



**HAL**  
open science

# Evolutionary game dynamics in a finite continental island population model and emergence of cooperation

Véronique Ladret

► **To cite this version:**

Véronique Ladret. Evolutionary game dynamics in a finite continental island population model and emergence of cooperation. 2013. hal-00921092v2

**HAL Id: hal-00921092**

**<https://hal.science/hal-00921092v2>**

Preprint submitted on 25 Mar 2014 (v2), last revised 28 Mar 2024 (v4)

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Evolutionary game dynamics in a finite continental island population model and emergence of cooperation

Véronique Ladret \*

Institut de mathématiques et modélisation de Montpellier  
I3M (UMR CNRS UM2) and Université Montpellier 1

March 25, 2014

We consider the continental island model for a finite haploid population with a total number of  $n$  demes consisting in one continent and  $n - 1$  islands. We assume viability differences in the population captured by a linear game within each deme as a result of pairwise interactions. Under the structured coalescent hypothesis, assuming weak selection and conservative migration, we derive the first order approximation for the fixation probability of a single mutant, initially introduced in the continent, with respect to the intensity of selection. When all the demes share the same game matrix, this result is applied to the case of the iterated Prisoner's Dilemma, when the resident strategy is always defect (AllD) and the mutant cooperative strategy is tit-for-tat (TFT). In this case, we investigate conditions under which selection favors the emergence of cooperation and we derive an extension of the "one-third law" of evolution. We find that the population subdivision of the continental island model weakens the one-third law that holds for a panmictic population. When the model is symmetric, i.e. when the continent and the islands are the same size, we compare this condition to the one obtained when the population structure is replaced by a Wright's finite island model with the same number of demes, the same deme sizes and the same expected total number of migrants per generation after population regulation as in the continental island model. We investigate under what conditions the continental island structure facilitates the emergence of the cooperative tit-for-tat strategy in comparison with its Wright's island model counterpart. And when the deme sizes differ, we investigate how the asymmetry in the deme sizes of the continental island model can better promote the evolution of tit-for-tat compared to its equal deme sizes model counterpart.

KEY WORDS : Evolutionary game theory; Finite population, Structured coalescent; Fixation probability; One third law; Iterated Prisoner's Dilemma

---

\**E-mail*: vladret@univ-montp2.fr; *Postal address*: V. Ladret, UM2, Case Courier 051, Place E. Bataillon, 34095 Montpellier CEDEX, France

# 1 Introduction

Evolutionary game theory has traditionally been considered in the context of infinitely large populations to study selection of traits with frequency dependent fitness (Maynard Smith, 1982; Nowak and Sigmund, 2004; Nowak, 2006a). Frequency-dependent selection means that the fitness values of the traits (strategies) depend on the relative abundance of the different strategies in the population. The trait of cooperation is a classic example under frequency dependent selection (Nowak, 2006b). When the populations are infinite, the evolution of the frequencies of the traits in the population are modelled by differential equations like, for example, the wide-spread replicator equations (Taylor and Jonker, 1978; Hofbauer and Sigmund, 1998) and mathematical predictions of the evolution of the system are based on the identification of equilibrium states and the analysis of their stability. Though, in nature, real populations are always finite and the evolutionary processes are prone to stochastic fluctuations, due, in particular, to the finiteness of the population.

When the population is finite, the evolution of a mutant trait may differ greatly from its evolutionary outcome predicted by the replicator dynamics in an infinite population. For instance, in a finite population, any mutant trait will either be eliminated or will take over the whole population due to random drift and there is always a non-zero probability that a mutant fixates in the population, even though it would always be wiped out by selection in an infinitely large population. When analysing evolutionary dynamics in a finite population, an interesting quantity is given by the fixation probability of a single mutant. It has been used by Nowak et al. (2004) to define of an advantageous mutation in a finite population : selection is said to favour the replacement of a resident strategy by a mutant strategy, if the fixation probability of a single mutant is greater than it would be under neutral selection.

From this definition, ensues the one-third law of evolution (which specifies conditions for weak selection to favour the emergence of a wild type strategy in a  $2 \times 2$  coordination game, in the limit of a large finite population). Considering a  $2 \times 2$  evolutionary game with two types  $A$  and  $B$  and payoff matrix

$$\begin{array}{c} A \quad B \\ A \quad \left( \begin{array}{cc} a & b \\ c & d \end{array} \right), \\ B \end{array}$$

where  $a$  is the payoff received by a type  $A$  individual when interacting with another type  $A$ ,  $b$  is the payoff a type  $A$  gets when interacting with a type  $B$ ,  $c$  is the payoff received by a type  $B$  individual when interacting with a type  $A$ , while two interacting type  $B$  individuals each receive  $d$ , and assuming the game is a coordination game, i.e.  $a > c$  and  $d > b$  (which means the two pure strategies  $A$  and  $B$  are strict Nash equilibria), Nowak et al. (2004) proved in the context of a frequency dependent Moran process, that weak selection favours a single type  $A$  mutant replacing a population of type  $B$  individuals in the limit of a large finite well-mixed population, if the unstable internal equilibrium frequency of the wild type strategy in the deterministic evolutionary dynamics in an infinite population,  $x^* = (d - b)/(a - b - c + d)$ ,

is smaller than  $1/3$ . In other words, for a coordination game, provided selection is weak enough and the population size is large enough, the fixation probability of a single mutant strategy is larger than under neutrality if the basin of attraction of the wild type strategy in the traditional replicator dynamics is less than one-third.

The one-third law does not only apply to the frequency dependent Moran process. It has also been obtained for the Wright-Fisher model (Imhof and Nowak, 2006; Lessard, 2005), for pairwise comparison updating (Traulsen et al., 2006), for exchangeable models within the domain of Kingman's coalescence (Lessard, 2007; Lessard and Ladret 2007), for games on graphs (Ohtsuki and Nowak, 2006; Ohtsuki et al., 2006, 2007); it has been extended to the  $d$ -player two-strategies game (Kurokawa and Ihara, 2009; Gokhale and Traulsen, 2010) in a Moran population; and the extension of the one-third law from the 2-player to the  $d$ -player game has been shown to be the same for all exchangeable models in the domain of application of the Kingman coalescent in the limit of a large neutral population (Lessard, 2011).

When applied to the Iterated Prisoner's Dilemma (IPD) game (Trivers, 1971; Axelrod and Hamilton, 1981), with the cooperative 'tit-for-tat' (TFT) strategy as a mutant strategy and 'always defect' (AllD) as a resident strategy, this result enables to investigate how the evolution of cooperation can be facilitated by selection in a finite population. Indeed, when the number of iterations of this game is large enough, it becomes a coordination game with both TFT and AllD being evolutionarily stable in the traditional sense. It then follows from the one-third law that in the limit of a large finite population size, weak selection favors a single mutant TFT cooperator taking over a population of resident AllD defectors if the proportion of TFT at the unstable equilibrium of the deterministic replicator's dynamics is smaller than one-third.

Population subdivision can weaken the one-third law, as it happens in the case of the symmetric Wright's finite island model (Ladret and Lessard, 2007) or the asymmetric two-deme model with unequal deme sizes and different migration rates (Ladret and Lessard, 2008). In this paper, we extend the latter asymmetric two-deme model to a continental island model with  $n$  demes, in which the population is subdivided into  $n$  subpopulations (or demes) : a central deme ("the continent") surrounded by  $n - 1$  identical islands of equal sizes, with migration only occurring between the continent and each island (with no direct island-to-island migration). The size of the continent is possibly different from that of each island and we allow for unequal migration rates from the continent to each island and from each island to the continent. We study the probability of fixation of a single mutant initially introduced into the continent under weak selection, conservative migration and the structured coalescent assumptions (Notohara, 1990; Herbots, 1994, 1997; Wilkinson-Herbots, 1998). The individuals in the population are haploid and we assume that two strategies are genetically encoded at a single locus by two alternative alleles. At the beginning of every generation, each individual produces an infinite equal number of offspring which then disperse independently among the  $n$  demes according to the migration pattern of a continental

island model (Wilkinson Herbots, 1998) (and references therein). Following migration, the offspring undergo viability differences, as a result of linear games within demes, described by deme-specific matrices. We assume that the islands share the same deme-specific matrix which may differ from the game matrix associated to the continent. Then, following selection, the deme sizes are restored by random sampling. We derive an approximation for the first-order effect of selection on the probability of fixation of a single mutant. The method is an extension of the coalescent approach used in Ladret and Lessard (2008), which is adapted from a direct Markov chain method proposed by Rousset (2003) and ascertained in Lessard and Ladret (2007). We then derive conditions under which a mutant strategy is selectively advantageous. When the two strategies are best replies to themselves with respect to both game matrices, we obtain an extension of the one-third law. When the game matrices are identical, we find that this condition can be written in terms of the unstable equilibrium frequency of the mutant type with respect to the deterministic evolutionary dynamics in an infinite panmictic population,  $x^*$ , in the form  $x^* < \Lambda$ , for some threshold  $\Lambda$  that depends on the population structure. We find that this condition is less stringent than the one-third law. We compare this condition in the asymmetric continental island model with its symmetric (equal deme sizes) model counterpart. We find that it is less stringent in the asymmetric model provided the proportion of the population living in the continent is smaller than  $1/n$  (which means that the continent is smaller than any island) if the number of demes is large enough, with the additional requirement that this proportion is not too small for a lesser number of demes. We also compare the symmetric  $n$ -deme continental island model to the Wright's island model (Ladret and Lessard, 2007) that has the same deme sizes, the same number of demes,  $n$ , and the same expected total number of migrants per generation after population regulation as the symmetric  $n$ -deme continental island model. When the total number of demes  $n$  is fixed, we find that the condition under which a mutant strategy is selectively advantageous in the continental island model is less stringent than in its Wright's island model counterpart provided a certain scaled migration rate  $M$  remains above an explicit threshold value that depend on the number of demes  $n$  (namely,  $\frac{\sqrt{2n+4}}{n}$ ). And when the scaled migration rate  $M$  is fixed, we find that it is always the case if  $M \geq \sqrt{10}/3$ , while it otherwise requires that the number of demes remains large enough (namely, larger than  $\frac{1+\sqrt{1+4M^2}}{M^2}$ ).

This paper is organized as follows. Section 2 describes the continental island linear game model. In section 3, we give a first order approximation for the probability of fixation of a single mutant with respect to the intensity of selection, under weak selection. In section 4, this approximation is derived under the structured coalescent assumptions and conservative migration. Section 5 investigates the deterministic evolutionary dynamics of the model in the case when the population is infinite. Section 6 goes back to the finite population model. Conditions under which a mutant strategy is favoured by selection to replace a resident strategy are derived and links with evolutionarily stable strategies in the traditional sense are given. In section 7, the question of the emergence of cooperation is considered in the case when the games are coordination games. When the game matrices are identical, the extended one-third law obtained for the continental island model is compared to the one associated with a panmictic population. The results derived for the asymmetric continental

island model are compared to their symmetric model counterparts. And comparisons are made between the symmetric continental island model and the Wright's finite island model. Discussion and conclusions are provided in section 8.

## 2 The continental island linear-game model

Consider a haploid population divided into a central subpopulation (which will be referred to as the continent) of size  $N_0$  surrounded by  $n - 1$  islands of equal sizes,  $N_1$ . Subpopulation 0 will stand for the continent, while the islands will be labeled subpopulations  $i = 1, \dots, n - 1$ . Suppose there are two types of individuals in the population, type  $A$  and type  $B$ , and that, at time  $t = 0$ , a single mutant  $A$  is introduced into the continent. We assume that the population evolves in discrete, non-overlapping generations represented by time steps  $t = 0, 1, 2, \dots$ . At the beginning of every generation, each individual in the population produces the same very large number of offspring (treated as if it were infinite), which then disperse independently. An offspring from the continent can migrate to each island and an offspring from each island can migrate to the continent. From each island, individuals can only migrate to the continent. Migration between islands is not possible in one step of the migration process, it implies passing through the continent. For  $i \in \{1, \dots, n - 1\}$ , let  $m_{0i}$  denote the proportion of offspring in the central island (the continent) that come from island  $i$ , in the previous generation, and assume those backward migration probabilities to be equal ( $m_{0i} = m_{01}, \forall i \in \{1, \dots, n - 1\}$ ). Thus, if  $m_0 = \sum_{i \in \{1, \dots, n - 1\}} m_{0i}$  denotes the probability that an individual from the continent emigrated from one of the islands (any of them) in the previous generation, it follows that for all  $i \in \{1, \dots, n - 1\}$ ,  $m_{0i} = m_{01} = m_0 / (n - 1)$ . For  $i \in \{1, \dots, n - 1\}$ , let  $m_{i0}$  stand for the proportion of offspring in island  $i$  that come from the continent and assume these migration probabilities from each island to the continent to be the same, i.e.  $m_{i0} = m_{10}, \forall i \in \{1, \dots, n - 1\}$ . For  $i \in \{1, \dots, n - 1\}$ , let  $m_i$  denote the probability that an individual from island  $i$  emigrated from another subpopulation in the previous generation. Since there is no direct island to island migration, this subpopulation has to be the continent and  $m_i = m_1 = m_{10}, \forall i \in \{1, \dots, n - 1\}$ . Moreover, we suppose all the backward migration probabilities to be constant over time. If for  $i \in \{0, \dots, n - 1\}$ ,  $x_i$  denotes the frequency of  $A$  in demes  $i$  at the beginning of a given generation, and  $\tilde{x}_i$  the frequency of  $A$  in the same deme after migration, then we have

$$\tilde{x}_0 = (1 - m_0)x_0 + \frac{m_0}{n - 1} \sum_{i=1}^{n-1} x_i, \quad (1)$$

$$\tilde{x}_i = m_1 x_0 + (1 - m_1)x_i, \quad \forall i \in \{1, \dots, n - 1\}. \quad (2)$$

Migration is then followed by selection among offspring within the same deme. We consider a linear game scenario in which we assume that the two types of individuals,  $A$  and  $B$ , are associated with mixed strategies  $\mathbf{p}_A$  and  $\mathbf{p}_B$ , respectively, these being frequency vectors whose components give the probabilities of using some pure strategies in a contest against an opponent. More precisely, pairwise interactions take place among the individuals within each deme and affect their viabilities (fitnesses) such that they are written, for the continent

and types  $A$  and  $B$ , respectively, in the form :

$$f_{A,0} = 1 + s\mathbf{p}_A \cdot W_0\bar{\mathbf{p}}_0, \quad f_{B,0} = 1 + s\mathbf{p}_B \cdot W_0\bar{\mathbf{p}}_0, \quad (3)$$

and for island  $i$  and type  $A$  and  $B$ , respectively, in the form

$$f_{A,i} = 1 + s\mathbf{p}_A \cdot W_i\bar{\mathbf{p}}_i, \quad f_{B,i} = 1 + s\mathbf{p}_B \cdot W_i\bar{\mathbf{p}}_i, \quad (4)$$

where  $W_0$  and  $W_1 = W_2 = \dots = W_{n-1}$  refer to some game matrices, parameter  $s > 0$  measures the selection intensity,  $\mathbf{x} \cdot \mathbf{y}$  denotes the scalar product of two vectors  $\mathbf{x}$  and  $\mathbf{y}$ , and  $\bar{\mathbf{p}}_i$  stands for the mean strategy in deme  $i$  before selection, that is,

$$\bar{\mathbf{p}}_i = \tilde{x}_i\mathbf{p}_A + (1 - \tilde{x}_i)\mathbf{p}_B = \tilde{x}_i(\mathbf{p}_A - \mathbf{p}_B) + \mathbf{p}_B, \quad (5)$$

for  $i = 0, \dots, n-1$ . Following selection, the frequency of  $A$  among the individuals in deme  $i$  becomes

$$x'_i = \frac{\tilde{x}_i f_{A,i}}{\tilde{x}_i f_{A,i} + (1 - \tilde{x}_i) f_{B,i}}. \quad (6)$$

The next generation is obtained by drawing at random  $N_0$  individuals in the continent and  $N_1$  individuals in each island, respectively, so that the number of  $A$  individuals in the continent follows a binomial distribution with parameters  $N_0$  and  $x'_0$ , and the number of  $A$  individuals in island  $i$ , a binomial distribution with parameters  $N_1$  and  $x'_i$ . After this population regulation, the frequency of  $A$  in deme  $i$  has mean  $x'_i$ , for  $i = \{0, \dots, n-1\}$ . Note that in the special case when the number of demes is reduced to  $n = 2$ , this model reduces to the asymmetric two-deme model studied in Ladret and Lessard (2008). Throughout the paper we focus on the weak selection scenario in which the selection intensity  $s$  is small, that is,  $s = o(1)$ .

### 3 Fixation coefficient

For  $i = 0, \dots, n-1$ , let  $X_i(t)$  denote the frequency of  $A$  in deme  $i$  at time  $t$  and define

$$X(t) = (X_0(t), \dots, X_{n-1}(t)).$$

Since a single mutant  $A$  is introduced into the continent at time  $t = 0$ , it follows that  $X(0) = \mathbf{x}_0$ , where

$$\mathbf{x}_0 = (1/N_0, 0, \dots, 0).$$

We are interested in  $u(s)$ , the probability of fixation of a single mutant  $A$  introduced at time  $t = 0$  into the continent in a population of  $B$  individuals when selection intensity is  $s$ . More specifically, we calculate the first order development of  $u(s)$  with respect to  $s$ , in the case of weak selection, given by

$$u(s) = u(0) + su'(0) + o(s),$$

where  $u'(0)$  is the derivative of  $u$  at  $s = 0$ . The term  $u'(0)$  will be referred to as the fixation coefficient. The method we use to calculate this coefficient is an adaptation of a method introduced by Rousset (2003) and ascertained in Lessard and Ladret (2007). In the case of a panmictic population (Lessard, 2005, 2007) or a symmetric population structure (Rousset and Billiard, 2000; Ladret and Lessard, 2007), it allows to calculate the fixation coefficient in terms of expected coalescence times, under neutrality of samples of individuals. In the case of a less symmetric population structure, like the one of the asymmetric two-deme linear-game model, this method had to be adapted using an extension of the direct Markov chain approach (Ladret and Lessard, 2008). Here, we adapt this latter approach to the continental island linear-game model.

Let  $E_0$  ( $P_0$ , respectively) denote the expectation (probability, respectively) in the case of neutrality, that is when  $s = 0$  and let  $E$  ( $P$ , respectively) denote the expectation (probability, respectively), in the case of an unspecified selection intensity  $s$ . Let  $\mathbf{M}$  be the matrix whose entries are the backward migration probabilities, that is

$$\mathbf{M} = \begin{bmatrix} m_{00} & m_{01} & \dots & m_{0n-1} \\ m_{10} & m_{11} & \dots & m_{1n-1} \\ \dots & & & \\ m_{i0} & m_{i1} & \dots & m_{in-1} \\ \dots & & & \\ m_{n-10} & m_{n-11} & \dots & m_{n-1n-1} \end{bmatrix},$$

which, in this continental island model, reduces to

$$\mathbf{M} = \begin{bmatrix} 1 - (n-1)m_{01} & m_{01} & m_{01} & \dots & m_{01} & m_{01} & m_{01} \\ m_{10} & 1 - m_{10} & 0 & \dots & 0 & 0 & 0 \\ m_{10} & 0 & 1 - m_{10} & 0 & \dots & 0 & 0 \\ m_{10} & 0 & 0 & 1 - m_{10} & 0 \dots & & 0 \\ \dots & & & & & & \\ m_{10} & 0 & \dots & & 0 & 1 - m_{10} & 0 \\ m_{10} & 0 & \dots & & 0 & 0 & 1 - m_{10} \end{bmatrix},$$

or, equivalently,

$$\mathbf{M} = \begin{bmatrix} 1 - m_0 & m_0/(n-1) & m_0/(n-1) & \dots & m_0/(n-1) & m_0/(n-1) & m_0/(n-1) \\ m_1 & 1 - m_1 & 0 & \dots & 0 & 0 & 0 \\ m_1 & 0 & 1 - m_1 & 0 & \dots & 0 & 0 \\ m_1 & 0 & 0 & 1 - m_1 & 0 \dots & & 0 \\ \dots & & & & & & \\ m_1 & 0 & \dots & & 0 & 1 - m_1 & 0 \\ m_1 & 0 & \dots & & 0 & 0 & 1 - m_1 \end{bmatrix}.$$

Let  $\mathbf{u} = (u_0, \dots, u_{n-1})$  with  $u_i > 0$ , for all  $i = 0, \dots, n-1$ , and  $\sum_{i=0}^{n-1} u_i = 1$  denote the



stationary distribution of  $\mathbf{M}$  so that

$$[u_0, \dots, u_{n-1}] \mathbf{M} = [u_0, \dots, u_{n-1}]. \quad (7)$$

Thus,

$$u_0 = \frac{m_1}{m_1 + m_0} \quad \text{and} \quad u_1 = u_2 = \dots = u_{n-1} = \frac{m_0}{(m_0 + m_1)(n-1)}.$$

As in Ladret and Lessard (2008), consider  $Z(t)$ , the weighted frequency of  $A$ , defined as

$$Z(t) = u_0 X_0(t) + \dots + u_{n-1} X_{n-1}(t), \quad (8)$$

for  $t \geq 0$ .

For any given selection intensity  $s \geq 0$ , the sequence  $Z(t)$ , for  $t = 0, 1, \dots$  is a discrete-time stochastic process on the state space  $\{u_0 \frac{k_0}{N_0} + u_1 \frac{k_1}{N_1} + \dots + u_{n-1} \frac{k_{n-1}}{N_1}\} = \{u_0 \frac{k_0}{N_0} + u_1 \frac{\sum_{i=1}^{n-1} k_i}{N_1}\}$ , where  $k_0 = 0, \dots, N_0$  and  $k_i = 0, \dots, N_1$  for  $i = 1, \dots, n-1$ , with initial state  $Z(0) = u_0/N_0$  and absorbing states  $z = 0$  and  $z = 1$ , while all other states are transient. This process converges in probability to a random variable  $Z(\infty)$ , which takes value 1 with probability  $u(s)$ , and 0 otherwise. In the neutral scenario ( $s = 0$ ), this process is a bounded martingale. By the stopping time theorem (see, e.g. Karlin and Taylor, 1975), we find that the probability of fixation of  $A$ , which occurs when the absorbing state  $z = 1$  is reached, is equal to

$$u(0) = E_0(Z(\infty)) = Z(0) = \frac{u_0}{N_0}.$$

In the general case when selection intensity is  $s$ , we can write following Rousset (2003) that

$$E[Z(\infty) - Z(0)] = \sum_{t \geq 0} E[Z(t+1) - Z(t)],$$

which is equivalent to

$$u(s) - u(0) = \sum_{t \geq 0} E[Z(t+1) - Z(t)].$$

After differentiating with respect to  $s$  and assuming the interchangeability of summation and derivation (see Lessard and Ladret, 2007, for a formal proof under mild regularity conditions) we get

$$u'(0) = \sum_{t \geq 0} \frac{d}{ds} E[Z(t+1) - Z(t)] \Big|_{s=0}$$

Conditioning on the values  $\mathbf{x} = (x_0, x_1, \dots, x_{n-1})$  taken by  $\mathbf{X}(t)$ , we can write

$$E[Z(t+1) - Z(t)] = \sum_{\mathbf{x}} E[Z(t+1) - Z(t) | \mathbf{X}(t) = \mathbf{x}] P(\mathbf{X}(t) = \mathbf{x}), \quad (9)$$

which implies that

$$\begin{aligned} \frac{d}{ds} E[Z(t+1) - Z(t)] &= \sum_{\mathbf{x}} \frac{d}{ds} E[Z(t+1) - Z(t) | \mathbf{X}(t) = \mathbf{x}] \Big|_{s=0} P_0(\mathbf{X}(t) = \mathbf{x}) \\ &+ \sum_{\mathbf{x}} E_0[Z(t+1) - Z(t) | \mathbf{X}(t) = \mathbf{x}] \frac{d}{ds} P(\mathbf{X}(t) = \mathbf{x}) \Big|_{s=0}. \end{aligned} \quad (10)$$

Under neutrality,  $Z(t)$  does not change in mean from one generation to the next, that is,

$$E_0[Z(t+1) - Z(t) | \mathbf{X}(t) = \mathbf{x}] = 0. \quad (11)$$

Therefore,

$$u'(0) = \sum_{t \geq 0} \sum_{\mathbf{x}} P_0(\mathbf{X}(t) = \mathbf{x}) \frac{d}{ds} E[Z(t+1) - Z(t) | \mathbf{X}(t) = \mathbf{x}] \Big|_{s=0}. \quad (12)$$

Since

$$E[Z(t+1) - Z(t) | \mathbf{X}(t) = \mathbf{x}] = \sum_{i=0}^{n-1} u_i \left( E[X_i(t+1) | \mathbf{X}(t) = \mathbf{x}] - x_i \right) \quad (13)$$

it follows from (??), and (??) that

$$\frac{d}{ds} E[Z(t+1) - Z(t) | \mathbf{X}(t) = \mathbf{x}] \Big|_{s=0} = \sum_{i=0}^{n-1} u_i \tilde{x}_i (1 - \tilde{x}_i) (\mathbf{p}_A - \mathbf{p}_B) \cdot W_i \bar{\mathbf{p}}_i \quad (14)$$

and from (??) that

$$\begin{aligned} \frac{d}{ds} E[Z(t+1) - Z(t) | \mathbf{X}(t) = \mathbf{x}] \Big|_{s=0} &= \sum_{i=0}^{n-1} u_i \tilde{x}_i^2 (1 - \tilde{x}_i) (p_A - p_B) \cdot W_i (p_A - p_B) \\ &\quad + \sum_{i=0}^{n-1} u_i \tilde{x}_i (1 - \tilde{x}_i) (p_A - p_B) \cdot W_i p_B. \end{aligned} \quad (15)$$

From (??), it follows that

$$\begin{aligned} \tilde{x}_0(1 - \tilde{x}_0) &= (1 - m_0)^2 x_0(1 - x_0) + m_0 \frac{(1 - m_0)}{n-1} \sum_{i=1}^{n-1} x_0(1 - x_i) + \frac{m_0(1 - m_0)}{n-1} (1 - x_0) \sum_{i=1}^{n-1} x_i \\ &\quad + \left( \frac{m_0}{n-1} \right)^2 \sum_{i,j \in \{1, \dots, n-1\}} x_i(1 - x_j); \end{aligned}$$

and

$$\begin{aligned} \tilde{x}_0^2(1 - \tilde{x}_0) &= (1 - m_0)^3 x_0^2(1 - x_0) + m_0 \frac{(1 - m_0)^2}{n-1} \sum_{i=1}^{n-1} x_0^2(1 - x_i) + 2(1 - m_0)^2 \frac{m_0}{n-1} x_0(1 - x_0) \sum_{i=1}^{n-1} x_i \\ &\quad + 2(1 - m_0) \left( \frac{m_0}{n-1} \right)^2 x_0 \sum_{i,j \in \{1, \dots, n-1\}} x_i(1 - x_j) + \left( \frac{m_0}{n-1} \right)^2 (1 - m_0)(1 - x_0) \sum_{i,j \in \{1, \dots, n-1\}} x_i x_j \\ &\quad + \left( \frac{m_0}{n-1} \right)^3 \sum_{i,j,k \in \{1, \dots, n-1\}} x_i x_j (1 - x_k). \end{aligned}$$

We derive from (??) that, for all  $i \in \{1, \dots, n-1\}$ ,

$$\tilde{x}_i(1 - \tilde{x}_i) = m_1^2 x_0(1 - x_0) + m_1(1 - m_1)x_0(1 - x_i) + m_1(1 - m_1)x_i(1 - x_0) + (1 - m_1)^2 x_i(1 - x_i);$$

and

$$\begin{aligned}\tilde{x}_i^2(1 - \tilde{x}_i) &= m_1^3 x_0^2(1 - x_0) + m_1^2(1 - m_1)x_0^2(1 - x_i) + 2m_1^2(1 - m_1)x_0(1 - x_0)x_i \\ &\quad + 2m_1(1 - m_1)^2 x_0 x_i(1 - x_i) + m_1(1 - m_1)^2 x_i^2(1 - x_0) + (1 - m_1)^3 x_i^2(1 - x_i).\end{aligned}$$

