Journal of Physics: Complexity

PAPER • **OPEN ACCESS**

Fixation and fluctuations in two-species cooperation

IOP Publishing J.Phys.Complex. 3 (2022) 015011 (15pp) https://doi.org/10.1088/2632-072X/ac52e7

Journal of Physics: Complexity

RECEIVED
19 September 2021

REVISED
1 February 2022

ACCEPTED FOR PUBLICATION
8 February 2022

PUBLISHED
28 February 2022

PAPER

Fixation and fluctuations in two-species cooperation

Jordi Piñero[1], S Redner[2,*] and Ricard Solé[3,*]

[1] ICREA-Complex Systems Lab, Universitat Pompeu Fabra, 08003 Barcelona, Spain
[2] Institut de Biologia Evolutiva (CSIC-UPF), Psg Maritim Barceloneta, 37, 08003 Barcelona, Spain
[3] Santa Fe Institute, 1399 Hyde Park Road, Santa Fe NM 87501, United States of America
[*] Authors to whom any correspondence should be addressed.

E-mail: jordi.pinero@upf.edu, redner@santafe.edu and ricard.sole@upf.edu

Keywords: fixation, extinction, first passage, stochastic processes, cooperation

Abstract

Cooperative interactions pervade in a broad range of many-body populations, such as ecological communities, social organizations, and economic webs. We investigate the dynamics of a population of two equivalent species A and B that are driven by cooperative and symmetric interactions between these species. For an isolated population, we determine the probability to reach fixation, where only one species remains, as a function of the initial concentrations of the two species, as well as the time to reach fixation. The latter scales exponentially with the population size. When members of each species migrate into the population at rate λ and replace a randomly selected individual, surprisingly rich dynamics ensues. Ostensibly, the population reaches a steady state, but the steady-state population distribution undergoes a unimodal to trimodal transition as the migration rate decreases below a critical value λ_c. In the low-migration regime, $\lambda < \lambda_c$, the steady state is not truly steady, but instead strongly fluctuates between near-fixation states, where the population consists of mostly A's or of mostly B's. The characteristic time scale of these fluctuations diverges as λ^{-1}. Thus in spite of the cooperative interaction, a typical snapshot of the population will contain almost all A's or almost all B's.

1. Introduction

Competitive interactions have played a prominent role in the literature of ecological and evolutionary dynamics, as well as in economics and sociology [1–3]. Resource limitations and their impact in defining the outcome of competition among species has shaped a large part of evolutionary thinking. A counterpoint to competition is cooperativity in which there are positive interactions and feedback loops between species. These mechanisms have received increasing attention recently [4, 5]. In fact, it is the presence of cooperative interactions, where positive reciprocal exchanges are at work, that appear to drive innovations in evolution and also maintain biodiversity in Nature [5].

Cooperation, or mutualism, has been part of mathematical models of populations since the formulation of Lotka–Volterra equations [2] and a variety of statistical physics models of human cooperation [6, 7]. In its most abstract form, two species (for example, A and B in figure 1(a)) 'help' each other by means of a mutual positive interaction; in some cases, both partners completely rely on one another for survival. This feature underlies the two-species system in figure 1(b), where a given species requires the other to replicate because each species needs a molecule that is produced by the partner species.

Recent experimental studies have shown that such cooperative populations can in fact be engineered. By following the scheme in figure 1(b), it is possible to create a completely symmetric pairwise dependence and make these mixed populations grow on a Petri dish [8–10]. Figure 1(c) shows the outcome of symmetric competition when each strain is marked with a different fluorescent protein: each strain locally out competes the other, thereby generating stripes of segregated domains. The cooperatively interacting population, on the other hand, constrains both species to remain in proximity, leading to a well-mixed population (figure 1(d)). These simple engineered, or synthetic, bacterial populations, which can be tuned so that they become virtually

