Differential migration in a finite-island model and the evolution of cooperation

Dhaker Kroumi [*]
Department of Mathematics, King Fahd University of Petroleum and Minerals, Dhahran 31261, Saudi Arabia

Sabin Lessard
Department of Mathematics and Statistics, University of Montreal, Montreal H3C 3J7, Canada

(Received 5 March 2025; accepted 21 May 2025; published 10 June 2025)

In this paper, we investigate the evolution of cooperation in a haploid population subdivided into d demes, each composed of $N \geqslant 2$ individuals interacting according to a Prisoner's Dilemma. Using a discrete-time Moran process with migration of offspring, we analyze the fixation probability ρ_C of a single cooperative mutant. Our model incorporates a baseline migration probability m, with cooperators exhibiting an additional migration tendency proportional to the local frequency of defectors. Using a two-time-scale approach, we derive a diffusion approximation valid for a large number of demes d to analytically determine ρ_C. We show that, in the absence of differential migration and for a single-round Prisoner's Dilemma, the fixation of cooperation remains generally disfavored by selection in the sense that $\rho_C < 1/(Nd)$ even as the baseline migration probability increases. In contrast, in a repeated Prisoner's Dilemma, selection favors the fixation of cooperation when the number of repetitions n exceeds a critical threshold n^*, which is inversely related to m and N. The probability of fixation ρ_C increases as the level of differential migration increases, setting a threshold beyond which the fixation of cooperation is favored even in a single-round game. This threshold is modulated by the size of the deme, the probability of baseline migration and the cost of cooperation, with larger demes and a higher probability of baseline migration requiring a higher differential migration level. Numerical simulations validate the accuracy of our diffusion approximation, showing strong agreement with the simulated fixation probabilities for large d. These findings highlight the pivotal role of repeated interactions, migration strategies, and population structure in promoting the evolution of cooperation.

DOI: 10.1103/qgdk-977g

I. INTRODUCTION

The evolution of cooperation has been a central focus in evolutionary biology and social dynamics, especially considering that cooperation between individuals can produce collective benefits that enhance group survival and success. However, cooperation brings an inherent dilemma: While cooperative acts benefit the community, they come at a cost to the individual [1–5]. This cost-benefit asymmetry makes cooperation vulnerable to exploitation by selfish individuals who can take advantage of the benefits of cooperation without incurring any cost.

In well-mixed populations, where individuals interact randomly with every other individual, defectors will frequently engage with cooperators, making it difficult for cooperative behavior to persist [6–8]. In such scenarios, selfish strategies tend to dominate and become evolutionarily stable, which ultimately leads to the collapse of cooperation over time [1,7]. This dynamic poses a significant barrier to the spread or maintenance of cooperative traits in populations without spatial or social structure under constant environmental conditions. However, in fluctuating environments, such as those with variable rewards, cooperative behavior can be promoted [9–12].

The situation is more complex in structured populations, where individuals interact locally rather than uniformly across the population. In these environments, cooperation can emerge due to the spatial assortment of individuals of the same type and localized interactions [13]. Theoretical models and empirical studies have shown that when cooperators are clustered and interact with each other, their vulnerability to defectors is significantly reduced. Mechanisms such as kin selection [14], group selection [15], and network reciprocity [16] show how structured interactions can promote cooperation by creating favorable conditions where cooperative behaviors are more likely to be reciprocated and passed on to future generations. Helbing and Yu [17] showed that success-driven movers quickly cluster, producing a rapid shift from universal defection to widespread cooperation. Later work enlarged this view: Yang *et al.* [18] found that migration guided by social norms (conformity, inequity aversion) already supports cooperation and does so even better when the two norms act together, while Yang and Zhang [19] showed that purely random moves can also help if the movers keep a few ties to former neighbors.

Still, the evolution of cooperation in structured populations faces specific challenges. High migration rates, for example, can destabilize cooperative clusters by allowing defectors to infiltrate and spread within groups of cooperators. Understanding the delicate balance between local interactions, migration, and selection is therefore essential, since these factors shape the success and stability of cooperative behaviors in structured environments [20,21].

[*]Contact author: dhaker.kroumi@kfupm.edu.sa

 ©2025 American Physical Society

A foundational model for structured populations is the island model, which divides a population into isolated demes that exchange migrants. However, analyzing these interactions, especially in finite populations, can be mathematically complex. Diffusion approximations offer a continuous-time framework to study allele frequency changes in structured populations. The two-time-scale approach, formalized by Ethier and Nagylaki [22], refines these models by distinguishing between faster and slower evolutionary processes. For instance, while group type frequencies may change rapidly, individual type changes may progress more gradually across demes. This separation of time scales facilitates the study of evolutionary dynamics, enabling the calculation of key quantities such as fixation probabilities, fixation times, and the stationary distribution of allele frequencies.

This approach has been applied to several population genetics models. For instance, Nagylaki [23] studied the effects of selection and strong migration in a Wright-Fisher model under ergodic migration. He later extended this approach to infinite populations subdivided into a lattice of colonies, assuming weak selection and migration [24,25]. Cherry [26,27] and Cherry and Wakeley [28] demonstrated the applicability of this approach to island models using numerical simulations, considering cases where the number of demes and the size of each deme are both large enough. Similarly, Whitlock [29] examined stepping-stone models, where populations are subdivided into demes of fixed size, each connected to neighboring demes through migration.

Wakeley [30] provided an analytical verification of the diffusion conditions proposed by Ethier and Nagylaki [22] for haploid populations structured into d demes, each containing $N \geqslant 2$ individuals. In this model, each deme follows a Wright-Fisher scheme, with migration occurring via a migrant pool to which all demes contribute equally. Additionally, a fraction $m > 0$ of individuals in each deme is replaced by offspring from this pool. Wakeley [30] examined the limiting behavior of the system as $d \to \infty$. In a follow-up study, Wakeley and Takahashi [31] further investigated the applicability of the diffusion approximation under a variety of conditions, assuming reproduction governed by a structured Moran process. In their model, at each time step, a random deme is chosen, and an individual is selected proportionally to its death rate to die. On the other hand, with probability $1 - m$, an individual from the same deme is selected to reproduce; otherwise, with probability m, the individual to reproduce is selected from another deme. The offspring produced replaces the removed individual. Their findings confirmed that, as $d \to \infty$, the dynamics of the type frequencies converges to a diffusion process.

Lessard [32] extended the diffusion approximation approach to kin selection models in diploid populations structured into $d \to \infty$ groups, each containing $2N$ individuals. In this framework, he explored a range of scenarios involving selection, mutation, and migration, while considering different assumptions about the timing of dispersal and recolonization in the event of group extinction. This method has since been applied to population genetics involving random viability rates [33] as well as evolutionary game theory with random payoffs [34].

Previous models typically assumed migration probabilities that are constant and independent of individual type, neglecting the ecological drivers of migration. In reality, cooperative behaviors are frequently threatened by defectors who exploit shared resources, leading to environmental degradation and higher extinction risks [35–38]. As a result, cooperators often migrate to more favorable environments, seeking reciprocal interactions to enhance survival [39–41]. This adaptive migration, influenced by expectations and risk containment, fosters supportive clusters and stable cooperative communities through spatial dynamics [42–47]. Therefore, the effects of these differential migration patterns, shaped by selective pressures, should be further explored in the context of the evolution of cooperation.

In this study, we explore how differential migration probabilities may affect the fixation of cooperative behavior within a population subdivided into a large number of demes, each with a size $N \geqslant 2$. The updating rule from one time step to the next follows a Moran model for a deme-structured population, where every new offspring has a baseline migration probability. However, an offspring of a cooperator experiences an additional migration probability that depends on the composition of the deme where the offspring was produced: in a deme with a high proportion of defectors, an offspring of a cooperator is more likely to migrate, whereas in a deme with a greater number of cooperators, this additional migration tendency is reduced. Using a two-time-scale diffusion approximation, we show that this migratory incentive enhances the fixation of cooperation across the entire population after being introduced as a single mutant. When it is sufficiently strong, this effect can lead to selection favoring the fixation of cooperation.

The structure of this paper is organized as follows. Section II presents a formulation of the model, including the key assumptions and dynamics. In Sec. III, we introduce a one-dimensional diffusion approximation for the stochastic population dynamics. Section IV is dedicated to deriving the fixation probability of cooperation introduced as a single mutant, where we explore the impact that various factors such as the deme size, the baseline migration probability, and the additional migration incentive for an offspring of a cooperator to migrate may have on the fixation of cooperation under different scenarios. In Sec. V, we evaluate the accuracy of the diffusion approximation by comparing its predictions with those obtained using stochastic simulations of the model. Finally, in Sec. VI, we conclude the paper with a discussion of our findings and their broader implications.

II. MODEL

Consider a haploid population subdivided into d demes, each containing $N \geqslant 2$ individuals. Each individual can be of type C for cooperation or type D for defection. From one time step to the next, random pairwise interactions occur in each deme according to the Prisoner's dilemma with the payoff matrix

$$\begin{array}{c} \\ C \\ D \end{array} \begin{array}{c} \begin{array}{cc} C & D \end{array} \\ \begin{pmatrix} R & S \\ T & P \end{pmatrix} \end{array}, \tag{1}$$

where $T > R > P > S$. As a result of these interactions, each individual accumulates an average payoff π that is transformed into a fitness $f = 1 + s\pi$, where $s = 1/(Nd)$ represents the selection intensity. Then a deme is randomly chosen and, within this deme, an individual is selected for reproduction with a probability proportional to its fitness. The offspring produced, denoted as I, replaces an individual chosen at random within the same deme with probability $1 - m_I$. With probability $m_I > 0$, the offspring migrates to one of the $d - 1$ other demes chosen at random and replaces a randomly selected individual in that deme.

To capture how cooperative behaviors are reinforced within a deme, consider a linear relationship where the tendency for an offspring of a cooperator to stay in its original deme decreases as the proportion of cooperators in the deme decreases. More precisely, if the deme contains i cooperators and $N - i$ defectors, then the probability for an offspring of a cooperator to migrate to another deme is

$$m_{C,i} = m + sM_{C,i} = m + s\alpha(1 - x_i), \tag{2}$$

where $x_i = i/N$. Here, $m > 0$ represents a baseline migration probability common to all offspring and $\alpha \geqslant 0$ is a constant coefficient. For an offspring of a defector, the probability to migrate is given only by $m_{D,i} = m$. This model aligns with the idea that cooperative interactions are sensitive to deme composition, with defectors creating a pressure that reduces the incentive for cooperators to remain in a defector-dominated deme [1]. In fully cooperative groups, cooperators experience the strongest incentive to stay since $M_{C,N} = 0$. For defection, we suppose that there is no additional migration term.

We identify the type of each deme by the number of individuals of type C within it. More precisely, a deme is of type i if it contains i individuals of type C and $N - i$ individuals of type D, for $i = 0, 1, \ldots, N$. Let $Z_i(t)$ be the frequency of demes of type i at time step $t \in \mathbb{N}$, for $i = 0, 1, \ldots, N$. Then we have a discrete-time Markov chain for the frequency vector of deme types over time represented by $\{\mathbf{Z}(t) = (Z_0(t), Z_1(t), \ldots, Z_N(t))^T\}_{t\in\mathbb{N}}$. Finally, we define the discrete-time process $\{X(t)\}_{t\in\mathbb{N}}$, where

$$X(t) = \frac{1}{N}\sum_{i=0}^{N} iZ_i(t) \tag{3}$$

is the overall frequency of type C in the population at time step $t \in \mathbb{N}$.

III. DIFFUSION APPROXIMATION

Let us now define the second time-scale process as the deviation of the process $\{\mathbf{Z}(t)\}_{t\in\mathbb{N}}$ with respect to the distribution $\mathbf{v}(X(t)) = (v_0(X(t)), v_1(X(t)), \ldots, v_N(X(t)))^T$ for $t \in \mathbb{N}$, where

$$v_i(x) = \binom{N}{i} \frac{B\left(\frac{Nmx}{1-m} + i, \frac{Nm(1-x)}{1-m} + N - i\right)}{B\left(\frac{Nmx}{1-m}, \frac{Nm(1-x)}{1-m}\right)}, \tag{4}$$

for $i = 0, 1, \ldots, N$. Here, $B(a, b)$ denotes the Beta function defined as

$$B(a, b) = \int_0^1 t^{a-1}(1 - t)^{b-1}dt. \tag{5}$$

This distribution represents the stationary distribution of the deme types when the population is subdivided into an infinite number of demes under neutrality ($s = 0$), assuming the overall frequency of cooperators remains constant at $x \in [0, 1]$. For further details, see Appendix A 3. More precisely, the second process is $\{\mathbf{Y}(t) = (Y_0(t), Y_1(t), \ldots, Y_N(t))^T\}_{t\in\mathbb{N}}$, where

$$Y_i(t) = Z_i(t) - v_i(X(t)), \tag{6}$$

for $i = 0, 1, \ldots, N$ and $t \in \mathbb{N}$.