Thus, from (??) and (??), we find that

$$\begin{aligned}u'(0) &= u_0 \sum_{t \geq 0} \lambda_0(\mathbf{m}_0, \mathbf{X}(t))(p_A - p_B) \cdot W_0(p_A - p_B) + \sum_{i=1}^{n-1} u_i \sum_{t \geq 0} \lambda_i(\mathbf{m}_1, \mathbf{X}(t))(p_A - p_B) \cdot W_1(p_A - p_B) \\ &\quad + u_0 \sum_{t \geq 0} \delta_0(\mathbf{m}_0, \mathbf{X}(t))(p_A - p_B) \cdot W_0 p_B + \sum_{i=1}^{n-1} u_i \sum_{t \geq 0} \delta_i(\mathbf{m}_1, \mathbf{X}(t))(p_A - p_B) \cdot W_1 p_B,\end{aligned}\quad (16)$$

where

$$\begin{aligned}\lambda_0(\mathbf{m}_0, \mathbf{X}(t)) &= (1 - m_0)^3 E_0[X_0(t)^2(1 - X_0(t))] + m_0 \frac{(1 - m_0)^2}{n - 1} \sum_{i=1}^{n-1} E_0[X_0^2(t)(1 - X_i(t))] \\ &\quad + 2 \frac{m_0(1 - m_0)^2}{n - 1} \sum_{i=1}^{n-1} E_0[X_0(t)(1 - X_0(t))X_i(t)] \\ &\quad + 2(1 - m_0) \left(\frac{m_0}{n - 1}\right)^2 \sum_{i \neq j \in \{1, \dots, n-1\}} E_0[X_0(t)X_i(t)(1 - X_j(t))] \\ &\quad + 2(1 - m_0) \left(\frac{m_0}{n - 1}\right)^2 \sum_{i \in \{1, \dots, n-1\}} E_0[X_0(t)X_i(t)(1 - X_i(t))] \\ &\quad + (1 - m_0) \left(\frac{m_0}{n - 1}\right)^2 \sum_{i=1}^{n-1} E_0[X_i^2(t)(1 - X_0(t))] \\ &\quad + (1 - m_0) \left(\frac{m_0}{n - 1}\right)^2 \sum_{i \neq j \in \{1, \dots, n-1\}} E_0[X_i(t)X_j(t)(1 - X_0(t))] \\ &\quad + \left(\frac{m_0}{n - 1}\right)^3 \sum_{i=1}^{n-1} E_0[X_i^2(t)(1 - X_i(t))] + \left(\frac{m_0}{n - 1}\right)^3 \sum_{i, k \in \{1, \dots, n-1\}, k \neq i} E_0[X_i^2(t)(1 - X_k(t))] \\ &\quad + 2 \left(\frac{m_0}{n - 1}\right)^3 \sum_{i, j \in \{1, \dots, n-1\}, j \neq i} E_0[X_i(t)X_j(t)(1 - X_i(t))] \\ &\quad + \left(\frac{m_0}{n - 1}\right)^3 \sum_{i, j, k \in \{1, \dots, n-1\}, i \neq j, j \neq k, k \neq i} E_0[X_i(t)X_j(t)(1 - X_k(t))];\end{aligned}$$

$$\begin{aligned}\lambda_i(\mathbf{m}_1, \mathbf{X}(t)) &= m_1^3 E_0[X_0(t)^2(1 - X_0(t))] + m_1^2(1 - m_1)E_0[X_0(t)^2(1 - X_i(t))] \\ &\quad + 2m_1^2(1 - m_1)E_0[X_0(t)(1 - X_0(t))X_i(t)] + 2m_1(1 - m_1)^2 E_0[X_0(t)X_i(t)(1 - X_i(t))] \\ &\quad + m_1(1 - m_1)^2 E_0[X_i(t)^2(1 - X_0(t))] \\ &\quad + (1 - m_1)^3 E_0[X_i^2(t)(1 - X_i(t))];\end{aligned}$$

$$\begin{aligned}
\delta_0(\mathbf{m}_0, \mathbf{X}(t)) &= (1 - m_0)^2 E_0[X_0(t)(1 - X_0(t))] + m_0 \frac{(1 - m_0)}{n - 1} \sum_{i=1}^{n-1} E_0[X_0(t)(1 - X_i(t))] \\
&+ \frac{m_0(1 - m_0)}{n - 1} \sum_{i=1}^{n-1} E_0[X_i(t)(1 - X_0(t))] + \left(\frac{m_0}{n - 1}\right)^2 \sum_{i=1}^{n-1} E_0[X_i(t)(1 - X_i(t))] \\
&+ \left(\frac{m_0}{n - 1}\right)^2 \sum_{i,j \in \{1, \dots, n-1\}, i \neq j} E_0[X_i(t)(1 - X_j(t))]
\end{aligned}$$

and

$$\begin{aligned}
\delta_i(\mathbf{m}_1, \mathbf{X}(t)) &= m_1^2 E_0[X_0(t)(1 - X_0(t))] + m_1(1 - m_1) E_0[X_0(t)(1 - X_i(t))] \\
&+ m_1(1 - m_1) E_0[X_i(t)(1 - X_0(t))] + (1 - m_1)^2 E_0[X_i(t)(1 - X_i(t))].
\end{aligned}$$

Since  $u_i = (1 - u_0)/(n - 1)$ , for all  $i \in \{1, \dots, n - 1\}$ , and due to the symmetry of the model (which implies that  $E_0[X_i(t)^2(1 - X_i(t))] = E_0[X_1(t)^2(1 - X_1(t))]$  for all  $i \in \{1, \dots, n - 1\}$ ), we get

$$\begin{aligned}
\lambda_0(\mathbf{m}_0, \mathbf{X}(t)) &= (1 - m_0)^3 E_0[X_0(t)^2(1 - X_0(t))] + m_0(1 - m_0)^2 E_0[X_0^2(t)(1 - X_1(t))] \\
&+ 2m_0(1 - m_0)^2 E_0[X_0(t)(1 - X_0(t))X_1(t)] \\
&+ 2(1 - m_0)m_0^2 \left(\frac{n - 2}{n - 1}\right) E_0[X_0(t)X_1(t)(1 - X_2(t))] \\
&+ 2(1 - m_0) \left(\frac{m_0^2}{n - 1}\right) E_0[X_0(t)X_1(t)(1 - X_1(t))] \\
&+ (1 - m_0) \left(\frac{m_0^2}{n - 1}\right) E_0[X_1^2(t)(1 - X_0(t))] \\
&+ (1 - m_0)m_0^2 \left(\frac{n - 2}{n - 1}\right) E_0[X_1(t)X_2(t)(1 - X_0(t))] \\
&+ \frac{m_0^3}{(n - 1)^2} E_0[X_1^2(t)(1 - X_1(t))] + \frac{m_0^3(n - 2)}{(n - 1)^2} E_0[X_1^2(t)(1 - X_2(t))] \\
&+ 2m_0^3 \frac{(n - 2)}{(n - 1)^2} E_0[X_1(t)X_2(t)(1 - X_1(t))] \\
&+ m_0^3 \frac{(n - 2)(n - 3)}{(n - 1)^2} E_0[X_1(t)X_2(t)(1 - X_3(t))], \tag{17}
\end{aligned}$$

$$\begin{aligned}
\lambda_i(\mathbf{m}_1, \mathbf{X}(t)) &= m_1^3 E_0[X_0(t)^2(1 - X_0(t))] + m_1^2(1 - m_1) E_0[X_0(t)^2(1 - X_1(t))] \\
&+ 2m_1^2(1 - m_1) E_0[X_0(t)(1 - X_0(t))X_1(t)] + 2m_1(1 - m_1)^2 E_0[X_0(t)X_1(t)(1 - X_1(t))] \\
&+ m_1(1 - m_1)^2 E_0[X_1(t)^2(1 - X_0(t))] + (1 - m_1)^3 E_0[X_1^2(t)(1 - X_1(t))] \tag{18}
\end{aligned}$$

and

$$\begin{aligned}
\delta_0(\mathbf{m}_0, \mathbf{X}(t)) &= (1 - m_0)^2 E_0[X_0(t)(1 - X_0(t))] + m_0(1 - m_0) E_0[X_0(t)(1 - X_1(t))] \\
&\quad + m_0(1 - m_0) E_0[X_1(t)(1 - X_0(t))] + \frac{m_0^2}{n - 1} E_0[X_1(t)(1 - X_1(t))] \\
&\quad + m_0^2 \binom{n - 2}{n - 1} E_0[X_1(t)(1 - X_2(t))], \tag{19}
\end{aligned}$$

$$\begin{aligned}
\delta_i(\mathbf{m}_1, \mathbf{X}(t)) &= m_1^2 E_0[X_0(t)(1 - X_0(t))] + m_1(1 - m_1) E_0[X_0(t)(1 - X_1(t))] \\
&\quad + m_1(1 - m_1) E_0[X_1(t)(1 - X_0(t))] + (1 - m_1)^2 E_0[X_1(t)(1 - X_1(t))]. \tag{20}
\end{aligned}$$

From (??) and (??), it follows that the calculation of coefficient  $u'(0)$  requires the calculation of

$$\sum_{t \geq 0} E_0[X_i(t)X_j(t)(1 - X_k(t))],$$

for  $i, j, k \in \{0, 1, 2, 3\}$  and

$$\sum_{t \geq 0} E_0[X_i(t)X_j(t)(1 - X_k(t))],$$

for  $i, j \in \{0, 1, 2\}$ . Therefore, we need to calculate of the following terms  $E_0[X_0^2(t)(1 - X_0(t))]$ ;  $E_0[X_0^2(t)(1 - X_1(t))]$ ;  $E_0[X_0(t)(1 - X_0(t))X_1(t)]$ ;  $E_0[X_0(t)X_1(t)(1 - X_2(t))]$ ;  $E_0[X_1^2(t)(1 - X_1(t))]$ ;  $E_0[X_1^2(t)(1 - X_0(t))]$ ;  $E_0[X_1(t)X_0(t)(1 - X_1(t))]$ ;  $E_0[X_1(t)X_2(t)(1 - X_0(t))]$ ;  $E_0[X_1(t)(1 - X_1(t))X_2(t)]$ ;  $E_0[X_1^2(t)(1 - X_2(t))]$ ;  $E_0[X_1(t)X_2(t)(1 - X_3(t))]$ ;  $E_0[X_0(t)(1 - X_0(t))]$ ;  $E_0[X_0(t)(1 - X_1(t))]$ ;  $E_0[X_1(t)(1 - X_0(t))]$ ;  $E_0[X_0(t)X_1(t)(1 - X_2(t))]$ ;  $E_0[X_1^2(t)(1 - X_1(t))]$ ;  $E_0[X_1^2(t)(1 - X_0(t))]$ ;  $E_0[X_1(t)(1 - X_1(t))]$ ; and  $E_0[X_1^2(t)(1 - X_2(t))]$ .

For each  $k \in \{1, \dots, N_1\}$  and  $i \in \{0, \dots, n - 1\}$ , let  $\xi_{k,i}$  denote the random variable that assigns the value 1 to the  $k$ -th individual in deme  $i$  if it is of type  $A$  and the value 0, otherwise :

$$\xi_{k,i}(t) = \begin{cases} 1 & \text{if individual } k \text{ in deme } i \text{ at time } t \text{ is of type } A, \\ 0 & \text{otherwise.} \end{cases} \tag{21}$$

Then, the frequency of  $A$  in the continent at time  $t$  is

$$X_0(t) = \frac{1}{N_0} \sum_{k=1}^{N_0} \xi_{k,0}(t). \tag{22}$$

It follows that

$$\begin{aligned}
E_0[X_0^2(t)(1 - X_0(t))] &= \frac{1}{N_0^3} E_0 \left[ \left( \sum_{k=1}^{N_0} \xi_{k,0}(t) \right)^2 \left( \sum_{l=1}^{N_0} (1 - \xi_{l,0}(t)) \right) \right] \\
&= \frac{1}{N_1^3} \sum_{k=1}^{N_1} \sum_{l=1}^{N_1} \sum_{m=1}^{N_1} E_0[\xi_{k,0}(t)\xi_{l,0}(t)(1 - \xi_{m,0}(t))] \\
&= \left(1 - \frac{1}{N_0}\right) \left(1 - \frac{2}{N_0}\right) \alpha_0(t) + \frac{1}{N_0} \left(1 - \frac{1}{N_0}\right) \beta_0(t), \quad (23)
\end{aligned}$$

where

$$\alpha_0(t) = P_0(\xi_{1,0}(t) = \xi_{2,0}(t) = 1, \xi_{3,0}(t) = 0), \quad (24)$$

and

$$\beta_0(t) = P_0(\xi_{1,0}(t) = 1, \xi_{2,0}(t) = 0). \quad (25)$$

Similarly, we get that

$$\begin{aligned}
E_0[X_1^2(t)(1 - X_1(t))] &= \left(1 - \frac{1}{N_1}\right) \left(1 - \frac{2}{N_1}\right) P_0(\xi_{1,1}(t) = \xi_{2,1}(t) = 1, \xi_{3,1}(t) = 0) \\
&\quad + \frac{1}{N_1} \left(1 - \frac{1}{N_1}\right) P_0(\xi_{1,1}(t) = 1, \xi_{2,1}(t) = 0), \quad (26)
\end{aligned}$$

$$\begin{aligned}
E_0[X_0^2(t)(1 - X_1(t))] &= \left(1 - \frac{1}{N_0}\right) P_0(\xi_{1,0}(t) = \xi_{2,0}(t) = 1, \xi_{1,1}(t) = 0) \\
&\quad + \frac{1}{N_0} P_0(\xi_{1,0}(t) = 1, \xi_{1,1}(t) = 0), \quad (27)
\end{aligned}$$

$$\begin{aligned}
E_0[X_1(t)^2(1 - X_0(t))] &= \left(1 - \frac{1}{N_1}\right) P_0(\xi_{1,1}(t) = \xi_{2,1}(t) = 1, \xi_{1,0}(t) = 0) \\
&\quad + \frac{1}{N_1} P_0(\xi_{1,1}(t) = 1, \xi_{1,0}(t) = 0), \quad (28)
\end{aligned}$$

$$E_0[X_0(t)X_1(t)(1 - X_0(t))] = \left(1 - \frac{1}{N_0}\right) P_0(\xi_{1,0}(t) = 1, \xi_{2,0}(t) = 0, \xi_{1,1}(t) = 1), \quad (29)$$

$$E_0[X_0(t)X_1(t)(1 - X_1(t))] = \left(1 - \frac{1}{N_1}\right) P_0(\xi_{1,0}(t) = \xi_{1,1}(t) = 1, \xi_{2,1}(t) = 0), \quad (30)$$

$$\begin{aligned}
E_0[X_1^2(t)(1 - X_2(t))] &= \left(1 - \frac{1}{N_1}\right) P_0(\xi_{1,1}(t) = \xi_{2,1}(t) = 1, \xi_{1,2}(t) = 0) \\
&\quad + \frac{1}{N_1} P_0(\xi_{1,1}(t) = 1, \xi_{1,2}(t) = 0), \quad (31)
\end{aligned}$$

$$E_0[X_2(t)X_1(t)(1 - X_1(t))] = \left(1 - \frac{1}{N_1}\right) P_0(\xi_{1,1}(t) = 1, \xi_{2,1}(t) = 0, \xi_{1,2}(t) = 1), \quad (32)$$

$$E_0[X_0(t)X_1(t)(1 - X_2(t))] = P_0(\xi_{1,0}(t) = \xi_{1,1}(t) = 1, \xi_{1,2}(t) = 0), \quad (33)$$

$$E_0[X_1(t)X_2(t)(1 - X_0(t))] = P_0(\xi_{1,1}(t) = \xi_{1,2}(t) = 1, \xi_{1,0}(t) = 0), \quad (34)$$

$$E_0[X_1(t)X_2(t)(1 - X_3(t))] = P_0(\xi_{1,1}(t) = \xi_{1,2}(t) = 1, \xi_{1,3}(t) = 0), \quad (35)$$

and

$$E_0[X_0(t)(1 - X_0(t))] = \left(1 - \frac{1}{N_0}\right)P_0(\xi_{1,0}(t) = 1, \xi_{2,0}(t) = 0), \quad (36)$$

$$E_0[X_1(t)(1 - X_1(t))] = \left(1 - \frac{1}{N_1}\right)P_0(\xi_{1,1}(t) = 1, \xi_{2,1}(t) = 0). \quad (37)$$

$$E_0[X_0(t)(1 - X_1(t))] = P_0(\xi_{1,0}(t) = 1, \xi_{1,1}(t) = 0). \quad (38)$$

$$E_0[X_1(t)(1 - X_0(t))] = P_0(\xi_{1,1}(t) = 1, \xi_{1,0}(t) = 0). \quad (39)$$

$$E_0[X_1(t)(1 - X_2(t))] = P_0(\xi_{1,1}(t) = 1, \xi_{1,2}(t) = 0). \quad (40)$$

Let us first focus on

$$\alpha_0(t) = P_0(\xi_{1,0}(t) = \xi_{2,0}(t) = 1, \xi_{3,0}(t) = 0), \quad (41)$$

which is the probability that individuals 1, 2 and 3 from deme 0 at time  $t$  are of respective types  $A$ ,  $A$  and  $B$ . Recalling that a single mutant  $A$  was introduced in the continent at time 0, this event will occur if the lineages of individuals 1 and 2 coalesce before time 0 and their common ancestor at time 0 is of type  $A$  and lives on the continent, while no coalescence event occurs before time 0 between the lineages of individual 3 on the one hand and those of individuals 1 and 2 on the other.

The ancestral process that describes the locations of the ancestors of 3 individuals (labeled 1, 2 and 3), assigning them the labels of the demes they live in and accounting for the number of distinct ancestors they have (describing whether they have three distinct ancestors, two distinct ancestors, in which case whether one is common to individuals 1 and 2, 2 and 3 or 1 and 3, or just a single common ancestor) is Markov. Its state space  $S$  can be partitioned into 5 ordered subsets, which, since all the islands are equivalent, can be reduced

to the following :

$$\begin{aligned}
S_{1,2,3} &= (000, 001, 010, 100, 011, 101, 110, 111, 012, 102, 120, 123, 112, 121, 211), \\
S_{12,3} &= (00, 01, 10, 11, 12), \\
S_{13,2} &= (00, 01, 10, 11, 12), \\
S_{23,1} &= (00, 01, 10, 11, 12), \\
S_{123} &= (0, 1).
\end{aligned}$$

The element 101 in  $S_{1,2,3}$ , for instance, means that the ancestors of individuals 1, 2, 3 are all different and in demes 1, 0, 1, respectively, the element 01 in  $S_{12,3}$  that individuals 1, 2 have a common ancestor in the continent different from the ancestor of individual 3, and this ancestor is in island 1, and finally the element 0 in  $S_{123}$  that individuals 1, 2, 3 have a common ancestor, and this ancestor is in the continent.

For  $\tau \geq 0$ , let  $\sigma(\tau)$  denote the state in  $S$  the ancestors of three individuals are in,  $\tau$  generations back. The transition matrix of this Markov chain takes a block form with respect to the above ordered subsets in the corresponding order. Without loss of generality, let

$$K = \begin{bmatrix} F & Q & W & J & R_1 \\ 0 & P & 0 & 0 & R \\ 0 & 0 & P & 0 & R \\ 0 & 0 & 0 & P & R \\ 0 & 0 & 0 & 0 & R_0 \end{bmatrix} \quad (42)$$

be this transition matrix under neutrality. The submatrices  $F$ ,  $Q$  and  $P$  are given in Appendix A.1. Note that the states in  $S_{123}$  are absorbing for this chain, while all other states are transient. The Perron-Frobënus theory for non-negative matrices ascertains in particular that the eigenvalues of  $F$  and  $P$  are all less than 1 in modulus.

Individuals 1, 2 and 3 in deme 0 at time  $t$  are of types  $A$ ,  $A$  and  $B$ , respectively, if a single coalescence occurs from time  $t$  to 0, this event being a coalescence between the lineages of individuals 1 and 2, and the two distinct ancestors at time 0 of these three individuals are of respective types  $A$  (for the shared ancestor of individuals 1 and 2) and  $B$  (for the ancestor of individual 3). This implies that  $\sigma(t)$  must either be in state 00 or 01 while the type of the ancestor common to individuals 1 and 2 is  $A$ . Thus,  $\alpha_0(t)$  is equal to the probability that the process  $\sigma()$ , lies in state 00 or 01 at time  $t$  given that it started in state 000 at time 0, times  $1/N_0$ , which is the frequency of  $A$  in the continent at time 0, which translates as follows

$$\alpha_0(t) = \sum_{\tau=1}^t \left( F^{\tau-1} Q P^{t-\tau} u \right)_{000},$$

where  $u$  is the column vector  $u = (1/N_0, 1/N_0, 0, 0, 0)$ , and index 000 refers to the vector's component that corresponds to the chain starting in state 000. As a consequence, since matrices  $I - F$  and  $I - P$  are invertible, where  $I$  refers to an identity matrix of appropriate



order, if we sum  $\alpha_0(t)$  over  $t \geq 0$ , we find that

$$\sum_{t \geq 0} \alpha_0(t) = \left( (I - F)^{-1} Q (I - P)^{-1} u \right)_{000}.$$

Similarly,

$$\beta_0(t) := P_0(\xi_{1,0}(t) = 1, \xi_{2,0}(t) = 0) = (P^t u)_{00}, \quad (43)$$

leading to

$$\sum_{t \geq 0} \beta_0(t) = \left( (I - P)^{-1} u \right)_{00}.$$

By (??), it implies that

$$\begin{aligned} \sum_{t \geq 0} E_0 [X_0(t)^2 (1 - X_0(t))] &= \left(1 - \frac{1}{N_0}\right) \left(1 - \frac{2}{N_0}\right) \left( (I - F)^{-1} Q (I - P)^{-1} u \right)_{000} \\ &\quad + \frac{1}{N_0} \left(1 - \frac{1}{N_0}\right) \left( (I - P)^{-1} u \right)_{00}. \end{aligned} \quad (44)$$

In the same way, we get the following expressions

$$\begin{aligned} \sum_{t \geq 0} E_0 [X_1(t)^2 (1 - X_1(t))] &= \left(1 - \frac{1}{N_1}\right) \left(1 - \frac{2}{N_1}\right) \left( (I - F)^{-1} Q (I - P)^{-1} u \right)_{111} \\ &\quad + \frac{1}{N_1} \left(1 - \frac{1}{N_1}\right) \left( (I - P)^{-1} u \right)_{11}, \end{aligned} \quad (45)$$

$$\begin{aligned} \sum_{t \geq 0} E_0 [X_0(t)^2 (1 - X_1(t))] &= \left(1 - \frac{1}{N_0}\right) \left( (I - F)^{-1} Q (I - P)^{-1} u \right)_{001} \\ &\quad + \frac{1}{N_0} \left( (I - P)^{-1} u \right)_{01}, \end{aligned} \quad (46)$$

$$\begin{aligned} \sum_{t \geq 0} E_0 [X_1(t)^2 (1 - X_0(t))] &= \left(1 - \frac{1}{N_1}\right) \left( (I - F)^{-1} Q (I - P)^{-1} u \right)_{110} \\ &\quad + \frac{1}{N_1} \left( (I - P)^{-1} u \right)_{10}, \end{aligned} \quad (47)$$

$$\sum_{t \geq 0} E_0 [X_0(t) X_1(t) (1 - X_0(t))] = \left(1 - \frac{1}{N_0}\right) \left( (I - F)^{-1} Q (I - P)^{-1} u \right)_{010}, \quad (48)$$

$$\sum_{t \geq 0} E_0 [X_0(t) X_1(t) (1 - X_1(t))] = \left(1 - \frac{1}{N_1}\right) \left( (I - F)^{-1} Q (I - P)^{-1} u \right)_{011}, \quad (49)$$

$$\begin{aligned} \sum_{t \geq 0} E_0[X_1(t)^2(1 - X_2(t))] &= \left(1 - \frac{1}{N_1}\right) \left( (I - F)^{-1} Q (I - P)^{-1} u \right)_{112} \\ &\quad + \frac{1}{N_1} \left( (I - P)^{-1} u \right)_{12}, \end{aligned} \quad (50)$$

$$\sum_{t \geq 0} E_0[X_1(t)X_2(t)(1 - X_1(t))] = \left(1 - \frac{1}{N_1}\right) \left( (I - F)^{-1} Q (I - P)^{-1} u \right)_{112}, \quad (51)$$

$$\sum_{t \geq 0} E_0[X_0(t)X_1(t)(1 - X_2(t))] = \left( (I - F)^{-1} Q (I - P)^{-1} u \right)_{012}, \quad (52)$$

$$\sum_{t \geq 0} E_0[X_1(t)X_2(t)(1 - X_0(t))] = \left( (I - F)^{-1} Q (I - P)^{-1} u \right)_{120}, \quad (53)$$

$$\sum_{t \geq 0} E_0[X_1(t)X_2(t)(1 - X_3(t))] = \left( (I - F)^{-1} Q (I - P)^{-1} u \right)_{123}, \quad (54)$$

$$E_0[X_0(t)(1 - X_0(t))] = \left(1 - \frac{1}{N_0}\right) \left( (I - P)^{-1} u \right)_{00}, \quad (55)$$

$$E_0[X_1(t)(1 - X_1(t))] = \left(1 - \frac{1}{N_1}\right) \left( (I - P)^{-1} u \right)_{11}, \quad (56)$$

$$E_0[X_0(t)(1 - X_1(t))] = \left( (I - P)^{-1} u \right)_{01}, \quad (57)$$

$$E_0[X_1(t)(1 - X_0(t))] = \left( (I - P)^{-1} u \right)_{10}, \quad (58)$$

$$E_0[X_1(t)(1 - X_2(t))] = \left( (I - P)^{-1} u \right)_{12}. \quad (59)$$

## 4 Fixation coefficient under the structured-coalescent assumptions

In this section, we focus on the structured coalescent scenario (Notohara, 1990; Herbots, 1994, 1997; Wilkinson-Herbots, 1998), in which the total population size,  $cN$ , is assumed to be large, with demes sizes  $N_1 = \dots = N_{n-1} = c_1N$  and  $N_0 = c_0N$  of order  $N$  and  $c = c_0 + (n-1)c_1$ , and backward migration rates of order inversely proportional to  $N$ ,  $m_{01} = M_0/(2N(n-1))$ ,  $m_0 = M_0/2N$  and  $m_1 = \dots = m_{n-1} = M_1/2N$ , so that  $c_0, c_1, c, M_0$  and  $M_1$  are all of order 1.

Using these notations, the stationary distribution  $(u_0, \dots, u_{n-1})$  of migration matrix  $M$  (see (??)), can be written as follows

$$u_0 = \frac{M_1}{M_0 + M_1} \quad \text{and} \quad u_i = \frac{M_0}{(M_0 + M_1)(n-1)} = \frac{(1-u_0)}{n-1}, \quad \text{for } i = 1, \dots, n-1.$$

In the remainder of this article we restrict to the case of conservative migration, which takes place when the relative size of each deme is maintained after migration. Thus, in this model, conservative migration requires that

$$c_0M_0 = c_1(n-1)M_1.$$

Let  $P = c_0/c$  denote the proportion of the total population that lives in the continent and let  $M$  be the "migration rate" defined as

$$M := c_1M_1 = \frac{c_0M_0}{(n-1)}.$$

Consequently, under the conservative migration assumptions, we have

$$M_0 = \frac{(n-1)M}{cP}, \quad M_1 = \frac{M(n-1)}{c(1-P)}, \quad c_0 = cP \quad c_1 = \frac{c(1-P)}{n-1}, \quad \text{and} \quad u_0 = P.$$

Let  $C_0$  denote the constant term of  $u'(0)$  as  $N$  goes to infinity so that

$$u'(0) = C_0 + O\left(\frac{1}{N}\right).$$

By (??) and equations (??) to (??), we find that  $C_0$  can be written as

$$\begin{aligned} C_0 &= (p_A - p_B) \cdot \left( u_0\gamma_0W_0 + (1-u_0)\gamma_1W_1 \right) (p_A - p_B) \\ &\quad + (p_A - p_B) \cdot \left\{ u_0\delta_0W_0 + (1-u_0)\delta_1W_1 \right\} p_B, \end{aligned} \tag{60}$$

where

$$\gamma_i = \lim_{N \rightarrow +\infty} \sum_{t \geq 0} E_0[X_i^2(t)(1 - X_i(t))] \quad \text{and} \quad \delta_i = \lim_{N \rightarrow +\infty} \sum_{t \geq 0} E_0[X_i(t)(1 - X_i(t))],$$

for  $i = 0, 1$ , since all the other terms in (??) are multiplied by migration rates of order  $1/N$  and are thus, of order  $1/N$  or smaller.

From (??) and (??), we obtain, for  $i = 0, 1$

$$\gamma_i = \lim_{N \rightarrow +\infty} \left(1 - \frac{1}{N_i}\right) \left(1 - \frac{2}{N_i}\right) \left( (I - F)^{-1} Q (I - P)^{-1} u \right)_{iii} + \frac{1}{N_i} \left(1 - \frac{1}{N_i}\right) \left( (I - P)^{-1} u \right)_{ii},$$

and from (??) and (??),

$$\delta_i = \lim_{N \rightarrow +\infty} \left(1 - \frac{1}{N_i}\right) \left( (I - P)^{-1} u \right)_{ii}, \quad \text{for } i = 0, 1.$$

The calculation of these limits as functions of the population parameters  $M, n$  and  $P$  can be done using a software like Maple and we find that

$$\delta_0 = \frac{-P + Mn + n - M}{Mn - M + 1 - 2P + nP}$$

$$\delta_1 = \frac{1 - P + Mn - M}{Mn - M + 1 - 2P + nP},$$

while the expressions for  $\gamma_0$  and  $\gamma_1$  are given in Appendix A.2.

Note that, as expected, in the special case when  $n = 2$ , using the proper parameter rescalings, these results match the first order approximation for the fixation probability of a single mutant in an asymmetric two-deme linear game model (Ladret and Lessard, 2008).

On the other hand, note that when  $n = 2$ , assuming identical game matrices ( $W_0 = W_1 = W$ ) and identical deme sizes ( $P = 1/2$ ), (??) leads to

$$C_0 = (p_A - p_B) \cdot W p_B + \left( \frac{1}{3} + \frac{1}{12(M+1)} \right) (p_A - p_B) \cdot W (p_A - p_B),$$

in accordance with the first order approximation of the fixation probability of a single mutant in a (symmetric) finite island linear game model with 2 demes (Ladret and Lessard, 2007, 2008).