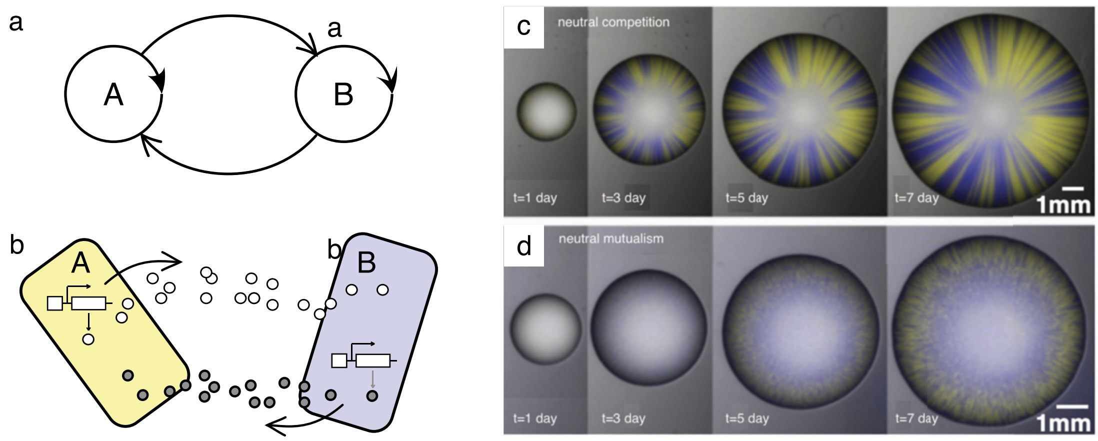

Figure 1. In pairwise cooperation (a) the replication (closed arrows) of a given species A requires the help of B and vice versa. This feedback occurs, for example, when two types of bacteria A and B each lack a metabolite for reproduction that is supplied by the other species (small circles in (b)). Such cooperative feedbacks are commonplace and help maintaining diversity. Experimental set-ups using engineered bacteria (c) and (d) reveal the marked difference between competitive and cooperative interactions. In (c), two equal competitors grow on a Petri dish and locally exclude each other, as shown by the growing bands that indicate the presence of only one strain. With cooperativity (d), the mutual dependency drives both strains to persist and mix.

symmetric, allow one to explore the fundamental dynamical features of interacting living consortia and also study the impact of stochasticity [11–13].

In this paper, we present an analytic approach to understand the role of stochasticity in a simple two-species stochastic model of cooperation. This model represents a special case of evolutionary game dynamics [14–17], with a specific and particularly simple form of the payoff matrix. We emphasize that we are not treating a general ecological model, but rather a simplified system in which only cooperative interactions occur. This scenario appears to be more relevant for the microbiome [18–20]. Current advances in microbial ecology involve experimental setups with a small number of interacting species [21, 22]. In this model, we treat both closed and open populations, in which there is either no migration or a finite rate of migration into the population, respectively. We define microscopic rules that incorporate both cooperativity, in which each species helps the other, as well as neutrality, in which neither species is preferred over the other. We first determine the steady state of the population in the absence of stochastic fluctuations. For a finite population, we then incorporate stochasticity and determine the time until fixation is reached for the situation where no migration can occur. When migration is allowed (with compensatory removal), the population now reaches a steady state; however, the character of this steady state dramatically changes as function of the migration rate. For large migration rate, both species are present in roughly equal abundances. However, for a sufficiently small migration rate, the population strongly fluctuates between consisting of nearly all A or all B. Thus a *typical* realization of the population has a completely different character that the *average* population. This change in behavior is mirrored by a bimodal to trimodal transition in the shape of the steady-state probability distribution of species abundances.

In section 2, we outline basic features of our two-species cooperation model in the absence of migration. We solve the model in the mean-field approximation and then include the role of stochasticity due to the finiteness of the population. For a finite population, we determine the fixation probability as a function of the initial population composition and the time until fixation, where only a single species remains. In section 5, we incorporate migratory inflow, with compensatory removal, so that the population size remains fixed and reaches a steady state. We discuss basic features of this steady state, including the intriguing feature that the species abundances can exhibit huge fluctuations, even though time-averaged properties are constant. We give some concluding remarks in section 6.

2. Two-species cooperation

We investigate a finite population of N particles, with n of species A and $N - n$ of species B. The population undergoes repeated reaction events in which each event consists of the following steps (figure 2):

(a) Pick a random pair of particles.

(b) If the pair is AB, one member of this pair reproduces; if the pair is AA or BB, nothing happens.

(c) The newly reproduced offspring replaces one randomly selected particle in the remainder of the population.

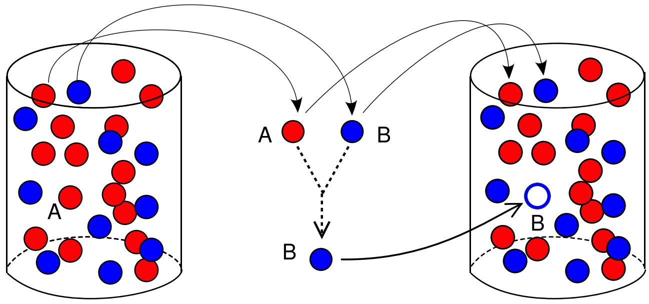

Figure 2. The reaction step in two-species cooperation. Two randomly selected particles happen to be from different species, namely A and B (red and blue). One of them reproduces (B, blue), an event that is aided by the presence of the other (A, red). The offspring replaces another randomly selected particle from the remainder of the population. In the example shown, the newly generated B replaces an A.

Thus interactions between members of different species are cooperative in nature, while members of the same species are non interacting. The replacement step (iii) ensures that the total population remains fixed. The lack of interactions between AA and BB pairs follows from the assumption of strict mutualism, i.e., replication occurs if and only if both species are present (as sketched in figure 2). After each update, time is incremented by $\frac{1}{N}$. This time increment corresponds to each particle undergoing, on average, steps (i)–(iii) in a single time unit. While this dynamics manifestly conserves the total number of particles, the composition of the population can change. When the population consists entirely of a single species—either all A's or all B's—there is no further dynamics and the population fixates.