Let $\mathbf{Z}(0) = \mathbf{z} = (z_0, z_1, \ldots, z_N)^T$, in which case $\mathbf{Y}(0) = \mathbf{y} = \mathbf{z} - \mathbf{v}(x)$ and $X(0) = x = (1/N)\sum_{i=0}^{N} iz_i$. It can be shown that the changes in X and $\mathbf{Y}$ from time 0 to 1, denoted by $\Delta X(0) = X(1) - X(0)$ and $\Delta\mathbf{Y}(0) = \mathbf{Y}(1) - \mathbf{Y}(0)$, respectively, satisfy the following conditions:

$$\mathbb{E}_{\mathbf{z}}[\Delta X(0)] = \frac{\mu(x, \mathbf{y})}{N^2d^2} + o(1/d^2), \tag{7a}$$

$$\mathbb{E}_{\mathbf{z}}[(\Delta X(0))^2] = \frac{\sigma^2(x, \mathbf{y})}{N^2d^2} + o(1/d^2), \tag{7b}$$

$$\mathbb{E}_{\mathbf{z}}[(\Delta X(0))^4] = o(1/d^2), \tag{7c}$$

$$\mathbb{E}_{\mathbf{z}}[\Delta Y_i(0)] = \frac{c_i(x, \mathbf{y})}{Nd} + o(1/d), \tag{7d}$$

$$\mathrm{Var}_{\mathbf{z}}[\Delta Y_i(0)] = o(1/d), \tag{7e}$$

as $d \to \infty$ uniformly in $\mathbf{z}$, where

$$\mu(x, \mathbf{y}) = \alpha\sum_{i=0}^{N}(v_i(x) + y_i)x_i(1 - x_i)(x_i - x) + \sum_{i=0}^{N}(v_i(x) + y_i)x_i(1 - x_i)(\pi_{C,i} - \pi_{D,i}), \tag{8a}$$

$$\sigma^2(x, \mathbf{y}) = 2(1 - m)\sum_{i=0}^{N}(v_i(x) + y_i)x_i(1 - x_i) + 2mx(1 - x), \tag{8b}$$

$$c_i(x, \mathbf{y}) = \sum_{j=0}^{N} y_j P^*_{j,i}(x) - y_i, \tag{8c}$$

with

$$P^*_{i-1,i}(x) = [(1 - m)x_{i-1} + mx](1 - x_{i-1}), \tag{9a}$$

$$P^*_{i+1,i}(x) = [(1 - m)(1 - x_{i+1}) + m(1 - x)]x_{i+1}, \tag{9b}$$

$$P^*_{i,i}(x) = 1 - [(1 - m)x_i + mx](1 - x_i) - [(1 - m)(1 - x_i) + m(1 - x)]x_i, \tag{9c}$$

and $P^*_{j,i}(x) = 0$ for $|j - i| \geqslant 2$, for $i = 0, 1, \ldots, N$. Here, $\mathbb{E}_{\mathbf{z}}$ and $\mathrm{Var}_{\mathbf{z}}$ denote the conditional expectation and conditional variance, respectively, given that $\mathbf{Z}(0) = \mathbf{z}$. Moreover, $\mathbf{0}$ is a globally asymptotically stable solution of the differential equation

$$\begin{cases} \dfrac{d}{d\tau}\mathbf{Y}(\tau, x, \mathbf{y}) = \mathbf{c}(x, \mathbf{Y}(\tau, x, \mathbf{y})), \\ \mathbf{Y}(0, x, \mathbf{y}) = \mathbf{y}, \end{cases} \tag{10}$$

where

$$\mathbf{c}(x, \mathbf{y}) = (c_0(x, \mathbf{y}), c_1(x, \mathbf{y}), \ldots, c_N(x, \mathbf{y}))^T. \tag{11}$$

The proofs of Eqs. (7a)–(7e) and the stability of the zero solution in Eq. (10) are given in Appendix A. Moreover, the additional condition $\mu(0, \mathbf{y}) = 0$ ensures the existence of a strongly continuous semigroup associated with a diffusion process (see Ethier [48]). Based on the expressions in Eqs. (8a)–(8c), it is evident that $\mu(x, \mathbf{y})$, $\sigma^2(x, \mathbf{y})$, and $\mathbf{c}(x, \mathbf{y})$ are smooth functions.

Noting that $1/d \to 0$ and $(1/d^2)/(1/d) \to 0$ as $d \to \infty$, Theorem 3.3 in Ethier and Nagylaki [22] leads to the following convergence result.

Theorem 1. Let $X(\lfloor \frac{N^2 d^2}{2(1-F)} \tau \rfloor)$ be the frequency of type C at time $\tau \geqslant 0$ measured in number of $\frac{N^2 d^2}{2(1-F)}$ time steps in a haploid population subdivided into d demes containing each N individuals. Here, $\lfloor x \rfloor$ denotes the integer part of x and $F = \frac{1-m}{1-m+mN}$. As $d \to \infty$, the process $\{X(\lfloor \frac{N^2 d^2}{2(1-F)} \tau \rfloor)\}_{\tau \geqslant 0}$ converges in distribution to a diffusion process $\{X^*(\tau)\}_{\tau \geqslant 0}$ on $[0, 1]$ with generator

$$\mathcal{L} = \mu(x) \frac{d}{dx} + \frac{\sigma^2(x)}{2} \frac{d^2}{dx^2},$$

where

$$\begin{aligned} \mu(x) &= \frac{\mu(x, \mathbf{0})}{2(1-F)} \\ &= \frac{(N-1)x(1-x)}{2N} \Big[\delta(x)(R-T) + (1-\delta(x))(S-P) \\ &\quad + \frac{\alpha}{2 + m(N-2)} (1-2x) \Big], \end{aligned} \tag{12a}$$

$$\sigma^2(x) = \frac{\sigma^2(x, \mathbf{0})}{2(1-F)} = x(1-x), \tag{12b}$$

with

$$\delta(x) = \frac{1 + m(N-2)x}{2 + m(N-2)}. \tag{13}$$

This result holds for any fixed deme size $N \geqslant 2$ and any baseline migration probability $m \in (0, 1]$.

The detailed derivation of $\mu(x, \mathbf{0})$ and $\sigma^2(x, \mathbf{0})$ can be found in Appendix B. In Appendix C, we show that F is the probability for two randomly selected individuals within the same deme, in a population structured into an infinite number of demes under neutrality ($s = 0$), to be identical by descent (i.e., to share a common ancestor). This corresponds to the probability for two ancestral lines in the same deme to coalesce before one of them migrates. This probability is commonly known in the literature as Wright's [49] fixation index (see, e.g., Wakeley [50], Rousset [51]).

For a detailed interpretation of the drift and diffusion functions in Eqs. (12a)–(12b), we refer the reader to the Discussion section.

IV. FIXATION PROBABILITIES

Consider the diffusion process $\{X^*(\tau)\}_{\tau \geqslant 0}$ as introduced in Theorem 1. It is obvious that states 0 and 1 are absorbing. Assuming that the process starts at a point $X^*(0) = x \in (0, 1)$, let $\rho(x)$ be the probability to reach state 1, representing fixation at state 1. Let h be sufficiently small such that the probability of reaching 0 or 1 in the time interval $[0, h]$ is $o(h)$. The existence of such an h is ensured by the condition

$$\lim_{h \to 0^+} \frac{\mathbb{P}(|X^*(h) - x| > \varepsilon \mid X^*(0) = x)}{h} = 0, \tag{14}$$

where $\varepsilon = \min\{x, 1-x\} > 0$. Applying the law of total probability, we obtain

$$\rho(x) = \mathbb{E}_x[\rho(X^*(h))] + o(h). \tag{15}$$

Here, $\mathbb{E}_x$ denotes the conditional expectation given that $X^*(0) = x$. Expanding $\rho(X^*(h))$ around x results in

$$\begin{aligned} \rho(X^*(h)) &= \rho(x) + \rho'(x)\Delta X^*(h) + \frac{\rho''(x)}{2!}(\Delta X^*(h))^2 \\ &\quad + \frac{\rho'''(\xi(h))}{3!}(\Delta X^*(h))^3, \end{aligned} \tag{16}$$

where $\Delta X^*(h) = X^*(h) - x$, and $\xi(h)$ is a point between x and $X^*(h)$ that depends on $X^*(h)$. Taking the expectation of this expansion and substituting it in Eq. (15), we have

$$\begin{aligned} \rho(x) &= \rho(x) + \rho'(x)\mathbb{E}_x[\Delta X^*(h)] + \frac{\rho''(x)}{2!}\mathbb{E}_x[(\Delta X^*(h))^2] \\ &\quad + o(h) \\ &= \rho(x) + \mu(x)\rho'(x)h + \frac{\sigma^2(x)}{2!}\rho''(x)h + o(h). \end{aligned} \tag{17}$$

Here, we have used the Cauchy-Schwarz inequality which guarantees that

$$\mathbb{E}_x[|\Delta X^*(h)|^3] \leqslant \sqrt{\mathbb{E}_x[(\Delta X^*(h))^2] \cdot \mathbb{E}_x[(\Delta X^*(h))^4]} = o(h).$$

Dividing in Eq. (17) by h and taking the limit as $h \to 0^+$, we find that the probability of reaching 1 before 0 satisfies the differential equation

$$0 = \mu(x)\rho'(x) + \frac{\sigma^2(x)}{2}\rho''(x) \tag{18}$$

with the boundary conditions $\rho(0) = 0$ and $\rho(1) = 1$. The solution of this equation is given by

$$\rho(x) = \frac{\int_0^x \Psi(y)\, dy}{\int_0^1 \Psi(y)\, dy}, \tag{19}$$

where

$$\Psi(y) = \exp\left(-2 \int_0^y \frac{\mu(x)}{\sigma^2(x)}\, dx \right). \tag{20}$$

For more details, see Sec. 14.3 in Karlin and Taylor [52].

Now, let ρ_C be the fixation probability of C following its introduction as a single mutant into a population consisting of d demes of N defectors. For a sufficiently large number of demes d, the discrete-time frequency process $\{X(t)\}_{t \in \mathbb{N}}$ for type C can be approximated by the diffusion process $\{X^*(\tau)\}_{\tau \geqslant 0}$ (see Theorem 1), yielding

$$\rho_C \approx \rho\left(\frac{1}{Nd}\right) = \frac{\int_0^{(Nd)^{-1}} \Psi(y) dy}{\int_0^1 \Psi(y) dy} \approx \frac{1}{Nd \int_0^1 \Psi(y) dy}, \tag{21}$$

where we have used the approximation

$$\int_0^{(Nd)^{-1}} \Psi(y) dy \approx \frac{\Psi(0)}{Nd} = \frac{1}{Nd}. \tag{22}$$

Similarly, let ρ_D denote the fixation probability of a single D mutant in a population of d demes of N cooperators. One finds

$$\rho_D \approx 1-\rho\left(1-\frac{1}{Nd}\right)=1-\frac{\int_0^{1-(Nd)^{-1}}\Psi(y)dy}{\int_0^1\Psi(y)dy}$$
$$\approx \frac{\int_{1-(Nd)^{-1}}^1\Psi(y)dy}{\int_0^1\Psi(y)dy}\approx\frac{1}{ND\int_0^1\Phi(y)dy}, \tag{23}$$

where $\Phi(y)=\frac{\Psi(y)}{\Psi(1)}=\exp(2\int_y^1(\mu(x)/\sigma^2(x))dx)$. In the absence of selection ($s=0$), we have $\rho_C=\rho_D=1/(Nd)$. In the presence of selection ($s>0$), we say that *selection favors (respectively, disfavors) the fixation of type X* if $\rho_X>1/(Nd)$ [respectively, $\rho_X<1/(Nd)$], for $X=C,D$ [53]. Hence, using Eqs. (21) and (25), selection favors the fixation of C when $\int_0^1\Psi(y)dy<1$, and disfavors the fixation of D when $\int_0^1\Phi(y)dy>1$.

Using Eqs. (12a)–(12b) and integrating, the functions Ψ and Φ can now be expressed as

$$\Psi(y)=\exp\left\{\frac{N-1}{N}\left(-\frac{y(1-y)}{2(2+\beta)}\alpha+\frac{2y+\beta y^2}{2(2+\beta)}(T-R)\right.\right.$$
$$\left.\left.+\frac{2y+\beta(2-y)y}{2(2+\beta)}(P-S)\right)\right\} \tag{24}$$

and

$$\Phi(y)=\exp\left\{-\frac{(N-1)(1-y)}{2N(2+\beta)}((2+\beta(1+y))(T-R)\right.$$
$$\left.+(2+\beta(1-y))(P-S)+\alpha y)\right\}, \tag{25}$$

where $\beta=m(N-2)$. For each fixed $y\in(0,1)$, both $\Psi(y)$ and $\Phi(y)$ are continuous in $\alpha\geqslant 0$ and strictly decrease as α increases, which in turn makes ρ_C and ρ_D strictly increasing functions of α. Moreover, since $\Phi(y)<1$ on $(0,1)$ for all $\alpha\geqslant 0$, we always have $\rho_D>1/(Nd)$. One also finds

$$\frac{\rho_C}{\rho_D}=\frac{1}{\Psi(1)}=\exp\left(\frac{N-1}{2N}(R+S-T-P)\right)<1,$$

a fraction independent of both α and m. Thus, while increasing α uniformly boosts both fixation probabilities, defectors always have the higher fixation probability ($\rho_C<\rho_D$), and m affects ρ_C and ρ_D in exactly the same way. Owing to this symmetry, it suffices to focus on ρ_C without loss of generality.

In the remainder of this paper, we will explore the effect of this additional migration term across various parameter settings within the context of the Prisoner's dilemma.

A. Case 1: No differential migration ($\alpha=0$)

Here, we consider a scenario where an offspring of a cooperator and an offspring of a defector have the same probability to migrate to a different deme chosen at random, given only by the baseline value m. In this case, we have

$$\Psi(y)=\Psi_0(y)=\exp\left\{\frac{N-1}{N}\left(\frac{2y+\beta y^2}{2(2+\beta)}(T-R)\right.\right.$$
$$\left.\left.+\frac{2y+\beta(2-y)y}{2(2+\beta)}(P-S)\right)\right\}>1 \tag{26}$$

for any $y\in(0,1)$, from which $\int_0^1\Psi_0(y)dy>1$. Therefore,

$$\rho_C\approx\frac{1}{Nd\int_0^1\Psi_0(y)dy}<\frac{1}{Nd}, \tag{27}$$

for $N\geqslant 2$ and $0<m\leqslant 1$. This means that selection disfavors the fixation of C.