## 5 Game dynamics in an infinite continental island model population

The case where the backward migration matrix is constant and the selection intensity,  $s$ , is of order  $1/N$  as  $N$  goes to infinity, corresponds to a strong-migration limit. In this case, it follows from Nagylaki (1980) that, up to a rescaling of time with respect to the effective population size, which depends on the pattern of migration, the frequency of mutant type  $A$  in each deme converges in probability to the frequency of  $A$  in the whole population averaged with respect to the stationary distribution of the migration matrix, denoted by  $Z$  (see equ. ??), and this averaged frequency converges in distribution to the usual diffusion process in a panmictic population.

In this section, we are concerned with the case where the sizes of the continent and of the islands are infinite. Moreover, we suppose that the relative size of the continent is equal to  $0 < P < 1$  and that the migration rates,  $m_0$  and  $m_1$ , are kept constant as the selection intensity  $s$  goes to 0, while time is measured in units of  $1/s$  generations. Note that in the special case when  $n = 2$ , this model corresponds to the infinite two-deme population model considered in Ladret and Lessard (2008). Here, we generalize their approach to study the infinite  $n$ -deme continental island model. Let  $\mathbf{x}(\tau) = (x_0(\tau), x_1(\tau), \dots, x_{n-1}(\tau))$  denote the frequency vector of the mutant type  $A$  in deme 0 and deme  $i$ ,  $i = 1, \dots, n-1$  at time  $\tau = ts$ . Then, it follows from (??) that

$$\mathbf{x}(\tau + s) = \mathbf{M}\mathbf{x}(\tau) + s\Phi(\mathbf{M}\mathbf{x}(\tau)) + o(s), \quad (61)$$

where

$$\Phi(\mathbf{x}) = \begin{bmatrix} \Phi_0(\mathbf{x}) \\ \Phi_1(\mathbf{x}) \\ \dots \\ \Phi_{n-1}(\mathbf{x}) \end{bmatrix}, \quad (62)$$

in which

$$\Phi_i(\mathbf{x}) = x_i(1 - x_i)\{x_i(\mathbf{p}_A - \mathbf{p}_B) \cdot W_i\mathbf{p}_A + (1 - x_i)(\mathbf{p}_A - \mathbf{p}_B) \cdot W_i\mathbf{p}_A\}, \quad (63)$$

for  $\mathbf{x} = (x_0, x_1, \dots, x_{n-1})$ , with  $0 \leq x_i \leq 1$ . In the  $s \rightarrow 0$  limit, we find, assuming continuity, that

$$\mathbf{x}(\tau) = \mathbf{M}\mathbf{x}(\tau). \quad (64)$$

Therefore, we obtain

$$\mathbf{x}(\tau) = z(\tau) \begin{bmatrix} 1 \\ 1 \\ \dots \\ 1 \end{bmatrix}, \quad (65)$$

where

$$z(\tau) = \sum_{i=0}^{n-1} u_i x_i(\tau)$$

is the averaged frequency of  $A$  with respect to the stationary distribution  $(u_0, u_1, \dots, u_{n-1})$  for the backward migration matrix  $\mathbf{M}$ . Now, dotting both sides of Equation (??) with  $(u_0, u_1, \dots, u_{n-1})$  yields

$$\begin{aligned} z(\tau + s) &= z(\tau) + sz(\tau)(1 - z(\tau)) \\ &\times \left\{ (1 - z(\tau))(\mathbf{p}_A - \mathbf{p}_B) \cdot \left( \sum_{i=0}^{n-1} u_i W_i \right) \mathbf{p}_B + z(\tau)(\mathbf{p}_A - \mathbf{p}_B) \cdot \left( \sum_{i=0}^{n-1} u_i W_i \right) \mathbf{p}_A \right\} \\ &+ o(s). \end{aligned} \tag{66}$$

Consequently, if we divide (??) by  $s$  and let  $s$  go to 0, we derive the following differential equation

$$\frac{dz(\tau)}{d\tau} = z(\tau)(1 - z(\tau)) \left\{ (\mathbf{p}_A - \mathbf{p}_B) \cdot \left( \sum_{i=0}^{n-1} u_i W_i \right) \mathbf{p}_B + z(\tau)(\mathbf{p}_A - \mathbf{p}_B) \cdot \left( \sum_{i=0}^{n-1} u_i W_i \right) (\mathbf{p}_A - \mathbf{p}_B) \right\}. \tag{67}$$

This is the modified replicator equation for two types in the case of an infinite  $n$ -deme continental island population. Note that in the special case when  $n = 2$ , (??) reduces to the modified replicator equation established in Ladret and Lessard (2008) in the case of an infinite two-deme population. Note that (??) corresponds to the classical replicator equation (see, e.g., Hofbauer and Sigmund, 1998, and references therein) for a linear game with game matrix  $\sum_{i=0}^{n-1} u_i W_i$  in a panmictic population with  $z(\tau)$  interpreted as the frequency of  $A$  in the population.

An evolutionarily stable strategy (ESS) as originally conceived by Maynard Smith and Price (Maynard-Smith and Price, 1973; Maynard-Smith, 1982) is a strategy  $\mathbf{p}$  which, when almost fixed in the population cannot be invaded by any alternative mutant strategy,  $\mathbf{p}' \neq \mathbf{p}$ , when in a low enough frequency. In fact,  $\mathbf{p}$  is an evolutionarily stable strategy (ESS) if the replicator equation that describes the population dynamics leads to the elimination of the mutant strategy. Therefore, we deduce from Equation (??) that strategy  $\mathbf{p}_B$  is evolutionarily stable for the current continental island linear game model, with game matrices  $(W_0, W_1, \dots, W_{n-1})$  and stationary distribution  $(u_0, u_1, \dots, u_{n-1})$  with respect to the backward migration matrix, if and only if

$$(i) \quad (\mathbf{p}_A - \mathbf{p}_B) \cdot \left( \sum_{i=0}^{n-1} u_i W_i \right) \mathbf{p}_B \leq 0, \tag{68}$$

or, in case of equality in (i),

$$(ii) \quad (\mathbf{p}_A - \mathbf{p}_B) \cdot \left( \sum_{i=0}^{n-1} u_i W_i \right) (\mathbf{p}_A - \mathbf{p}_B) < 0, \tag{69}$$

for every  $\mathbf{p}_A \neq \mathbf{p}_B$ . Thus, owing to (??), we find that  $\mathbf{p}_B$  is an ESS for the game matrix  $W = \sum_{i=0}^{n-1} u_i W_i$  (Maynard Smith and Price, 1973; Maynard Smith, 1974). Note that  $\mathbf{p}_B$  is also convergence stable (Christiansen, 1991, and references therein) which means that strategies successively closer to it can invade a population using any nearby strategy.

## 6 Conditions against replacement in a finite continental island model population

Let us now go back to the finite population model under weak selection. As defined in Nowak et al. (2004), selection favours (opposes, respectively)  $A$  replacing  $B$  if the fixation probability of a single mutant  $A$  is greater (lower, respectively) than its value in the absence of selection, that is,  $u(s) > u_0/N_0$  ( $u(s) < u_0/N_0$ , respectively) for  $s > 0$ . It follows from Equation (??), that provided selection is weak enough and  $\mathbf{p}_A$  is different but close enough to  $\mathbf{p}_B$ , selection opposes  $A$  replacing  $B$  if and only if either

$$(i) \quad (\mathbf{p}_A - \mathbf{p}_B) \cdot (u_0\delta_0W_0 + (1 - u_0)\delta_1W_1)\mathbf{p}_B \leq 0, \quad (70)$$

or, in the case of equality in (i),

$$(ii) \quad (\mathbf{p}_A - \mathbf{p}_B) \cdot (u_0\gamma_0W_0 + (1 - u_0)\gamma_1W_1)(\mathbf{p}_A - \mathbf{p}_B) < 0. \quad (71)$$

This condition generalizes the one obtained in the case of a continental island model with  $n = 2$  demes (Ladret and Lessard, 2008) to the case of an  $n$ -demes continental island model.

Note that if condition (??-??) holds for every  $\mathbf{p}_A \neq \mathbf{p}_B$ , close enough to  $\mathbf{p}_B$ , it will be met for every  $\mathbf{p}_A \neq \mathbf{p}_B$ . However, if (??-??) holds for every  $\mathbf{p}_A \neq \mathbf{p}_B$ , it is not sufficient to ensure that the probability of fixation of  $A$  remains lower than  $u_0/N_0$ , for all  $\mathbf{p}_A \neq \mathbf{p}_B$ . This will occur if  $\mathbf{p}_B$  has all positive components, since in this case, an equality holds in (i) for all  $\mathbf{p}_A \neq \mathbf{p}_B$ . On the other hand, if  $\mathbf{p}_B$  has some null components, then a strict inequality in (i) can hold for some  $\mathbf{p}_A \neq \mathbf{p}_B$  which has at least one positive component corresponding to a null component of  $\mathbf{p}_B$ . Therefore, in case the expression on the left-hand side of inequality (??) is positive,  $\mathbf{p}_A$  has to be close enough to  $\mathbf{p}_B$  to make sure that

$$(\mathbf{p}_A - \mathbf{p}_B) \cdot (u_0\delta_0W_0 + (1 - u_0)\delta_1W_1)\mathbf{p}_B + (\mathbf{p}_A - \mathbf{p}_B) \cdot (u_0\gamma_0W_0 + (1 - u_0)\gamma_1W_1)(\mathbf{p}_A - \mathbf{p}_B) < 0. \quad (72)$$

In other words, as long as selection is weak enough, a resident strategy  $\mathbf{p}_B$  is selectively favored against replacement by any alternative mutant strategy  $\mathbf{p}_A$ , if  $\mathbf{p}_B$  satisfies condition (??-??) and makes use of all pure strategies; or by any alternative mutant strategy  $\mathbf{p}_A$  using the same subset of pure strategies as  $\mathbf{p}_B$  and any mutant strategy  $\mathbf{p}_A$  close enough to  $\mathbf{p}_B$  using some new pure strategies, if  $\mathbf{p}_B$  meets condition (??-??) and uses a strict subset of pure strategies.

Note that, in general,  $u_0\gamma_0W_0 + (1 - u_0)\gamma_1W_1$  differs from  $u_0\delta_0W_0 + (1 - u_0)\delta_1W_1$ . Thus, as already mentioned in Ladret and Lessard (2008) in the particular case of a two-

deme continental island model, if condition (??-??) is met for every  $\mathbf{p}_A \neq \mathbf{p}_B$ , this does not mean that  $\mathbf{p}_B$  is an ESS for some game matrix. In particular, this does not imply that  $\mathbf{p}_B$  is an ESS for the infinite- $n$ -deme continental island linear-game model with game matrices  $(W_0, W_1, \dots, W_1)$  and stationary distribution of the backward migration matrix  $(u_0, \dots, u_{n-1})$ , or equivalently, for the panmictic linear game model with game matrix  $u_0 W_1 + \sum_{i=1}^{n-1} u_i W_1 = u_0 W_0 + (1 - u_0) W_1$  in a panmictic population. (See Ladret and Lessard (2008), for an example of a two-deme continental island model in a finite population with a resident strategy,  $\mathbf{p}_B$ , whose replacement by any other strategy  $\mathbf{p}_A \neq \mathbf{p}_B$  is opposed by weak selection, even though  $\mathbf{p}_B$  is not an ESS for the game matrix  $u_0 W_0 + (1 - u_0) W_1$ ). Conversely, a resident strategy  $\mathbf{p}_B$  that is an ESS for  $u_0 W_0 + (1 - u_0) W_1$ , does not necessarily meet the two-fold condition (??-??) and thus, is not necessarily selectively favored against replacement by any close enough mutant in the case of a finite population following an  $n$ -deme continental island model, as illustrated in another example in Ladret and Lessard (2008) in the special case when  $n = 2$ .

As a consequence, as already specified in the case of a two-deme continental island model with distinct game matrices (Ladret and Lessard, 2008), and unlike what happens in the case of a panmictic population (Lessard, 2005; Lessard and Ladret, 2007) or in the case of a finite island model with identical game matrices, deme sizes and migration rates (Ladret and Lessard, 2007), conditions for strategies to be favored against replacement by weak selection in an  $n$ -deme continental island linear-game model for a finite population, even large, do not generally translate, unless some specific conditions are satisfied as identical game matrices ( $W_0 = W_1 = W$ ), into classical ESS conditions based on a single game matrix.

Note that, even though in the context of the infinite  $n$ -deme continental island model, a strategy  $\mathbf{p}_B$  that is an ESS for  $u_0 W_0 + (1 - u_0) W_1$  is convergence stable (Christiansen, 1991), in the finite population scenario with distinct game matrices ( $W_0 \neq W_1$ ), it may occur that selection even opposes the fixation of mutants that would bring the population closer to a strategy satisfying (??-??) in a population using a close enough strategy. This is illustrated in an example in Ladret and Lessard (2008) in the framework of a two-deme continental island linear games model. On the other hand, if a strategy  $\mathbf{p}_B$  is an ESS for the game matrix  $u_0 W_0 + (1 - u_0) W_1$ , then weak selection will favor a mutant strategy  $\mathbf{p}_A$  closer to it replacing a resident nearby strategy  $\mathbf{p}_C$ , while it may favor  $\mathbf{p}_A$  replacing  $\mathbf{p}_B$ . This is also illustrated in another example in Ladret and Lessard (2008), in the context of a two-deme continental island linear games model.

## 7 Extension of the one-third law

In this section, we focus on the case when both strategies are best replies to themselves (i.e., when they are strict Nash equilibria) with respect to the game matrices, and we investigate conditions under which the replacement of a resident strategy by a mutant strategy is selectively favored. In particular, we see how the one-third law (Nowak et al., 2004) and its



two-deme continental island linear game model extension (Ladret and Lessard, 2008) can be generalized to a continental island linear game model with a total number of  $n$  demes.

## 7.1 Coordination games

Suppose that  $\mathbf{p}_A$  and  $\mathbf{p}_B$  are pure strategies of the form  $\mathbf{p}_A = (1, 0)$  and  $\mathbf{p}_B = (0, 1)$  respectively, with respect to the  $2 \times 2$  game matrices

$$W_0 = \begin{pmatrix} a_0 & b_0 \\ c_0 & d_0 \end{pmatrix} \quad \text{and} \quad W_1 = \begin{pmatrix} a_1 & b_1 \\ c_1 & d_1 \end{pmatrix}, \quad (73)$$

with  $a_0 > c_0$ ,  $a_1 > c_1$ ,  $d_0 > b_0$ , and  $d_1 > b_1$ , which means that  $\mathbf{p}_A$  and  $\mathbf{p}_B$  are best replies to themselves for both  $W_0$  and  $W_1$ . We know from the deterministic replicator equation in an infinite population, (??), that there exists a unique unstable equilibrium at a frequency of  $A$  averaged with respect to the stationary distribution  $(u_0, \dots, u_{n-1})$  of the backward migration matrix, which is equal to

$$z^* = \frac{u_0(d_0 - b_0) + u_1(d_1 - b_1)}{u_0(a_0 - c_0 - b_0 + d_0) + u_1(a_1 - c_1 - b_1 + d_1)}. \quad (74)$$

When the population is finite, selection, provided it is weak enough, will favor  $A$  replacing  $B$  if inequality (??) is reversed, which is equivalent to

$$\tilde{z} = \frac{u_0\delta_0(d_0 - b_0) + (1 - u_0)\delta_1(d_1 - b_1)}{u_0\gamma_0(a_0 - c_0 - b_0 + d_0) + (1 - u_0)\gamma_1(a_1 - c_1 - b_1 + d_1)} < 1. \quad (75)$$

This inequality extends the one-third law to the case of a continental-island linear-game model with  $n$  demes. In the special case  $n = 2$ , (??) reduces to the extended one-third law for an asymmetric two-deme linear game model (Ladret and Lessard, 2008).

Note that, as already mentioned in the special case of a continental island model with  $n = 2$  demes (Ladret and Lessard, 2008), condition (??) does not translate in terms of the unstable frequency of  $A$  in an infinite population,  $z^*$ , in the form  $z^* < C$ , in contrast to what happens in a panmictic population (we refer to Nowak et al., 2004, for the case of a Moran model; Lessard, 2005, and Imhof and Nowak, 2006, for a Wright-Fisher population model; Lessard, 2007, for a more general model of reproduction in which a fraction of the population is replaced at discrete time steps; Traulsen et al., 2006, for pairwise comparison updating; Lessard and Ladret, 2007, for an extended Cannings exchangeable model), or in the case of a symmetric island model with identical game matrices (Ladret and Lessard, 2007).

Consequently, in general, when considering an  $n$ -deme continental island model with different game matrices,  $W_0 \neq W_1$ , and two strategies  $\mathbf{p}_A$  and  $\mathbf{p}_B$  that are best replies to themselves with respect to both  $W_0$  and  $W_1$ , the conditions under which a single type

A mutant is selectively advantageous in a finite population can not be directly related to conditions regarding the unstable equilibrium frequency of  $A$  (or, equivalently, in terms of the basin of attraction of  $B$ ) in an infinite population.

## 7.2 Coordination games with identical game matrices

Let us focus on the case when the game matrices are identical, that is,

$$W_0 = W_1 = W = \begin{pmatrix} a & b \\ c & d \end{pmatrix}, \quad (76)$$

with strategies  $\mathbf{p}_A = (1, 0)$  and  $\mathbf{p}_B = (0, 1)$  that are best replies to themselves with respect to  $W$  ( $a > c$  and  $d > b$ ). In this context,  $z^*$  is equal to  $x^* = (d - b)/(a - c - b + d)$ , the equilibrium frequency of  $A$  in an infinite panmictic population with game matrix  $W$ , and  $z^*$  is proportional to  $\tilde{z}$  so that (??) becomes

$$z^* = x^* = \frac{(d - b)}{(a - b - c + d)} < \frac{u_0\gamma_0 + (1 - u_0)\gamma_1}{u_0\delta_0 + (1 - u_0)\delta_1} = \Lambda. \quad (77)$$

Note that this inequality generalizes the one obtained in the case of a two-deme continental island model (Ladret and Lessard, 2008) to include a total number of  $n$  demes. This inequality means that if condition (??) is met, then selection favors a single mutant  $A$  taking over the whole population.

### 7.2.1 Symmetric structure with identical deme sizes

The particular case of a symmetric population structure with identical deme sizes ( $P = 1/n$ ,  $N_0 = N_1 = \dots = N_{n-1}$ ) will be referred to as the symmetric continental island model. Note that in this case, for any island  $i \in \{1, \dots, n-1\}$ , the backward migration probabilities  $m_{0i}$  and  $m_{i0}$  are identical and satisfy  $m_{0i} = m_{i0} = \frac{nM}{2cN}$ . In this context, Equation (??) can be written as

$$x^* = \frac{(d - b)}{(a - b - c + d)} < \Lambda = \frac{1}{3} + \frac{(n - 1)}{3n^2} \frac{f(M, n)}{g(M, n)} \quad (78)$$

where

$$f(M, n) = n^2(n + 2)(n - 1)M^3 + n(5n^2 + 4n + 4)M^2 + (10n^2 + 8n + 8)M + 8n, \quad (79)$$

$$g(M, n) = n^2(n + 2)M^4 + 2(n^3 + 5n^2 + 4n + 4)M^3 + 4(n + 2)(3n + 2)M^2 + 8(4 + 3n)M + 16. \quad (80)$$

Note that when  $n = 2$ , (??) reduces to  $x^* < \frac{1}{3} + \frac{1}{12(M+1)}$ , as given in Ladret and Lessard (2007 and 2008); whereas when evaluated at  $n = 1$ , (??) corresponds to the one-third law,  $x^* < 1/3$ , which holds in a panmictic population and which can be obtained from Lessard

(2005) and Imhof and Nowak (2006). On the other hand, note from their definition that functions  $f$  and  $g$  are nonnegative. Consequently, it follows from (??) that  $\Lambda \geq \frac{1}{3}$ . This shows that the modified one-third law for the symmetric  $n$ -deme continental island model is always less stringent than the one-third law.

### 7.2.2 Comparison of the symmetric continental island model with its Island Model counterpart

In this section, we compare the symmetric continental island linear game model with identical game matrices  $W_0 = W_1 = W$  to the Wright's island linear game model given in Ladret and Lessard (2007). The two models share the same game matrix,  $W$ , and the same geographic structure ( $n$  demes of equal sizes equal to  $\frac{cN}{n}$ ). They only differ in the migration scenario, which, in the latter model, occurs according to a Wright's island scenario. More precisely, in the Wright's island linear game model, let  $m_{Isl}$  denote the probability that an individual from any given deme emigrated from a different deme in the previous generation and let  $m_{ij, Isl}$  stand for the probability that an individual from deme  $i$  emigrated from deme  $j \neq i$ , in the previous generation. Then, for any  $i, j \in \{1, \dots, n\}$ , with  $i \neq j$ ,  $m_{ij, Isl} = m_{Isl}/(n-1)$ . Moreover, we suppose that the Wright's island model and the symmetric continental island model have the same average expected number of migrants per generation after population regulation. Since in the continental island model this number is  $(n-1)M$ , where  $M$  refers to the migration rate in the continental island model, it implies that  $m_{Isl} = \frac{(n-1)M}{Nc}$ . Using this notation, we know that in the context of a Wright's Island linear game model with game matrix  $W$ , the condition for selection to favor a mutant strategy A replacing a resident strategy B, takes the form

$$x^* < \Lambda_{WIsland} = \frac{1}{3} + \frac{1}{6(M+1)} \left(1 - \frac{1}{n}\right), \quad (81)$$

as shown in Ladret and Lessard (2007). In the context of the symmetric continental island model with game matrix  $W$ , we have seen that this condition takes the form

$$x^* < \Lambda,$$

where  $\Lambda$  is given in (??). This enables us to compare the two models by comparing their respective threshold frequency values,  $\Lambda$  and  $\Lambda_{WIsland}$ . From (??) and (??), we derive that

$$\begin{aligned} \Lambda &= \frac{1}{3} + \frac{1}{6(M+1)} \left(1 - \frac{1}{n}\right) + \frac{(n-1)(n-2)M(2M+2+Mn)\phi(M, n)}{6n^2(M+1)\psi(M, n)} \\ &= \Lambda_{WIsland} + \frac{(n-1)(n-2)M(2M+2+Mn)\phi(M, n)}{6n^2(M+1)\psi(M, n)}, \end{aligned} \quad (82)$$

where

$$\begin{aligned} \psi(M, n) &= 2M^3n^3 + M^4n^3 + 10M^3n^2 + 12M^2n^2 + 2M^4n^2 + 24Mn + 32M^2n + 8M^3n \\ &\quad + 16 + 16M^2 + 32M + 8M^3, \end{aligned} \quad (83)$$

and

$$\phi(M, n) = (M^2 n^2 - 2n - 4). \quad (84)$$

Since,

$$\phi(M, n) = M^2 \left( n - \frac{1 + \sqrt{1 + 4M^2}}{M^2} \right) \left( n + \frac{\sqrt{1 + 4M^2} - 1}{M^2} \right), \quad (85)$$

we can write

$$\Lambda - \Lambda_{WIsland} = (n - 1)(n - 2) \left( n - \frac{1 + \sqrt{1 + 4M^2}}{M^2} \right) \gamma(M, n), \quad (86)$$

where  $\gamma(M, n)$  is given by

$$\gamma(M, n) = \frac{M^3(2M + 2 + Mn) \left( n + \frac{\sqrt{1 + 4M^2} - 1}{M^2} \right)}{6n^2(M + 1)\psi(M, n)}.$$

Note that  $\gamma(M, n)$  only takes positive values. Thus,  $\Lambda - \Lambda_{WIsland}$  is proportional to, and has the same sign as  $(n - 1)(n - 2) \left( n - \frac{1 + \sqrt{1 + 4M^2}}{M^2} \right)$ .

As a consequence, when  $n = 1$  or  $n = 2$ , it follows from Equation (??), that  $\Lambda = \Lambda_{WIsland}$ . This is not surprising, since in both cases, the continental island model and the Wright's island model are the exact same model : when  $n = 1$ , they correspond to a linear game in a panmictic population and  $\Lambda = \frac{1}{3}$  (Lessard, 2005; Imhoff and Nowak, 2006; Lessard and Ladret, 2007); while when  $n = 2$ , the continental island model with two demes and the Wright's island model with two demes are identical and  $\Lambda = \frac{1}{3} + \frac{1}{12(M+1)}$ , as given in Ladret and Lessard (2007, 2008).

Now, let us focus on the case when  $n \geq 3$ . In this case, for any migration rate  $M > 0$  and any integer  $n$ ,  $\Lambda - \Lambda_{WIsland}$  has the same sign as (and is proportional to)  $n - \frac{1 + \sqrt{1 + 4M^2}}{M^2}$ . Thus, we have

$$\begin{cases} \Lambda > \Lambda_{WIsland} & \Leftrightarrow n > n_0(M) := \frac{1 + \sqrt{1 + 4M^2}}{M^2} \\ \Lambda = \Lambda_{WIsland} & \Leftrightarrow n = n_0(M), \end{cases}$$

or, equivalently

$$\begin{cases} \Lambda > \Lambda_{WIsland} & \Leftrightarrow M > M_0(n) := \frac{\sqrt{2n + 4}}{n}, \\ \Lambda = \Lambda_{WIsland} & \Leftrightarrow M = M_0(n). \end{cases}$$

Note, from their definition, that  $n_0(M)$  and  $M_0(n)$  are strictly decreasing functions of  $M$  and  $n$ , respectively (Fig. ??). Moreover,  $n_0(M) < 3$  if and only if  $M > \sqrt{10}/3$ . As a consequence, when the migration rate  $M$  is fixed and is greater than  $\sqrt{10}/3$ , we find that  $\Lambda$  is always larger than  $\Lambda_{WIsland}$  (whatever the number of demes,  $n \geq 3$ , is). On the other hand, when  $M \leq \frac{\sqrt{10}}{3}$ , then  $n_0(M) \geq 3$ . Thus, for a fixed  $M$ , only when  $M \leq \frac{\sqrt{10}}{3}$ , is  $n_0(M)$  a critical value for  $n$  (with  $\Lambda > \Lambda_{WIsland}$  if and only if  $n > n_0(M)$ ). Moreover, the only case when there exists a number of demes  $n$ , other than 1 or 2, for which  $\Lambda = \Lambda_{WIsland}$ , is

when both  $M \leq \frac{\sqrt{10}}{3}$  and  $n_0(M)$  is an integer (in which case this number is  $n = n_0(M)$ ). On the other hand, for any fixed number of demes  $n \geq 3$ ,  $M_0(n)$  always plays the role of a threshold value for  $M$  (such that  $\Lambda > \Lambda_{WIsland}$  if and only if  $M > M_0(n)$ ). Moreover, when  $n \geq 3$  is fixed, there always exists a migration rate at which  $\Lambda = \Lambda_{WIsland}$ , namely  $M = M_0(n)$ .

	$n_0(M)$
M=0.01	$10000 + 200\sqrt{2501} \simeq 20001.9998$
M=0.1	$100 + 20\sqrt{26} \simeq 201.98$
M=0.5	$4 + 4\sqrt{2} \simeq 9.657$
M=1	$\sqrt{5} + 1 \simeq 3.236$
M=10	$1/100 + (1/100)\sqrt{401} \simeq 0.21025$
M=100	$\simeq 0.02$

Table 1: Values taken by  $n_0(M)$  for different values of  $M \in [0.01, 100]$ .

	$M_0(n)$
n=3	$(1/3)\sqrt{10} \simeq 1.054092553$
n=10	$(1/5)\sqrt{6} \simeq 0.4898979486$
n= 50	$(1/25)\sqrt{26} \simeq 0.2039607806$
n=100	$(1/50)\sqrt{51} \simeq 0.1428285686$
n=500	$(1/250)\sqrt{251} \simeq 0.06337191808$

Table 2: Values taken by  $M_0(n)$  for different values of  $n \in [3, 500]$ .

This result is illustrated in Fig. ?? which represents numerical evaluations of the threshold frequency  $\Lambda$  in the case of the continental island model with equal deme sizes ( $P = 1/n$ ) as a function of the migration rate,  $M$ , and the number of demes,  $n$ , versus the threshold value for its Wright's island model counterpart,  $\Lambda_{WIsland}$ . Fig. ?? plots a projection of Fig. ?? for four fixed values of  $M$  :  $M = 0.1$ ,  $M = 0.5$ ,  $M = \sqrt{10}/3$  and  $M = 10$ . For those values of  $M$  (as for any  $M > 0$ ),  $\Lambda = \Lambda_{WIsland}$  when  $n = 1$  or  $n = 2$ . When  $M = \frac{\sqrt{10}}{3} \simeq 1.054$ , the critical value for the total number of demes is  $n_0(\frac{\sqrt{10}}{3}) = 3$ , and  $\Lambda > \Lambda_{WIsland}$  if and only if  $n > 3$ ; while  $\Lambda = \Lambda_{WIsland}$  when  $n = 1$ ,  $n = 2$  or  $n = 3$ . When  $M = 10$ , there is no critical value for  $n$  since  $M > \frac{\sqrt{10}}{3}$ . In this case,  $\Lambda > \Lambda_{WIsland}$  for all  $n \geq 3$ ; When  $M = 0.1$ , the critical value for  $n$  is  $n_0(0.1) \simeq 201.98$  (Table ??) which means that  $\Lambda > \Lambda_{WIsland}$  if and only if  $n \geq 202$ . Moreover, since  $n_0(0.1)$  is not an integer,  $\Lambda = \Lambda_{WIsland}$  only occurs when  $n = 1$  or  $n = 2$ . Similarly, when  $M = 0.5$ , since  $n_0(0.5) \simeq 9.657$  (Table ??), we find that  $\Lambda > \Lambda_{WIsland}$  if and only if  $n \geq 10$ , and  $\Lambda = \Lambda_{WIsland}$  if and only if  $n = 1$  or  $n = 2$  (since  $n_0(0.5)$  is not an integer). Fig. ?? represents a projection of Fig. ?? for three fixed values of  $n$  :  $n = 3$ ,  $n = 10$  and  $n = 50$ . In those cases, the corresponding critical value for  $M$  (which exists for any  $n \geq 3$ ) is  $M_0(3) = \sqrt{10}/3 \simeq 1.054$ ,  $M_0(10) = \sqrt{6}/5 \simeq 0.49$  and  $M_0(50) = \sqrt{26}/25 \simeq 0.204$ , respectively (Table ??) . We see that given those number of demes (as it is the case for any  $n \geq 3$ ),  $\Lambda > \Lambda_{WIsland}$  if and only if  $M$  is greater than the corresponding critical value,  $M_0(n)$  (with an equality if and only if  $M = M_0(n)$ ).