If an A reproduces in a single interaction, then with probability $1-\frac{n}{N}\equiv 1-x$, the A offspring replaces a B and $n\to n+1$. Conversely, with probability $x=\frac{n}{N}$, the A offspring replaces an existing A so that n does not change. The probability a_n at which $n\to n+1$ therefore is

$$a_n = 2x(1-x)\times\frac{1}{2}\times(1-x) = x(1-x)^2. \tag{1a}$$

Throughout, we use the variables n and $x=\frac{n}{N}$ interchangeably. In (1a), the factor $2x(1-x)$ gives the probability a randomly selected pair is AB, the factor $\frac{1}{2}$ gives the probability that the A in this pair reproduces, and the factor $1-x$ gives the probability that the A offspring replaces a B. By the same reasoning, when a B reproduces in an interaction, the probability at which $n\to n-1$ is

$$b_n = 2x(1-x)\times\frac{1}{2}\times x = x^2(1-x). \tag{1b}$$

Here the trailing factor x in (1b) accounts for the probability that the B offspring replaces an A. Finally, the probability that the number of A's and B's does not change is given by

$$x^2+(1-x)^2+2x(1-x)\left[\frac{1}{2}x+\frac{1}{2}(1-x)\right] = 1-x(1-x) = 1-a_n-b_n. \tag{1c}$$

The terms $x^2+(1-x)^2$ give the probability to pick either an AA or BB pair, for which no change in n occurs. For the last term, $2x(1-x)$ is again the probability of picking an AB pair, while the factor in the square brackets is the probability that the offspring (either A or B with probability $\frac{1}{2}$) replaces its own kind so that n does not change.

3. Mean-field approaches

3.1. Rate equation

Using the probabilities in equation (1), the rate equation for the average number of A's is

$$\dot{n} = N\left(a_n-b_n\right) = Nx(1-x)(1-2x), \tag{2a}$$

or equivalently,

$$\dot{x} = x(1-x)(1-2x). \tag{2b}$$

To keep the notation simple, n and x refer to average values in this section; that is, we do not write angle brackets. This rate equation has a stable fixed point at $x = \frac{1}{2}$ and unstable fixed points at $x = 0$ and $x = 1$. The stability of the fixed point at $x = \frac{1}{2}$ arises because the transition probabilities (1) tend to reduce population imbalances. Thus the steady state in this continuum description is a static population that consists of equal densities of A's and B's. That is, cooperativity promotes diversity in the mean-field description.

The solution to the rate equation (2b) may be straightforwardly obtained by first performing a partial fraction decomposition:

$$\mathrm{d}t = \frac{\mathrm{d}x}{x(1-x)(1-2x)} = \mathrm{d}x\left(\frac{1}{x} - \frac{1}{1-x} + \frac{4}{1-2x}\right),$$

from which

$$t = \int_{x_0}^{x} \mathrm{d}y\left(\frac{1}{y} - \frac{1}{1-y} + \frac{4}{1-2y}\right) = -4\ \ln\ \frac{x(1-x)(1-2x)}{x_0(1-x_0)(1-2x_0)}.$$

We then obtain $x(t)$ by solving the resulting cubic equation. For $t \to \infty$, the limiting behavior is

$$x(t) \simeq \frac{1}{2} - 2x_0(1-x_0)(1-2x_0)\mathrm{e}^{-t/4}, \tag{3}$$

so that the stable fixed point $x^* = \frac{1}{2}$ is approached exponentially quickly in time.

3.2. Master equation and its moments

In the stochastic dynamics where n is a discrete variable, the true fixed points are at $x = 0$ and $x = 1$. Even though the fixed point at $x^* = \frac{1}{2}$ is stable in the continuum limit, stochastic fluctuations allow the population to explore the full state space and eventually get trapped at either $x = 0$ or $x = 1$. This behavior is analogous to the extinction phenomena that arise, for example, in the logistic birth-death process, $A \to 2A$ and $2A \to 0$, as well as other reactions of this genre [23–26]. In these reactions, the rate equation predicts a steady population, N_s, which is determined by the balance between the birth and death rates. However, in the true stochastic dynamics, the population fluctuates around N_s, which actually is the *quasi* steady-state value. Ultimately, a sufficiently large fluctuation occurs that leads to extinction, from which there can be no escape, with an extinction time that scales exponentially in N_s [23–27].

To understand the stochastic dynamics for two-species cooperation, we study $P_n(t)$, the probability that the population consists of n A's and $(N-n)$ B's at time t. The time dependence of this probability distribution is given by

$$\dot{P}_n = N\left[a_{n-1}P_{n-1} + b_{n+1}P_{n+1} - (a_n + b_n)\,P_n\right]. \tag{4}$$

When the number of particles N is small, the set of equation (4) can be straightforwardly solved. For the initial condition of equal numbers of A's and B's, both $P_0(t)$ and $P_N(t)$ approach $\frac{1}{2}$ as $t \to \infty$, while all the other $P_n(t)$ vanish exponentially quickly in time. This direct approach quickly becomes tedious as N increases, however, and to gain insight into the long-time dynamics for general N, it is useful to study low-order moments of P_n. From equation (4) and using a_n and b_n from equation (1), the first moment obeys

$$\begin{aligned}\langle\dot{x}\rangle = \frac{1}{N}\sum_n n\dot{P}_n &= \sum_{1\leqslant n\leqslant N}\left\{na_{n-1}P_{n-1} + nb_{n+1}P_{n+1} - n\,(a_n + b_n)\,P_n\right\}\\ &= \sum_{1\leqslant n\leqslant N}\left\{(n+1)a_nP_n + (n-1)b_nP_n - n\,(a_n+b_n)\,P_n\right\}\\ &= \sum_{1\leqslant n\leqslant N}(a_n - b_n)P_n = \langle x(1-x)(1-2x)\rangle.\end{aligned} \tag{5a}$$

Here we now explicitly write angle brackets to denote average values. Under the assumption of no correlations, that is, $\langle x^k\rangle = \langle x\rangle^k$, (5a) reproduces the rate equation (2b).