Furthermore, we have

$$\frac{\partial}{\partial\beta}\left(\frac{2y+\beta y(2-y)}{2(\beta+2)}\right)=\frac{\partial}{\partial\beta}\left(\frac{2y+\beta y^2}{(\beta+2)}\right)$$
$$=-\frac{y(1-y)}{(\beta+2)^2}<0, \tag{28}$$

for any $y\in(0,1)$. This implies that increasing the baseline migration probability m will reduce the weights of $T-R$ and $P-S$ in the expression of Ψ_0, thereby increasing the fixation probability ρ_C without making it possible for selection to favor the fixation of C.

Let us summarize these findings in our first result.

Result 1: With no differential migration, selection disfavors the fixation of C for any deme size $N\geqslant 2$ and any baseline migration probability $0<m\leqslant 1$. Additionally, increasing m will increase the fixation probability ρ_C, but this is not enough to make selection favor the fixation of C.

B. Case 2: Demes of size $N=2$

Here, we analyze a situation where each deme contains exactly $N=2$ individuals. In this case, we have

$$\Psi(y)=\exp\left(-\frac{y(1-y)}{4}\alpha+\frac{y}{4}\gamma\right) \tag{29}$$

for any $y\in(0,1)$, where $\gamma=T+P-S-R>0$. It is important to note that the baseline migration probability m does not affect $\Psi(y)$. Now, we will study the effect of α and the possibility of the existence of a threshold from which selection will favor the fixation of C.

As shown before, an increase in α will result in a decrease in $\Psi(y)$ for any $y\in(0,1)$. Consequently, $\alpha\mapsto\int_0^1\Psi(y)dy$ is a strictly decreasing continuous function on $[0,\infty)$. Additionally, $\lim_{\alpha\to\infty}\Psi(y)=0$ and $\Psi(y)\leqslant\Psi_0(y)$ for any $y\in(0,1)$ and any $\alpha>0$, where $\int_0^1\Psi_0(y)\,dy<\infty$. By the dominated convergence theorem, we have

$$\lim_{\alpha\to\infty}\int_0^1\Psi(y)\,dy=0. \tag{30}$$

Combining this with $\int_0^1\Psi(y)\,dy|_{\alpha=0}=\int_0^1\Psi_0(y)\,dy>1$, we conclude the following: there exists a unique threshold $\alpha^*>0$ such that $\int_0^1\Psi(y)dy>1$ for $\alpha<\alpha^*$, which corresponds to $\rho_C<1/(Nd)$, and $\int_0^1\Psi(y)\,dy<1$ for $\alpha>\alpha^*$, which corresponds to $\rho_C>1/(Nd)$. Furthermore, $\int_0^1\Psi_0(y)\,dy$ increases as γ increases, since this is the case of $\Psi(y)$ for any $y\in(0,1)$, which implies a decrease of ρ_C. Figure 1 illustrates these conclusions.

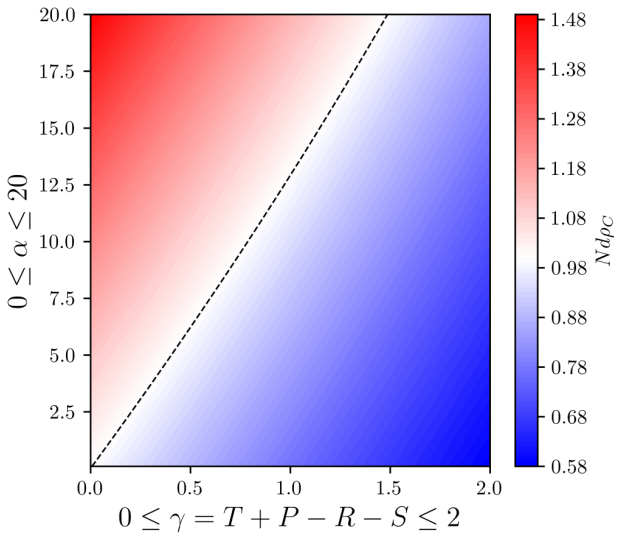


FIG. 1. The graph illustrates the effect of $\gamma = T + P - R - S$ as well as the differential migration coefficient α on the fixation probability ρ_C, for a deme size $N = 2$. An increase in γ makes it more difficult for selection to favor the fixation of C. Conversely, an increase in α, which indicates that cooperators are more prone to migrate than defectors, will increase ρ_C. The dashed curve represents the curve of the threshold α^* as a function of γ, which α must exceed for selection to favor the fixation of C. The red region indicates that selection favors the fixation of C, while the blue region indicates that selection disfavors it.

Finding a closed-form expression for α^* is challenging. However, it can be expressed implicitly. Note that

$$\int_0^1 \Psi(y)dy = \frac{2}{\sqrt{\alpha}} e^{\frac{(\alpha-\gamma)^2}{16\alpha}} \int_{\frac{\gamma-\alpha}{4\sqrt{\alpha}}}^{\frac{\gamma+\alpha}{4\sqrt{\alpha}}} e^{s^2} ds. \tag{31}$$

Then, α^* is the unique solution to the equation $g(t) = 0$ on $(0, \infty)$, where

$$g(t) = \sqrt{t} e^{-\frac{(t-\gamma)^2}{16t}} - 2 \int_{\frac{\gamma-t}{4\sqrt{4t}}}^{\frac{\gamma+t}{4\sqrt{t}}} e^{s^2} ds. \tag{32}$$

C. Case 3: Simplified Prisoner's dilemma

Consider the simplified Prisoner's dilemma where a cooperator pays a cost $c > 0$ for its partner to receive a benefit $b > c$. Then, we have $R = b - c$, $S = -c$, $T = b$ and $P = 0$. As shown in the general case, with no differential migration ($\alpha = 0$), selection will disfavor the fixation of C. However, introducing differential migration so that cooperators are more prone to leave demes dominated by defectors increases the fixation probability ρ_C. In the remainder of this subsection, we will show the existence of a threshold α^* beyond which selection favors the fixation of C and then we will examine how the parameters c, N, and m impact this threshold.

Inserting the new values for R, S, T, P in Eq. (24) and using the notation $\beta = m(N-2)$, we have

$$\Psi(y) = \exp\left(\frac{N-1}{N} cy - uy(1-y)\right), \tag{33}$$

where $u = \frac{(N-1)\alpha}{N(2+\beta)}$. Note that

$$\frac{\partial}{\partial u} \int_0^1 \Psi(y)dy = -\int_0^1 y(1-y) \exp\left(\frac{N-1}{N} cy - uy(1-y)\right) dy < 0. \tag{34}$$

In addition, the dominated convergence theorem guarantees that

$$\lim_{u\to 0^+} \int_0^1 \Psi(y)dy = \int_0^1 \exp\left(\frac{N-1}{N} cy\right) dy > 1, \tag{35a}$$

$$\lim_{u\to \infty} \int_0^1 \Psi(y)dy = 0. \tag{35b}$$

Therefore, there exists a unique solution $u^*(\frac{N-1}{N}c)$ to the equation

$$\int_0^1 \Psi(y)\, dy = \int_0^1 \exp\left(\frac{N-1}{N} cy - uy(1-y)\right) dy = 1, \tag{36}$$

so that $\int_0^1 \Psi(y)dy > 1$ for $u < u^*(\frac{N-1}{N}c)$ and $\int_0^1 \Psi(y)dy < 1$ for $u > u^*(\frac{N-1}{N}c)$. Consequently, the threshold for α, beyond which selection favors the fixation of C, and below which it disfavors the fixation of C, is given by

$$\alpha^* = \frac{N(2+\beta)}{N-1} u^*\left(\frac{N-1}{N} c\right). \tag{37}$$

With a fixed deme-scaled migration rate $\nu = mN$ as $N \to \infty$ and $m \to 0$, this threshold becomes

$$\alpha^* = (2 + \nu)u^*(c). \tag{38}$$

Here, the function $u^*(c)$ increases as c increases. Moreover, the effect of ν on α^* is linear: as ν increases, a higher coefficient α is required for selection to favor the fixation of C. In other words, the optimal scenario for the fixation of C occurs when the structure effect is dominant ($\nu \to 0$), while the worst scenario arises when the structure effect is minimal ($\nu \to \infty$).

Let us summarize.

Result 2: In the simplified Prisoner's dilemma as $N \to \infty$ and $m \to 0$ with the deme-scaled baseline migration rate $\nu = Nm$ kept fixed, selection favors the fixation of C when the differential migration coefficient α is above some threshold $\alpha^* = (2 + \nu)u^*(c)$, and it disfavors it when it is below. An increase in the cost c or in ν will increase the threshold, making it more difficult for selection to favor the fixation of C.

Figure 2 illustrates how the population-scaled fixation probability $Nd\rho_C$ depends on c and α for $N = 5, 10, 20$ and $m = 0.1, 0.2$. A higher α or a lower c increases $Nd\rho_C$, with the most favorable scenario occurring at $N = 5$ and $m = 0.1$, and the least favorable at $N = 20$ and $m = 0.2$. This indicates that increasing N or m raises the threshold α^*. Figure 3 confirms this trend by numerically evaluating the population-scaled fixation probability $Nd\rho_C$ when $N \to \infty$ and $m \to 0$ as a function of $\nu = Nm$ and α for $c = 0.1, 0.2, 0.3, 0.4, 0.5$, and identifying the curve where $Nd\rho_C = 1$. Figure 4 extracts

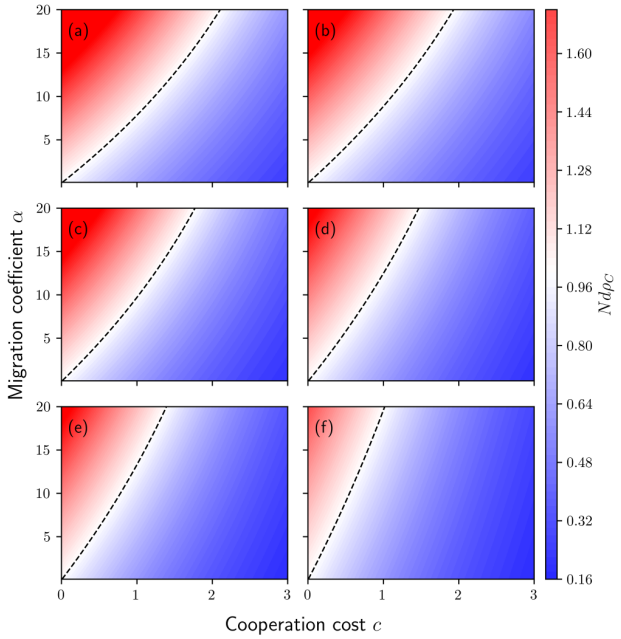


FIG. 2. The graphs illustrate the relationship between $Nd\rho_C$, the cost parameter c, and the differential migration coefficient α for various deme sizes and baseline migration probabilities. For $m = 0.1$, (a), (c), and (e) correspond to $N = 5$, $N = 10$ and $N = 20$, respectively. Similarly, for $m = 0.2$, (b), (d), and (f) correspond to $N = 5$, $N = 10$ and $N = 20$, respectively. A decrease in c or an increase in α leads to a higher population-scaled fixation probability $Nd\rho_C$. The dashed curve represents the neutral selection point where $Nd\rho_C = 1$. Above this curve (red region), selection favors the fixation of C, while below it (blue region), selection disfavors the fixation of C. This curve corresponds to the critical threshold α^* as a function of c. Comparing columns (where m is constant and N varies) or rows (where N is constant and m varies) shows that a decrease in either N or m favors the fixation of C.

the threshold α^* from Fig. 2 as a function of c. It shows that increasing c raises α^*. These results demonstrate that the fixation of C is favored by selection only with a sufficiently large α, while increases in cooperation cost, deme size, or baseline migration probability reduce the probability of C to fix, highlighting the intricate interplay between these parameters.

D. Case 4: General Prisoner's dilemma under weak selection

In the general context of the Prisoner's dilemma, we consider the scenario where the payoffs R, S, T, P as well as the differential migration coefficient α are all of the same small order. Under this assumption and using $e^u \approx 1 + u$ for u small enough, $\Psi(y)$ can be approximated as

$$\Psi(y) \approx 1 + \frac{N-1}{N}\left[-\frac{y(1-y)}{2(2+\beta)}\alpha + \frac{2y+\beta y^2}{2(2+\beta)}(T-R) + \frac{2y+\beta(2-y)y}{2(2+\beta)}(P-S)\right] \tag{39}$$

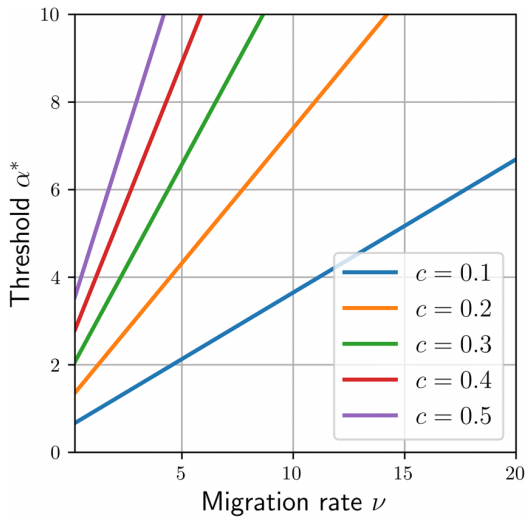


FIG. 3. The graph shows the relationship between the threshold α^* and the deme-scaled baseline migration rate $\nu = mN$ when $N \to \infty$ and $m \to 0$ for various values of the cost $c = 0.1, 0.2, 0.3, 0.4, 0.5$. For every value of c, we have a straight line of slope $u^*(c)$ with respect to ν that takes the value $2u^*(c)$ at $\nu = 0$, consistent with the analytical result in Eq. (38).

for any $y \in (0, 1)$, from which we have

$$\int_0^1 \Psi(y)dy \approx 1 + \frac{N-1}{6(2+\beta)N}[(3+2\beta)(T-R) + (3+\beta)(P-S) - \alpha] \tag{40}$$

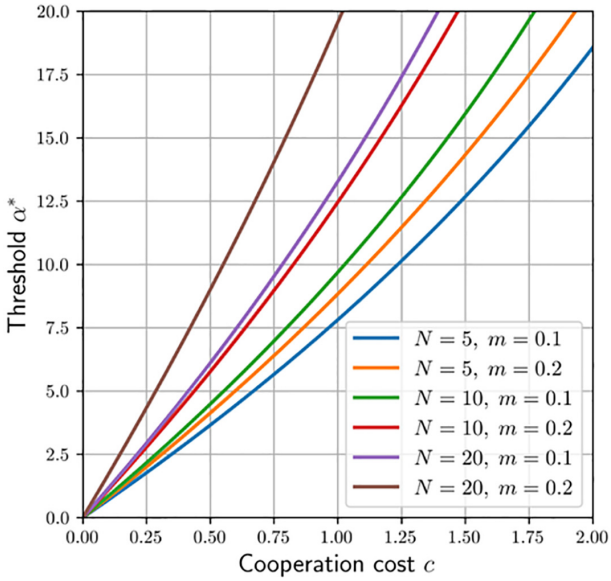


FIG. 4. The graphs depict the relationship between the threshold α^* and the cost c across various deme sizes $N = 5, 10, 20$ and baseline migration probabilities $m = 0.1, 0.2$. It is observed that the threshold α^* exhibits an increasing trend with respect to c, N, and m.

and

$$\rho_C \approx \frac{1}{Nd} + \frac{N-1}{6N^2d(2+\beta)}[\alpha + (3+2\beta)(R-T) + (3+\beta)(S-P)], \tag{41}$$

where $\beta = m(N-2)$. It follows that selection favors the fixation of C as long as

$$\alpha + (3+2\beta)(R-T) + (3+\beta)(S-P) > 0. \tag{42}$$

Then the threshold that α must exceed for this to hold is

$$\alpha^* = (3+2\beta)(T-R) + (3+\beta)(P-S), \tag{43}$$

which is clearly positive. Furthermore, increasing the deme size N or the baseline migration probability m will increase this threshold, making it more challenging for selection to favor the fixation of C. This result aligns with the conclusion in the context of the simplified Prisoner's dilemma.