As reminded earlier, in the context of the IPD with TFT versus ALLD, the frequency  $x^*$  decreases to zero with the number of rounds (see, e.g., Nowak et al., 2004). As a consequence, a symmetric continental island model with  $n \geq 3$  demes and a fixed migration rate  $M$  greater than  $\frac{\sqrt{10}}{3}$  will require less rounds of the game than its Wright's Island model counterpart to ensure that cooperation fixates in the population with a selective advantage; whereas, in the case of a fixed migration rate  $M \leq \frac{\sqrt{10}}{3}$ , this will be the case provided the number of demes is large enough (namely,  $n > n_0(M) \geq 3$ ). On the other hand, if the number of demes  $n \geq 3$  is fixed, a symmetric continental island model will require less rounds of the game than its Wright's Island model counterpart provided  $M$  is not too small (namely,  $M > M_0(n)$ ), to ensure that selection favors cooperation taking over the whole population.

### 7.2.3 Numerical comparisons

In this section, still in the framework of identical game matrices ( $W = W_0 = W_1$ ), we compare the asymmetric continental island model with unequal deme sizes to its symmetric model counterpart with equal deme sizes. Both models share the same migration rate,  $M$ , the same number of demes,  $n$ , and the same game matrix,  $W$ . They only differ in  $P$ , the proportion of the total population that lives in the continent ( $P = 1/n$ , in the case of equal deme sizes; while  $P \neq 1/n$ , in the case of unequal deme sizes). In this context, condition (??) for selection to favor a mutant strategy  $A$  replacing a resident strategy  $B$ , takes the form  $x^* < \Lambda$ , as defined in (??), for the general model, and in (??), in the case of equal deme sizes. Thus, here again, the two models are compared by comparing their respective threshold value,  $\Lambda$ . On the other hand, comparisons are also made with the one-third law,  $x^* < 1/3$ , which holds when the population is panmictic.

Fig. ?? shows numerical evaluations of the threshold frequency  $\Lambda$  as a function of  $P$ , the proportion of individuals that live on the continent and  $n$ , the number of demes, for three different values of the migration rate  $M$  ( $M = 0.1$ ,  $M = 1$  and  $M = 10$ ) and  $n$  ranging from 2 to 50, versus the reference value of  $1/3$  obtained in the case of a panmictic population and the value taken by  $\Lambda$  in the case of a symmetric population structure corresponding to  $P = 1/n$ . On the other hand, Figs. ?? to ?? plot projections of Fig. ?? that represent  $\Lambda$  as a function of  $P$ , for fixed values of  $n$  ranging between 2 and 8, when  $M = 1$ ,  $M = 10$  and  $M = 0.1$ , respectively. Whereas, Figs. ?? to ?? plot projections of Fig. ?? for fixed values of  $n$  ( $n = 3$ ,  $n = 4$  and  $n = 10$ , respectively), for  $M$  ranging from 0.01 to 10.

In the remainder of this section, we shall denote by  $\Lambda(M, n, P)$ , the threshold frequency value  $\Lambda$  corresponding to an asymmetric continental island model with  $n$  demes, in which the proportion of individuals in the population living on the continent is  $P$ , and the migration rate is  $M$ ; while  $\Lambda(M, n, 1/n)$  will refer to the threshold frequency value of its symmetric model counterpart.

### 7.2.4 Comparisons with the symmetric model

The case when the continental island model consists in only two demes, one continent and one island ( $n = 2$ ), which is treated in Ladret and Lessard (2008), is illustrated again here in Fig. ?? (a), Fig. ?? (a) and Fig. ?? (a), in the case when  $M = 1$ ,  $M = 10$  and  $M = 0.1$ , respectively. In this context, as shown in Ladret and Lessard (2008), numerical evaluations performed for  $M$  ranging from 0.1 to 10 (Fig. ?? and ??) suggest that for any  $M$  in that range, the following equation

$$\Lambda(M, 2, P) - \Lambda(M, 2, 1/2) = 0 \quad (87)$$

has exactly two roots  $P \in [0, 1] : P = 1/2$  and  $P = P^*(M, 2)$ , for some  $0 \leq P^*(M, 2) < 1/2$ . Moreover, if we consider a model in which the proportion of individuals living on the continent,  $P$ , lies in  $]P^*(M, 2), 1/2[$ , then its threshold frequency  $\Lambda = \Lambda(M, 2, P)$  is larger than the one of its symmetric model counterpart, i.e.  $\Lambda(M, 2, P) > \Lambda(M, 2, 1/2)$ . Whereas, if  $P$  is either strictly smaller than  $P^*(M, 2)$  or strictly larger than  $1/2$ , then the reverse is true, i.e.  $\Lambda(M, 2, P) < \Lambda(M, 2, 1/2)$ . On the other hand, for any fixed value of  $M$  in that range,  $\Lambda$  is maximized at some  $P = P^{**}(M, 2) \in ]P^*(M, 2), 1/2[$  and  $\Lambda$  increases with  $P$  from  $1/3$  until it reaches its maximum (which is smaller than  $1/2$ ) at  $P = P^{**}(M, 2)$ , and then  $\Lambda$  decreases with  $P$ , back to  $1/3$  (see Ladret and Lessard, 2008).

In the more general case of a continental island model with a total number of demes equal to  $n$ , numerical evaluations performed for  $M$  ranging from 0.01 to 10 and  $n$  ranging from 2 to 50 (Figs. ?? to ??), suggest that for any fixed value of  $M$  in that range, as long as  $n$  remains small enough, i.e. less than or equal to some integer  $n^*(M) \geq 2$ , the following equation

$$\Lambda(M, n, P) - \Lambda(M, n, 1/n) = 0 \quad (88)$$

has exactly two roots  $P \in [0, 1] : P = 1/n$  and  $P = P^*(M, n)$ , for some  $0 \leq P^*(M, n) < 1/n$ . In this case (i.e.,  $n \leq n^*(M)$ ), we find that the threshold frequency  $\Lambda$  is larger than its symmetric model counterpart, i.e.  $\Lambda(M, n, P) > \Lambda(M, n, 1/n)$ , if and only if the proportion of individuals living on the continent,  $P$ , lies in  $]P^*(M, n), 1/n[$ . Moreover,  $\Lambda$  is maximized at some  $P = P^{**}(M, n) \in ]P^*(M, n), 1/n[$  and  $\Lambda$  increases from  $\Lambda(M, n, 0) = 1/3 + \frac{(n-2)}{3(M+2)(n-1)}$  to its maximum (which is smaller than  $1/2$ ) until  $P$  reaches  $P^{**}(M, n)$ , and then  $\Lambda$  decreases with  $P$ , back to  $1/3$ . Note that numerical evaluations performed for  $M$  ranging from 0.1 to 10 suggest that  $n^*(M)$  is a decreasing function of  $M$  and that  $2 \leq n^*(M) \leq 4$ , for  $M$  in that range (Figs. ??, ??, and Figs. ?? to ?? and Table ??).

On the other hand, if the number of demes  $n$  is larger than  $n^*(M)$ , then Equation (??) has a single root in  $[0, 1]$ , namely,  $P = 1/n$ . In this case,  $\Lambda$  is greater than its symmetric model counterpart, i.e.  $\Lambda(M, n, P) > \Lambda(M, n, 1/n)$ , if and only if the size of the continent is smaller than the size of each island, i.e.  $P < 1/n$ . Furthermore, when  $n > n^*(M)$ , numerical evaluations suggest that when both the number of demes,  $n$ , and the migration rate,  $M$ , are fixed, depending on whether the total number of demes is small enough or not, there are two types of variation modes for  $\Lambda(M, n, P)$ , as a function of  $P$ . More precisely,

there exists an integer  $n^{**}(M) \geq n^*(M)$ , such that if  $n \in [n^*(M); n^{**}(M)]$ , then  $\Lambda(M, n, \cdot)$  increases with  $P$  from  $\Lambda(M, n, 0) = 1/3 + \frac{(n-2)}{3(M+2)(n-1)}$  until it reaches its maximum located at  $P = P^{**}(M, n) \in ]0; 1/n[$ , and then it decreases back to  $1/3$ ; whereas, when the number of demes  $n$  is greater than  $n^{**}(M)$ ,  $\Lambda(M, n, \cdot)$  is a decreasing function of  $P$  that decreases from  $\Lambda(M, n, 0) = 1/3 + \frac{(n-2)}{3(M+2)(n-1)}$  to  $1/3$  (and the maximum for  $\Lambda$  is located at  $P^{**}(M, n) = 0$ ). Numerical evaluations suggest that  $n^{**}(M)$  is a decreasing function of  $M$  that does not exceed 7 when  $M$  ranges between 0.1 and 10 (Fig. ?? and Table ??).

Let us focus on the case  $M = 1$  (Fig. ??a, ??b, ??, and Figs. ??c, ??c and ??c). In this context, we find that  $n^*(1) = 2$ , which means that the only value of  $n$  for which Equation (??) has two roots,  $1/n$  and  $P^*(1, n) < 1/n$ , is  $n = 2$ . In this case ( $M = 1, n = 2$ ), as noted in Ladret and Lessard (2008), this extra root is  $P = P^*(1, 2) \simeq 0.224$  and we have  $\Lambda(1, 2, P) > \Lambda(1, 2, 1/2) = 3/8$ , which means a larger threshold value for  $x^*$  than in the symmetric model, if and only if  $0.224 < P < 0.5$ . Moreover in this case, numerical evaluations show that the maximum value for  $\Lambda$  is reached when  $P = P^{**} = P^{**}(1, 2) \simeq 0.349$  and is equal to  $\Lambda(1, 2, P^{**}) \simeq 0.379$ . On the other hand, as soon as  $n \geq 3$ , we find that  $P = 1/n$  is the only root of Equation (??) and  $\Lambda(1, n, P) > \Lambda(1, n, 1/n)$  if and only if  $P < 1/n$ : for  $n = 3$ , the value of  $\Lambda$  at  $P = 1/3$ , which corresponds to the case of equal deme sizes, is  $\Lambda(1, 3, 1/3) = 5887/15147 \simeq 0.388$  and the maximum for  $\Lambda$  is reached at  $P^{**}(1, 3) \simeq 0.137$  and is equal to  $\Lambda(1, 3, P^{**}) \simeq 0.395$ . Moreover,  $\Lambda$  increases with  $P$  from  $7/18 \simeq 0.3889$ , which is the limit of  $\Lambda(1, 3, P)$  as  $P$  goes to 0 and whose value is slightly larger than  $\Lambda(1, 3, 1/3) \simeq 0.38866$ , to 0.395 until  $P$  reaches  $P^{**}(1, 3)$  and then decreases back to  $1/3$ . When the model has  $n = 4$  demes, the value of  $\Lambda$  at  $P = 1/4$  (symmetric model) is  $\Lambda(1, 4, 1/4) = 2153/5424 \simeq 0.39694$  and the maximum for  $\Lambda$  is reached at  $P^{**}(1, 4) \simeq 0.0182$  and is equal to  $\Lambda(1, 4, P^{**}) \simeq 0.40753$ . Moreover,  $\Lambda$  increases with  $P$  from  $\Lambda(1, 4, 0) = 11/27 \simeq 0.4074$  (which is greater than  $\Lambda(1, 4, 1/4)$ ) until it reaches  $\Lambda(1, 4, P^{**})$  when  $P$  hits  $P^{**}(1, 4)$  and then it decreases with  $P$  back to  $1/3$ ; On the other hand, as soon as  $n \geq 5$ , numerical evaluations suggest that  $\Lambda(1, n, P)$  is a decreasing function of  $P$  whose maximum is reached at the limit when  $P$  goes to zero ( $P^{**}(1, n) \simeq 0$ ). As a consequence, when  $M = 1$ , numerical evaluations indicate that  $n^{**}(1) = 4$ .

Similarly, in the case  $M = 0.1$ , we find that  $n^*(0.1) = 4$  and  $n^{**}(0.1) = 7$ ; while in the case  $M = 10$ , numerical evaluations suggest that  $n^*(10) = 2$  and  $n^{**}(10) = 3$ . When  $M = 0.01$ , we have  $n^*(0.01) = 7$  and  $n^{**}(0.01) = 17$ ; while, when  $M = 100$ , we have  $n^*(100) = 2$  and  $n^{**}(100) = 3$  (Table ??).

	$n^*(M)$	$n^{**}(M)$
M=0.01	7	17
M=0.1	4	7
M=1	2	4
M=10	2	3
M=100	2	3

Table 3: Values taken by  $n^*(M)$  and  $n^{**}(M)$ , for different values of  $M \in [0.01, 100]$ .



More generally, numerical evaluations made for  $M$  ranging from 0.1 to 10 (Fig. ??, ?? and Fig. ?? to ??), suggest that as long as  $n$  remains small enough, i.e.  $2 \leq n \leq n^*(M)$  (with  $n^*(M) \leq 4$  for  $M \in [0.1; 10]$ ), condition for the replacement of the resident strict Nash equilibria  $\mathbf{p}_B$  by the other mutant strict Nash equilibria  $\mathbf{p}_A$  is less stringent in an asymmetric model ( $P \neq 1/n$ ), in which the size of the islands exceeds the size of the continent, provided it is not too much larger ( $P^*(M, n) < P < 1/n$ ); On the other hand, for larger values of  $n$ , i.e.  $n > n^*(M)$ , this condition is met as soon as the size of the islands exceeds the size of the continent ( $P < 1/n$ ). Numerical evaluations show that similar results hold when  $M$  ranges from 0.01 to 100 (with  $n^*(M) \leq 7$  for  $M \in [0.01; 100]$ ) (not show).

Furthermore, numerical evaluations (Fig. ?? and Figs. ?? to ??) suggest that for any fixed migration rate,  $M$ , and for any number of demes,  $n$ , provided it remains small enough (namely, provided  $n \leq n^{**}(M)$ ), there exists an "optimal" relative continent size,  $P = P^{**}(M, n) \in ]0, 1/n[$ , at which the condition for the replacement of the resident strict Nash equilibria  $\mathbf{p}_B$  by the other mutant strict Nash equilibria  $\mathbf{p}_A$  is the least stringent. In other words, given a migration rate  $M$  and a total number of demes  $n \leq n^{**}(M)$ , the model in which the proportion of individuals living on the continent is equal to  $P = P^{**}(M, n)$ , is the one for which the threshold value for  $x^*$  is maximized, i.e.  $\Lambda(M, n, P^{**}(M, n)) = \max_{P \in [0, 1]} \Lambda(M, n, P)$ . On the other hand, if the fixed number of demes,  $n$ , is larger than  $n^{**}(M)$ , then the maximum for  $\Lambda(M, n, P)$  is reached at the limit when  $P$  goes to zero, i.e.  $P^{**}(M, n) = 0$ , since in this case, numerical evaluations suggest that  $\Lambda(M, n, P)$  is a decreasing function of the relative continent size,  $P$ .

Fig. ?? (a) shows numerical evaluations of the maximum threshold frequency  $\Lambda^{**}(M, n) = \max_{P \in [0, 1]} \Lambda(M, n, P) = \Lambda(M, n, P^{**}(M, n))$  as a function of the number of demes  $n$  and the proportion of individuals living on the continent at which it is reached,  $P^{**}(M, n)$ , for three different values of the migration rate ( $M = 0.1$ ,  $M = 1$  and  $M = 10$ ), versus  $\Lambda(M, n, 1/n)$ , the value taken by  $\Lambda$  in its symmetric model counterpart that corresponds to  $P = 1/n$ . On the other hand, Fig. ?? (b) plots projections of Fig. ?? (a) that represent  $\Lambda^{**}(M, n)$  as a function of  $n$ , for  $M = 0.1$ ,  $M = 1$  and  $M = 10$ ; while ?? (c) plots projections of Fig. ?? (a) that represent  $P^{**}(M, n)$  as a function of  $n$ , for the same three values of  $M$ . We see that for a fixed number of demes,  $n$ , the maximum value taken by  $\Lambda$  when the migration parameter is  $M$ ,  $\Lambda^{**}(M, n)$ , and the value taken by  $\Lambda$  in its symmetric model counterpart,  $\Lambda(M, n, 1/n)$ , are both decreasing functions of  $M$  (Fig. ?? and Fig. ??). Moreover, we find that as  $n$  goes to infinity,  $\Lambda^{**}(M, n)$  and  $\Lambda(M, n, 1/n)$  share the same limit, which is  $\frac{1}{3} + \frac{1}{3(M+2)}$ . More generally, numerical evaluations (Fig. ??) show that for any fixed number of demes,  $n$ , and any fixed  $P$  provided it remains small enough (at least for  $P$  smaller than or equal to  $1/n$ ),  $\Lambda$  is a decreasing function of  $M$ . Note, however, that this behavior does not hold for larger values of  $P \in [0, 1]$  (Fig. ??).

Let us go back to IPD with TFT versus AllD. Since in this framework,  $x^*$  decreases to zero with the number of rounds (see, e.g., Nowak et al., 2004), we find that for a fixed migration rate,  $M$ , depending on the value of the number of demes,  $n$ , there are two types

of conditions, regarding  $P$ , under which an asymmetric continental island population model will require less rounds of the game than its symmetric counterpart to ensure that cooperation will take over the population with a selective advantage : if the number of demes remains small enough (i.e.  $n \leq n^*(M)$ ), this will happen if  $P$  is not too small and remains below  $1/n$ , as it is already known in the case of a two-deme ( $n = 2$ ) continental island model (see Ladret and Lessard, 2008); whereas for larger values of  $n$  (i.e.  $n > n^*(M)$ ), it will only require  $P$  to be smaller than  $1/n$  (i.e., it will only require the continent's size to be smaller than that of the islands).

### 7.2.5 Comparisons with the one-third law

For a fixed number of demes  $n \geq 2$ , numerical evaluations performed for  $M$  ranging from 0.1 to 10 (Figs. ?? to ??) indicate that  $1/3 < \Lambda(M, n, P) \leq 1/2$  for every  $P \in ]0, 1[$ . Similar results hold when  $M$  ranges between 0.01 to 100 (not shown). In particular, this implies that condition (??) is less stringent than the one-third law.

Going back to the case of TFT versus ALLD, a continental island linear-game model with identical game matrices will require less rounds of the game than a panmictic model to ensure that cooperation invades the whole population with a selective advantage.

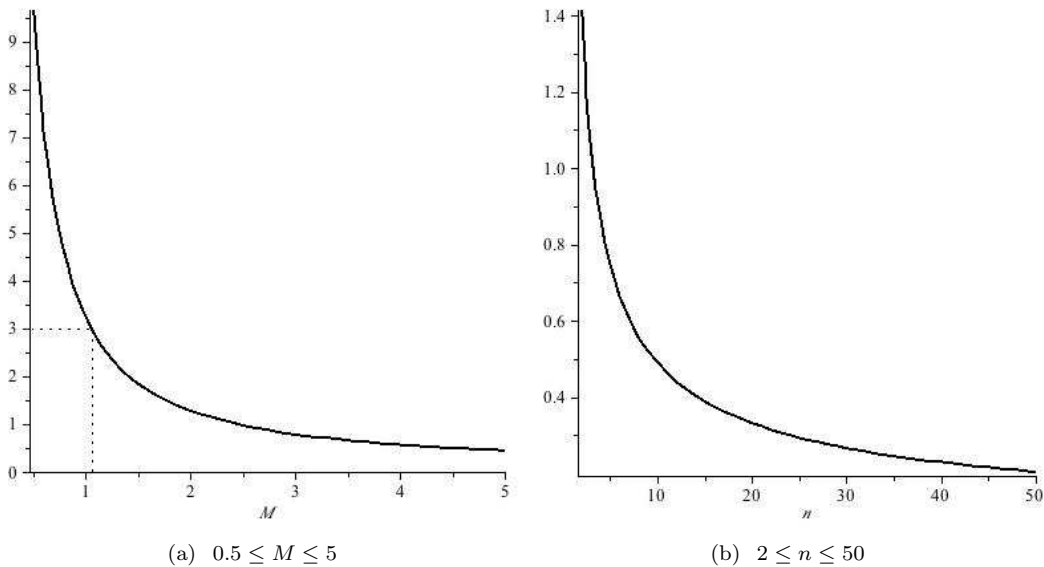


Figure 1: (a) Critical value  $n_0(M) = (1 + \sqrt{1 + 4M^2})/M^2$  as a function of  $M$ , with  $M$  ranging from 0.5 to 5. The point on the graph whose  $M$ -value is  $M = \sqrt{10}/3$  (vertical dots) satisfies  $n_0(M) = 3$  (horizontal dots); (b) Critical value  $M_0(n) = \sqrt{\frac{2}{n} + \frac{4}{n^2}}$  as a function of  $n$ , with  $n$  ranging from 2 to 50.

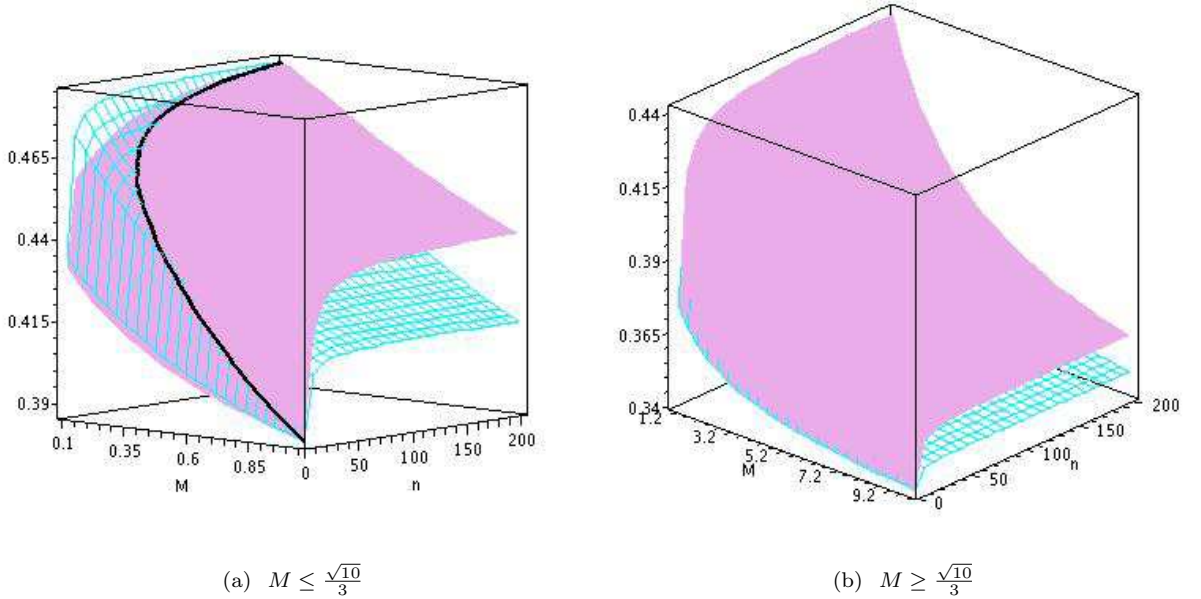


Figure 2: Threshold  $\Lambda$  in the case of equal deme sizes, i.e.  $P = 1/n$ , (solid purple surface) as a function of  $M$  and  $n$  versus its Wright's island model counterpart (cyan lines), i.e.  $\frac{1}{3} + \frac{1}{6(M+1)}(1 - \frac{1}{n})$ : (a)  $M$  ranges from 0.1 to  $\sqrt{10}/3$  and  $n$  from 2 to 200. The black curve is for the case when  $n = n_0(M) = (1 + \sqrt{1 + 4M^2})/M^2$ ; (b)  $M$  ranges from  $\sqrt{10}/3$  to 10 and  $n$  from 2 to 200.

## 8 Discussion

The probability that a single individual using strategy  $A$  takes over a population consisting of individuals using another strategy,  $B$ , in the context of a Wright-Fisher panmictic population of large size,  $N$ , under weak frequency dependent selection based on a linear game with game matrix  $W$  and selection intensity  $s \ll 1/N \ll 1$ , can be approximated using a diffusion approximation (Lessard, 2005) by

$$u(s) = \frac{1}{N} + s\{(\mathbf{p}_A - \mathbf{p}_B) \cdot W\mathbf{p}_B + \frac{1}{3}(\mathbf{p}_A - \mathbf{p}_B) \cdot W(\mathbf{p}_A - \mathbf{p}_B)\} + O(s/N).$$

When the population size  $N$  is fixed and the selection intensity  $s$  is small enough, a more precise approximation, derived from a Markov Chain method introduced by Rousset (2003) that allows to write the first order effect of selection on the probability of fixation as a function of the expected coalescence time, under neutrality of samples of two and three individuals, is given by (Lessard and Ladret, 2007)

$$u(s) = \frac{1}{N} + s(1 - \frac{1}{N})\{(\mathbf{p}_A - \mathbf{p}_B) \cdot W\mathbf{p}_B + \frac{N}{3N-2}(\mathbf{p}_A - \mathbf{p}_B) \cdot W(\mathbf{p}_A - \mathbf{p}_B)\} + o(s),$$

We also refer to Imhof and Nowak (2006) for an alternative approach, in the context of the Wright-Fisher model. Note that the fixation probability in the Moran model can be calculated explicitly and approximated for  $s$  small enough (Nowak et al., 2004; Lessard, 2005). See also Lambert (2006) for a general branching process approach and Lessard (2007) for

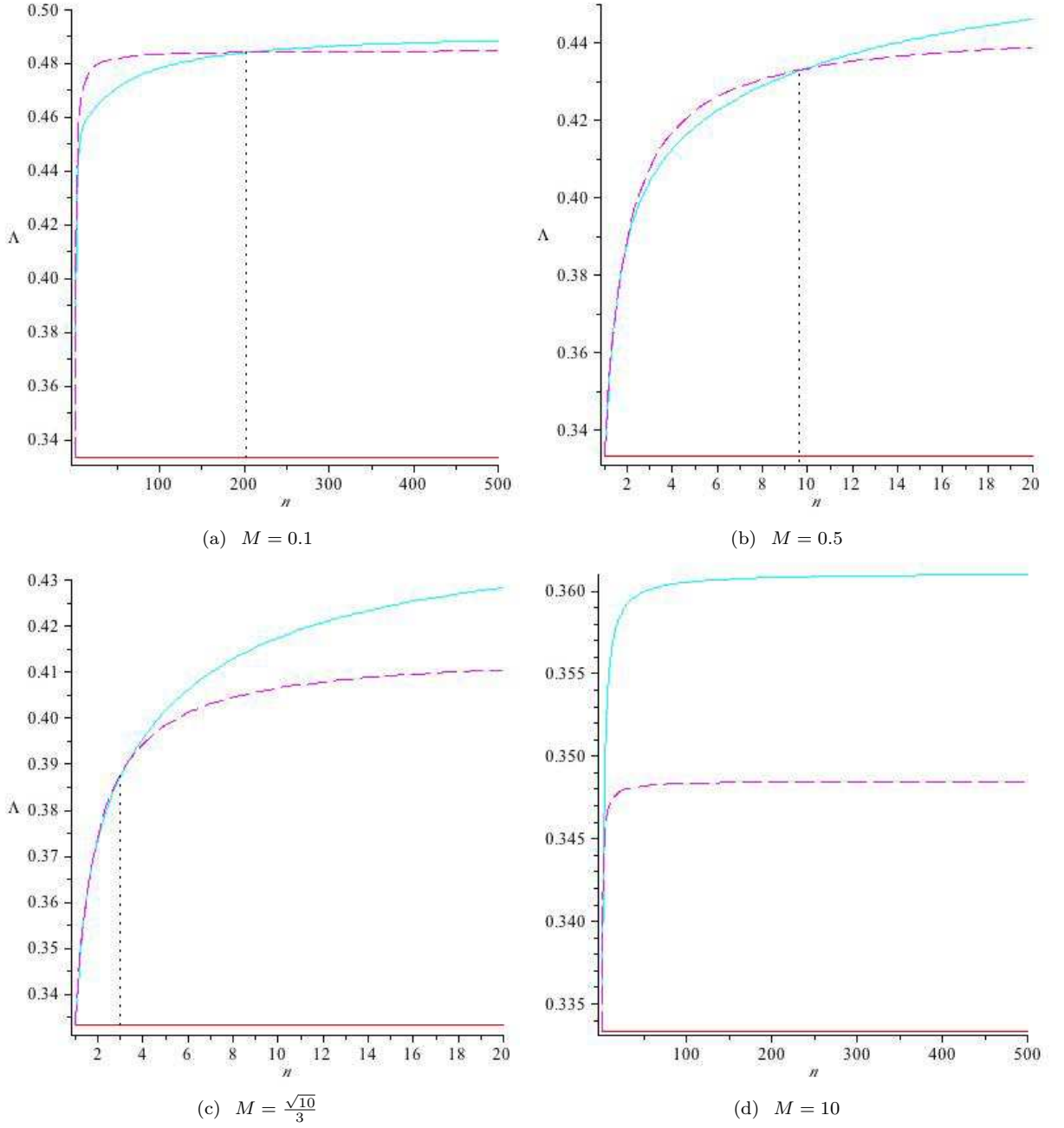


Figure 3: Effect of the number of demes,  $n$ , on coefficient  $\Lambda$  (solid blue surface) in the case of equal deme sizes for different values of  $M$  : (a)  $M = 0.1$  with  $n$  ranging from 1 to 500. The vertical black dotted line corresponds to  $n_0(0.1) \simeq 201.98$  ; (b)  $M = 0.5$  with  $n$  ranging from 1 to 20. The vertical black dotted line is at  $n = n_0(0.5) \simeq 9.66$  ; (c)  $M = \sqrt{10}/3$  with  $n$  ranging from 1 to 20. The black dotted line corresponds to  $n_0(\sqrt{10}/3) = 3$ ; (d)  $M = 10$  with  $n$  ranging from 1 to 500. The value  $\Lambda_{Wright} = 1/3 + 1/(6(M+1))(1 - \frac{1}{n})$  (magenta dots) is the value of threshold  $\Lambda$  in the case of a symmetric Wright's island Model and  $\Lambda = 1/3$  (horizontal red line) for the panmictic scenario.