Similarly, the equation of motion for the second moment is

$$
\begin{aligned}
\langle \dot{x}^2 \rangle = \frac{1}{N^2}\sum_n n^2 \dot{P}_n &= \frac{1}{N}\sum_{1\leqslant n\leqslant N}\left\{n^2 a_{n-1}P_{n-1} + n^2 b_{n+1}P_{n+1} - n^2\left(a_n+b_n\right)P_n\right\} \\
&= \frac{1}{N}\sum_{1\leqslant n\leqslant N}\left\{(n+1)^2 a_n P_n + (n-1)^2 b_n P_n - n^2\left(a_n+b_n\right)P_n\right\} \\
&= \frac{1}{N}\langle x(1-x)\rangle + 2\langle x^2(1-x)(1-2x)\rangle. \qquad (5b)
\end{aligned}
$$

It is more convenient to express (5a) and (5b) in terms of $z \equiv 2x - 1$, which lies in the range $[-1, 1]$. Doing so, we obtain

$$
\begin{aligned}
\langle \dot{z} \rangle &= -\frac{1}{2}\langle z(1-z^2)\rangle \\
\langle \dot{z}^2 \rangle &= \langle (1-z^2)\left(\frac{1}{N} - z^2\right)\rangle,
\end{aligned} \qquad (6)
$$

which are both symmetric about $z = 0$. If we make the assumption of no correlations, that is, $\langle z^k\rangle = \langle z\rangle^k$, then the first equation reproduces the result that $z = 0$ is a stable fixed point. The second equation then predicts that the width of the distribution initially grows and eventually 'sticks' at the value $\sqrt{N}$. To check this point, we numerically integrated equation (4) for small values of N. From the resulting solution, we find that the width of the probability distribution initially grows with time and later approaches a nearly fixed value that is proportional to $\sqrt{N}$. However, at very long times, there is slow leakage of the probability distribution to the true stochastic fixed points at $z = \pm 1$. Thus the probability distribution eventually approaches two delta-function peaks at these fixed points. This behavior cannot be captured by low-order moment equations, such as (6). Instead, we need to study the full stochastic dynamics; this is done in the following section.

4. Fixation probability and fixation time

We now turn to two quantities of primary interest in the stochastic dynamics, namely, (i) the fixation (or exit) probability E_n, and (ii) the fixation time T_n. The fixation probability E_n is defined as the probability that a population of size N that initially contains n particles of type A reaches the static fixation state of all A's. We use the backward Kolmogorov equation [28, 29] to compute the fixation probability. In this approach, E_n satisfies the recursion

$$
E_n = a_n E_{n+1} + b_n E_{n-1} + \left(1 - a_n - b_n\right)E_n. \qquad (7)
$$

Since the process renews itself after each event, we can express the fixation probability from the state that contains n A's in terms of the appropriately weighted average of the fixation probabilities after a single step to the states $n-1$, n, and $n+1$. The weights are merely the hopping probabilities to these respective states. Equation (7) is subject to the boundary conditions $E_0 = 0$ and $E_N = 1$. The first condition corresponds to the impossibility of reaching a population of all A's if the initial state contains no A's, while the second condition corresponds to the initial state coinciding with the desired final state of all A's.

The solution to (7) is (see appendix A for the calculational details)

$$
E_n = \sum_{m=0}^{n-1}\left[\binom{N-1}{m}\right]^{-1} \Big/ \sum_{m=0}^{N-1}\left[\binom{N-1}{m}\right]^{-1}. \qquad (8)
$$

Neither of these sums has a closed form, but for $N\to\infty$ the denominator approaches 2 [30]. For $N \gg 1$, E_n and its continuum counterpart $E(x)$ (see also appendix A) are nearly independent of $x = \frac{n}{N}$ when x is not close to 0 or 1. Figure 3 shows this dependence of E_n on n. Also shown are the corresponding results from discrete simulations of the fixation process. Simulations are carried out by setting a finite size (N) array of particles with two possible states. Each iteration the particles interact following rules (ii) and (iii) with the interaction rates given by expressions (1a) and (1b). Time is updated by $\Delta t = \left[N\left(a_n + b_n\right)\right]^{-1}$ after each iteration.

The anti-sigmoidal shape of E_n arises from the underlying drift that tends to drive any initial population towards $x = \frac{1}{2}$. Eventually, a large and rare stochastic fluctuation causes the population to escape this effective potential well and reach fixation. This anti-sigmoidal shape also strongly contrasts with the Moran process [31], which is symmetric (neutral), but non-cooperative. Here an AB pair equiprobably converts to either AA or to BB. As a result of this lack of cooperativity, the fixation probability in the strictly neutral Moran process is simply the linear function $E(x) = x$ [29, 31–33].