Result 3: For the general Prisoner's dilemma where the payoffs as well as the differential migration coefficient are of the same small order, there exists a positive threshold α^* above which selection will favor the fixation of C and under which selection will disfavor the fixation of C. Furthermore, increasing the deme size N or the baseline migration probability m will increase this threshold, making it more difficult for selection to favor the fixation of C.

E. Case 5: Repeated Prisoner's dilemma

Here, we consider the Prisoner's dilemma repeated n times with the following alternative strategies in use: an individual either defects in all rounds ($AllD$), or begins by cooperating and then imitates the opponent's previous action for the next $n-1$ rounds (TFT). Two individuals using TFT always cooperate, and the fixation of TFT is equivalent to the fixation of cooperation. The advantages of the TFT strategy lie in its ability to encourage cooperation by rewarding it and in its effectiveness in discouraging strategies that attempt to benefit from defecting without repercussions.

In this case, the payoff matrix is given by

$$\begin{array}{c} \\ TFT \\ AllD \end{array} \begin{array}{c} \begin{array}{cc} TFT & AllD \end{array} \\ \begin{pmatrix} nR & S+(n-1)P \\ T+(n-1)P & nP \end{pmatrix} \end{array}. \tag{44}$$

By substituting these entries of the payoff matrix in Eq. (24), we have

$$\Psi(y) = \exp\left(\frac{N-1}{N}\left[-\frac{y(1-y)}{2(2+\beta)}\alpha + \frac{(2y+\beta y^2)}{2(2+\beta)}[T-R-(n-1)(R-P)] + \frac{(2y+\beta(2-y)y)}{2(2+\beta)}(P-S)\right]\right). \tag{45}$$

As expected, increasing the number of repetitions n reduces $\psi(y)$ for any $y \in (0,1)$, thereby increasing the fixation probability ρ_C. This suggests a negative correlation between the threshold α^* and the number of repetitions n (see Fig. 5).

Suppose that the payoffs R, S, T, P as well as the differential migration coefficient α are of the same small order.

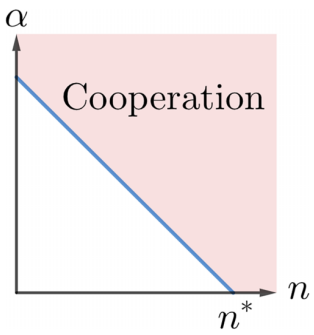


FIG. 5. The graph illustrates the relationship between the critical threshold α^* and the number of repetitions n in the repeated Prisoner's dilemma, represented by the blue curve. In the region above the curve, selection favors the fixation of C, while in the white region below, selection disfavors it. The intersection of the curve with the n axis is the threshold n^* above which selection favors the fixation of C, irrespective of the presence of any additional migration probability for cooperators.

By substituting the updated entries of the payoff matrix in Eq. (43), the threshold α^* takes the form

$$\alpha^* = (3+2\beta)(T-R) + (3+\beta)(P-S) - (n-1)(3+2\beta)(R-P). \tag{46}$$

The effect of the number of repetitions n on α^* is additive. In the first round, where cooperators help anyone, the dynamics work against the fixation of C, requiring α to exceed the threshold in the case of the Prisoner's dilemma given by $(3+2\beta)(T-R)+(3+\beta)(P-S)$. In subsequent rounds, cooperators continue helping only cooperators; such scenario will lower the required threshold by $(3+2\beta)(R-P)$ for each of the next $n-1$ rounds. Notably, increasing the average number of repetitions n reduces this threshold, thereby promoting conditions under which selection can favor the fixation of C. Furthermore, if the number of repetitions satisfies

$$n > n^* = 1 + \frac{(3+2\beta)(T-R)+(3+\beta)(P-S)}{(3+2\beta)(R-P)}, \tag{47}$$

selection will favor the fixation of C irrespective of any differential migration. Figure 5 illustrates these findings.

Note that

$$\frac{\partial}{\partial\beta}\left(\frac{(3+2\beta)(T-R)+(3+\beta)(P-S)}{(3+2\beta)(R-P)}\right) = -\frac{3(P-S)}{(R-P)(3+2\beta)^2} < 0, \tag{48}$$

indicating that increasing the deme size N or the baseline migration probability m lowers the threshold α^*, which plays in favor of the fixation of C.

Result 4: In the context of the repeated Prisoner's dilemma, the threshold α^* shows a negative correlation with the number of repetitions n. Furthermore, there is a threshold n^* beyond which selection favors the fixation of C regardless of the magnitude of the additional migration probability for cooperators. Increasing the deme size N or the baseline migration m will lower this threshold, making it easier for selection to favor the fixation of C.

TABLE I. The number of observed fixations of C following its introduction as a single mutant over 10^5 stochastic simulations for different values of N, d and α.

α	Number of occurrences[a]	Number of occurrences[b]	Number of occurrences[c]	Number of occurrences[d]
0.0	880	415	418	204
0.4	945	403	451	211
0.8	925	414	423	218
1.2	996	465	442	212
1.6	998	470	500	214
2.0	1062	447	495	221
2.4	1112	477	479	227
2.8	1086	420	479	233
3.2	1046	474	496	231
3.6	1063	491	476	236
4.0	1052	496	498	238

[a] $N = 2$ and $d = 50$.
[b] $N = 4$ and $d = 50$.
[c] $N = 2$ and $d = 100$.
[d] $N = 4$ and $d = 100$.

V. COMPUTER SIMULATIONS

In this section, we will test the robustness of the approximation by a diffusion process using stochastic simulations. The state of the population is represented by an array containing d arrays, each of size N. At the start of each simulation, we initialize each deme with N individuals of type D. Then, we randomly select a deme and replace one individual with an individual of type C to initiate the simulation. As long as fixation has not yet occurred, we choose a random deme, calculate the fitness of each individual in that deme using its payoff, and select an individual to reproduce based on these fitness values. If the offspring migrates, it replaces an individual chosen at random in one of the other $d - 1$ demes chosen at random; otherwise, it replaces a random individual in the current deme. Finally, the code returns the type that reaches fixation, either C or D. We repeat this process 10^5 times and count how many times the fixation of C occurs.

We consider $d = 50, 100$ demes, each of size $N = 2, 4$, connected by a baseline migration probability $m = 0.1$. Furthermore, there is a linear migration term for C offspring with a differential migration coefficient $\alpha = 0, 0.4, 0.8, \ldots, 4$. For the payoff matrix, we use $R = 0.5, S = -0.5, T = 1$, and $P = 0$ (this typically represents the simplified Prisoner's dilemma with $b = 1$ and $c = 0.5$). The results of these stochastic simulations are presented in Table I.

We used the simulated values to evaluate the predictive accuracy of the diffusion process. In Fig. 6, the simulated data are presented in conjunction with the predicted values obtained from the diffusion model using the approximation in Eq. (21). For $d = 50$ and $N = 2$, the analysis reveals discrepancies between the simulated and predicted values. However, when we increase $Nd = 200$ with $d = 50$ and $N = 4$ or $d = 100$ and $N = 2$, this gap significantly decreases, resulting in minor discrepancies. For $N = 100$ and $d = 4$ making $Nd = 400$, the difference between the observed and predicted values becomes markedly smaller. These results reinforce the efficacy of the diffusion process as a reliable approximation of the underlying dynamics.

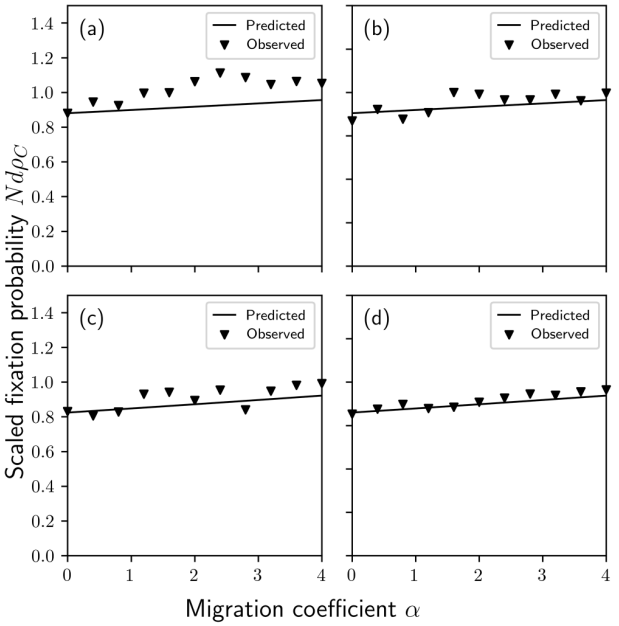


FIG. 6. These graphs present a comparison between the population-scaled fixation probability $Nd\rho_C$ obtained from stochastic simulations (observed values) and the one obtained using the diffusion approximation. For $N = 2$, (a) and (b) correspond to $d = 50$ and $d = 100$, respectively. Similarly, for $N = 4$, (c) and (d) represent $d = 50$ and $d = 100$. The simulations are performed with $m = 0.1$, using payoffs $R = 0.5$, $S = -0.5$, $T = 1$, and $P = 0$.

VI. DISCUSSION

A. Summary of the main results

In this study, we have investigated conditions for the evolution of cooperation, represented by C, in a haploid population subdivided into d demes, each containing $N \geqslant 2$ individuals. Individuals within each deme interact according to a Prisoner's dilemma. Reproduction follows a discrete-time Moran process. At each time step, we choose one deme at random and one individual in that deme to reproduce with a probability proportional to its fitness. The offspring produced either replaces a randomly selected individual within the same deme or migrates to a different deme with a baseline probability $0 < m \leqslant 1$. Additionally, an offspring of a cooperator has an extra migration probability of $s\alpha(1 - x_i)/(Nd)$ for some differential migration coefficient α, where s is the intensity of selection and x_i is the frequency of cooperators in the deme. Following methods similar to the one used in Wakeley and Takahashi [31], we show that the changes from one time step to the next, given the frequency of deme types and the overall frequency of cooperators in the population, meet key conditions for a Markov chain with two time scales outlined in Ethier and Nagylaki [22] as d approaches infinity. This allows us to approximate the stochastic dynamics of the total frequency of C with a one-dimensional diffusion process that has two absorbing states, 0 and 1. Using this, we have calculated

the fixation probability of C following its introduction as a single mutant in the population, denoted by ρ_C. We have also examined how migration, deme size, and payoff differences influence this fixation probability.

In the absence of any additional migration probability for cooperators, selection disfavors the fixation of C in the sense that $\rho_C < 1/(Nd)$. While raising the baseline migration probability m increases ρ_C, it is not sufficient to reverse the inequality for selection to favor the fixation of C. However, in a repeated Prisoner's dilemma with $AllD$ and TFT starting with C as strategies, an increase in the average number of repetitions boosts the probability for C to fix. If the number of repetitions surpasses some critical threshold n^*, then the fixation of C will become favored by selection. Furthermore, increasing the deme size N or the baseline migration probability m lowers this threshold, thereby relaxing the conditions for C to fix.

In a general Prisoner's dilemma model, adding an extra migration probability for cooperators driven by a differential migration coefficient α will increase the fixation probability ρ_C making it possible for selection to favor the fixation of C. This scenario, which is similar to the opting-out strategy [54–57], but different since the probability to leave is here type-specific, will allow cooperators to more readily move away from defector-dominated demes increasing their chances of aggregating with other cooperators within demes, thereby enhancing the probability for C to fix. In demes with a high frequency of cooperators, this extra migration probability will be reduced, which will decrease the risk of joining defector-dominated demes outweighing potential benefits. This aligns with findings by Lehmann and Keller [58], who suggest that stable demes of cooperators provide a secure environment, reducing the incentive for cooperators to migrate. There exists a threshold α^*, dependent on N, m, and the payoffs, beyond which selection favors the fixation of C and below which selection disfavors the fixation of C. In a simplified Prisoner's dilemma with a cost c to a cooperator and benefit $b > c$ to its partner, our diffusion approximation suggests that b does not affect the stochastic dynamics. However, an increase in c lowers the fixation probability ρ_C. This shows that c and b do not affect the dynamics in the same way. Furthermore, we have $\alpha^* = (2+\nu)u^*(c)$, where $u^*(c)$ is increasing with respect to c and $Nm \to \nu$ as $N \to \infty$ and $m \to 0$. This implies that a large value of ν raises the threshold α^*, making the fixation of C more challenging.