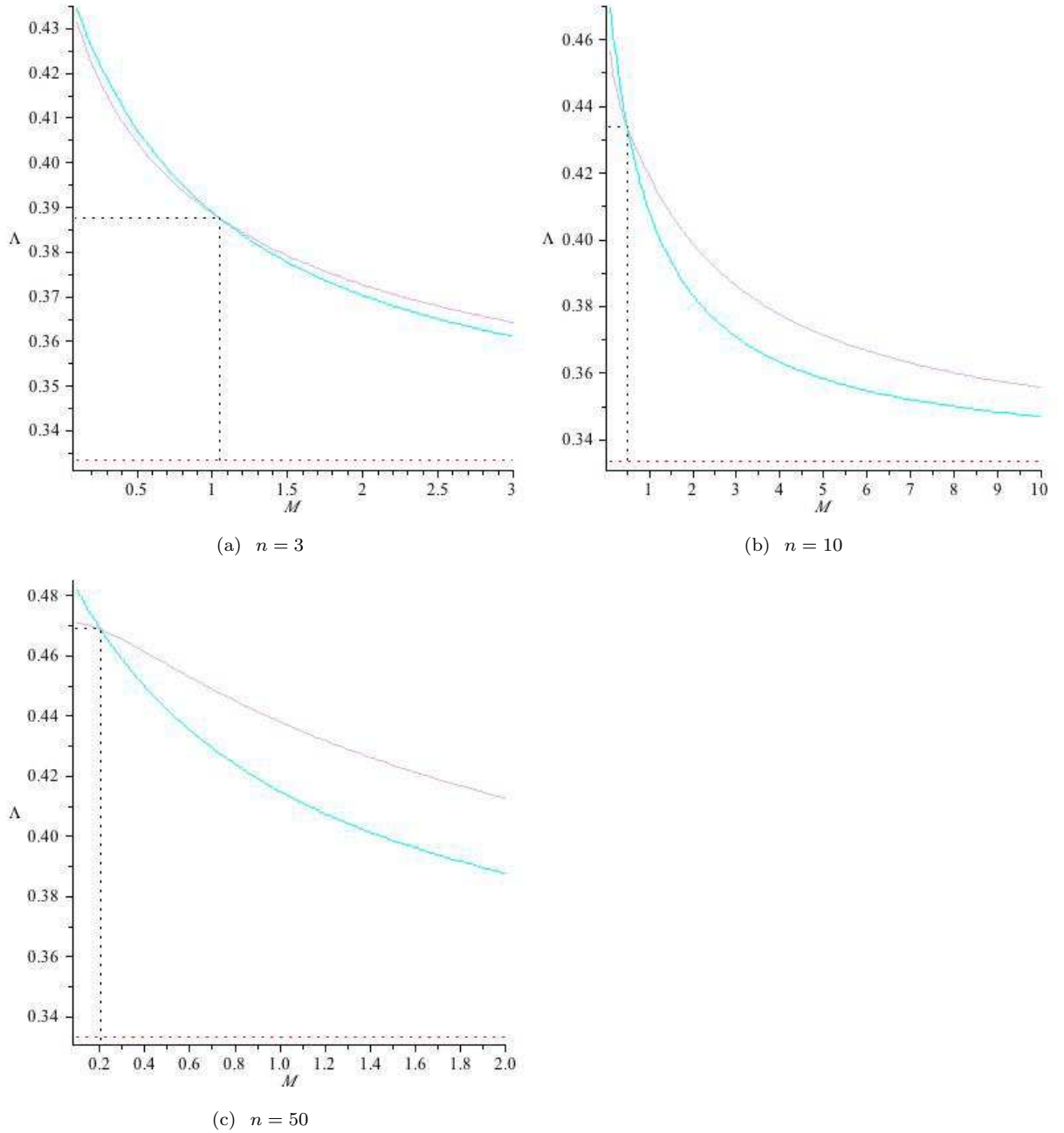


Figure 4: Effect of the scaled migration rate,  $M$ , on threshold  $\Lambda$  in the case of equal deme sizes, i.e.  $P = 1/n$ , (solid purple curve) versus  $\Lambda = \frac{1}{3} + \frac{1}{6(M+1)}(1 - \frac{1}{n})$  in the Wright's island scenario (cyan curve) for various values of the number of demes,  $n$ ; the vertical black dashed line corresponds to  $M = M_0(n) = \sqrt{\frac{2}{n} + \frac{4}{n^2}}$  and the horizontal red dotted line is  $1/3$ , the value of  $\Lambda$  in the panmictic scenario: (a)  $n = 3$  and  $M$  ranges from 0.1 to 10; (b)  $n = 10$  and  $M$  ranges from 0.1 to 10; (c)  $n = 50$  and  $M$  ranges from 0.1 to 2.

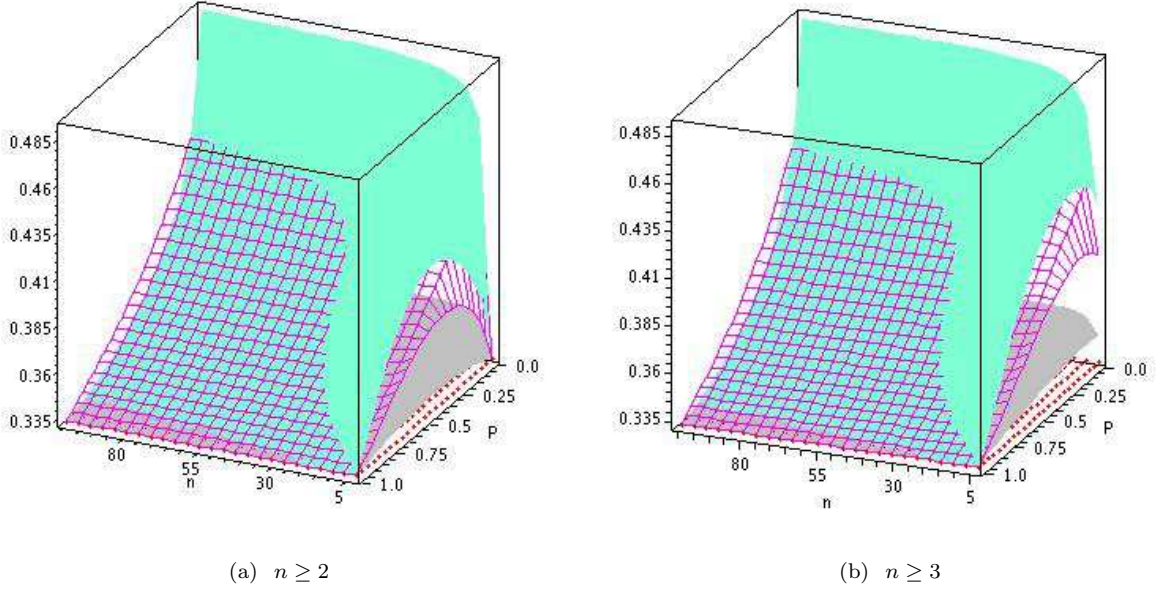


Figure 5: Threshold  $\Lambda$  as a function of  $n$  and  $P$  for three different values of  $M$  :  $M = 0.1$  (green solid surface);  $M = 1$  (magenta lines);  $M = 10$  (grey solid surface) versus  $1/3$  (red dots), the value taken by  $\Lambda$  in the case of a panmictic population. (a)  $P$  ranges from 0 to 1 and  $n$  from 2 to 100; (b)  $P$  ranges from 0 to 1 and  $n$  from 3 to 100.

more general discrete-time reproduction schemes.

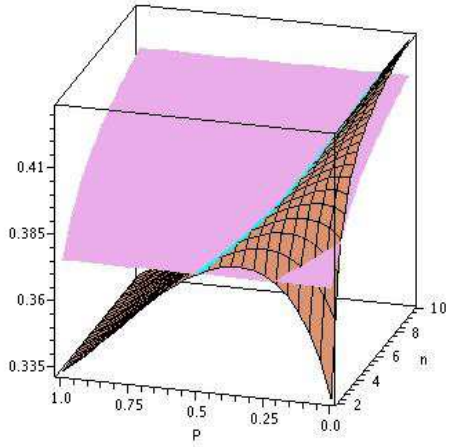
In the case of a finite island model with  $n \geq 3$  demes of equal size,  $N$ , and the same game matrix within each deme,  $W$ , the generalization of the previous Markov chain method leads to the following approximation for the fixation probability (Ladret and Lessard, 2007)

$$u(s) = \frac{1}{nN} + s \left\{ \left( \frac{1}{3} + \frac{\gamma}{\delta} \right) (\mathbf{p}_A - \mathbf{p}_B) \cdot W(\mathbf{p}_A - \mathbf{p}_B) + \left( 1 - \frac{1}{nN} \right) (\mathbf{p}_A - \mathbf{p}_B) \cdot W\mathbf{p}_B \right\} + o(s), \quad (89)$$

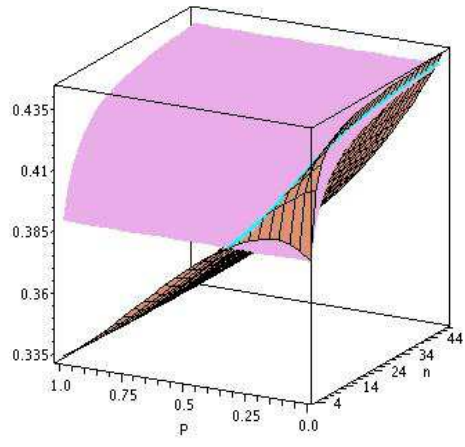
where  $\gamma/\delta$  is a coefficient that depends on the population structure parameters,  $N$ ,  $m$ , and  $n$  and whose value ranges from 0 to  $1/6$ . Moreover, in the case where the structured coalescent (SC) assumptions are met, which requires the deme size,  $N$ , to be large and the backward migration probability,  $m$ , to be in the form  $m = M(n-1)/Nn$ , where  $M$  is of order 1 so that  $m$  is of order  $1/N$ , and under the additional assumption that  $s \ll 1/(Nn)$ , this approximation reduces to (Ladret and Lessard, 2007)

$$u(s) = \frac{1}{nN} + s \left\{ \left( \frac{1}{3} + \frac{1}{6(M+1)} \left( 1 - \frac{1}{n} \right) \right) (\mathbf{p}_A - \mathbf{p}_B) \cdot W(\mathbf{p}_A - \mathbf{p}_B) + (\mathbf{p}_A - \mathbf{p}_B) \cdot W\mathbf{p}_B \right\} + o(s). \quad (90)$$

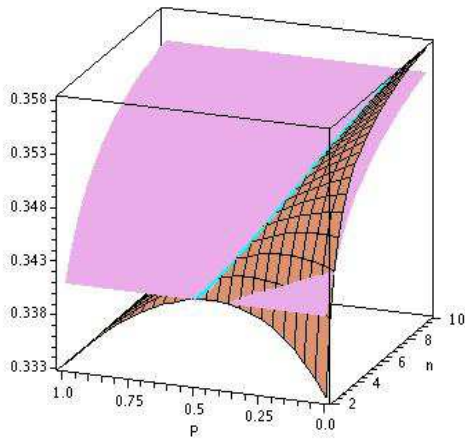
When the model consists in an asymmetric two-deme linear-game model, in which the population is subdivided into two demes of respective size  $N_1$  and  $N_2$ , with respective game matrix  $W_1$  and  $W_2$ , under the structured coalescent scenario (Notohara, 1990; Herbots, 1994, 1997; Wilkinson-Herbots, 1998) which holds when the backward migration rates  $m_1$



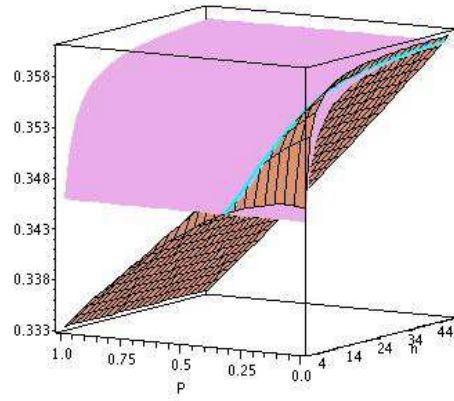
(a)  $M = 1, n \geq 2$



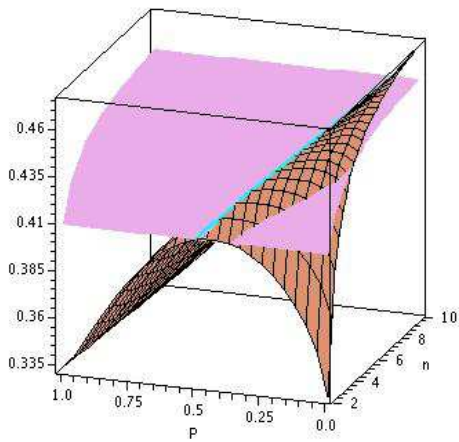
(b)  $M = 1, n \geq 3$



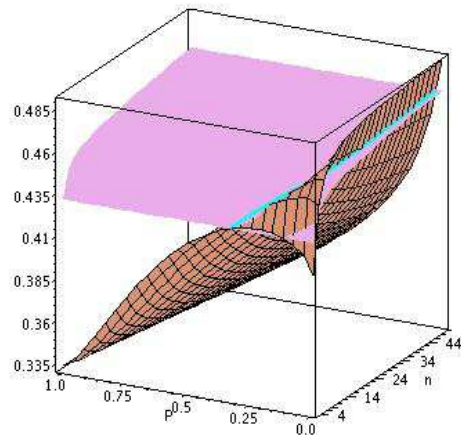
(c)  $M = 10, n \geq 2$



(d)  $M = 10, n \geq 3$

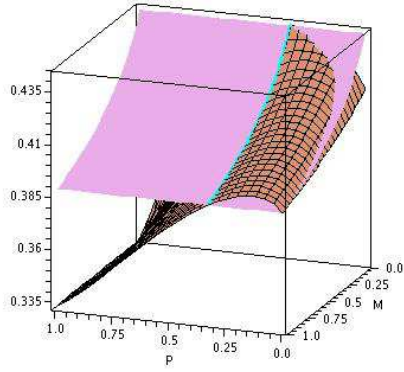


(e)  $M = 0.1, n \geq 2$

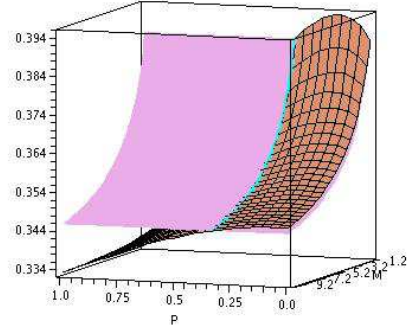


(f)  $M = 0.1, n \geq 3$

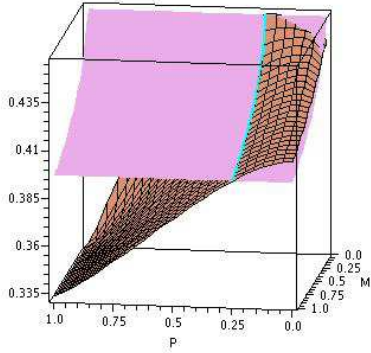
Figure 6: Threshold  $\Lambda$  (brown patched surfaces) as a function of  $n$  and  $P$  versus its value in the case of equal deme sizes (solid purple surface) for different values of  $M$ : (a)  $M = 1$  and  $n$  ranges from 2 to 10; (b)  $M = 1$  and  $n$  ranges from 3 to 50; (c)  $M = 10$  and  $n$  ranges from 2 to 10; (d)  $M = 10$  and  $n$  ranges from 3 to 50; (e)  $M = 0.1$  and  $n$  ranges from 2 to 10; (f)  $M = 0.1$  and  $n$  ranges from 3 to 50. The cyan curve represents the intersection of the two surfaces when  $P = 1/n$  (i.e. when the deme sizes are equal).



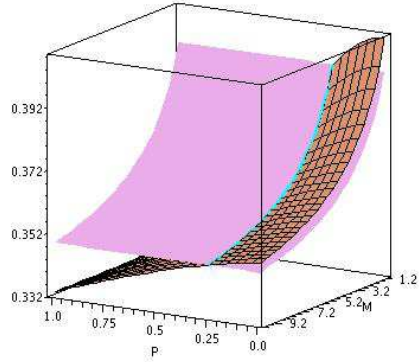
(a)  $n = 3$ ,  $M = 0.01 \dots 1$



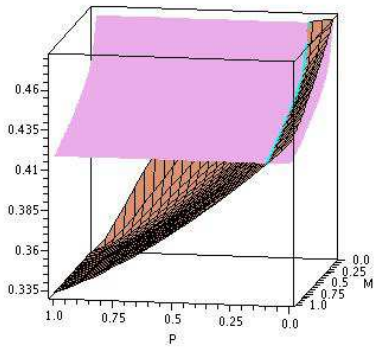
(b)  $n = 3$ ,  $M = 1 \dots 10$



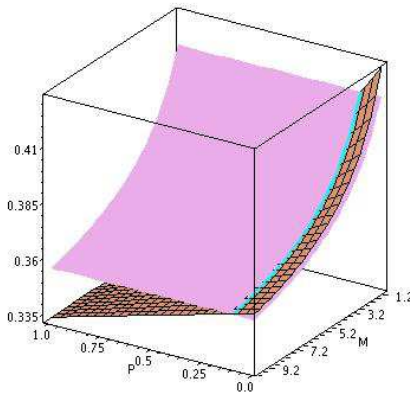
(c)  $n = 4$ ,  $M = 0.01 \dots 1$



(d)  $n = 4$ ,  $M = 1 \dots 10$



(e)  $n = 10$ ,  $M = 0.01 \dots 1$



(f)  $n = 10$ ,  $M = 1 \dots 10$

Figure 7: Threshold  $\Lambda$  (brown patched surface) versus its value in the case of equal deme sizes (solid purple surface) for different values of  $n$ : (a)  $n = 3$  and  $M$  ranges from 0.01 to 1; (b)  $n = 3$  and  $M$  ranges from 1 to 10; (c)  $n = 4$  and  $M$  ranges from 0.01 to 1; (d)  $n = 4$  and  $M$  ranges from 1 to 10; (e)  $n = 10$  and  $M$  ranges from 0.01 to 1; (f)  $n = 3$  and  $M$  ranges from 1 to 10. The cyan curve represents the intersection of the two surfaces when  $P = 1/n$  (i.e. when the deme sizes are equal).



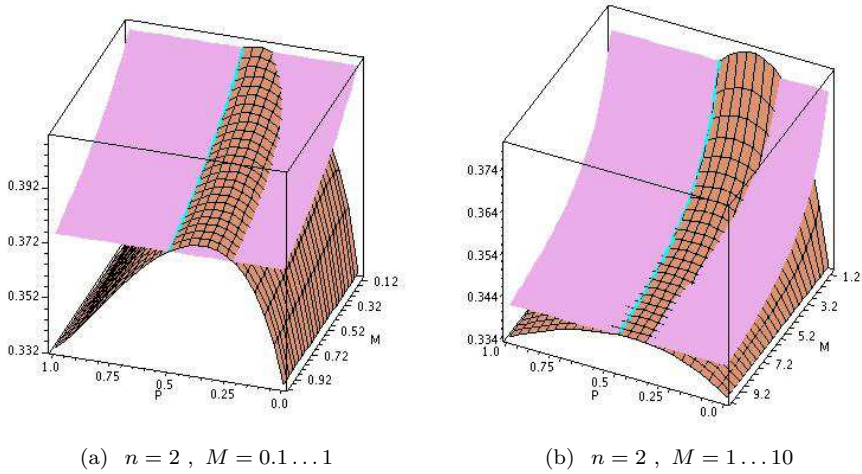


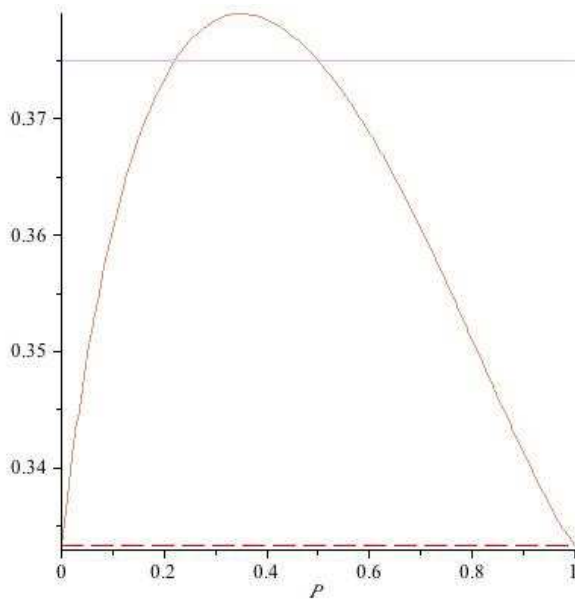
Figure 8: Threshold  $\Lambda$  (brown patched surface) versus its value in the case of equal deme sizes (solid purple surface) in the case when  $n = 2$ : (a)  $M$  ranges from 0.1 to 1; (b)  $M$  ranges from 1 to 10.

and  $m_2$  are of the form  $m_1 = M_1/(4N)$  and  $m_2 = M_2/(4N)$ , where  $N$  stands for half the total population size,  $2N = N_1 + N_2$ , and  $N$  is assumed to be large, with  $N_1, N_2$  of order  $N$  and  $M_1, M_2$  of order 1, the fixation probability of a single mutation  $A$  initially introduced in deme 1 can be approximated (Ladret and Lessard, 2008), when selection is weak ( $s \ll 1/N$ ), by

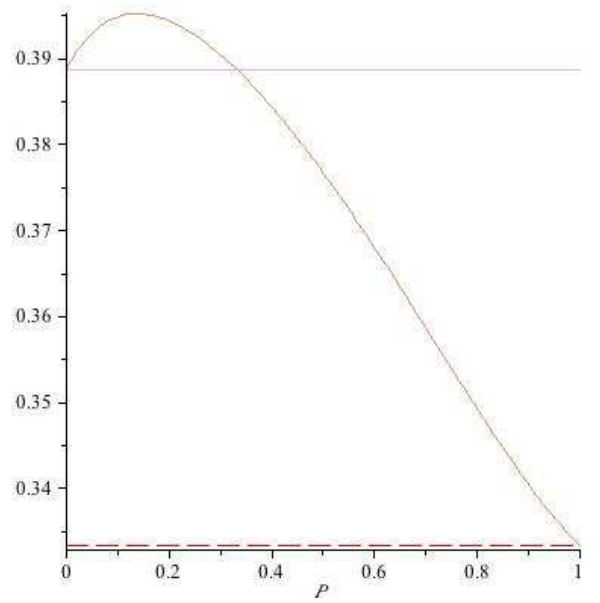
$$u(s) = \frac{u_1}{N_1} + s \{(\mathbf{p}_A - \mathbf{p}_B) \cdot (\mu_1 W_1 + \mu_2 W_2)(\mathbf{p}_A - \mathbf{p}_B) + (\mathbf{p}_A - \mathbf{p}_B) \cdot (\nu_1 W_1 + \nu_2 W_2) \mathbf{p}_B\} + o(s), \quad (91)$$

where  $u_1$  denotes the first coordinate of  $(u_1, u_2)$  of the stationary distribution of the backward migration matrix  $\mathbf{M}$ , and coefficients  $\mu_i, \nu_i$  are functions which depend on the migration rates,  $M_1, M_2$ , and the proportion of individuals from deme 1 to the whole population,  $e_1 := N_1/(2N)$ . This result is derived from an adaptation of the direct Markov chain method and the proof consists in writing the derivative of  $u(s)$  evaluated at  $s = 0$  as a function of the terms  $\sum_{t \geq 0} E_0(X_i^2(t)(1 - X_i(t)))$  and  $\sum_{t \geq 0} E_0(X_i(t)(1 - X_i(t))^2)$ , for  $i = 1, 2$ , where  $X_i(t)$  is the proportion of individuals using strategy  $A$  in deme  $i$  at time  $t$  and  $E_0$  stands for the expectation under neutrality ( $s = 0$ ), which are then calculated using a method based on the coalescent (Kingman, 1982) for a structured population (Notohara, 1990).

