4. Kingman, J.F.C. On the genealogy of large populations. *J. Appl. Probab.* **1982**, *19*, 27–43.
5. Kingman, J.F.C. Exchangeability and the evolution of large populations. In *Exchangeability in Probability and Statistics*; Koch, G., Spizzichino, F., Eds.; North-Holland Publishing Company: Amsterdam, The Netherlands, 1982; pp. 97–112.
6. Kingman, J.F.C. The coalescent. *Stoch. Process. Their Appl.* **1982**, *13*, 235–248.
7. Axelrod, R.M.; Hamilton, W.D. The evolution of cooperation. *Science* **1981**, *211*, 1390–1396.
8. Szabò, G.; Tóke, C. Evolutionary prisoner’s dilemma on a square lattice. *Phys. Rev. E* **1998**, *58*, 68.
9. Traulsen, A.; Röhl, T.; Schuster, H.G. Stochastic gain in population dynamics. *Phys. Rev. Lett.* **2004**, *93*, 028701.
10. Santos, F.C.; Pacheco, J.M. Scale-free networks provide a unifying framework for the emergence of cooperation. *Phys. Rev. Lett.* **2005**, *95*, 098104.
11. Hauert, C.; Szabò, G. Game theory and physics. *Am. J. Phys.* **2005**, *73*, 405–414.
12. Nowak, M.A. Five rules for the evolution of cooperation. *Science* **2006**, *314*, 1560–1563.
13. Nowak, M.A. *Super Cooperators*; Free Press: New York, NY, USA, 2011.
14. Nowak, M.A.; Sigmund, K. How populations cohere: Five rules for cooperation. In *Theoretical Ecology*, 3rd ed.; May, R., McLean, A., Eds.; Oxford Press: Oxford, UK, 2007.
15. Ohtsuki, H.; Nowak, M.A.; Pacheco, J.M. Breaking the symmetry between interaction and replacement in evolutionary dynamics on graphs. *Phys. Rev. Lett.* **2007**, *98*, 108106.
16. Szabò, G.; Fáth, G. Evolutionary games on graphs. *Phys. Rep.* **2012**, *446*, 97–216.
17. Hamilton, W.D. The genetic of social behaviour I. *J. Theor. Biol.* **1964**, *7*, 1–16.
18. Hamilton, W.D. The genetic of social behaviour II. *J. Theor. Biol.* **1964**, *7*, 17–32.
19. Trivers, R.L. The evolution of reciprocal altruism. *Q. Rev. Biol.* **1971**, *46*, 35–37.
20. Frank, S.A. *Foundations of Social Evolution*; Princeton University Press: Princeton, NJ, USA, 1998.
21. Nowak, M.A.; Sigmund, K. Evolution of indirect reciprocity by image scoring. *Nature* **1998**, *393*, 573–577.
22. Nowak, M.A.; Sigmund, K. The dynamics of indirect reciprocity. *J. Theor. Biol.* **1998**, *194*, 561–571.
23. Nowak, M.A. *Evolutionary Dynamics*; The Belknap Press of Harvard University Press: Cambridge, MA, USA, 2006.
24. Traulsen, A.; Nowak, M.A. Evolution of cooperation by multi-level selection. *Proc. Natl. Acad. Sci. USA* **2006**, *103*, 10952–10955.
25. Kroumi, D.; Lessard, S. Evolution of cooperation in a multidimensional phenotype space. *Theor. Popul. Biol.* **2015**, *102*, 60–75.
26. Tarnita, C.E.; Antal, T.; Ohtsuki, H.; Nowak, M.A. Evolutionary dynamics in set structured populations. *Proc. Natl. Acad. Sci. USA* **2009**, *106*, 8601–8604.
27. Hammerstein, P. The role of asymmetries in animal contests. *Anim. Behav.* **1981**, *29*, 578–581.
28. Wilson, E.O. *Sociobiology: The New Synthesis*; Belknap Press: Cambridge, MA, USA, 1975.

29. Tao, Y.; Xu, J.J.; Li, C.; Cressman, R. Dominance hierarchies induce a population's full cooperation. *Dyn. Games Appl.* **2014**, *4*, 432–447.
30. Kroumi, D.; Lessard, S. Conditions for cooperation to be more abundant than defection in a hierarchically structured population. *Dyn. Games Appl.* **2014**, *5*, 239–262.
31. Möhle, M. A convergence theorem for Markov chains arising in population genetics and the coalescent with selfing. *Adv. Appl. Probab.* **1973**, *30*, 493–512.
32. Nowak, M.A.; Tarnita, C.E.; Antal, T. Evolutionary dynamics in structured populations. *Philos. Trans. R. Soc. B* **2010**, *365*, 19–30.
33. Moran, P.A.P. Random processes in genetics. *Proc. Camb. Philos. Soc.* **1958**, *54*, 60–71.
34. Moran, P.A.P. *Statistical Processes of Evolutionary Theory*; Clarendon Press: Oxford, UK, 1962.
35. Krebs, J.R.; Davies, N.B. *An Introduction to Behaviour Ecology*; Blackwell Publishing: Oxford, UK, 1982.
36. Walters, J.R.; Seyfarth, R.M. Conflict and cooperation. In *Primate Societies*; Smuts, B.B., Cheney, D.L., Seyfarth, R.M., Wrangham, R.W., Struhsaker, T., Eds.; University of Chicago Press: Chicago, IL, USA, 1982; pp. 306–317.
37. Axelrod, R.M. *The Evolution of Cooperation*; Basic Books: New York, NY, USA, 1984.
38. Linhart, S. Some thoughts on the Ken game in Japan: From the viewpoint of comparative civilization studies. *Senri Ethnol. Stud.* **1995**, *40*, 101–124.
39. Harsanyi, J.C.; Selten, R. *A General Theory of Equilibrium Selection in Games*; MIT Press: Cambridge, MA, USA, 1988.
40. Lessard, S. Effective game matrix and inclusive payoff in group-structured populations. *Dyn. Games Appl.* **2011**, *2*, 301–318.
41. Culver, W.J. On the existence and uniqueness of the real logarithm of a matrix. *Proc. Am. Math. Soc.* **1973**, *17*, 1146–1151.
42. Higham, N.J. *Functions of Matrices: Theory and Computation*; Society for Industrial and Applied Mathematics (SIAM): Philadelphia, PA, USA, 2008.

© 2015 by the authors; licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution license (<http://creativecommons.org/licenses/by/4.0/>).