In a repeated Prisoner's dilemma, an inverse relationship arises between the threshold α^* and the number of repetitions n. As n increases, the threshold α^* decreases, lowering the barrier for the fixation of C. This effect highlights that repeated interactions, combined with demographic structure, can promote the evolution of cooperative behavior within populations [20,59].

Numerical simulations were performed to evaluate the accuracy of the diffusion approximation to calculate the fixation probability ρ_C for a finite number of demes d. The results show that the theoretical predictions closely match the observed fixation probabilities when $d \gg 1$. For a simplified Prisoner's dilemma with $b = 1$, $c = 0.1$, and $m = 10^{-2}$, simulations were conducted for $d = 50$ and $d = 100$, with deme sizes $N = 2$ and $N = 4$. The values calculated with the diffusion approximation closely matches the simulated values, with minor discrepancies for $Nd = 100$ that diminish when $Nd = 200$. For $Nd = 400$, the agreement is nearly perfect. These findings demonstrate the reliability of the diffusion approximation under appropriate conditions.

In our model, we assume a constant baseline migration probability m for all individuals in each deme. Allowing m to vary across generations—provided these probabilities are independent and identically distributed—does not substantially affect our main result which then depends on the mean migration probability and on the greater tendency for cooperators to leave defector-rich demes. Indeed, because the transition probabilities $P_{i,i+1}(\mathbf{z})$ and $P_{i,i-1}(\mathbf{z})$ depend linearly on m (see Appendix A) and the diffusion coefficients μ and σ^2 depend linearly on those transition probabilities, heterogeneity in m merely adds variance to individual dispersal events without changing the infinitesimal mean or diffusion. Thus, moderate variation in migration propensities preserves the advantage of cooperation and validates our choice of a uniform m in the main analysis.

B. Influence of various factors on the diffusion dynamics

The diffusion process under examination is affected by three primary forces: genetic drift, arising from random fluctuations in allele frequencies due to the stochastic processes of reproduction and replacement; selection, driven by the payoff disparities between cooperators and defectors, and variations in migration probabilities.

Each force contributes additively to the infinitesimal mean and infinitesimal variance of the process. For instance, consider the infinitesimal mean

$$\mu(x) = \frac{(N-1)x(1-x)}{2N}\left[\delta(x)(R-T) + (1-\delta(x))(S-P) + \frac{\alpha}{2+m(N-2)}(1-2x)\right], \tag{49}$$

where $\delta(x) = \frac{1+m(N-2)x}{2+m(N-2)}$. In this formulation, genetic drift exerts no direct effect on the mean. Payoff differences, however, influence the infinitesimal mean through the term

$$\frac{(N-1)x(1-x)}{2N}[\delta(x)(R-T) + (1-\delta(x))(S-P)].$$

This term shows that the game is equivalent to a game in a well-mixed population with nonuniform interactions, where $\delta(x)$ is the probability of interacting with a cooperator, and $1-\delta(x)$ is the probability of interacting with a defector, when the frequency of cooperators in the population is x. The term

$$\frac{(N-1)x(1-x)}{2N}\left[\frac{\alpha}{2+m(N-2)}(1-2x)\right] \tag{50}$$

reflects the differential migration probabilities for cooperators versus defectors. For the infinitesimal variance

$$\sigma^2(x) = x(1-x), \tag{51}$$

the effects of payoff differences and migration coefficients are minimal, while genetic drift is represented by the classic variance term $x(1-x)$, consistent with the Wright-Fisher model [60]. This is obtained, however, by taking $\frac{N^2d^2}{2(1-F)}$ time steps

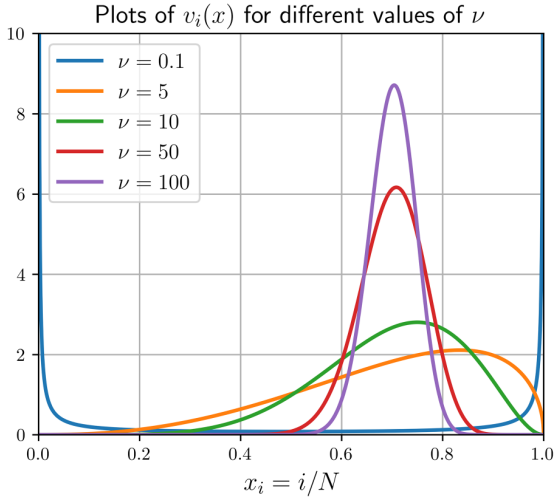


FIG. 7. These plots use the approximation in Appendix D. They illustrate the impact of the parameter ν on the distribution of $v_i(x)$, representing the frequency of demes composed of $x_i = i/N$ cooperators and $1 - x_i$ defectors, for $i = 0, 1, \ldots, N$, where the overall frequency of C is $x = 0.7$. For a low value of ν (e.g., $\nu = 0.01$), the peaks are near 0 and 1, indicating that demes are mostly homogeneous, consisting of either all cooperators or all defectors. As ν increases to moderate values (e.g., $\nu = 5$ or 10), the distribution broadens, favoring demes with more mixed compositions. At high values of ν (e.g., $\nu = 50$ or 100), the distribution becomes sharply centered around $x = 0.7$, suggesting that demes with cooperator frequencies near x tend to dominate. This shift reflects how ν acts as a mixing parameter, with low values promoting polarization and high values encouraging homogeneity around the global frequency. Additionally, increasing ν reduces variability, suppressing extreme deme compositions and leading to a stronger alignment with the overall cooperator frequency.

as unit of time as $d \to \infty$, where F is Wright's [49] fixation index (see, e.g., Wakeley [50], Rousset [51]).

When $\nu = Nm \to \infty$ as $N \to \infty$ and $m \to 0$, dispersal tends to eliminate the population structure, decreasing relatedness among individuals and driving the population dynamics towards a well-mixed model with an infinitesimal mean given by

$$\mu(x) = \frac{x(1-x)}{2}[x(R-T) + (1-x)(S-P)]. \tag{52}$$

In this scenario, the additional migration force benefiting cooperators becomes negligible compared to the effect of the payoff differences. This can be understood from Eq. (8a), where this migration force is modulated mainly by $x_i - x$, the deviation of the local cooperator frequency from the global frequency in the population, for $i = 0, 1, \ldots, N$.

As ν reaches high values, this deviation tends towards zero (as shown in Fig. 7 for $\nu = 100$), reflecting a population dominated by demes with cooperator frequencies near x. Consequently, when cooperators move to different demes, they gain no advantage since the composition of their new demes mirrors that of the ones they left.

On the other hand, when $\nu = Nm \to 0$ as $N \to \infty$ and $m \to 0$, the deme-scaled baseline migration probability is minimal, allowing genetic drift and localized interactions to maintain the population structure. Under these conditions, the infinitesimal mean becomes

$$\mu(x) = \frac{x(1-x)}{4}[R - T + S - P + \alpha(1-2x)]. \tag{53}$$

In this scenario, the influence of differential migration is maximized, exerting a substantial effect on the diffusion dynamics. This occurs because the deviation of the local cooperator frequencies from the global frequency is at its greatest reflecting that the population homogeneity is at its highest level. This is visually supported by Fig. 7 for $\nu = 0.01$. As a result, when cooperators switch demes, they gain a greater advantage due to the shifting composition of their new demes.

In this case, differential migration probabilities can balance or offset selection pressures from payoff differences, enabling cooperators to avoid competitive disadvantages. This mechanism enhances the spread of cooperation up to fixation, even in environments where defectors might otherwise prevail.

ACKNOWLEDGMENTS

S.L. was supported by the Natural Sciences and Engineering Research Council of Canada (Grant No. 8833). We thank the anonymous reviewers for their constructive feedback, which helped improve the clarity and quality of the manuscript.

D.K. contributed to Conceptualization, Methodology, Formal Analysis, Investigation, Validation, Writing – Original Draft, Writing – Review & Editing, Performed computer simulations, Analyzed Data. S.L. contributed to Conceptualization, Validation and Writing – Review & Editing.

The authors declare no competing financial interest.

DATA AVAILABILITY

The data that support the findings of this article are openly available on GitHub [61] and preserved on Zenodo [62].

APPENDIX A: PROOFS OF EQS. (7) AND (10)

This Appendix is organized as follows. In Appendix A 1, we derive the probabilities for a deme of type i to become either a deme of type $i-1$ or a deme of type $i+1$. Using these probabilities, Appendix A 2 establishes the first three conditions in Eqs. (7a)–(7c). Appendix A 3 presents the expression for $\mathbf{v}(x)$, followed by an analysis of its conditional change in Appendix A 4. These results are then applied in Appendices A 5 and A 6 to derive the remaining two conditions in Eq. (7d)–(7e). Finally, Appendix A 7 demonstrates the stability of the zero solution of Eq. (10).

Let the frequency vector of deme types at time step $t = 0$ be represented by $\mathbf{z} = (z_0, z_1, \ldots, z_N)^T$. Demes of type i are randomly assigned positions from 1 to $d_i = dz_i$, for $i = 0, 1, \ldots, N$, in a such way that each deme can be identified by the pair (i, k), where i and k indicate the deme type and its position among demes of the same type, respectively. From time step $t = 0$ to time step $t = 1$, a deme of type i can become a deme of type $i-1$, a deme of type $i+1$, or remain of the same type.

Let $X_{(i,k)}(0)$ and $X_{(i,k)}(1)$ denote the frequency of type C in deme (i, k) at the time steps $t = 0$ and $t = 1$, respectively.

Then, the change in the average frequency of type C across all demes is given by

$$\Delta X(0) = \frac{1}{d}\sum_{i=0}^{N}\sum_{k=1}^{d_i}\Delta X_{(i,k)}(t), \tag{A1}$$

where

$$\Delta X_{(i,k)}(0) = X_{(i,k)}(1) - X_{(i,k)}(0) \tag{A2}$$

can take values $-1/N, 0, 1/N$. Let $P_{i,i-1}(\mathbf{z})$ and $P_{i,i+1}(\mathbf{z})$ denote the probabilities that $\Delta X_{(i,k)}(0) = -1/N$ and $\Delta X_{(i,k)}(0) = 1/N$, respectively.

1. Derivation of $P_{i,i+1}(\mathbf{z})$ and $P_{i,i-1}(\mathbf{z})$

Deme (i,k), which is of type i at time step $t=0$, will become of type $i+1$ at time step $t=1$ if an offspring of a cooperator replaces a defector picked up at random in deme (i,k). This offspring will come from a deme of type j with probability $m_{C,j}/(d-1)$, for $j \neq i$. Alternatively, the offspring can be originally from deme (i,k) with a probability of $1-m_{C,i}$. Thus, deme (i,k) will become of type $i+1$ with probability

$$P_{i,i+1}(\mathbf{z}) = \sum_{j=0}^{N}\left(z_j - \frac{\delta_{ij}}{d}\right)\frac{x_j(1+s\pi_{C,j})}{1+s\tilde{\pi}_j}\times\frac{m_{C,j}}{d-1}\times(1-x_i) + \frac{1}{d}\times\frac{x_i(1+s\pi_{C,i})}{1+s\tilde{\pi}_i}\times(1-m_{C,i})\times(1-x_i), \tag{A3}$$

where $\tilde{\pi}_j = x_j\pi_{C,j} + (1-x_j)\pi_{D,j}$ is the average fitness in a deme of type j, and δ_{ij} is equal to 1 if $i=j$ and 0 otherwise. Using the facts that $x = \sum_{i=0}^{N} z_i x_i$ and $\sum_{i=0}^{N} z_i = 1$, we obtain

$$\begin{aligned}
P_{i,i+1}(\mathbf{z}) &= \sum_{j=0}^{N}(z_j - \delta_{ij}Ns)x_j[1+s(\pi_{C,j}-\tilde{\pi}_j)](m+sM_{C,j})(Ns+N^2s^2)(1-x_i)\\
&\quad + Nsx_i[1+s(\pi_{C,i}-\tilde{\pi}_i)](1-m-sM_{C,i})(1-x_i) + \mathcal{O}(s^3)\\
&= Ns(1-x_i)[(1-m)x_i + mx] + Ns^2(1-x_i)\Bigg(mN(x-x_i) + \sum_{j=0}^{N} z_j x_j[M_{C,j} + m(\pi_{C,j}-\tilde{\pi}_j)]\\
&\quad + x_i[(1-m)(\pi_{C,i}-\tilde{\pi}_i) - M_{C,i}]\Bigg) + \mathcal{O}(s^3),
\end{aligned} \tag{A4}$$

where $\mathcal{O}(h)$ denotes a term of order h.