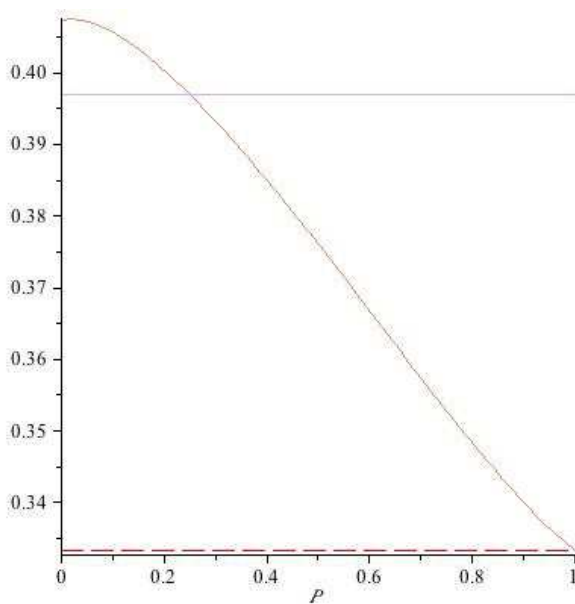
In this paper, we have generalized the previous asymmetric two-deme linear game model to an  $n$ -deme continental island linear game model, for which we have calculated the first order approximation of the fixation probability of a single mutant  $A$  initially introduced into the continent, with respect to the intensity of selection, in the context of weak selection and conservative migration. In this model, the population is subdivided into a continent of size  $N_0 = c_0 N$ , with game matrix  $W_0$ , surrounded by  $n - 1$  islands of identical size  $N_1 = c_1 N$  and identical game matrix  $W_1$ , with symmetric migration occurring between the continent and each island (with no direct island-to-island migration in one step of the migration process). The total population size denoted by  $cN$ , where  $c = c_0 + (n - 1)c_1$ , is assumed to be



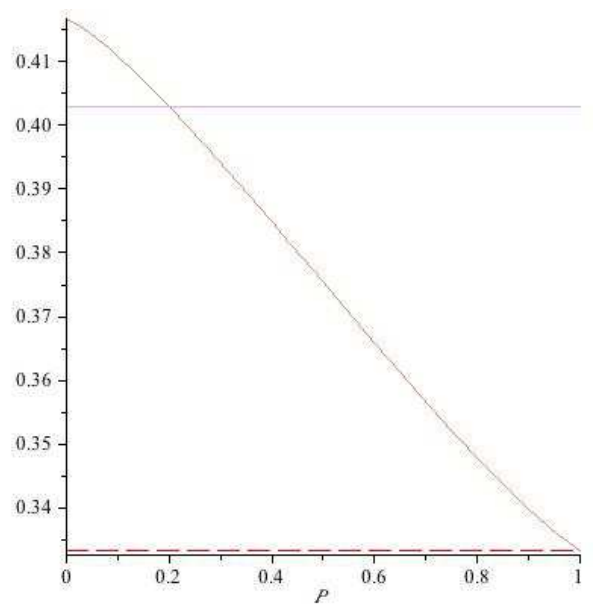
(a)  $M = 1, n = 2$



(b)  $M = 1, n = 3$

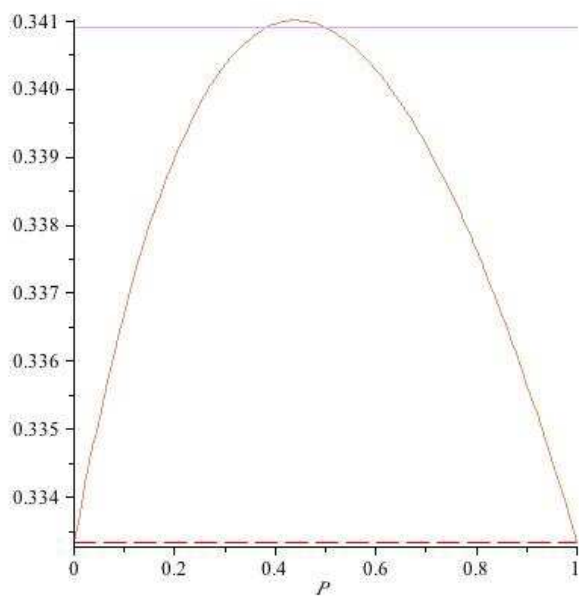


(c)  $M = 1, n = 4$

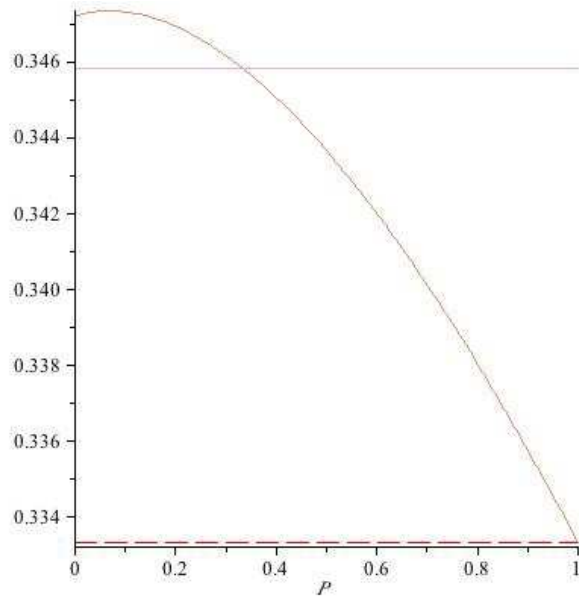


(d)  $M = 1, n = 5$

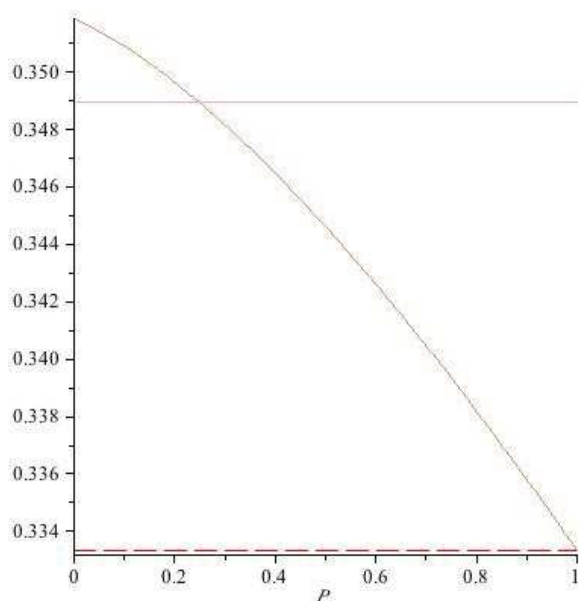
Figure 9: Threshold  $\Lambda$  (brown curve) versus its value in the case of equal deme sizes (purple curve) in the case when  $M = 1$ : (a)  $n = 2$ ; (b)  $n = 3$ ; (c)  $n = 4$ ; (d)  $n = 5$ .



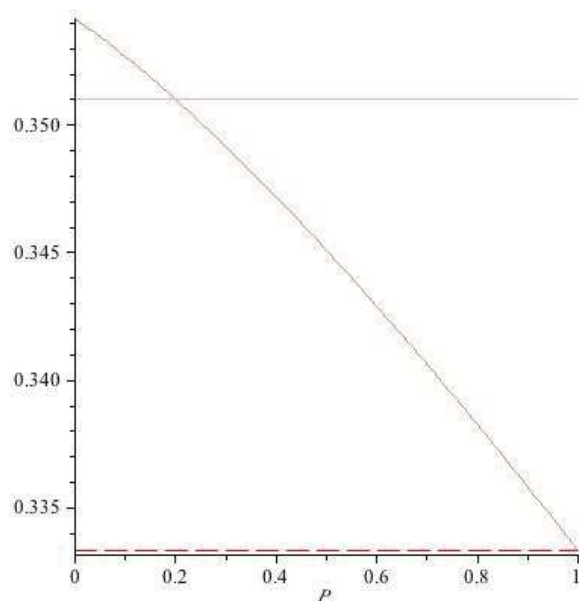
(a)  $M = 10, n = 2$



(b)  $M = 10, n = 3$

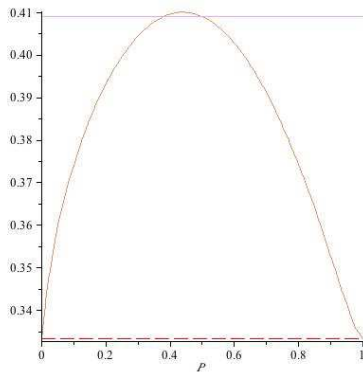


(c)  $M = 10, n = 4$

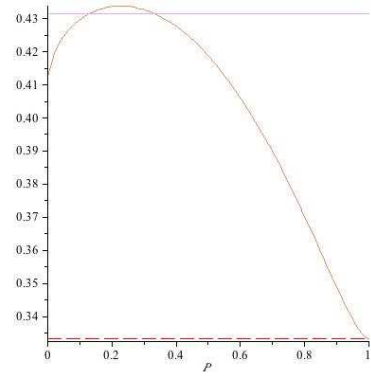


(d)  $M = 10, n = 5$

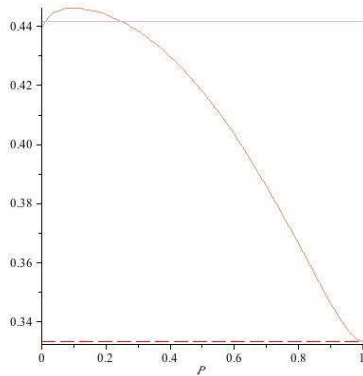
Figure 10: Threshold  $\Lambda$  (brown curve) versus its value in the case of equal deme sizes (purple curve) in the case when  $M = 10$  : (a)  $n = 2$ ; (b)  $n = 3$ ; (c)  $n = 4$ ; (d)  $n = 5$ .



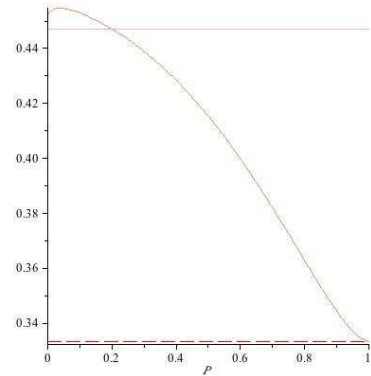
(a)  $M = 0.1$ ,  $n = 2$



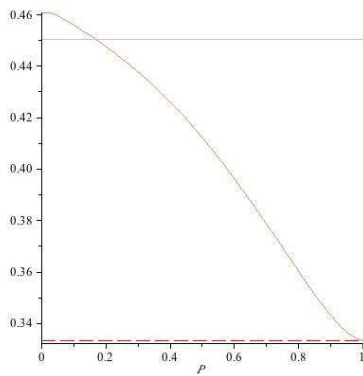
(b)  $M = 0.1$ ,  $n = 3$



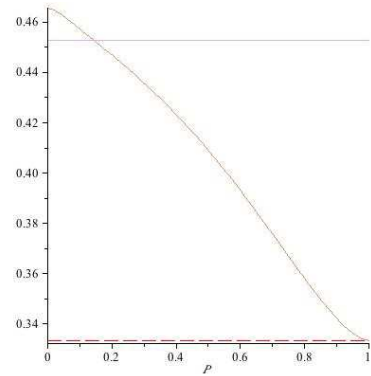
(c)  $M = 0.1$ ,  $n = 4$



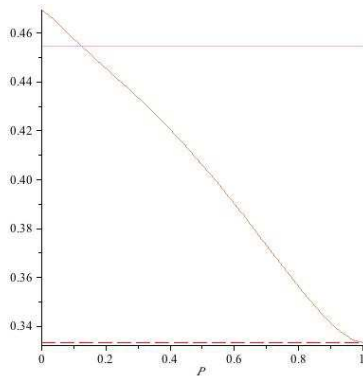
(d)  $M = 0.1$ ,  $n = 5$



(e)  $M = 0.1$ ,  $n = 6$

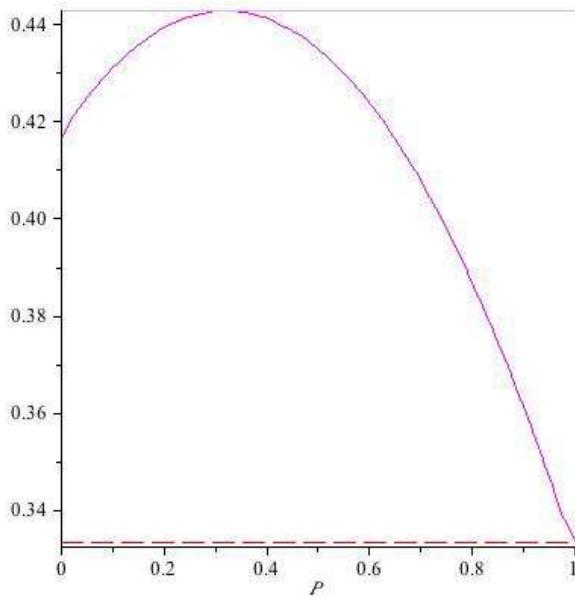


(f)  $M = 0.1$ ,  $n = 7$

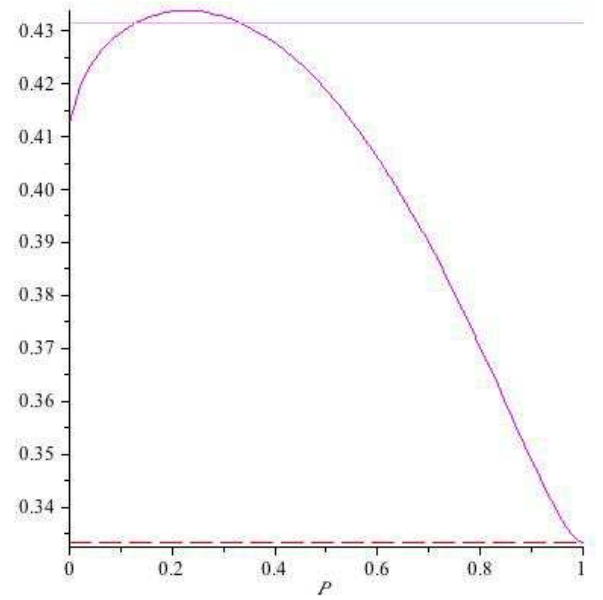


(g)  $M = 0.1$ ,  $n = 8$

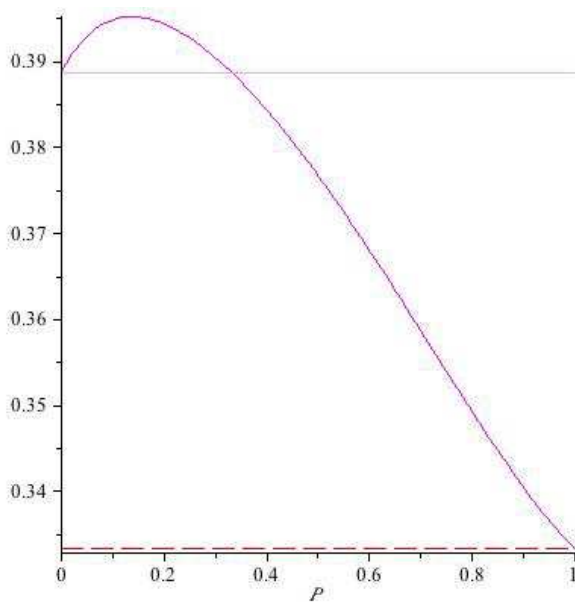
Figure 11: Threshold  $\Lambda$  (brown curve) versus its value in the case of equal deme sizes (purple curve) in the case when  $M = 0.1$ : (a)  $n = 2$ ; (b)  $n = 3$ ; (c)  $n = 4$ ; (d)  $n = 5$ ; (e)  $n = 6$ ; (f)  $n = 7$ ; (g)  $n = 8$ .



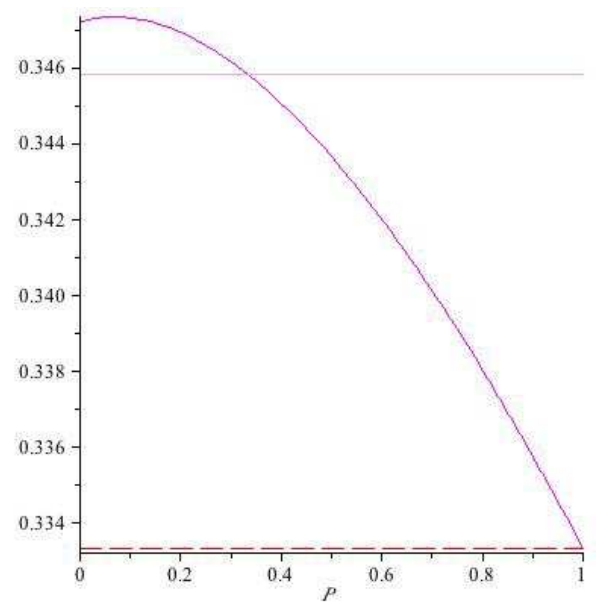
(a)  $n = 3$ ,  $M = 0.01$



(b)  $n = 3$ ,  $M = 0.1$

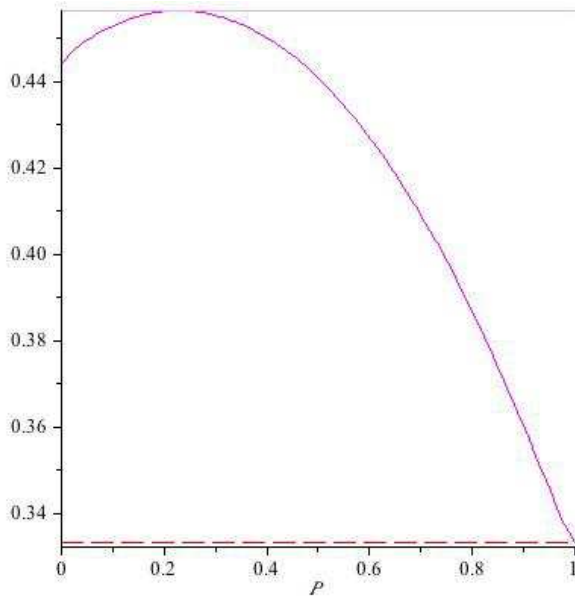


(c)  $n = 3$ ,  $M = 1$

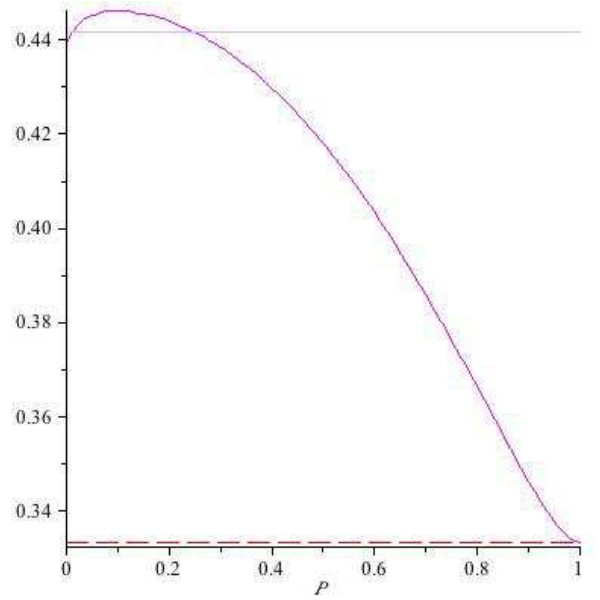


(d)  $n = 3$ ,  $M = 10$

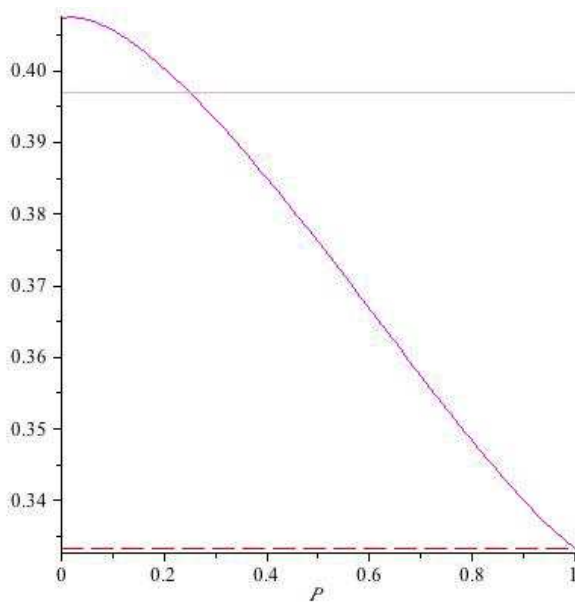
Figure 12: Threshold  $\Lambda$  (magenta curve) versus its value in the case of equal deme sizes (purple curve) in the case when  $n = 3$ : (a)  $M = 0.01$ ; (b)  $M = 0.1$ ; (c)  $M = 1$ ; (d)  $M = 10$ .



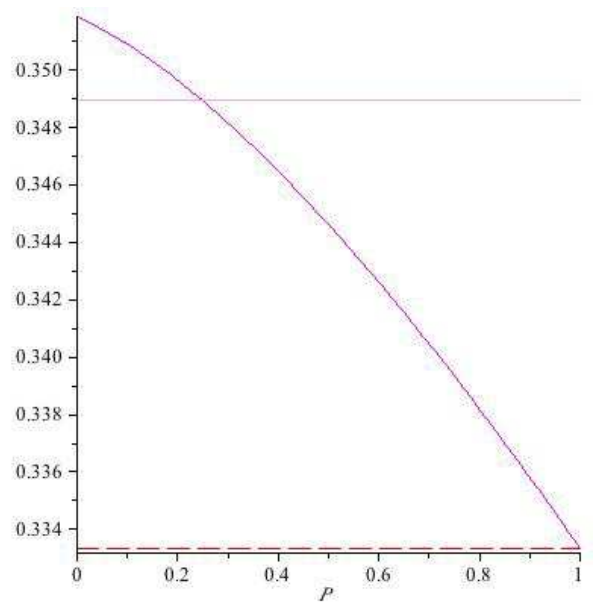
(a)  $n = 4$ ,  $M = 0.01$



(b)  $n = 4$ ,  $M = 0.1$

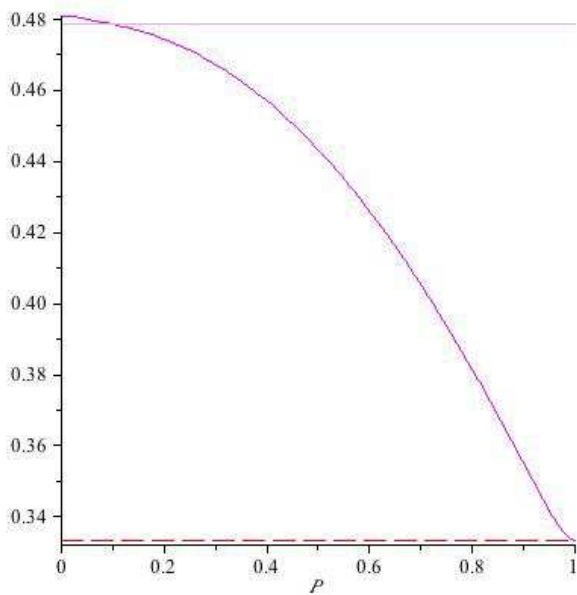


(c)  $n = 4$ ,  $M = 1$

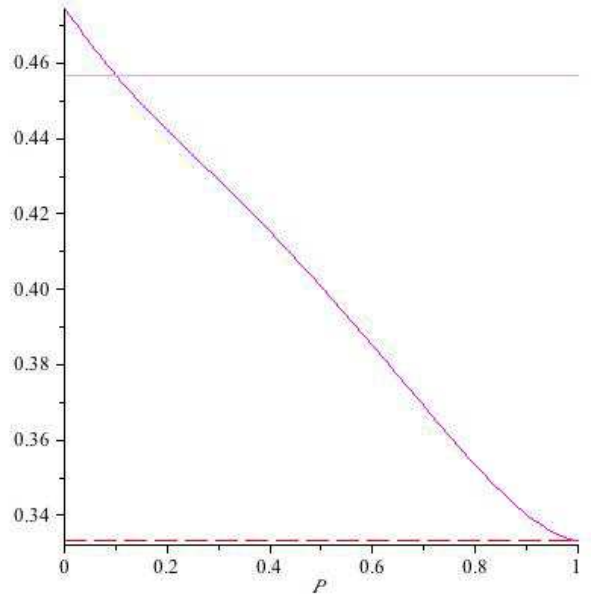


(d)  $n = 4$ ,  $M = 10$

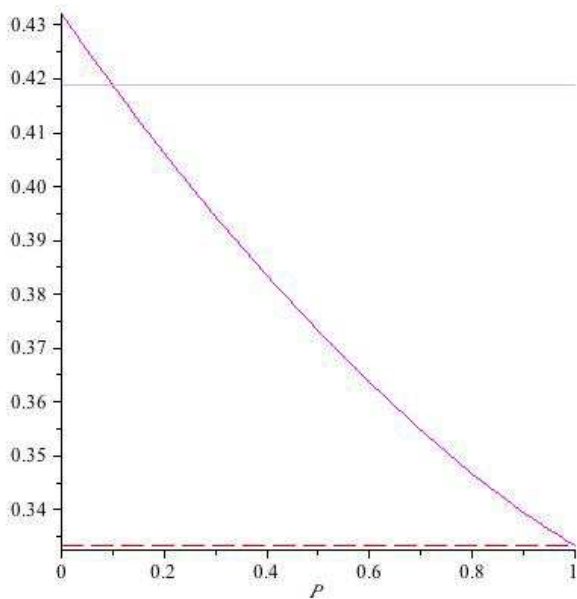
Figure 13: Threshold  $\Lambda$  (magenta curve) versus its value in the case of equal deme sizes (purple curve) in the case when  $n = 4$ : (a)  $M = 0.01$ ; (b)  $M = 0.1$ ; (c)  $M = 1$ ; (d)  $M = 10$ .



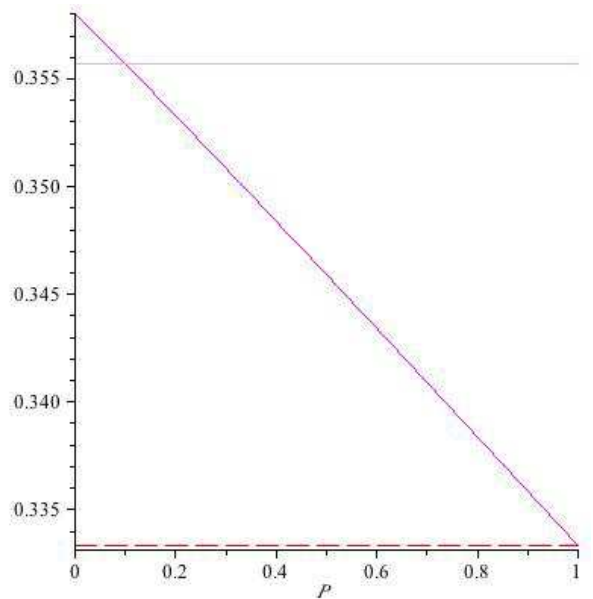
(a)  $n = 10$ ,  $M = 0.01$



(b)  $n = 10$ ,  $M = 0.1$



(c)  $n = 10$ ,  $M = 1$



(d)  $n = 10$ ,  $M = 10$

Figure 14: Threshold  $\Lambda$  (magenta curve) versus its value in the case of equal deme sizes (purple curve) in the case when  $n = 10$ : (a)  $M = 0.01$ ; (b)  $M = 0.1$ ; (c)  $M = 1$ ; (d)  $M = 10$ .

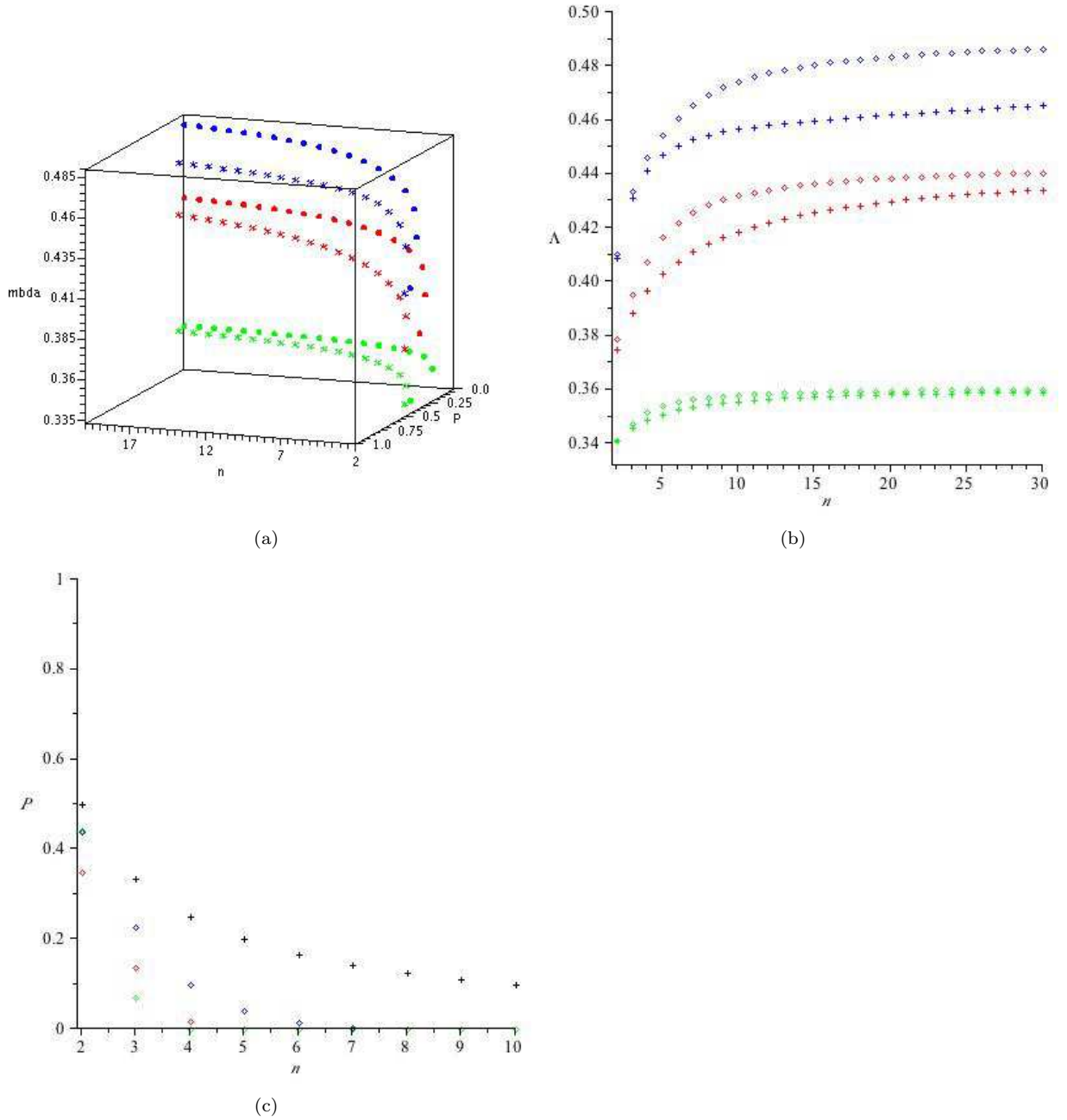


Figure 15: (a) Maximum value taken by threshold  $\Lambda$  (diamonds) versus the value taken by  $\Lambda$  in the case of equal deme sizes (diagonal crosses) as functions of the number of demes,  $n$ , for 3 different values of the migration rate,  $M$ :  $M = 0.1$  and  $n$  ranges from 2 to 20 (blue);  $M = 1$  and  $n$  ranges from 2 to 20 (red);  $M = 10$  and  $n$  ranges from 2 to 20 (green). For any fixed  $n$ , the maximum value of  $\Lambda$  is reached at some  $P = P^{**}(M, n)$ , where  $P$  denotes the proportion of individuals living on the continent. In the case of equal deme sizes,  $P$  is always equal to  $1/n$ . (b) Maximum value taken by threshold  $\Lambda$  (diamonds) versus its value in the case of equal deme sizes (crosses) as a function of the number of demes,  $n$ , for 3 different values of the migration rate,  $M$ :  $M = 0.1$  and  $n$  ranges from 2 to 30 (blue);  $M = 1$  and  $n$  ranges from 2 to 30 (red);  $M = 10$  and  $n$  ranges from 2 to 30 (green); (c) Value taken by  $P$  in which the maximum value for  $\Lambda$  is reached (i.e.  $P^{**}(M, n)$ ) as a function of the number of demes,  $n$ , for 3 different values of the migration rate,  $M$ :  $M = 0.1$  and  $n$  ranges from 2 to 10 (blue);  $M = 1$  and  $n$  ranges from 2 to 10 (red);  $M = 10$  and  $n$  ranges from 2 to 10 (green). The black crosses correspond to the value taken by  $P$  in the case of equal deme sizes ( $P = 1/n$ ).



large, and the backward migration rates are of the form  $m_{0i} = M_0/(2N(n-1))$ , for the probability that an individual from the continent emigrated from island  $i$  in the previous generation, and  $m_{i0} = M_1/(2N)$ , for the backward probability that an individual from island  $i$  emigrated from the continent, and  $M_0$  and  $M_1$  are of order 1, which holds under the structured coalescent assumptions (Notohara, 1990; Herbots, 1994, 1997; Wilkinson-Herbots, 1998). Moreover, assuming conservative migration, the parameter  $M_0$  is related to  $M_1$  by  $c_1 M_1 = c_0 M_0/(n-1) = M$ , where the "migration rate"  $M$  is of order 1. In the context of weak selection ( $s \ll 1/(cN) \ll 1$ ), we have shown that the fixation probability can be approximated by

$$u(s) = \frac{u_0}{N_0} + s \left\{ (p_A - p_B) \cdot \left( u_0 \gamma_0 W_0 + (1 - u_0) \gamma_1 W_1 \right) (p_A - p_B) \right. \\ \left. + (p_A - p_B) \cdot \left\{ u_0 \delta_0 W_0 + (1 - u_0) \delta_1 W_1 \right\} p_B \right\} + o(s), \quad (92)$$

where  $u_0$  denotes the first coordinate of  $(u_0, u_1, \dots, u_1)$ , the stationary distribution of the backward migration matrix  $\mathbf{M}$ , and coefficients  $\delta_i, \gamma_i, i = 0, 1$ , are explicit functions that depend on the migration rate  $M$ , the proportion of the total population living on the continent,  $P = c_0/c$ , and the total number of demes,  $n$ . Note that in the special case when  $n = 2$ , up to the proper parameter rescaling, this model matches exactly the asymmetric two-deme linear game model (Ladret and Lessard, 2008) and in this case, approximation (??) unsurprisingly reduces to (??). The proof in the general case  $n \geq 3$ , is directly adapted from the Markov chain method used in the asymmetric two-deme model (Ladret and Lessard, 2008). It relies on the calculation of the first order effect of selection on the probability of fixation as a function of the expressions  $\sum_{t \geq 0} E_0(X_i^2(t)(1 - X_i(t)))$  and  $\sum_{t \geq 0} E_0(X_i(t)(1 - X_i(t)))$ , with  $i \in \{0, \dots, n-1\}$ , where  $X_i(t)$  denotes the proportion of type  $A$  individuals in deme  $i$  at time  $t$  and  $E_0$  stands for the expectation under neutrality ( $s = 0$ ). These terms are then calculated using a coalescent (Kingman, 1982) based approach adapted to the population structure of the continental island model.