We now investigate the fixation time T_n, which is defined as the average time for the population of N particles to first reach *either* of the two fixation states, $n = 0$ or $n = N$, when the population initially contains

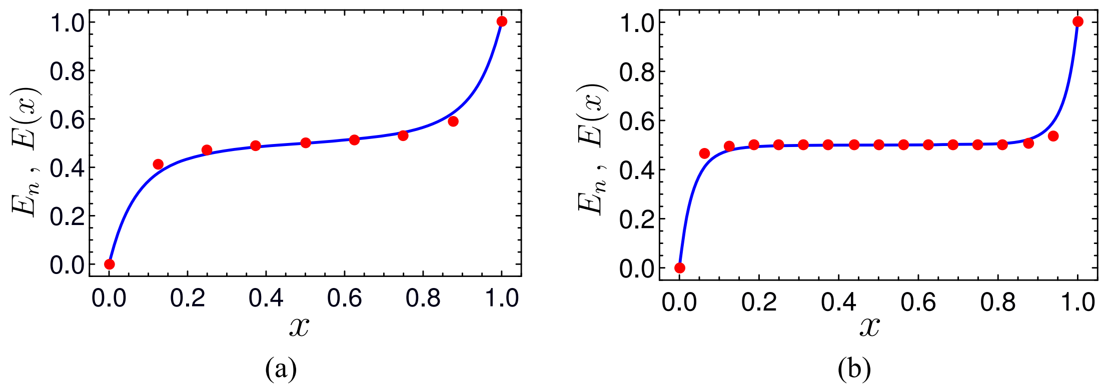

Figure 3. Dependence of the discrete and continuum fixation probabilities, E_n and $E(x)$, versus x for the cases $N = 8$ and 16. The smooth curves represent $E(x)$ from equation (A.5) and the dots represent simulation results.

n A's. Within the same backward Kolmogorov framework as that used for the fixation probability, the fixation time satisfies [28, 29]

$$T_n = a_n T_{n+1} + b_n T_{n-1} + (1 - a_n - b_n)\, T_n + \delta t. \tag{9}$$

Again T_n may be expressed as the appropriately weighted average of the fixation time after a single hopping event to the states $n-1$, n, and $n+1$, plus the time $\delta t = \frac{1}{N}$ required for this single step. The latter corresponds to each particle being updated once, on average, in a single time unit. The equation for T_n is subject to the boundary conditions $T_0 = T_N = 0$; namely, if the population starts in a fixation state, the time to reach this state is 0. The result for the fixation time is (see appendix B)

$$T_n = E_n \sum_{m=1}^{N-1} Q_m - \sum_{m=1}^{n-1} Q_m, \tag{10}$$

where

$$Q_n \equiv \alpha_n + r_n \alpha_{n-1} + r_n r_{n-1}\, \alpha_{n-2} + \cdots + r_n\, r_{n-1} \ldots r_2 \alpha_1,$$

with $Q_0 = 0$, $r_n = b_n/a_n$, and $\alpha_n \equiv \delta t/a_n$.

It does not seem possible to reduce (10) to a compact form, but the main feature of this exact expression is that the fixation time scales exponentially in N and is nearly independent of n (or, equivalently, x), except for n close to 0 or to N (figure 4). The exponential dependence on N again arises because of the existence of an effective potential well, whose depth grows linearly with N, which draws the population toward the state $x = \frac{1}{2}$. The near independence of the fixation time on the initial condition is a consequence of the population being drawn toward the bottom of this potential well, where the concentrations of A and B are equal. As a result, the value of the fixation time for any initial value of x is close to the fixation time when the population starts from the bottom of the potential well at $x = \frac{1}{2}$.

It is possible, however, to obtain an analytical expression for the average fixation time by the WKB method [23–26]. The idea of this approach is that the probability distribution settles into a quasi-steady state that assumes an exponential large-deviation form. From equation (4), there is a slow leakage from this quasi-steady state to the fixation state whose rate, $\Gamma(N)$, is given by

$$\Gamma(N)\delta t = b_1 \widetilde{P}_1 + a_{N-1} \widetilde{P}_{N-1} = 2 b_1 \widetilde{P}_1, \tag{11}$$

i.e., the flux from states that are one step away from fixation to the fixation states. Here the tilde denotes the steady-state distribution and we also use the symmetry $n \leftrightarrow N - n$. We then identify the inverse of this leakage rate with the fixation time.

We obtain an approximate equation for the continuum probability distribution $\widetilde{P}_n \to \widetilde{P}(x)$ by setting the time derivative in the master equation (4) to zero, and writing $n \pm 1$ as $x \pm \delta x$ to give

$$a(x - \delta x)\widetilde{P}(x - \delta x) + b(x + \delta x)\widetilde{P}(x + \delta x) = [a(x) + b(x)]\, \widetilde{P}(x). \tag{12}$$

We now assume that $\widetilde{P}(x)$ has the exponential form $\widetilde{P}(x) \sim \mathrm{e}^{S(x)/\delta x} = \mathrm{e}^{NS_0(x) + S_1(x) + \cdots}$ and substitute this form into (12) to give (up to $O(1)$)

$$S_0(x) = \int^x \mathrm{d}z\, \log\left[\frac{a(z)}{b(z)}\right], \quad S_1(x) = -\frac{1}{2}\log\left[a(x)b(x)\right]. \tag{13}$$