The other possible change in the composition of deme (i,k) is to have less cooperators in the next generation. This occurs if a cooperator in deme (i,k) is replaced by an offspring of a defector, coming from a deme of type j with probability $m_{j,D}/(d-1)$ for $j \neq i$ or originally from deme (i,k) with probability $1-m_{i,D}$. This scenario holds with probability

$$\begin{aligned}
P_{i,i-1}(\mathbf{z}) &= \sum_{j=0}^{N}\left(z_j - \frac{\delta_{ij}}{d}\right)\frac{(1-x_j)(1+s\pi_{D,j})}{1+s\tilde{\pi}_j}\times\frac{m_{D,j}}{d-1}\times x_i + \frac{1}{d}\times\frac{(1-x_i)(1+s\pi_{D,i})}{1+s\tilde{\pi}_i}\times(1-m_{D,i})\times x_i,\\
&= Nsx_i\sum_{j=0}^{N}(z_j - \delta_{ij}Ns)(1-x_j)[m + ms(N+\pi_{D,j}-\tilde{\pi}_j)] + Nsx_i(1-x_i)[1-m+s(1-m)(\pi_{D,i}-\tilde{\pi}_i)] + \mathcal{O}(s^3)\\
&= Nsx_i[(1-m)(1-x_i) + m(1-x)] + Ns^2x_i\Bigg(mN(x_i-x) + m\sum_{j=0}^{N} z_j(1-x_j)(\pi_{D,j}-\tilde{\pi}_j)\\
&\quad + (1-m)(1-x_i)(\pi_{D,i}-\tilde{\pi}_i)\Bigg) + \mathcal{O}(s^3),
\end{aligned} \tag{A5}$$

for any $i = 1, \ldots, N$.

2. Derivation of the moments of $\Delta X(0)$

In this section, our aim is to determine the moments of the conditional change in the total frequency of C from time step $t=0$ to time step $t=1$, denoted by $\Delta X(0)$. This change can be expressed as

$$\Delta X(0) = \frac{1}{d}\sum_{i=0}^{N}\sum_{k=1}^{dz_i}\Delta X_{(i,k)}(0). \tag{A6}$$

Then the expected conditional change is given by

$$\begin{aligned}
\mathbb{E}_{\mathbf{z}}[\Delta X(0)] &= \frac{1}{d}\sum_{i=0}^{N}\sum_{k=1}^{dz_i}\mathbb{E}_{\mathbf{z}}[\Delta X_{(i,k)}(0)]\\
&= \frac{1}{d}\sum_{i=0}^{N}\sum_{k=1}^{dz_i}\frac{P_{i,i+1}(z) - P_{i,i-1}(z)}{N}.
\end{aligned} \tag{A7}$$

Using Eqs. (A4) and (A5), we obtain

$$\mathbb{E}_{\mathbf{z}}[\Delta X(0)] = s^2\Bigg[\sum_{i=0}^{N} z_i x_i(x_i - x)M_{C,i} + \sum_{i=0}^{N} z_i x_i(1-x_i)(\pi_{C,i} - \pi_{D,i})\Bigg] + \mathcal{O}(s^3). \tag{A8}$$

This shows the condition in Eq. (7a).

In the structured Moran model considered in this paper, only one deme can change at a time. This means that $\Delta X_{(i,k)}(0) = 0$ for all (i,k) except perhaps for $(i,k) = (i_0, k_0)$. This implies that

$$(\Delta X(0))^k = \frac{1}{d^k}\sum_{i=0}^{N}\sum_{k=1}^{dz_i}(\Delta X_{(i,k)}(0))^k \tag{A9}$$

for any $k \geqslant 1$. In particular, the second moment of the conditional change in the frequency of type C is

$$\begin{aligned}\mathbb{E}_{\mathbf{z}}[(\Delta X(0))^2] &= \frac{1}{d^2}\sum_{i=0}^{N}\sum_{k=1}^{dz_i}\mathbb{E}_{\mathbf{z}}\big[(\Delta X_{(i,k)}(0))^2\big] \\ &= \frac{1}{d^2}\sum_{i=0}^{N}\sum_{k=1}^{dz_i}\frac{P_{i,i-1}(\mathbf{z}) + P_{i,i+1}(\mathbf{z})}{N^2} \\ &= 2s^2\Bigg[(1-m)\sum_{i=0}^{N} z_i x_i(1-x_i) + mx(1-x)\Bigg] \\ &\quad + \mathcal{O}(s^3).\end{aligned} \tag{A10}$$

Moreover, using Eq. (A9) with $k = 4$ and the fact that $|\Delta X_{(i,k)}(0)| \leqslant 1/N$ for any (i,k) lead to

$$\mathbb{E}_{\mathbf{z}}[(\Delta X(0))^4] = \mathcal{O}\big(s^3\big). \tag{A11}$$

This completes the conditions in Eqs. (7b)–(7c).

3. Derivation of $\mathbf{v}(x)$

Next, we will develop the expression of the distribution $\mathbf{v}(x)$. It is the stationary distribution of the deme types in a structured population consisting of infinitely many demes, each of size $N \geqslant 2$ under neutrality, where we assume no selection or additional migration probabilities for cooperators, maintaining a constant overall frequency of C at x.

When a replacement occurs in a deme of type i, it may become of type $i+1$, type $i-1$, or remain of the same type. Let $P^*_{i,i+1}(x)$, $P^*_{i,i-1}(x)$, and $P^*_{i,i}(x)$ be the probabilities of these respective events. Their expressions are given by

$$\begin{aligned}P^*_{i,i+1}(x) &= [(1-m)x_i + mx](1-x_i), \\ P^*_{i,i-1}(x) &= [(1-m)(1-x_i) + m(1-x)]x_i, \\ P^*_{i,i}(x) &= 1 - P^*_{i,i+1}(x) - P^*_{i,i-1}(x).\end{aligned} \tag{A12}$$

Extending these transition probabilities by taking $P^*_{i,j}(x) = 0$ for $|j-i| \leqslant 2$, the frequency vector $\mathbf{v}(x) = (v_0(x), v_1(x), \ldots, v_N(x))^T$ with $\sum_{i=0}^{N} v_i(x) = 1$ is the solution of

$$\mathbf{v}(x)^T = \mathbf{v}(x)^T\mathbf{P}^*(x), \tag{A13}$$

where $\mathbf{P}^*(x) = (P^*_{i,j}(x))_{0\leqslant i,j\leqslant N}$. This defines $\mathbf{v}(x)$ as the stationary distribution of a discrete-time birth-death process on the state space $\{0, 1, \ldots, N\}$ with birth probabilities $P^*_{i,i+1}(x)$ for $i = 0, 1, \ldots, N-1$ and death probabilities $P^*_{i,i-1}(x)$ for $i = 1, \ldots, N$. Using a standard argument (see, e.g., Karlin and Taylor [63]), the solution is found to be

$$v_i(x) = v_0(x) \times \frac{P^*_{0,1}(x)P^*_{1,2}(x)\cdots P^*_{i-1,i}(x)}{P^*_{1,0}(x)P^*_{2,1}(x)\cdots P^*_{i,i-1}(x)}, \tag{A14}$$

where

$$v_0(x) = \left(1 + \sum_{i=1}^{N}\frac{P^*_{0,1}(x)P^*_{1,2}(x)\cdots P^*_{i-1,i}(x)}{P^*_{1,0}(x)P^*_{2,1}(x)\cdots P^*_{i,i-1}(x)}\right)^{-1}, \tag{A15}$$

for $i \geqslant 1$.

Lemma A 1. We have

$$v_i(x) = \binom{N}{i}\frac{B\big(\frac{Nmx}{1-m} + i, \frac{Nm(1-x)}{1-m} + N - i\big)}{B\big(\frac{Nmx}{1-m}, \frac{Nm(1-x)}{1-m}\big)} \tag{A16}$$

for $i = 0, 1, \ldots, N$.

Proof. Recall the Gamma function, defined as

$$\Gamma(t) = \int_0^\infty u^{t-1}e^{-u}\,du,$$

which satisfies $\frac{\Gamma(t+i)}{\Gamma(t)} = \prod_{k=0}^{i-1}(t+k)$, for any $t > 0$ and any integer $i \geqslant 1$. Using this identity, we obtain

$$\begin{aligned}&P^*_{0,1}(x)P^*_{1,2}(x)\cdots P^*_{i-1,i}(x) \\ &= \prod_{k=0}^{i-1}\left[(1-m)\frac{k}{N} + mx\right]\left(1 - \frac{k}{N}\right) \\ &= \frac{(1-m)^i}{N^{2i}}N(N-1)\cdots(N-i+1)\prod_{k=0}^{i-1}\left[\frac{Nmx}{1-m} + k\right] \\ &= \frac{(1-m)^i}{N^{2i}}i!\binom{N}{i}\frac{\Gamma\big(\frac{Nmx}{1-m} + i\big)}{\Gamma\big(\frac{Nmx}{1-m}\big)}.\end{aligned} \tag{A17}$$

Similarly, we have

$$\begin{aligned}&P^*_{1,0}(x)P^*_{2,1}(x)\cdots P^*_{i,i-1}(x) \\ &= \frac{(1-m)^i}{N^{2i}}i!\frac{\Gamma\big(\frac{Nm(1-x)}{1-m} + N\big)}{\Gamma\big(\frac{Nm(1-x)}{1-m} + N - i\big)}.\end{aligned} \tag{A18}$$

Dividing Eq. (A17) by Eq. (A18) leads to

$$\begin{aligned}&\frac{P^*_{0,1}(x)P^*_{1,2}(x)\cdots P^*_{i-1,i}(x)}{P^*_{1,0}(x)P^*_{2,1}(x)\cdots P^*_{i,i-1}(x)} \\ &= \binom{N}{i}\frac{\Gamma\big(\frac{Nmx}{1-m} + i\big)\Gamma\big(\frac{Nm(1-x)}{1-m} + N - i\big)}{\Gamma\big(\frac{Nmx}{1-m}\big)\Gamma\big(\frac{Nm(1-x)}{1-m} + N\big)} \\ &= \binom{N}{i}\frac{B\big(\frac{Nmx}{1-m} + i, \frac{Nm(1-x)}{1-m} + N - i\big)}{B\big(\frac{Nmx}{1-m}, \frac{Nm(1-x)}{1-m} + N\big)}.\end{aligned} \tag{A19}$$

On the other hand, we have the identity

$$\sum_{i=0}^{N}\binom{N}{i}B(a+i, b+N-i) = B(a,b), \tag{A20}$$

which is a consequence of the binomial formula. Inserting Eq. (A19) into Eq. (A15) and using the above identity with $a = \frac{Nmx}{1-m}$ and $b = \frac{Nm(1-x)}{1-m}$, we get

$$v_0(x) = \frac{B\left(\frac{Nmx}{1-m}, \frac{Nm(1-x)}{1-m} + N\right)}{B\left(\frac{Nmx}{1-m}, \frac{Nm(1-x)}{1-m}\right)}. \tag{A21}$$

Now, inserting Eq. (A21) into Eq. (A14) and using again Eq. (A19), we conclude that

$$v_i(x) = \binom{N}{i} \frac{B\left(\frac{Nmx}{1-m} + i, \frac{Nm(1-x)}{1-m} + N - i\right)}{B\left(\frac{Nmx}{1-m}, \frac{Nm(1-x)}{1-m}\right)} \tag{A22}$$

for $i = 0, 1, \ldots, N$. ■

4. Derivation of the moments of $\Delta v_i(0) = v_i(X(1)) - v_i(x)$

Given that $X(0) = x$, we have

$$\begin{aligned}
v_i(X(1)) &= \frac{\binom{N}{i}\Gamma\left(\frac{Nm}{1-m}\right)}{\Gamma\left(\frac{N}{1-m}\right)} \frac{\Gamma\left(\frac{NmX(1)}{1-m} + i\right)}{\Gamma\left(\frac{NmX(1)}{1-m}\right)} \frac{\Gamma\left(\frac{Nm(1-X(1))}{1-m} + N - i\right)}{\Gamma\left(\frac{Nm(1-X(1))}{1-m}\right)} \\
&= \frac{\binom{N}{i}\Gamma\left(\frac{Nm}{1-m}\right)}{\Gamma\left(\frac{N}{1-m}\right)} \prod_{k=0}^{i-1} \left(\frac{NmX(1)}{1-m} + k\right) \prod_{l=0}^{N-i-1} \left(\frac{Nm(1-X(1))}{1-m} + l\right) \\
&= \frac{\binom{N}{i}\Gamma\left(\frac{Nm}{1-m}\right)}{\Gamma\left(\frac{N}{1-m}\right)} \prod_{k=0}^{i-1} \prod_{l=0}^{N-i-1} \left(\frac{Nm(x+\Delta X(0))}{1-m} + k\right)\left(\frac{Nm(1-x-\Delta X(0))}{1-m} + l\right) \\
&= \frac{\binom{N}{i}\Gamma\left(\frac{Nm}{1-m}\right)}{\Gamma\left(\frac{N}{1-m}\right)} \prod_{k=0}^{i-1} \prod_{l=0}^{N-i-1} \left[\left(\frac{Nmx}{1-m} + k\right)\left(\frac{Nm(1-x)}{1-m} + l\right) + \frac{Nm}{1-m}\left(\frac{Nm(1-2x)}{1-m} + k - l\right)\Delta X(0)\right. \\
&\quad \left. + \left(\frac{Nm}{1-m}\right)^2 (\Delta X(0))^2\right] \\
&= v_i(x) + \sum_{j=1}^{2N} C(i, j, N, m, x)(\Delta X(0))^j,
\end{aligned} \tag{A23}$$

where $C(i, j, N, m, X(0))$ is a bounded constant. Using Eq. (A9), we find that

$$\mathbb{E}_{\mathbf{z}}[\Delta v_i(0)] = \sum_{j=1}^{2N} C(i, j, N, m, x)\mathbb{E}_{\mathbf{z}}[(\Delta X(0))^j] = \mathcal{O}(1/d^2) \tag{A24}$$

and

$$\mathbb{E}_{\mathbf{z}}[(\Delta v_i(0))^2] = \sum_{j,l=1}^{2N} C(i, j, N, m, x)C(i, l, N, m, x)\mathbb{E}_{\mathbf{z}}[(\Delta X(0))^{j+l}] = \mathcal{O}(1/d^2) \tag{A25}$$

for $i = 0, 1, \ldots, N$. Combining (A24) and (A25) leads to

$$\mathrm{Var}_{\mathbf{z}}[\Delta v_i(0)] = \mathcal{O}(1/d^2) \tag{A26}$$

for $i = 0, 1, \ldots, N$.