Assuming an infinitely large population subdivided into a continent and  $n - 1$  islands with a constant backward migration matrix,  $\mathbf{M}$ , and weak selection, we have obtained a modified replicator equation which describes how the relative abundance (frequency) of a mutant type in the entire population, averaged with respect to the stationary distribution of the backward migration matrix  $M$ , changes over time. This differential equation is an extension of the modified replicator equation for an infinite asymmetric two-deme population (Ladret and Lessard, 2008). Note that it should provide a good approximation of the dynamics in the case  $1/N \ll s \ll 1$ . Moreover, in this case of an infinite  $n$ -deme continental island linear game model, we have derived that evolutionary stable strategies (ESS) correspond to ESS in the usual context of a linear game for an infinite panmictic population with game matrix  $u_0 W_0 + (n-1)u_1 W_1$ .

Going back to the finite population continental island model with  $n \geq 3$  demes under weak selection, we have investigated conditions under which a resident strategy is selectively

favoured against replacement by a mutant strategy as defined in Nowak et al. (2004). In the particular case of identical game matrices  $W_0 = W_1$ , this condition corresponds to the traditional ESS condition in an infinite population, provided the mutant strategy remains close enough to the resident strategy if the mutant strategy uses some new pure strategies. This also holds when the model is reduced to a single deme (Lessard, 2005; Lessard and Ladret, 2007), when the population structure follows a symmetric finite island model with  $n \geq 2$  demes (Ladret and Lessard, 2007), or when the continental island model is reduced to only  $n = 2$  demes, which corresponds to the case of the asymmetric two-deme model (Ladret and Lessard, 2008). Note however, that in the case of non identical game matrices ( $W_0 \neq W_1$ ), the ESS condition is neither a necessary nor a sufficient condition for a strategy to be selectively advantageous against replacement in the continental island model, as already shown in the special case  $n = 2$  via various counterexamples in Ladret and Lessard (2008).

In the case of two pure strategies  $\mathbf{p}_A = (1, 0)$  and  $\mathbf{p}_B = (0, 1)$  that are best replies to themselves with respect to both the  $2 \times 2$  game matrices  $W_0$  and  $W_1$ ,

$$W_0 = \begin{pmatrix} a_0 & b_0 \\ c_0 & d_0 \end{pmatrix} \quad \text{and} \quad W_1 = \begin{pmatrix} a_1 & b_1 \\ c_1 & d_1 \end{pmatrix},$$

which means that  $a_0 > c_0$ ,  $d_0 > b_0$ ,  $a_1 > c_1$  and  $d_1 > b_1$ , we have specified a condition for the resident strategy to be selectively favoured against replacement by the mutant strategy. Moreover, when the two game matrices are identical ( $W_0 = W_1$ ), we have shown that this condition translates into a criterion of the form  $z^* < \Lambda$  where  $z^*$  denotes the unstable equilibrium frequency of the mutant, averaged with respect to the stationary distribution  $(u_0, \dots, u_n)$  of the backward migration matrix, and where the coefficient  $\Lambda$  is an explicit function of the population parameters  $M$ ,  $n$  and  $P$ . This inequality extends the one-third law to the  $n$ -deme continental island model and when  $n = 2$ , it reduces to the modified one-third law for an asymmetric two-deme model. We have found that this condition is less stringent than the one-third law.

Still in the context of identical game matrices, we have compared the modified one-third law for an asymmetric continental island model with unequal deme sizes ( $P \neq 1/n$ ) with its equivalent for its symmetric (in the sense of equal deme sizes) model counterpart ( $P = 1/n$ ). The two models share the same population parameters except for the proportion of individuals living on the continent, which is equal to  $P \neq 1/n$  in the latter model and  $P = 1/n$  in the former. We have found that when the migration rate,  $M$ , is fixed, there exists a threshold value,  $n^*(M) \geq 2$ , for the number of demes,  $n$ , such that, if  $n \leq n^*(M)$ , then the condition for the asymmetric model is less stringent than the one for its symmetric model counterpart, provided the relative size of the continent,  $P$ , is not too small and remains less than  $1/n$ , as it is already the case for  $n = 2$  (Ladret and Lessard, 2008); while, if the number of demes exceeds  $n^*(M)$ , it only requires the relative size of the continent to be less than  $1/n$ . Moreover, we have found that numerical evaluations suggest that  $n^*(M)$  does not exceed 4 when  $M$  ranges from 0.1 to 10, and does not exceed 7 when  $0.01 \leq M \leq 100$ .

On the other hand, we have compared the  $n$ -deme symmetric continental island model with its symmetric finite island model counterpart. In these two models, there are  $n$  demes of equal size  $\frac{cN}{n}$  with identical game matrix  $W$  within each deme and both models have the same expected total number of migrants per generation after population regulation, only the migration scenario differs. When the number of demes,  $n$ , is larger than 3 (otherwise both models are equivalent) and the migration rate,  $M$ , is fixed, we have found that as long as  $M$  remains larger than  $\sqrt{10}/3$ , the extended one-third law for the continental island model is always less stringent than the extended one-third law for its island model counterpart. On the other hand, for smaller values of  $M$  (i.e.  $M \leq \sqrt{10}/3$ ), we have found that this remains true, provided the number of demes is large enough (more precisely, provided it remains larger than  $n_0(M) = \frac{1+\sqrt{1+4M^2}}{M^2}$ ). Equivalently, when the number of demes,  $n$ , is fixed, we have shown that the extended one-third law for the continental island model is less stringent than its Wright's island counterpart, provided the migration rate,  $M$ , remains large enough (namely, larger than  $M_0(n) = \frac{\sqrt{2n+4}}{n}$ ). Applying these results to the IPD with the strategies TFT versus AllD, we have found that the condition for the cooperative TFT strategy to fixate in the population with a selective advantage is less stringent in a continent island model than in a panmictic model, in the sense that it requires less repetitions of the game. Moreover, the population structure of the symmetric continental island model can facilitate the evolution of cooperation compared to its finite island model counterpart. On the other hand, the asymmetry in the population structure of a continental island model with unequal deme sizes can outperform its symmetric model counterpart when promoting the emergence of cooperation.

## Appendix A.1: Transition probabilities

### A.1.1 Matrix $F$

$$F_{000,000} = (1 - m_0)^3 \left(1 - \frac{1}{N_0}\right) \left(1 - \frac{2}{N_0}\right) = 1 - \frac{3}{N_0} - 3m_0 + O\left(\frac{1}{N_0^2}\right)$$

$$F_{000,001} = (1 - m_0)^2 m_0 \left(1 - \frac{1}{N_0}\right)$$

$$F_{000,010} = (1 - m_0)^2 m_0 \left(1 - \frac{1}{N_0}\right)$$

$$F_{000,100} = (1 - m_0)^2 m_0 \left(1 - \frac{1}{N_0}\right)$$

$$F_{000,011} = (1 - m_0) m_0 \frac{m_0}{n-1} \left(1 - \frac{1}{N_1}\right)$$

$$F_{000,101} = (1 - m_0) m_0 \frac{m_0}{n-1} \left(1 - \frac{1}{N_1}\right)$$

$$F_{000,110} = (1 - m_0) m_0 \frac{m_0}{n-1} \left(1 - \frac{1}{N_1}\right)$$

$$F_{000,111} = m_0 \left(\frac{m_0}{n-1}\right)^2 \left(1 - \frac{1}{N_1}\right) \left(1 - \frac{2}{N_1}\right)$$

$$F_{000,012} = (1 - m_0) m_0^2 \left(1 - \frac{1}{n-1}\right)$$

$$\begin{aligned}
F_{000,102} &= (1 - m_0)m_0^2\left(1 - \frac{1}{n-1}\right) \\
F_{000,120} &= (1 - m_0)m_0^2\left(1 - \frac{1}{n-1}\right) \\
F_{000,123} &= m_0^3\left(1 - \frac{1}{n-1}\right)\left(1 - \frac{2}{n-1}\right) \\
F_{000,112} &= m_0^2\frac{m_0}{n-1}\left(1 - \frac{1}{n-1}\right)\left(1 - \frac{1}{N_1}\right) \\
F_{000,121} &= m_0^2\frac{m_0}{n-1}\left(1 - \frac{1}{n-1}\right)\left(1 - \frac{1}{N_1}\right) \\
F_{000,211} &= m_0^2\frac{m_0}{n-1}\left(1 - \frac{1}{n-1}\right)\left(1 - \frac{1}{N_1}\right) \\
F_{001,000} &= (1 - m_0)^2m_1\left(1 - \frac{1}{N_0}\right)\left(1 - \frac{2}{N_0}\right) \\
F_{001,001} &= (1 - m_0)^2(1 - m_1)\left(1 - \frac{1}{N_0}\right) \\
F_{001,010} &= m_0m_1(1 - m_0)\left(1 - \frac{1}{N_0}\right) \\
F_{001,100} &= m_0m_1(1 - m_0)\left(1 - \frac{1}{N_0}\right) \\
F_{001,011} &= (1 - m_0)\frac{m_0}{n-1}(1 - m_1)\left(1 - \frac{1}{N_1}\right) \\
F_{001,101} &= (1 - m_0)\frac{m_0}{n-1}(1 - m_1)\left(1 - \frac{1}{N_1}\right) \\
F_{001,110} &= m_1m_0\frac{m_0}{n-1}\left(1 - \frac{1}{N_1}\right) \\
F_{001,111} &= \left(\frac{m_0}{n-1}\right)^2(1 - m_1)\left(1 - \frac{1}{N_1}\right)\left(1 - \frac{2}{N_1}\right) \\
F_{001,012} &= (1 - m_0)(1 - m_1)m_0\left(1 - \frac{1}{n-1}\right) \\
F_{001,102} &\text{ (cf } F_{001,201}\text{)} \\
F_{001,201} &= (1 - m_0)(1 - m_1)m_0\left(1 - \frac{1}{n-1}\right) \\
F_{001,120} &= m_1m_0m_0\left(1 - \frac{1}{n-1}\right) \\
F_{001,123} &= (1 - m_1)m_0\left(1 - \frac{1}{n-1}\right)m_0\left(1 - \frac{2}{n-1}\right) \\
F_{001,112} &= F_{001,221} = (1 - m_1)m_0\left(1 - \frac{1}{n-1}\right)\frac{m_0}{n-1}\left(1 - \frac{1}{N_1}\right) \\
F_{001,121} &= (1 - m_1)\frac{m_0}{n-1}m_0\left(1 - \frac{1}{n-1}\right)\left(1 - \frac{1}{N_1}\right) \\
F_{001,211} &= (1 - m_1)\frac{m_0}{n-1}m_0\left(1 - \frac{1}{n-1}\right)\left(1 - \frac{1}{N_1}\right)
\end{aligned}$$

$$F_{010,000} = m_1(1 - m_0)^2(1 - \frac{1}{N_0})(1 - \frac{2}{N_0})$$

$$F_{010,001} = m_0m_1(1 - m_0)(1 - \frac{1}{N_0})$$

$$F_{010,010} = (1 - m_0)^2(1 - m_1)(1 - \frac{1}{N_0})$$

$$F_{010,100} = m_0m_1(1 - m_0)(1 - \frac{1}{N_0})$$

$$F_{010,011} = \frac{m_0}{n-1}(1 - m_0)(1 - m_1)(1 - \frac{1}{N_1})$$

$$F_{010,101} = m_0m_1\frac{m_0}{n-1}(1 - \frac{1}{N_1})$$

$$F_{010,110} = \frac{m_0}{n-1}(1 - m_0)(1 - m_1)(1 - \frac{1}{N_1})$$

$$F_{010,111} = (1 - m_1)(1 - \frac{1}{N_1})(1 - \frac{2}{N_1})(\frac{m_0}{n-1})^2$$

$$F_{010,012} = (1 - m_0)(1 - m_1)m_0(1 - \frac{1}{n-1})$$

$$F_{010,102} = m_1m_0m_0(1 - \frac{1}{n-1})$$

$$F_{010,120} = 0$$

$$F_{010,210} = (1 - m_0)(1 - m_1)m_0(1 - \frac{1}{n-1})$$

$$F_{010,123} = (1 - m_1)m_0(1 - \frac{1}{n-1})m_0(1 - \frac{2}{n-1})$$

$$F_{010,112} = \frac{m_0}{n-1}(1 - m_1)m_0(1 - \frac{1}{N_1})(1 - \frac{1}{n-1})$$

$$F_{010,121} = (1 - m_1)m_0(1 - \frac{1}{n-1})\frac{m_0}{n-1}(1 - \frac{1}{N_1})$$

$$F_{010,211} = (1 - m_1)\frac{m_0}{n-1}(1 - \frac{1}{N_1})m_0(1 - \frac{1}{n-1})$$

$$F_{100,000} = m_1(1 - m_0)^2(1 - \frac{1}{N_0})(1 - \frac{2}{N_0})$$

$$F_{100,001} = m_0m_1(1 - m_0)(1 - \frac{1}{N_0})$$

$$F_{100,010} = m_0m_1(1 - m_0)(1 - \frac{1}{N_0})$$

$$F_{100,100} = (1 - m_0)^2(1 - m_1)(1 - \frac{1}{N_0})$$

$$F_{100,011} = m_0m_1(1 - \frac{1}{N_1})\frac{m_0}{n-1}$$

$$F_{100,101} = (1 - m_0)(1 - m_1)(1 - \frac{1}{N_1})\frac{m_0}{n-1}$$

$$F_{100,110} = \frac{m_0}{n-1}(1 - m_0)(1 - m_1)(1 - \frac{1}{N_1})$$

$$F_{100,111} = \left(\frac{m_0}{n-1}\right)^2 (1 - m_1) \left(1 - \frac{1}{N_1}\right) \left(1 - \frac{2}{N_1}\right)$$

$$F_{100,012} = m_0^2 m_1 \left(1 - \frac{1}{n-1}\right)$$

$$F_{100,102} = (1 - m_1) (1 - m_0) m_0 \left(1 - \frac{1}{n-1}\right)$$

$$F_{100,120} = (1 - m_1) (1 - m_0) m_0 \left(1 - \frac{1}{n-1}\right)$$

$$F_{100,123} = (1 - m_1) m_0 \left(1 - \frac{1}{n-1}\right) m_0 \left(1 - \frac{2}{n-1}\right)$$

$$F_{100,112} = (1 - m_1) \frac{m_0}{n-1} m_0 \left(1 - \frac{1}{n-1}\right) \left(1 - \frac{1}{N_1}\right)$$

$$F_{100,121} = (1 - m_1) \frac{m_0}{n-1} m_0 \left(1 - \frac{1}{n-1}\right) \left(1 - \frac{1}{N_1}\right)$$

$$F_{100,211} = (1 - m_1) m_0 \left(1 - \frac{1}{n-1}\right) \frac{m_0}{n-1} \left(1 - \frac{1}{N_1}\right)$$

$$F_{011,000} = m_1^2 (1 - m_0) \left(1 - \frac{1}{N_0}\right) \left(1 - \frac{2}{N_0}\right)$$

$$F_{011,001} = (1 - m_0) (1 - m_1) m_1 \left(1 - \frac{1}{N_0}\right)$$

$$F_{011,010} = (1 - m_0) (1 - m_1) m_1 \left(1 - \frac{1}{N_0}\right)$$

$$F_{011,100} = m_0 m_1^2 \left(1 - \frac{1}{N_0}\right)$$

$$F_{011,011} = (1 - m_0) (1 - m_1)^2 \left(1 - \frac{1}{N_1}\right)$$

$$F_{011,101} = \frac{m_0}{n-1} (1 - m_1) m_1 \left(1 - \frac{1}{N_1}\right)$$

$$F_{011,110} = (1 - m_1) \frac{m_0}{n-1} m_1 \left(1 - \frac{1}{N_1}\right)$$

$$F_{011,111} = \frac{m_0}{n-1} (1 - m_1)^2 \left(1 - \frac{1}{N_1}\right) \left(1 - \frac{2}{N_1}\right)$$

$$F_{011,012} = 0$$

$$F_{011,102} = (1 - m_1) m_1 m_0 \left(1 - \frac{1}{n-1}\right)$$

$$F_{011,120} = (1 - m_1) m_1 m_0 \left(1 - \frac{1}{n-1}\right)$$

$$F_{011,123} = 0$$

$$F_{011,112} = 0$$

$$F_{011,121} = 0$$

$$F_{011,211} = (1 - m_1)^2 m_0 \left(1 - \frac{1}{n-1}\right) \left(1 - \frac{1}{N_1}\right)$$

$$F_{101,000} = (1 - m_0) m_1^2 \left(1 - \frac{1}{N_0}\right) \left(1 - \frac{2}{N_0}\right)$$

$$F_{101,001} = (1 - m_1) m_1 (1 - m_0) \left(1 - \frac{1}{N_0}\right)$$

$$F_{101,010} = m_1^2 m_0 \left(1 - \frac{1}{N_0}\right)$$

$$F_{101,100} = (1 - m_0) (1 - m_1) m_1 \left(1 - \frac{1}{N_0}\right)$$

$$F_{101,011} = (1 - m_1) m_1 \frac{m_0}{n-1} \left(1 - \frac{1}{N_1}\right)$$

$$F_{101,101} = (1 - m_1)^2 (1 - m_0) \left(1 - \frac{1}{N_1}\right)$$

$$F_{101,110} = (1 - m_1) \frac{m_0}{n-1} m_1 \left(1 - \frac{1}{N_1}\right)$$

$$F_{101,111} = (1 - m_1)^2 \frac{m_0}{n-1} \left(1 - \frac{1}{N_1}\right) \left(1 - \frac{2}{N_1}\right)$$

$$F_{101,012} = m_1 (1 - m_1) m_0 \left(1 - \frac{1}{n-1}\right)$$

$$F_{101,102} = 0$$

$$F_{101,120} = (1 - m_1) m_1 m_0 \left(1 - \frac{1}{n-1}\right)$$

$$F_{101,123} = 0$$

$$F_{101,112} = 0$$

$$F_{101,121} = (1 - m_1)^2 m_0 \left(1 - \frac{1}{n-1}\right) \left(1 - \frac{1}{N_1}\right)$$

$$F_{101,211} = 0$$

$$F_{110,000} = F_{101,000} = F_{011,000} = (1 - m_0) m_1^2 \left(1 - \frac{1}{N_0}\right) \left(1 - \frac{2}{N_0}\right)$$

$$F_{110,001} = m_1^2 m_0 \left(1 - \frac{1}{N_0}\right)$$

$$F_{110,010} = (1 - m_1) (1 - m_0) m_1 \left(1 - \frac{1}{N_0}\right)$$

$$F_{110,100} = (1 - m_1) (1 - m_0) m_1 \left(1 - \frac{1}{N_0}\right)$$

$$F_{110,011} = (1 - m_1) m_1 \frac{m_0}{n-1} \left(1 - \frac{1}{N_1}\right)$$

$$F_{110,101} = (1 - m_1) m_1 \frac{m_0}{n-1} \left(1 - \frac{1}{N_1}\right)$$

$$F_{110,110} = (1 - m_1)^2 (1 - m_0) \left(1 - \frac{1}{N_1}\right)$$

$$F_{110,111} = (1 - m_1)^2 \frac{m_0}{n-1} \left(1 - \frac{1}{N_1}\right) \left(1 - \frac{2}{N_1}\right)$$

$$F_{110,012} = (1 - m_1) m_1 m_0 \left(1 - \frac{1}{n-1}\right)$$

$$F_{110,102} = (1 - m_1) m_1 m_0 \left(1 - \frac{1}{n-1}\right)$$

$$F_{110,120} = 0$$

$$F_{110,123} = 0$$

$$F_{110,112} = (1 - m_1)^2 m_0 \left(1 - \frac{1}{n-1}\right) \left(1 - \frac{1}{N_1}\right)$$

$$F_{110,121} = 0$$

$$F_{110,211} = 0$$

$$F_{111,000} = m_1^3 \left(1 - \frac{1}{N_0}\right) \left(1 - \frac{2}{N_0}\right)$$

$$F_{111,001} = m_1^2 (1 - m_1) \left(1 - \frac{1}{N_0}\right)$$

$$F_{111,010} = m_1^2 (1 - m_1) \left(1 - \frac{1}{N_0}\right)$$

$$F_{111,100} = m_1^2 (1 - m_1) \left(1 - \frac{1}{N_0}\right)$$

$$F_{111,011} = m_1 (1 - m_1)^2 \left(1 - \frac{1}{N_1}\right)$$

$$F_{111,101} = m_1 (1 - m_1)^2 \left(1 - \frac{1}{N_1}\right)$$

$$F_{111,110} = m_1 (1 - m_1)^2 \left(1 - \frac{1}{N_1}\right)$$

$$F_{111,111} = (1 - m_1)^3 \left(1 - \frac{1}{N_1}\right) \left(1 - \frac{2}{N_1}\right)$$

$$F_{111,012} = 0$$

$$F_{111,102} = 0$$

$$F_{111,120} = 0$$

$$F_{111,123} = 0$$

$$F_{111,112} = 0$$

$$F_{111,121} = 0$$



$$F_{111,211} = 0$$

$$F_{012,000} = (1 - m_0)m_1^2(1 - \frac{1}{N_0})(1 - \frac{2}{N_0})$$

$$F_{012,001} = F_{012,002} = (1 - m_0)m_1(1 - m_1)(1 - \frac{1}{N_0})$$

$$F_{012,010} = (1 - m_0)m_1(1 - m_1)(1 - \frac{1}{N_0})$$

$$F_{012,100} = m_1^2m_0(1 - \frac{1}{N_0})$$

$$F_{012,011} = 0$$

$$F_{012,101} = \frac{m_0}{n-1}(1 - m_1)m_1(1 - \frac{1}{N_1})$$

$$F_{012,110} = \frac{m_0}{n-1}(1 - m_1)m_1(1 - \frac{1}{N_1})$$

$$F_{012,111} = 0$$

$$F_{012,012} = (1 - m_0)(1 - m_1)^2$$

$$F_{012,102} = m_1m_0(1 - \frac{1}{n-1})(1 - m_1)$$

$$F_{012,120} = m_1(1 - m_1)m_0(1 - \frac{1}{n-1})$$

$$F_{012,123} = (1 - m_1)^2m_0(1 - \frac{2}{n-1})$$

$$F_{012,112} = \frac{m_0}{n-1}(1 - m_1)^2(1 - \frac{1}{N_1})$$

$$F_{012,121} = (1 - m_1)^2\frac{m_0}{n-1}(1 - \frac{1}{N_1})$$

$$F_{012,211} = 0$$

$$F_{102,000} = m_1^2(1 - m_0)(1 - \frac{1}{N_0})(1 - \frac{2}{N_0})$$

$$F_{102,001} = (1 - m_1)(1 - m_0)m_1(1 - \frac{1}{N_0})$$

$$F_{102,010} = m_1^2m_0(1 - \frac{1}{N_0})$$

$$F_{102,100} = (1 - m_1)(1 - m_0)m_1(1 - \frac{1}{N_0})$$

$$F_{102,011} = m_1(1 - m_1)\frac{m_0}{n-1}(1 - \frac{1}{N_1})$$

$$F_{102,101} = 0$$

$$F_{102,110} = (1 - m_1)\frac{m_0}{n-1}m_1(1 - \frac{1}{N_1})$$

$$F_{102,111} = 0$$

$$F_{102,012} = m_1(1 - m_1)m_0(1 - \frac{1}{n-1})$$

$$F_{102,102} = (1 - m_1)^2(1 - m_0)$$

$$F_{102,120} = (1 - m_1)m_0(1 - \frac{1}{n-1})m_1$$

$$F_{102,123} = (1 - m_1)^2m_0(1 - \frac{2}{n-1})$$

$$F_{102,112} = (1 - m_1)^2\frac{m_0}{n-1}(1 - \frac{1}{N_1})$$

$$F_{102,121} = 0$$

$$F_{102,211} = (1 - m_1)^2\frac{m_0}{n-1}(1 - \frac{1}{N_1})$$

$$F_{120,000} = m_1^2(1 - m_0)(1 - \frac{1}{N_0})(1 - \frac{2}{N_0})$$

$$F_{120,001} = m_1^2m_0(1 - \frac{1}{N_0})$$

$$F_{120,010} = m_1(1 - m_0)(1 - m_1)(1 - \frac{1}{N_0})$$

$$F_{120,100} = m_1(1 - m_0)(1 - m_1)(1 - \frac{1}{N_0})$$

$$F_{120,011} = m_1(1 - m_1)\frac{m_0}{n-1}(1 - \frac{1}{N_1})$$

$$F_{120,101} = (1 - m_1)\frac{m_0}{n-1}m_1(1 - \frac{1}{N_1})$$

$$F_{120,110} = 0$$

$$F_{120,111} = 0$$

$$F_{120,012} = m_1(1 - m_1)m_0(1 - \frac{1}{n-1})$$

$$F_{120,102} = (1 - m_1)m_1m_0(1 - \frac{1}{n-1})$$

$$F_{120,120} = (1 - m_1)^2(1 - m_0)$$

$$F_{120,123} = (1 - m_1)^2m_0(1 - \frac{2}{n-1})$$

$$F_{120,112} = 0$$

$$F_{120,121} = (1 - m_1)^2\frac{m_0}{n-1}(1 - \frac{1}{N_1})$$

$$F_{120,211} = (1 - m_1)^2 \frac{m_0}{n-1} \left(1 - \frac{1}{N_1}\right)$$

$$F_{123,000} = m_1^3 \left(1 - \frac{1}{N_0}\right) \left(1 - \frac{2}{N_0}\right)$$

$$F_{123,001} = (1 - m_1) m_1^2 \left(1 - \frac{1}{N_0}\right)$$

$$F_{123,010} = m_1^2 (1 - m_1) \left(1 - \frac{1}{N_0}\right)$$

$$F_{123,100} = m_1^2 (1 - m_1) \left(1 - \frac{1}{N_0}\right)$$

$$F_{123,011} = 0$$

$$F_{123,101} = 0$$

$$F_{123,110} = 0$$

$$F_{123,111} = 0$$

$$F_{123,012} = m_1 (1 - m_1)^2$$

$$F_{123,102} = F_{123,103} = m_1 (1 - m_1)^2$$

$$F_{123,120} = m_1 (1 - m_1)^2$$

$$F_{123,123} = (1 - m_1)^3$$

$$F_{123,112} = 0$$

$$F_{123,121} = 0$$

$$F_{123,211} = 0$$

$$F_{112,000} = m_1^3 \left(1 - \frac{1}{N_0}\right) \left(1 - \frac{2}{N_0}\right)$$