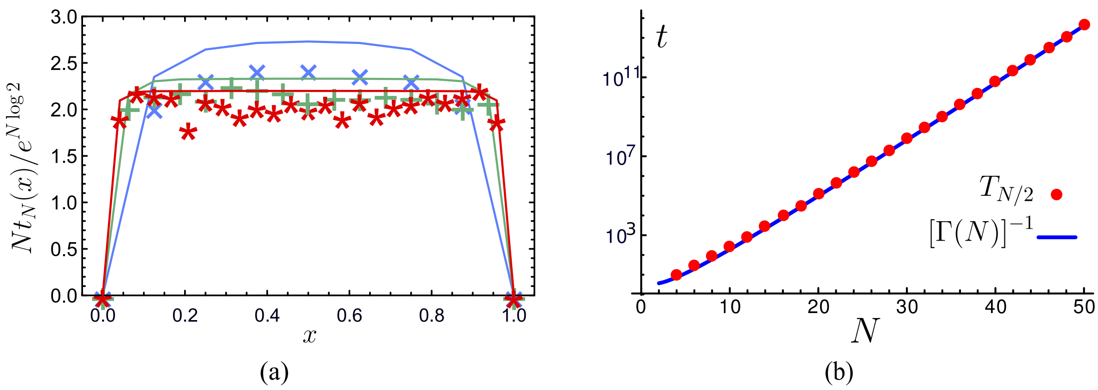

Figure 4. (a) Dependence of the fixation time T_n versus $x = \frac{n}{N}$ using a data-collapse scheme by resetting the scale $T_n \to N\,\mathrm{e}^{-N\log 2}T_n$ for each N value. This scaling factor corresponds to the prediction made in equation (16). The solid lines follow the predictions obtained from equation (10). The plot marks $\{\times, + \text{ and } *\}$ correspond to simulated average fixation times for $N = 8, 16$ and 24 in blue, green and red colors, respectively. Each data point is obtained by averaging 10^3 simulated fixation processes at corresponding values of n and N. (b) Dependence of the fixation time from the symmetric initial state, $T_{N/2}$ (red dots) computed by (10) and the WKB prediction for the inverse leakage rate from the quasi-steady state (blue line), following (16).

Now using (1a) and (1b) for $a(x)$ and $b(x)$, we have

$$\widetilde{P}(x) \sim \frac{\mathrm{e}^{NS_0(x)}}{x^{3/2}(1-x)^{3/2}}, \tag{14}$$

with $S_0(x) = -x\log x - (1-x)\log(1-x)$. Note that the action $S_0(x)$ is peaked at the quasi-state state $x = \frac{1}{2}$. We normalize $\widetilde{P}(x)$ by using the Laplace method for $N \to \infty$ [34],

$$\int_0^1 \frac{\mathrm{e}^{NS_0(x)}}{x^{3/2}(1-x)^{3/2}}\,\mathrm{d}x \approx \sqrt{\frac{32}{N}}\,\mathrm{e}^{N\,\log\,2},$$

so that

$$\widetilde{P}(x) \simeq \frac{N}{\sqrt{32\pi}}\,\frac{\mathrm{e}^{N\left[-x\,\log\,x-(1-x)\log(1-x)-\log\,2\right]}}{x^{3/2}(1-x)^{3/2}}. \tag{15}$$

We now compute the fixation rate Γ by substituting

$$\widetilde{P}\left(\frac{1}{N}\right) \simeq \frac{N^3}{\sqrt{32\pi}}\,\mathrm{e}^{1-N\,\log\,2} \quad \text{and} \quad b\left(\frac{1}{N}\right) \simeq \frac{1}{N^2},$$

into equation (11) to give

$$\Gamma(N) = \frac{N\,\mathrm{e}}{\sqrt{8\pi}}\,\mathrm{e}^{-N\,\log\,2}. \tag{16}$$

We now identify the inverse of this rate with the average fixation time. As shown in figure 4(b), this inverse rate accurately matches the simulation data for the fixation time.

5. Two-species cooperation with migration

We now incorporate migration into the dynamics, in which particles of either species migrate into the population at the same fixed rate λ, and each new particle replaces a randomly selected existing particle. Because migration is accompanied by replacement, the population size remains fixed, which is the physically most relevant case. Now the population is driven to a steady state rather than to fixation and we want to understand the nature of this steady state.

5.1. Probability distribution

For a population that consists of n A's and $(N-n)$ B's, suppose that the migrant is an A. With probability $\frac{1}{2}(1-x)$, the A migrant replaces a B and $n \to n+1$, while with probability $\frac{1}{2}x$, the A migrant replaces an A, and the composition of the population remains the same. Similar reasoning applies when the migrant is a B. As a result of a migration event, the average change in the number of A's is $\frac{1}{2}(1-x) - \frac{1}{2}x$. The rate equation for n now is (compare with equation (2))

$$\langle \dot{n} \rangle = N(1-\lambda)\left[x(1-x)(1-2x)\right] + \frac{1}{2}N\lambda(1-2x). \tag{17}$$

For $\lambda > 0$, $x = 0$ and $x = 1$ are no longer fixed points and only the remaining fixed point at $x = \frac{1}{2}$ is stable. In the absence of fluctuations, the population is thus driven to a steady-state distribution, $P_n(t \to \infty)$, that is peaked about $x = \frac{1}{2}$. Because there is no absorbing state in the stochastic dynamics, we might anticipate a similar behavior for $P_n(t \to \infty)$ when stochasticity is accounted for. We will show, however, that within the Fokker–Planck approximation the steady-state distribution can either be unimodal or trimodal in shape and the latter case corresponds to a steady state that is not truly steady.