5. Derivation of the moments of $\Delta Z_i(0) = Z_i(1) - z_i$

Given $\mathbf{Z}(0) = \mathbf{z}$, the random variable $Z_i(1)$ can take the value $z_i + 1/d$ if the replacement occurs in a deme of type $i-1$, where an offspring of a cooperator replaces a defector, or in a deme of type $i+1$, where an offspring of a defector replaces a cooperator. This implies that

$$\begin{aligned}
&\mathbb{P}_{\mathbf{z}}\left[Z_i(1) = z_i + \frac{1}{d}\right] \\
&= d_{i+1}P_{i+1,i}(\mathbf{z}) + d_{i-1}P_{i-1,i}(\mathbf{z}) \\
&= z_{i+1}P^*_{i+1,i}(x) + z_{i-1}P^*_{i-1,i}(x) + \mathcal{O}(1/d).
\end{aligned} \tag{A27}$$

Here, $\mathbb{P}_{\mathbf{z}}$ is the conditional probability given that $\mathbf{Z}(0) = \mathbf{z}$. On the other hand, $Z_i(1)$ can take the value $z_i - 1/d$ if the replacement occurs in a deme of type i, where either an offspring of a defector replaces a cooperator or an offspring of a cooperator replaces a defector. This occurs with probability

$$\begin{aligned}
\mathbb{P}_{\mathbf{z}}\left[Z_i(1) = z_i - \frac{1}{d}\right] &= d_iP_{i,i+1}(\mathbf{z}) + d_iP_{i,i-1}(\mathbf{z}) \\
&= z_i\left(P^*_{i,i+1}(x) + P^*_{i,i-1}(x)\right) + \mathcal{O}(1/d) \\
&= z_i\left(1 - P^*_{i,i}(x)\right) + \mathcal{O}(1/d).
\end{aligned} \tag{A28}$$

Finally, note that $Z_i(1) = z_i$ with probability

$$1 - \mathbb{P}_{\mathbf{z}}\left[Z_i(1) = z_i + \frac{1}{d}\right] - \mathbb{P}_{\mathbf{z}}\left[Z_i(1) = z_i - \frac{1}{d}\right].$$

Then, the conditional expected value of the change $\Delta Z_i(0) = Z_i(1) - z_i$ can be expressed as

$$\mathbb{E}_{\mathbf{z}}[\Delta Z_i(0)] = \frac{1}{d}[z_{i+1}P^*_{i+1,i}(x) + z_{i-1}P^*_{i-1,i}(x) - z_i(1 - P^*_{i,i}(x))] + \mathcal{O}(1/d^2) = \frac{1}{d}\left[\sum_{j=0}^{N} z_j P^*_{j,i}(x) - z_i\right] + \mathcal{O}(1/d^2), \tag{A29}$$

and the second conditional moment as

$$\begin{aligned}&\mathbb{E}_{\mathbf{z}}[(\Delta Z_i(0))^2]\\ &= \frac{1}{d^2}\left(\mathbb{P}_{\mathbf{z}}\left[Z_i(1) = z_i + \frac{1}{d}\right] + \mathbb{P}_{\mathbf{z}}\left[Z_i(1) = z_i - \frac{1}{d}\right]\right)\\ &= \mathcal{O}(1/d^2).\end{aligned} \tag{A30}$$

Combining Eqs. (A29) and (A30), the conditional variance is given by

$$\begin{aligned}\mathrm{Var}_{\mathbf{z}}[\Delta Z_i(0)] &= \mathbb{E}_{\mathbf{z}}[(\Delta Z_i(0))^2] - (\mathbb{E}_{\mathbf{z}}[\Delta Z_i(0)])^2\\ &= \mathcal{O}(1/d^2),\end{aligned} \tag{A31}$$

for $i = 0, 1, \ldots, N$.

6. Derivation of the moments of $\Delta Y_i(0) = Y_i(1) - y_i$

Now, we are ready to derive the first two conditional moments of $\Delta Y_i(0)$. First, combining Eqs. (A24) and (A29), we obtain

$$\begin{aligned}&\mathbb{E}_{\mathbf{z}}[\Delta Y_i(0)]\\ &= \mathbb{E}_{\mathbf{z}}[\Delta Z_i(0)] + \mathbb{E}_{\mathbf{z}}[\Delta v_i(0)]\\ &= \frac{1}{d}\left[\sum_{j=0}^{N}\big(y_j + v_j(x)\big)P^*_{j,i}(x) - y_i - v_i(x)\right] + \mathcal{O}(1/d^2)\\ &= \frac{1}{d}\left[\sum_{j=0}^{N} y_j P^*_{j,i}(x) - y_i\right] + \mathcal{O}(1/d^2).\end{aligned} \tag{A32}$$

In addition, combining Eqs. (A26) and (A31), the conditional variance satisfies

$$\mathrm{Var}_{\mathbf{z}}[\Delta Y_i(0)] \leqslant 2\mathrm{Var}_{\mathbf{z}}[\Delta Z_i(0)] + 2\mathrm{Var}_{\mathbf{z}}[\Delta v_i(0)] = \mathcal{O}(1/d^2). \tag{A33}$$

This completes the proof of the conditions in Eqs. (7d)–(7e).

7. Stability of the zero solution of Eq. (10)

Consider the function

$$\mathbf{c}(x, \mathbf{y})^T = \mathbf{y}^T(\mathbf{P}^*(x) - \mathbf{I}), \tag{A34}$$

where $\mathbf{I}$ denotes the identity matrix of size $N+1$. Notably, when $\mathbf{y} = \mathbf{0}$, the function satisfies $\mathbf{c}(x, \mathbf{0}) = \mathbf{0}$ for any value of x.

Next, examine the differential equation

$$\begin{aligned}\frac{d}{dt}\mathbf{Y}(t, x, \mathbf{y})^T &= \mathbf{c}(x, \mathbf{Y}(t, x, \mathbf{y}))^T\\ &= \mathbf{Y}(t, x, \mathbf{y})^T(\mathbf{P}^*(x) - \mathbf{I}),\end{aligned} \tag{A35}$$

with the initial condition $\mathbf{Y}(0, x, \mathbf{y}) = \mathbf{y} = (y_0, y_1, \ldots, y_N)^T$, where we assume that $\sum_{i=0}^{N} y_i = 0$. This linear system has the explicit solution

$$\mathbf{Y}(t, x, \mathbf{y})^T = \mathbf{y}^T e^{t(\mathbf{P}^*(x) - \mathbf{I})}, \tag{A36}$$

for any $t \geqslant 0$. Here, the matrix $\mathbf{P}^*(x) - \mathbf{I}$ serves as the infinitesimal generator of a continuous-time birth-death process on the state space $\{0, 1, \ldots, N\}$, where the birth rates are $P^*_{i,i+1}(x)$ for $i = 0, 1, \ldots, N-1$, and the death rates $P^*_{i,i-1}(x)$ for $i = 1, \ldots, N$. Furthermore, since $\mathbf{v}(x)$ satisfies $\mathbf{v}(x)^T = \mathbf{v}(x)^T\mathbf{P}^*(x)$, we deduce that $\mathbf{v}(x)$ is the stationary distribution of this process. According to Theorem 3.6.2 in Norris [64], the solution of the differential equation satisfies

$$\begin{aligned}&\lim_{t\to\infty}\mathbf{Y}(t, x, \mathbf{y})\\ &= \left(v_0(x)\sum_{i=0}^{N} y_i, v_1(x)\sum_{i=0}^{N} y_i, \ldots, v_N(x)\sum_{i=0}^{N} y_i\right)^T = \mathbf{0}.\end{aligned}$$

This shows that the zero solution is the globally asymptotically stable equilibrium of the system.

APPENDIX B: PROOF OF EQS. (12)

Our aim in this Appendix is to derive the infinitesimal mean $\mu(x, \mathbf{0})$ and infinitesimal variance $\sigma^2(x, \mathbf{0})$ stated in Theorem 1. To achieve this goal, we need to calculate the first three moments of K, the frequency of C, in a randomly chosen deme in a population subdivided into an infinite number of demes under neutrality, where the overall frequency of cooperators x remains constant. The distribution of deme types is denoted by $\mathbf{v}(x)$.

We begin by considering the first moment of K. Using the fact that $\frac{i}{N}\binom{N}{i} = \binom{N-1}{i-1}$ and the identity in Eq. (A20), we have

$$\begin{aligned}&\sum_{i=0}^{N}\frac{i}{N}\binom{N}{i}\Gamma(a+i)\Gamma(b+N-i)\\ &= \frac{a}{a+b}\frac{\Gamma(a)\Gamma(b)\Gamma(a+b+N)}{\Gamma(a+b)}.\end{aligned} \tag{B1}$$

Using this for $a = \frac{Nmx}{1-m}$ and $b = \frac{Nm(1-x)}{1-m}$, the first moment of K is

$$\mathbb{E}_{\mathbf{v}(\mathbf{x})}[K] = \sum_{i=0}^{N} x_i v_i(x) = \Gamma\left(\frac{Nm}{1-m}\right)\frac{\sum_{i=0}^{N}\frac{i}{N}\binom{N}{i}\Gamma\left(\frac{Nmx}{1-m}+i\right)\Gamma\left(\frac{Nm}{1-m}(1-x)+N-i\right)}{\Gamma\left(\frac{Nmx}{1-m}\right)\Gamma\left(\frac{Nm}{1-m}(1-x)\right)\Gamma\left(\frac{N}{1-m}\right)} = x. \tag{B2}$$

For the second moment of K, note first that

$$\sum_{i=0}^{N} \frac{i^2}{N^2}\binom{N}{i}\Gamma(a+i)\Gamma(b+N-i) = \frac{(b+N+aN)a}{N(a+b+1)(a+b)}\frac{\Gamma(a)\Gamma(b)\Gamma(a+b+N)}{\Gamma(a+b)}, \tag{B3}$$

which is a simple consequence of $\frac{i^2}{N^2}\binom{N}{i} = \frac{N-1}{N}\binom{N-2}{i-2} + \frac{1}{N}\binom{N-1}{i-1}$ and Eq. (A20). Then, we obtain

$$\mathbb{E}_{\mathbf{v}(\mathbf{x})}[K^2] = \sum_{i=0}^{N} x_i^2 v_i(x) = \Gamma\left(\frac{Nm}{1-m}\right)\frac{\sum_{i=0}^{N}\frac{i^2}{N^2}\binom{N}{i}\Gamma\left(\frac{Nmx}{1-m}+i\right)\Gamma\left(\frac{Nm}{1-m}(1-x)+N-i\right)}{\Gamma\left(\frac{Nmx}{1-m}\right)\Gamma\left(\frac{Nm}{1-m}(1-x)\right)\Gamma\left(\frac{N}{1-m}\right)} = \frac{1+mx(N-1)}{1-m+mN}x. \tag{B4}$$

Similarly, we have the identity

$$\sum_{i=0}^{N} \frac{i^3}{N^3}\binom{N}{i}\Gamma(a+i)\Gamma(b+N-i) = \frac{(-ab+b^2+3bN+3abN+2N^2+3aN^2+a^2N^2)a}{N^2(a+b+2)(a+b+1)(a+b)}\frac{\Gamma(a)\Gamma(b)\Gamma(a+b+N)}{\Gamma(a+b)}, \tag{B5}$$

from which the third moment of K can be expressed as

$$\begin{aligned}\mathbb{E}_{\mathbf{v}(\mathbf{x})}[K^3] &= \sum_{i=0}^{N} x_i^3 v_i(x)\\ &= \Gamma\left(\frac{Nm}{1-m}\right)\frac{\sum_{i=0}^{N}\frac{i^3}{N^3}\binom{N}{i}\Gamma\left(\frac{Nmx}{1-m}+i\right)\Gamma\left(\frac{Nm}{1-m}(1-x)+N-i\right)}{\Gamma\left(\frac{Nmx}{1-m}\right)\Gamma\left(\frac{Nm}{1-m}(1-x)\right)\Gamma\left(\frac{N}{1-m}\right)}\\ &= \frac{(N-1)(N-2)m^2x^2+3(N-1)mx+2-m}{(2+m(N-2))(1-m+mN)}x.\end{aligned} \tag{B6}$$

The infinitesimal mean and variance in Theorem 1 are obtained by inserting the first three moments of K in Eqs. (8a)–(8b), which yields

$$\begin{aligned}\mu(x,\mathbf{0}) &= \alpha\sum_{i=0}^{N} v_i(x)x_i(1-x_i)(x_i-x) + \sum_{i=0}^{N} v_i(x)x_i(1-x_i)(\pi_{C,i}-\pi_{D,i})\\ &= (R-T)\mathbb{E}_{\mathbf{v}(x)}\left[K^2(1-K)\right] + (S-P)\mathbb{E}_{\mathbf{v}(x)}\left[K(1-K)^2\right] + \alpha\mathbb{E}_{\mathbf{v}(x)}[K(1-K)(K-x)]\\ &= \left(\frac{2mN}{1-m+mN}\right)\frac{(N-1)x(1-x)}{2N}\left[\frac{1+m(N-2)x}{2+m(N-2)}(R-T) + \frac{1+m(N-2)(1-x)}{2+m(N-2)}(S-P)\right.\\ &\quad \left. + \frac{\alpha}{2+m(N-2)}(1-2x)\right],\end{aligned} \tag{B7}$$

$$\begin{aligned}\sigma^2(x,\mathbf{0}) &= 2(1-m)\sum_{i=0}^{N} v_i(x)x_i(1-x_i) + 2mx(1-x) = 2(1-m)\mathbb{E}_{\mathbf{v}(x)}[K-K^2] + 2mx(1-x)\\ &= \left(\frac{2mN}{1-m+mN}\right)x(1-x).\end{aligned} \tag{B8}$$

See the next Appendix for an interpretation of the term $2mN/(1-m+mN)$.