$$F_{112,001} = m_1^2 (1 - m_1) \left(1 - \frac{1}{N_0}\right)$$

$$F_{112,010} = m_1^2 (1 - m_1) \left(1 - \frac{1}{N_0}\right)$$

$$F_{112,100} = m_1^2 (1 - m_1) \left(1 - \frac{1}{N_0}\right)$$

$$F_{112,011} = 0$$

$$F_{112,101} = 0$$

$$F_{112,110} = (1 - m_1)^2 m_1 \left(1 - \frac{1}{N_1}\right)$$

$$F_{112,111} = 0$$

$$F_{112,012} = m_1(1 - m_1)^2$$

$$F_{112,102} = (1 - m_1)^2 m_1$$

$$F_{112,120} = 0$$

$$F_{112,123} = 0$$

$$F_{112,112} = (1 - m_1)^3 \left(1 - \frac{1}{N_1}\right)$$

$$F_{112,121} = 0$$

$$F_{112,211} = 0$$

$$F_{121,000} = m_1^3 \left(1 - \frac{1}{N_0}\right) \left(1 - \frac{2}{N_0}\right)$$

$$F_{121,001} = m_1^2 (1 - m_1) \left(1 - \frac{1}{N_0}\right)$$

$$F_{121,010} = m_1^2 (1 - m_1) \left(1 - \frac{1}{N_0}\right)$$

$$F_{121,100} = (1 - m_1) m_1^2 \left(1 - \frac{1}{N_0}\right)$$

$$F_{121,011} = 0$$

$$F_{121,101} = (1 - m_1)^2 m_1 \left(1 - \frac{1}{N_1}\right)$$

$$F_{121,110} = 0$$

$$F_{121,111} = 0$$

$$F_{121,012} = m_1(1 - m_1)^2$$

$$F_{121,102} = 0$$

$$F_{121,120} = (1 - m_1)^2 m_1$$

$$F_{121,123} = 0$$

$$F_{121,112} = 0$$

$$F_{121,121} = (1 - m_1)^3 \left(1 - \frac{1}{N_1}\right)$$

$$F_{121,211} = 0$$

$$F_{211,000} = m_1^3 \left(1 - \frac{1}{N_0}\right) \left(1 - \frac{2}{N_0}\right)$$

$$F_{211,001} = m_1^2 (1 - m_1) \left(1 - \frac{1}{N_0}\right)$$

$$F_{211,010} = m_1^2 (1 - m_1) \left(1 - \frac{1}{N_0}\right)$$

$$F_{211,100} = (1 - m_1) m_1^2 \left(1 - \frac{1}{N_0}\right)$$

$$F_{211,011} = m_1 (1 - m_1)^2 \left(1 - \frac{1}{N_1}\right)$$

$$F_{211,101} = 0$$

$$F_{211,110} = 0$$

$$F_{211,111} = 0$$

$$F_{211,012} = 0$$

$$F_{211,102} = (1 - m_1)^2 m_1$$

$$F_{211,120} = (1 - m_1)^2 m_1$$

$$F_{211,123} = 0$$

$$F_{211,112} = 0$$

$$F_{211,121} = 0$$

$$F_{211,211} = (1 - m_1)^3 \left(1 - \frac{1}{N_1}\right)$$

### A.1.2 Matrix $Q$

$$Q_{000,00} = (1 - m_0)^3 \left(1 - \frac{1}{N_0}\right) \frac{1}{N_0}$$

$$Q_{000,01} = (1 - m_0)^2 m_0 \frac{1}{N_0}$$

$$Q_{000,10} = (1 - m_0) m_0 \frac{m_0}{n-1} \frac{1}{N_1}$$

$$Q_{000,11} = m_0 \left(\frac{m_0}{n-1}\right)^2 \frac{1}{N_1} \left(1 - \frac{1}{N_1}\right)$$

$$\begin{aligned}
Q_{000,12} &= m_0 \frac{m_0}{n-1} m_0 \left(1 - \frac{1}{n-1}\right) \frac{1}{N_1} \\
Q_{001,00} &= (1 - m_0)^2 \frac{1}{N_0} m_1 \left(1 - \frac{1}{N_0}\right) \\
Q_{001,01} &= (1 - m_0)^2 (1 - m_1) \frac{1}{N_0} \\
Q_{001,10} &= m_0 \frac{m_0}{n-1} \frac{1}{N_1} m_1 \\
Q_{001,11} &= \left(\frac{m_0}{n-1}\right)^2 (1 - m_1) \frac{1}{N_1} \\
Q_{001,12} &= m_0 \left(1 - \frac{1}{n-1}\right) \frac{m_0}{n-1} (1 - m_1) \frac{1}{N_1} \\
Q_{010,00} &= (1 - m_0)^2 \frac{1}{N_0} m_1 \left(1 - \frac{1}{N_0}\right) \\
Q_{010,01} &= (1 - m_0) m_1 m_0 \frac{1}{N_0} \\
Q_{010,10} &= (1 - m_0) (1 - m_1) \frac{m_0}{n-1} \frac{1}{N_1} \\
Q_{010,11} &= (1 - m_1) \left(\frac{m_0}{n-1}\right)^2 \frac{1}{N_1} \left(1 - \frac{1}{N_1}\right) \\
Q_{010,12} &= (1 - m_1) \frac{m_0}{n-1} m_0 \left(1 - \frac{1}{n-1}\right) \frac{1}{N_1} \\
Q_{100,00} &= m_1 (1 - m_0)^2 \frac{1}{N_0} \left(1 - \frac{1}{N_0}\right) \\
Q_{100,01} &= m_1 (1 - m_0) m_0 \frac{1}{N_0} \\
Q_{100,10} &= (1 - m_1) \frac{m_0}{n-1} (1 - m_0) \frac{1}{N_1} \\
Q_{100,11} &= (1 - m_1) \left(\frac{m_0}{n-1}\right)^2 \frac{1}{N_1} \left(1 - \frac{1}{N_1}\right) \\
Q_{100,12} &= (1 - m_1) \frac{m_0}{n-1} m_0 \left(1 - \frac{1}{n-1}\right) \frac{1}{N_1} \\
Q_{011,00} &= (1 - m_0) m_1^2 \left(1 - \frac{1}{N_0}\right) \frac{1}{N_0} \\
Q_{011,01} &= (1 - m_0) (1 - m_1) m_1 \frac{1}{N_0} \\
Q_{011,10} &= (1 - m_1) \frac{m_0}{n-1} m_1 \frac{1}{N_1} \\
Q_{011,11} &= (1 - m_1)^2 \frac{m_0}{n-1} \frac{1}{N_1} \left(1 - \frac{1}{N_1}\right) \\
Q_{011,12} &= 0 \\
Q_{101,00} &= (1 - m_0) m_1^2 \frac{1}{N_0} \left(1 - \frac{1}{N_0}\right) \\
Q_{101,01} &= (1 - m_1) (1 - m_0) m_1 \frac{1}{N_0}
\end{aligned}$$

$$\begin{aligned}
Q_{101,10} &= (1 - m_1) \frac{m_0}{n-1} m_1 \frac{1}{N_1} \\
Q_{101,11} &= (1 - m_1)^2 \frac{m_0}{n-1} \frac{1}{N_1} (1 - \frac{1}{N_1}) \\
Q_{101,12} &= 0 \\
Q_{110,00} &= (1 - m_0) m_1^2 \frac{1}{N_0} (1 - \frac{1}{N_0}) \\
Q_{110,01} &= m_1^2 m_0 \frac{1}{N_0} \\
Q_{110,10} &= (1 - m_1)^2 (1 - m_0) \frac{1}{N_1} \\
Q_{110,11} &= (1 - m_1)^2 \frac{m_0}{n-1} \frac{1}{N_1} (1 - \frac{1}{N_1}) \\
Q_{110,12} &= (1 - m_1)^2 m_0 (1 - \frac{1}{n-1}) \frac{1}{N_1} \\
Q_{111,00} &= m_1^3 \frac{1}{N_0} (1 - \frac{1}{N_0}) \\
Q_{111,01} &= m_1^2 (1 - m_1) \frac{1}{N_0} \\
Q_{111,10} &= (1 - m_1)^2 m_1 \frac{1}{N_1} \\
Q_{111,11} &= (1 - m_1)^3 \frac{1}{N_1} (1 - \frac{1}{N_1}) \\
Q_{111,12} &= 0 \\
Q_{012,00} &= (1 - m_0) m_1^2 \frac{1}{N_0} (1 - \frac{1}{N_0}) \\
Q_{012,01} &= (1 - m_0) m_1 (1 - m_1) \frac{1}{N_0} \\
Q_{012,10} &= (1 - m_1) \frac{m_0}{n-1} m_1 \frac{1}{N_1} \\
Q_{012,11} &= 0 \\
Q_{012,12} &= (1 - m_1)^2 \frac{m_0}{n-1} \frac{1}{N_1} \\
Q_{102,00} &= (1 - m_0) m_1^2 \frac{1}{N_0} (1 - \frac{1}{N_0}) \\
Q_{102,01} &= (1 - m_0) (1 - m_1) m_1 \frac{1}{N_0} \\
Q_{102,10} &= m_1 (1 - m_1) \frac{m_0}{n-1} \frac{1}{N_1} \\
Q_{102,11} &= 0
\end{aligned}$$

$$Q_{102,12} = (1 - m_1)^2 \frac{m_0}{n-1} \frac{1}{N_1}$$

$$Q_{120,00} = (1 - m_0) m_1^2 \frac{1}{N_0} \left(1 - \frac{1}{N_0}\right)$$

$$Q_{120,01} = m_1^2 m_0 \frac{1}{N_0}$$

$$Q_{120,10} = 0$$

$$Q_{120,11} = 0$$

$$Q_{120,12} = 0$$

$$Q_{123,00} = m_1^3 \frac{1}{N_0} \left(1 - \frac{1}{N_0}\right)$$

$$Q_{123,01} = m_1^2 (1 - m_1) \frac{1}{N_0}$$

$$Q_{123,10} = 0$$

$$Q_{123,11} = 0$$

$$Q_{123,12} = 0$$

$$Q_{112,00} = m_1^3 \frac{1}{N_0} \left(1 - \frac{1}{N_0}\right)$$

$$Q_{112,01} = m_1^2 (1 - m_1) \frac{1}{N_0}$$

$$Q_{112,10} = (1 - m_1)^2 m_1 \frac{1}{N_1}$$

$$Q_{112,11} = 0$$

$$Q_{112,12} = (1 - m_1)^3 \frac{1}{N_1}$$

$$Q_{121,00} = m_1^3 \frac{1}{N_0} \left(1 - \frac{1}{N_0}\right)$$

$$Q_{121,01} = (1 - m_1) m_1^2 \frac{1}{N_0}$$

$$Q_{121,10} = 0$$

$$Q_{121,11} = 0$$

$$Q_{121,12} = 0$$

$$Q_{211,00} = m_1^3 \frac{1}{N_0} \left(1 - \frac{1}{N_0}\right)$$

$$Q_{211,01} = m_1^2 (1 - m_1) \frac{1}{N_0}$$



$$Q_{211,10} = 0$$

$$Q_{211,11} = 0$$

$$Q_{211,12} = 0$$

### A.1.2 Matrix $P$

$$P_{00,00} = (1 - m_0)^2(1 - \frac{1}{N_0})$$

$$P_{00,01} = (1 - m_0)m_0$$

$$P_{00,10} = (1 - m_0)m_0$$

$$P_{00,11} = m_0 \frac{m_0}{n-1} (1 - \frac{1}{N_1})$$

$$P_{00,12} = m_0 m_0 (1 - \frac{1}{n-1})$$

$$P_{01,00} = (1 - m_0)m_1(1 - \frac{1}{N_0})$$

$$P_{01,01} = (1 - m_0)(1 - m_1)$$

$$P_{01,10} = m_0 m_1$$

$$P_{01,11} = (1 - m_1) \frac{m_0}{n-1} (1 - \frac{1}{N_1})$$

$$P_{01,12} = (1 - m_1)m_0(1 - \frac{1}{n-1})$$

$$P_{10,00} = (1 - m_0)m_1(1 - \frac{1}{N_0})$$

$$P_{10,01} = m_1 m_0$$

$$P_{10,10} = (1 - m_1)(1 - m_0)$$

$$P_{10,11} = (1 - m_1) \frac{m_0}{n-1} (1 - \frac{1}{N_1})$$

$$P_{10,12} = (1 - m_1)m_0(1 - \frac{1}{n-1})$$

$$P_{11,00} = m_1^2(1 - \frac{1}{N_0})$$

$$P_{11,01} = m_1(1 - m_1)$$

$$P_{11,10} = (1 - m_1)m_1$$

$$P_{11,11} = (1 - m_1)^2(1 - \frac{1}{N_1})$$

$$P_{11,12} = 0$$

$$P_{12,00} = m_1^2(1 - \frac{1}{N_0})$$

$$P_{12,01} = (1 - m_1)m_1$$

$$P_{12,10} = (1 - m_1)m_1$$

$$P_{12,11} = 0$$

$$P_{12,12} = (1 - m_1)^2$$

### A.1.2 Matrix $R$

$$R_{00,0} = (1 - m_0)\frac{1}{N_0}$$

$$R_{00,1} = m_0 \frac{m_0}{n-1} \frac{1}{N_1}$$

$$R_{01,0} = (1 - m_0)m_1 \frac{1}{N_0}$$

$$R_{01,1} = \frac{m_0}{n-1}(1 - m_1) \frac{1}{N_1}$$

$$R_{10,0} = m_1(1 - m_0) \frac{1}{N_0}$$

$$R_{10,1} = (1 - m_1) \frac{m_0}{n-1} \frac{1}{N_1}$$

$$R_{11,0} = m_1^2 \frac{1}{N_0}$$

$$R_{11,1} = (1 - m_1)^2 \frac{1}{N_1}$$

$$R_{12,0} = m_1^2 \frac{1}{N_0}$$

$$R_{12,1} = 0$$

## Appendix A.2: approximations under the structured coalescent assumptions

$$\begin{aligned} \gamma_0 = & 1/3[16P^5 + 6M^5Pn - 71M^2nP + 29M^2n^2P - 39M^2n^2P^2 + 11M^2nP^2 - 59M^2n^2P^3 + \\ & 22Mn^2P + M^4P^2 - 13M^3P^2 + 3M^5n - 2M^5P + 10M^2P^2 + 11M^2n^2P^4 + 2M^2n^2P^5 - \\ & 52Mn^2P^3 + 4Mn^2P^4 + 22MnP^4 + 4Mn^3P^2 + 72MP^2 + 56M^3nP^3 - 28MP^3 - 12P^3n^2 - \\ & 34M^3n^2P^2 + 116MnP^3 - 11M^3Pn + 6M^3n^2 - 10M^4P^3 + 11M^3P - 36nP^2 - 15M^3n - \\ & 20MP - 24P^4M^2 - 40M^2P^3 + M^5n^3 - 3M^5n^2 - 44MP^4 + 16MP^5 + 2M^3n^3 - 4M^3P^4 + \\ & 113M^2nP^3 + 20M^4nP^3 - 16P^2 - 4n^2P^4 + 12nP - 8M^2nP^5 + 36nP^4 + 3M^4n - 20MP^5n + \\ & 8Mn^2P^2 + 6M^2n^3P + 10M^2n^3P^2 + 12n^2P^2 + 39M^3nP^2 + 13M^3n^3P + 6MP^5n^2 + 8M^3n^3P^2 - \\ & 32M^3n^2P^3 + 4M^3n^2P^4 - 6M^5n^2P + 2M^5n^3P - M^5 - 6M^2 + 8M^2n^2 + 9M^4n^3P + 2M^4n^3P^2 - \\ & 23M^4n^2P - 3M^4n^2P^2 - 10M^4n^2P^3 + 19M^4nP + 7M^3 - M^2n - 13M^3n^2P - 24M^3P^3 - \\ & 5M^4P + 6Mn - 8MnP - 104MnP^2 + 56P^3 - 56P^4 + 36M^2P + 8M^2P^5 + 4M^2nP^4 + 4n^2P^5 - \\ & 16nP^5 + 3M^4n^3 - 6M^4n^2 + 4nP^3][Mn - M + 1 - 2P + nP]^{-1}[4M + 6M^2nP + 6M^2n^2P + \\ & 6M^2n^2P^2 - 20M^2nP^2 + 8P + 16M^2P^2 - 20MP^2 - 8M^3nP^3 + 32MP^3 + 2M^3n^2P^2 - 20MnP^3 - \\ & 11M^3Pn + 2M^3n^2 + 4M^3P + 8nP^2 - M^3n - 12MP + 16M^2P^3 - 16M^2nP^3 - 24P^2 - 2M^4n + \\ & 4Mn^2P^2 - 2M^3nP^2 + M^4 - 4M^2 + 2M^4n^2P - 4M^4nP - M^3 + 6M^2n + 7M^3n^2P + 8M^3P^3 + \\ & 2M^4P + 16MnP - 4MnP^2 + 16P^3 - 16M^2P + M^4n^2 - 8nP^3]^{-1} ; \end{aligned}$$

$$\begin{aligned} \gamma_1 = & 1/(3(n-1))[-16P^5 - 8M - 8M^5Pn + 93M^2nP - 92M^2n^2P - 2M^2n^2P^2 + 47M^2nP^2 + \\ & 96M^2n^2P^3 + 36Mn^2P - 16P - M^4P^2 + 29M^3P^2 + 8M^3n^4P^2 + 2M^5n^4P - 4M^5n + 2M^5P - \\ & 26M^2P^2 - 26M^3n^3P^3 + 4M^3n^3P^4 + 9M^3n^4P + 29M^2n^2P^4 - 10M^2n^2P^5 + 62Mn^2P^3 + \\ & 14Mn^2P^4 - 38MnP^4 + 18Mn^3P^2 - 100MP^2 - 50M^3nP^3 - 4M^3nP^4 + 36MP^3 - 33M^2n^3P^3 - \\ & 28P^3n^2 + 101M^3n^2P^2 - M^2n^3P^4 + 2M^2n^3P^5 - 58MnP^3 + 42M^3Pn - 20Mn^3P^3 + 2Mn^3P^4 - \\ & 35M^3n^2 + 10M^4P^3 - 17M^3P - 68nP^2 + M^5n^4 + 34M^3n + 60MP + 32P^4M^2 + 24M^2P^3 - \\ & 4M^5n^3 + 6M^5n^2 + 12MP^4 + 12M^3n^3 + 4M^3P^4 - 87M^2nP^3 - 10M^4n^3P^3 + M^4nP^2 - \\ & 30M^4nP^3 + 72P^2 + 20n^2P^4 + 12nP + 16M^2nP^5 - 76nP^4 + 3M^4n^4 - 3M^4n + 4MP^5n - \\ & 110Mn^2P^2 + 33M^2n^3P - 25M^2n^3P^2 + 12n^2P^2 - 90M^3nP^2 - 10M^3n^3P - 2MP^5n^2 - \\ & 48M^3n^3P^2 + 64M^3n^2P^3 - 4M^3n^2P^4 + 12M^5n^2P - 8M^5n^3P + M^5 + 18M^2 + 15M^2n^2 - \\ & 32M^4n^3P - 5M^4n^3P^2 + 42M^4n^2P + 3M^4n^2P^2 + 30M^4n^2P^3 - 24M^4nP - 11M^3 + 6M^2n^4P^2 - \\ & 33M^2n - 24M^3n^2P + 12M^3P^3 + 5M^4P + 6Mn - 86MnP + 172MnP^2 - 112P^3 + 72P^4 + \\ & 9M^4n^4P + 2M^4n^4P^2 - 34M^2P - 8M^2P^5 - 60M^2nP^4 - 4n^2P^5 + 16nP^5 - 9M^4n^3 + 9M^4n^2 + \\ & 116nP^3][4M + 6M^2nP + 6M^2n^2P + 6M^2n^2P^2 - 20M^2nP^2 + 8P + 16M^2P^2 - 20MP^2 - \\ & 8M^3nP^3 + 32MP^3 + 2M^3n^2P^2 - 20MnP^3 - 11M^3Pn + 2M^3n^2 + 4M^3P + 8nP^2 - M^3n - \\ & 12MP + 16M^2P^3 - 16M^2nP^3 - 24P^2 - 2M^4n + 4Mn^2P^2 - 2M^3nP^2 + M^4 - 4M^2 + 2M^4n^2P - \\ & 4M^4nP - M^3 + 6M^2n + 7M^3n^2P + 8M^3P^3 + 2M^4P + 16MnP - 4MnP^2 + 16P^3 - 16M^2P + \\ & M^4n^2 - 8nP^3]^{-1}[Mn - M + 1 - 2P + nP]^{-1} ; \end{aligned}$$

$$\delta_0 = \frac{-P + Mn + n - M}{Mn - M + 1 - 2P + nP} ; \delta_1 = \frac{1 - P + Mn - M}{Mn - M + 1 - 2P + nP} ;$$

$$\begin{aligned} (1 - u_0)\gamma_1 + u_0\gamma_0 = & (-8M + 4Mn^3P^2 + 41M^2nP^3 - 12MP^4 + 8MP^5 + M^4n^3 + 12n^2P^2 - \\ & 50MnP + 39M^2nP^2 - 16P + 56P^2 + 12nP - 52nP^2 + 44nP^3 + 4nP^4 + 24Mn^2P + 6Mn + \\ & 5M^3P^2 + 2MnP^2 + 6Mn^2P^5 + 10M^2 - 21M^3n^2P - 56P^3 - 10M^3n^2P^3 - 23M^2n^2P^2 - \end{aligned}$$

$$\begin{aligned}
& 25M^2n^2P^3 + 8M^2n^2P^4 - 10M^3nP^2 + 20M^3nP^3 + 8M^2n^3P + 4M^2n^3P^2 - M^4 - 5M^3P + \\
& 3M^4n - 10M^3P^3 - 26Mn^2P^3 + 2M^4n^3P - 14Mn^2P^2 - 2Mn^2P^4 - 5M^2n^2P - 12M^2nP^4 - \\
& 13M^2nP - 12P^3n^2 + 8M^3n^3P - 16MP^3 - 2M^4P + 10M^2P - 7M^3n^2 + 62MP^3n + 4MP^4n - \\
& 12MP^5n - 6M^4Pn^2 + 6M^4Pn + 9M^2n^2 + 5M^3n - 19M^2n + 36MP - 12MP^2 + 5M^3n^2P^2 - \\
& 20M^2P^2 - 16M^2P^3 - M^3 + 16P^4 + 4n^2P^5 - 8nP^5 + 3M^3n^3 - 3M^4n^2 + 4P^4M^2 + 18M^3nP - \\
& 4n^2P^4) \left[ 3(n-1)(4M - 16M^2nP^3 + 16MnP - 20M^2nP^2 + 8P - 24P^2 + 8nP^2 - 8nP^3 - \right. \\
& 4MnP^2 - 4M^2 + 7M^3n^2P + 16P^3 + 6M^2n^2P^2 - 2M^3nP^2 - 8M^3nP^3 + M^4 + 4M^3P - \\
& 2M^4n + 8M^3P^3 + 4Mn^2P^2 + 6M^2n^2P + 6M^2nP + 32MP^3 + 2M^4P - 16M^2P + 2M^3n^2 - \\
& 20MP^3n + 2M^4Pn^2 - 4M^4Pn - M^3n + 6M^2n - 12MP - 20MP^2 + 2M^3n^2P^2 + 16M^2P^2 + \\
& \left. 16M^2P^3 - M^3 + M^4n^2 - 11M^3nP) \right]^{-1}
\end{aligned}$$

and

$$(1 - u_0)\delta_1 + u_0\delta_0 = 1.$$

## References

- AXELROD, R., HAMILTON, W.D. 1981. The evolution of cooperation. *Science* 211, 1390–1396.
- CHRISTIANSEN, F.B. 1991. On conditions for evolutionary stability for a continuously varying character. *Am. Nat.* 128, 37–50.
- GOKHALE, C.S., TRAUlsen, A. 2010. Evolutionary games in the multiverse. *Proc Natl Acad Sci USA* 107, 5500–5504.
- HOFBAUER, J., SIGMUND, K. 1998. *Evolutionary game dynamics*. Cambridge University Press.
- HERBOTS, H. M. 1994. *Stochastic models in population genetics: genealogy and genetic differentiation in structured populations*. Ph.D. Thesis, University of London, London.
- HERBOTS, H. M. 1997. The structured coalescent. In: Donnelly, P., Tavaré, S. (Eds), *Progress in Population Genetics and Human Evolution, IMA Volumes in Mathematics and its Applications*, vol. 87.. Springer, New-York, pp. 231–255.
- IMHOF, L. A., NOWAK, M. A., 2006. Evolutionary game dynamics in a Wright-Fisher process, *J. Math. Biol.* 52 (2006), pp. 667–681.
- KARLIN, S., TAYLOR, H.M. 1975. *A First Course in Stochastic Processes*. Academic Press, New York.
- KINGMAN, J. F. C. 1982. The coalescent. *Stochastic Process. Appl.* 13, 235–248.
- KINGMAN, J. F. C. 1982. On the genealogy of large populations. *J. Appl. Probab. Essays in statistical science. Special Vol.* 19A, 27–43.
- KUROKAWA, S., IHARA, Y. 2009. Emergence of cooperation in public goods games. *Proc R Soc B* 276, 1379–1384.
- LADRET, V., LESSARD, S. 2007. Fixation probability for a beneficial allele and a mutant strategy in a linear game under weak selection in a finite island model. *Theor. Pop. Biol.* 72 (3), 409–425.
- LADRET, V., LESSARD, S. 2008. Evolutionary game dynamics in a finite asymmetric two-deme population and emergence of cooperation, *J. Theor. Biol.* 255, 137–151.
- LAMBERT, A., 2006. Probability of fixation under weak selection: a branching process unifying approach. *Theoret. Pop. Biol.* 69 419–441.
- LESSARD, S. 2005. Long-term stability from fixation probabilities in finite populations: New perspectives for ESS theory. *Theoret. Popul. Biol.* 68, 19–27.

- LESSARD, S. 2007. Cooperation is less likely to evolve in a finite population with a highly skewed distribution of family size. *Proc. Biol. Sci.* 274 (1620), 1861–1865.
- LESSARD, S. 2011. On the Robustness of the Extension of the One-Third Law of Evolution to the Multi-Player Game, *Dyn. Games Appl.* 1, 408–418.
- LESSARD, S., LADRET, V. 2007. The probability of fixation of a single mutant in an exchangeable selection model. *J. Math. Biol.* 54, 721–744.
- MAYNARD SMITH, J. 1982. *Evolution and the Theory of Games*. Cambridge University Press.
- MAYNARD SMITH, J., PRICE, G.R. 1973. The logic of animal conflict. *Nature* 246 15–18.
- NAGYLAKI, T. 1980. The strong-migration limit in geographically structured populations. *J. Math. Biol.* 9, 101–114.
- NOTOHARA, M. 1990. The coalescent and the genealogical process in geographically structured population *J. Math. Biol.* 29, 59–75.
- NOWAK, M. A., SASAKI, A., TAYLOR, C., FUDENBERG, D. 2004. Emergence of cooperation and evolutionary stability in finite populations. *Nature* 428, 646–650.
- NOWAK, M.A., 2006a. *Evolutionary Dynamics*. Harvard University Press, MA.
- NOWAK, M.A., 2006b. Five rules for the evolution of cooperation. *Science* 314, 1560–1563.
- NOWAK, M. A., SIGMUND, K. 2004. *Evolutionary Dynamics of Biological Games*. *Science* 303, 793-799.
- OHTSUKI, H., NOWAK, M.A. 2007. Evolutionary games on cycles. *Proc. R. Soc. B* 273, 2249–2256.
- OHTSUKI, H., HAUERT, C., LIEBERMAN, E., NOWAK, M.A. 2006. A simple rule for the evolution of cooperation on graphs and social networks. *Nature* 441, 502–505.
- OHTSUKI, H., PACHECO, J.M., NOWAK, M.A. 2007. Evolutionary graph theory: Breaking the symmetry between interaction and replacement. *J. Theor. Biol.* 246, 681–694.
- ROUSSET, F. 2003. A minimal derivation of convergence stability measures. *J. Theor. Biol.* 221, 665–668.
- ROUSSET, F., BILLIARD, S. 2000. A theoretical basis for measures of kin selection in subdivided populations: Finite populations and localized dispersal. *J. Evol. Biol.* 13, 814–825.
- TAYLOR, P.D., JONKER, L.B. 1978. Evolutionary stable strategies and game dynamics. *Math. Biosciences* 40, 145–156.
- TRAULSEN, A., PACHECO, J.M., IMHOF, L.A. 2006. Stochasticity and evolutionary stability. *Phys. Rev. E* 74, 021905.
- TRIVERS, R. 1971. The evolution of reciprocal altruism. *Q. Rev. Biol.* 46, 35–57.
- WILD, G., TAYLOR, P. D. 2004. Fitness and evolutionary stability in game theoretic models of finite populations. *Proceedings of the Royal Society, Series B* 271, 2345–2349.
- WILKINSON-HERBOTS, H. 1998. Genealogy and subpopulation in differentiation under various models of population structure. *Journal of Mathematical Biology*, 37, pp. 535-585.
- WRIGHT, S. 1931. Evolution in Mendelian populations. *Genetics* 16, 97–159.