The probability distribution P_n is now governed by the master equation

$$\begin{aligned}\dot{P}_n &= N(1-\lambda)\left[a_{n-1}P_{n-1} + b_{n+1}P_{n+1} - (a_n + b_n)\,P_n\right]\\ &\quad + N\lambda\left[c_{n-1}P_{n-1} + d_{n+1}P_{n+1} - (c_n + d_n)\,P_n\right],\end{aligned} \tag{18}$$

with hopping probabilities due to migration that are given by

$$c_n = \frac{1}{2}\left(1 - \frac{n}{N}\right) \qquad d_n = \frac{1}{2}\frac{n}{N}. \tag{19}$$

We now determine the continuum probability distribution in the Fokker–Planck approximation. As we shall see, this continuum expression for the probability distribution matches simulation data quite well, thus justifying the Fokker–Planck approximation *ex post facto* as a way to probe steady-state properties.

In terms of $x = \frac{n}{N}$, $\mathrm{d}x = \frac{1}{N}$, $P_n \to P(x)$, we expand (18) in a Taylor series up to second order. This gives the Fokker–Planck equation [28, 35]

$$\begin{aligned}P_t &= -\left\{(1-2x)\left[(1-\lambda)x(1-x) + \frac{\lambda}{2}\right]P(x,t)\right\}_x + \frac{1}{2N}\left\{\left[(1-\lambda)x(1-x) + \frac{\lambda}{2}\right]P(x,t)\right\}_{xx}\\ &\equiv -\{v(x)P(x,t)\}_x + \{D(x)P(x,t)\}_{xx},\end{aligned} \tag{20}$$

where the subscripts denote partial derivatives.

The steady state is defined by solving this equation with the left-hand side set to zero. Integrating once gives $(DP)_x - vP = B$, where B is a constant. We determine the constant by evaluating this equation at the symmetry point $x = \frac{1}{2}$. Because the probability distribution is symmetric about $x = \frac{1}{2}$, $P_x(x = \frac{1}{2}) = 0$. Moreover, at $x = \frac{1}{2}$, $v = 0$ and $D_x = 0$, which implies that $B = 0$. Thus we only need to solve $(DP)_x - vP = 0$, whose solution is

$$\begin{aligned}P(x) &= C\,\exp\left\{\int^x \mathrm{d}y\,\frac{v(y) - D_y(y)}{D(y)}\right\} = C\,\exp\left\{-\log D(x) + \int^x \mathrm{d}y\,\frac{v(y)}{D(y)}\right\}\\ &= \frac{C}{D(x)}\,\exp\left\{\int^x \mathrm{d}y\;2N(1-2y)\right\}\\ &= C'\left[\frac{1}{(1-\lambda)x(1-x) + \frac{\lambda}{2}}\right]\mathrm{e}^{2Nx(1-x)},\end{aligned} \tag{21}$$

where the constant C' is determined by normalization.

For $\lambda \to 0$, $P(x)$ is concentrated near $x = 0$ and near $x = 1$; these peaks correspond to the near-fixation states. Because of rare fluctuations, however, the population stochastically switches between states where almost all particles are of type A to states where almost all particles are of type B. Naively, one therefore anticipates that the steady-state distribution should be bimodal, with a peak at each of the two near-fixation states. Unexpectedly, there always remains a peak at $x = \frac{1}{2}$ (which may be vanishingly small), so that the this distribution is *trimodal* in the small-λ regime. As λ increases beyond a critical value, the steady-state distribution undergoes a trimodal to unimodal transition (figure 5).

For fixed N, we determine the transition between trimodality and unimodality by finding the point(s) where $P'(x) = 0$. This calculation gives, after straightforward algebra,

$$P'(x) \propto (1-2x)\,\mathrm{e}^{2Nx(1-x)}\left[2N - \frac{(\lambda - 1)}{D(x)^2}\right],$$

where again $D(x)$ is the diffusion coefficient defined by equation (20). The leading factor of $1 - 2x$ equals 0 at $x = \frac{1}{2}$ and corresponds to the extremum in the distribution at $x = \frac{1}{2}$.