APPENDIX C: IDENTITY-BY-DESCENT MEASURES

Consider a neutral population subdivided into an infinite number of demes of size N. Here we are interested in the probability $f_{I_1\ldots I_n}$ for n individuals $I_1,\ldots,I_n$ chosen at random without replacement in the same deme to be identical by descent. Tracing the ancestral lines backward in time, a change can occur when an offspring replaced an individual chosen at random in the deme one time step back. Then, three scenarios are possible:

(1) The offspring was a migrant from another deme (probability m) and is one of the individuals among $I_1,\ldots,I_n$ (probability n/N), in which case this individual is not identical by descent to the others.

(2) The offspring was produced in the deme by one of the individuals among $I_1,\ldots,I_n$ and replaced another individual among $I_1,\ldots,I_n$ [probability $(1-m)n(n-1)/N^2$], in which case this individual is identical by descent to one of the others.

(3) Otherwise, the offspring did not change the lines of the individuals $I_1,\ldots,I_n$.

Thus, we have the recurrence equation

$$\begin{aligned}f_{I_1\ldots I_n} &= \frac{mn}{N}\times 0 + \frac{(1-m)n(n-1)}{N^2}\times f_{I_1\ldots I_{n-1}}\\ &\quad + \left(1-\frac{mnN+(1-m)n(n-1)}{N^2}\right)\times f_{I_1\ldots I_n},\end{aligned} \tag{C1}$$

from which we obtain

$$\begin{aligned} f_{I_1\ldots I_n} &= \frac{(n-1)(1-m)}{(n-1)(1-m)+mN} f_{I_1\ldots I_{n-1}} \\ &= \frac{(n-1)!(1-m)^{n-1}}{\prod_{k=1}^{n-1}(k(1-m)+mN)} \\ &= \frac{(n-1)!\Gamma\left(1+\frac{mN}{1-m}\right)}{\Gamma\left(n+\frac{mN}{1-m}\right)}. \end{aligned} \tag{C2}$$

In particular, for two individuals, this probability becomes

$$f_{I_1I_2} = \frac{1-m}{1-m+mN}. \tag{C3}$$

The first expression represents the probability for two ancestral lines in the same deme to coalesce before one of them migrates, denoted by F and called Wright's [49] fixation index (see, e.g., Wakeley [50] and Rousset [51]). Therefore, we have

$$2(1-F) = 2\left(1-f_{I_1I_2}\right) = \frac{2mN}{1-m+mN}. \tag{C4}$$

APPENDIX D: BETA APPROXIMATION FOR $(v_i(x))_{i=0}^N$

Using the Stirling formula

$$\Gamma(z) \sim \sqrt{\frac{2\pi}{z}}\left(\frac{z}{e}\right)^z \tag{D1}$$

for $z \to \infty$, we obtain the approximation

$$\frac{\Gamma(a+z)}{\Gamma(b+z)} \sim z^{a-b}. \tag{D2}$$

Let $i, N \to \infty$ and $m \to 0$ be such that $y = i/N \in (0,1)$ and $\nu = Nm$ remain constant. Using Eq. (D2), we have

$$\begin{aligned} &\binom{N}{i} B\left(\frac{Nmx}{1-m}+i, \frac{Nm(1-x)}{1-m}+N-i\right) \\ &\sim \binom{N}{i} \frac{\Gamma(\nu x+i)\Gamma(\nu(1-x)+N-i)}{\Gamma(\nu+N)} \\ &= \frac{\Gamma(\nu x+i)}{\Gamma(1+i)} \frac{\Gamma(\nu(1-x)+N-i)}{\Gamma(1+N-i)} \frac{\Gamma(1+N)}{\Gamma(\nu+N)} \\ &\sim i^{\nu x-1}(N-i)^{\nu(1-x)-1}N^{1-\nu} \\ &= \frac{1}{N} y^{\nu x-1}(1-y)^{\nu(1-x)-1} \end{aligned} \tag{D3}$$

for $i = 0, 1, \ldots, N$. Combining Eqs. (4) and (D3) shows that the discrete distribution $(v_i(x))_{i=0}^N$ converges to a continuous Beta distribution on $(0,1)$ with density

$$f(y) = \frac{y^{\nu x-1}(1-y)^{\nu(1-x)-1}}{B(\nu x, \nu(1-x))}. \tag{D4}$$

[1] R. Axelrod and W. D. Hamilton, The evolution of cooperation, Science **211**, 1390 (1981).

[2] J. Maynard Smith, *Evolution and the Theory of Games* (Cambridge University Press, Cambridge, 1982).

[3] R. Axelrod, *The Evolution of Cooperation* (Basic Books, New York, 1984).

[4] W. Poundstone, *Prisoner's Dilemma* (Doubleday, New York, 1992).

[5] M. A. Nowak and R. Highfield, *SuperCooperators: Altruism, Evolution, and Why We Need Each Other to Succeed* (Free Press, Los Angeles, CA, 2011).

[6] R. Boyd and P. J. Richerson, An evolutionary model of social learning: The effects of spatial and temporal variation, in *Social Learning: Psychological and Biological Perspectives*, edited by T. R. Zentall and B. G. Galef, Jr. (Lawrence Erlbaum Associates, Inc., Mahwah, NJ, 1988), pp. 29–48.

[7] M. A. Nowak, *Evolutionary Dynamics: Exploring the Equations of Life* (Harvard University Press, Cambridge, MA, 2006).

[8] M. A. Nowak and K. Sigmund, How populations cohere: Five rules for cooperation, Science **314**, 1560 (2006).

[9] D. Kroumi and S. Lessard, The effect of variability in payoffs on average abundance in two-player linear games under symmetric mutation, J. Theor. Biol. **513**, 110569 (2021).

[10] D. Kroumi and S. Lessard, Average abundance of cooperation in multi-player games with random payoffs, J. Math. Biol. **85**, 27 (2022).

[11] D. Kroumi, É. Martin, C. Li, and S. Lessard, Effect of variability in payoffs on conditions for the evolution of cooperation in a small population, Dyn. Games Appl. **11**, 803 (2021).

[12] D. Kroumi, É. Martin, and S. Lessard, Evolution of cooperation with respect to fixation probabilities in multi-player games with random payoffs, Theor. Popul. Biol. **145**, 1 (2022).

[13] M. A. Nowak and R. M. May, Evolutionary games and spatial chaos, Nature (London) **359**, 826 (1992).

[14] W. D. Hamilton, The genetical evolution of social behaviour. I, J. Theor. Biol. **7**, 1 (1964).

[15] D. S. Wilson, A theory of group selection, Proc. Natl. Acad. Sci. USA **72**, 143 (1975).

[16] H. Ohtsuki, C. Hauert, E. Lieberman, and M. A. Nowak, A simple rule for the evolution of cooperation on graphs and social networks, Nature (London) **441**, 502 (2006).

[17] D. Helbing and W. Yu, The outbreak of cooperation among success-driven individuals under noisy conditions, Proc. Natl. Acad. Sci. USA **106**, 3680 (2009).

[18] Z. Yang, Z. Li, and L. Wang, Oscillation and burst transition of human cooperation, Nonlinear Dyn. **108**, 4599 (2022).

[19] Z. Yang and L. Zhang, Random migration with tie retention promotes cooperation in the prisoner's dilemma game, Chaos **33**, 043126 (2023).

[20] L. Lehmann and F. Rousset, How life history and demography promote or inhibit the evolution of helping behaviours, Philos. Trans. R. Soc. B **365**, 2599 (2010).

[21] P. D. Taylor, Altruism in viscous populations - an inclusive fitness model, Evol. Ecol. **6**, 352 (1992).

[22] S. N. Ethier and T. Nagylaki, Diffusion approximations of Markov chains with two time scales and applications to population genetics, Adv. Appl. Probab. **12**, 14 (1980).

[23] T. Nagylaki, The strong-migration limit in geographically structured populations, J. Math. Biol. **9**, 101 (1980).
[24] T. Nagylaki, The diffusion model for migration and selection in a dioecious population, J. Math. Biol. **34**, 334 (1996).
[25] T. Nagylaki, The diffusion model for migration and selection in a plant population, J. Math. Biol. **35**, 409 (1997).
[26] J. L. Cherry, Selection in a subdivided population with dominance or local frequency dependence, Genetics **163**, 1511 (2003).
[27] J. L. Cherry, Selection in a subdivided population with local extinction and recolonization, Genetics **164**, 789 (2003).
[28] J. L. Cherry and J. Wakeley, A diffusion approximation for selection and drift in a subdivided population, Genetics **163**, 421 (2003).
[29] M. C. Whitlock, Fixation probability and time in subdivided populations, Genetics **164**, 767 (2003).
[30] J. Wakeley, Polymorphism and divergence for island-model species, Genetics **163**, 411 (2003).
[31] J. Wakeley and T. Takahashi, The many-demes limit for selection and drift in a subdivided population, Theor. Popul. Biol. **66**, 83 (2004).
[32] S. Lessard, Diffusion approximations for one-locus multi-allele kin selection, mutation and random drift in group-structured populations: A unifying approach to selection models in population genetics, J. Math. Biol. **59**, 659 (2009).
[33] D. Kroumi and S. Lessard, Stochastic viability in an island model with partial dispersal: Approximation by a diffusion process in the limit of a large number of islands, Theor. Popul. Biol. **158**, 170 (2024).
[34] D. Kroumi and S. Lessard, Evolutionary game with stochastic payoffs in a finite island model, J. Dyn. Games **11**, 295 (2024).
[35] G. Hardin, The tragedy of the commons, Science **162**, 1243 (1968).
[36] E. Ostrom, J. Burger, C. B. Field, R. B. Norgaard, and D. Policansky, Revisiting the commons: Local lessons, global challenges, Science **284**, 278 (1999).
[37] D. J. Rankin, K. Bargum, and H. Kokko, The tragedy of the commons in evolutionary biology, Trends Ecol. Evol. **22**, 643 (2007).
[38] A. J. Waite, C. Cannistra, and W. Shou, Defectors can create conditions that rescue cooperation, PLoS Comput. Biol. **11**, e1004645 (2015).
[39] X. Chen, A. Szolnoki, and M. Perc, Risk-driven migration and the collective-risk social dilemma, Phys. Rev. E **86**, 036101 (2012).
[40] L. A. Dugatkin and D. S. Wilson, Rover: A strategy for exploiting cooperators in a patchy environment, Am. Nat. **138**, 687 (1991).
[41] Z. Xiao, X. Chen, and A. Szolnoki, Leaving bads provides better outcome than approaching goods in a social dilemma, New J. Phys. **22**, 023012 (2020).
[42] M. Enquist and O. Leimar, The evolution of cooperation in mobile organisms, Anim. Behav. **45**, 747 (1993).
[43] J. Y. Wakano, M. A. Nowak, and C. Hauert, Spatial dynamics of ecological public goods, Proc. Natl. Acad. Sci. USA **106**, 7910 (2009).
[44] P. Buesser, M. Tomassini, and A. Antonioni, Opportunistic migration in spatial evolutionary games, Phys. Rev. E **88**, 042806 (2013).
[45] T. Wu, F. Fu, Y. Zhang, and L. Wang, Expectation-driven migration promotes cooperation by group interactions, Phys. Rev. E **85**, 066104 (2012).
[46] C. P. Roca and D. Helbing, Emergence of social cohesion in a model society of greedy, mobile individuals, Proc. Natl. Acad. Sci. USA **108**, 11370 (2011).
[47] C. Efferson, C. P. Roca, S. Vogt, and D. Helbing, Sustained cooperation by running away from bad behavior, Evol. Hum. Behav. **37**, 1 (2016).
[48] S. Ethier, A class of degenerate diffusion processes occurring in population genetics, Commun. Pure Appl. Math. **29**, 483 (1976).
[49] S. Wright, Isolation by distance, Genetics **28**, 114 (1943).
[50] J. Wakeley, Segregating sites in Wright's island model, Theor. Popul. Biol. **53**, 166 (1998).
[51] F. Rousset, Inbreeding and relatedness coefficients: What do they measure? Heredity **88**, 371 (2002).
[52] S. Karlin and H. M. Taylor, *A Second Course in Stochastic Processes* (Academic Press, New York, 1981).
[53] M. A. Nowak, A. Sasaki, S. Taylor, and D. Fudenberg, Emergence of cooperation and evolutionary stability in finite populations, Nature (London) **428**, 646 (2004).
[54] B. Y. Zhang, S. J. Fan, C. Li, X. D. Zheng, J. Z. Bao, R. Cressman, and Y. Tao, Opting out against defection leads to stable coexistence with cooperation, Sci. Rep. **6**, 35902 (2016).
[55] V. Křivan and R. Cressman, Defectors' intolerance of others promotes cooperation in the repeated public goods game with opting out, Sci. Rep. **10**, 19511 (2020).
[56] C. Li and S. Lessard, The effect of the opting-out strategy on conditions for selection to favor the evolution of cooperation in a finite population, J. Theor. Biol. **510**, 110543 (2021).
[57] R. Cressman and V. Křivan, Using chemical reaction network theory to show stability of distributional dynamics in game theory, J. Dyn. Games **9**, 351 (2022).
[58] L. Lehmann and L. Keller, The evolution of cooperation and altruism - a general framework and a classification of models, J. Evol. Biol. **19**, 1365 (2006).
[59] O. Leimar and P. Hammerstein, Evolution of cooperation through indirect reciprocity, Proc. R. Soc. London B **268**, 745 (2001).
[60] J. F. Crow and M. Kimura, *An Introduction to Population Genetics Theory* (Harper & Row, New York, 1970).
[61] https://github.com/kroumidhaker/structured-moran-model.
[62] K. Dhaker, "kroumidhaker/structured-moran-model: Initial public release (v1.0.1) (2025)" Zenodo, doi:10.5281/zenodo.15588250.
[63] S. Karlin and H. M. Taylor, *A First Course in Stochastic Processes*, 2nd ed. (Academic Press, New York, 1975).
[64] J. Norris, *Markov Chains* (Cambridge University Press, Cambridge, 1997).