However, there are additional extrema at the points where the factor in the square brackets equals 0. To determine these extrema, we first determine the zero of this factor at $x = 0, 1$. Thus we have the condition $2N = (1-\lambda)/(\lambda/2)^2$. Since we will find that $\lambda \ll 1$, we also neglect λ compared to 1 to give

$$\lambda_c = \sqrt{\frac{2}{N}}. \tag{22}$$

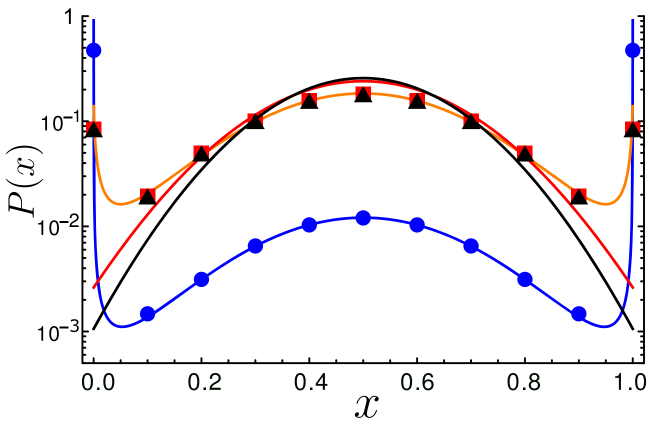

Figure 5. Steady-state probability distributions for $N = 10$ on a semi-logarithmic scale for values of $\lambda/\lambda_c = 10^{-4}, 10^{-2}, 10^{0}$ and 10^{1} in blue, orange, red and black, respectively. The respective data points correspond to simulation results.

For $\lambda < \lambda_c$, the distribution $P(x)$ has three extrema. To find the location of the two secondary extrema for $\lambda \lesssim \lambda_c$, it is convenient to now use the variable $y = x - \frac{1}{2}$ with $y \to \pm\frac{1}{2}$, which corresponds to x close to zero or to 1. Now the condition that the factor in the square brackets equals zero gives

$$2N \simeq \frac{1}{(2\epsilon + \lambda/2)^2},$$

from which we obtain

$$\epsilon \simeq \frac{1}{4}(\lambda_c - \lambda) \quad \text{for } \lambda < \lambda_c. \tag{23}$$

In the regime $\lambda < \lambda_c$, the distribution $P(x)$ is necessarily trimodal because the point $x = \frac{1}{2}$ is always a local maximum of $P(x)$. To verify this point, we compute the second derivative of $P(x)$ at $x = \frac{1}{2}$,

$$P''(x) = \mathrm{e}^{2Nx(1-x)}\left[(1-2x)^2 - 2 + \frac{2(1-\lambda)}{D(x)^3}D'(x)\right],$$

which is indeed negative at $x = \frac{1}{2}$.

5.2. Macroscopic fluctuations in the steady state

An intriguing feature of two-species cooperation with migration is that the steady state is not strictly steady, especially when λ is small (figure 6). For $\lambda \ll \lambda_c$, substantial time ranges exist during which little or no immigration occurs. During these periods, the population will tend to approach one of the fixation states. Even if the population does reach a state of all A's or all B's, immigration eventually drives the population towards an equal number of A's and B's. The competing effects of cooperation and immigration therefore cause the population to wander stochastically from one near-fixation state to the other, with returns to the equal-concentration point $x = \frac{1}{2}$ controlled by the immigration rate (figure 6). Related phenomenology occurs in noisy voter models [36–39]. The extended time periods during which the population is close to fixation corresponds to the large weight in the secondary peaks of the probability distribution in figure 5. In contrast, when the immigration rate is much larger than λ_c, the rapid inflow of equal numbers of A's and B's ensures that the population contains roughly equal numbers of both species.

One way to quantify these fluctuations is by the *return time RT*, which we define as the time interval between successive points where $x = \frac{1}{2}$. As suggested by figure 6, this return time will be short for $\lambda \gg \lambda_c$ and become long for $\lambda \ll \lambda_c$. The latter behavior is a harbinger of large composition fluctuations in the population. We can again use the backward Kolmogorov approach to determine the λ dependence of the return time. Let τ_n denote the average time to reach the balanced state of equal numbers of A's and B's when starting from a state where the number of A's equals $n > \frac{1}{2}N$. Then

$$RT = \frac{1}{N} + \tau_{1+N/2}. \tag{24}$$

That is, starting from the balanced state, the average return time to this state is the average time for a single event, $\frac{1}{N}$, where the population now consists of $\frac{N}{2} + 1$ A's and $\frac{N}{2} - 1$ B's, plus the average time to reach the balanced state when starting from this minimally imbalanced state.

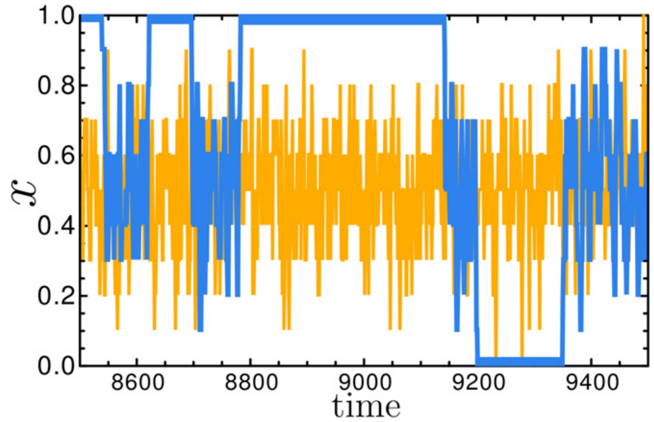

Figure 6. Typical population trajectories in composition space for $\lambda/\lambda_c = 2 \times 10^{-3}$ (blue) and $\lambda/\lambda_c = 2$ (orange) for the case $N = 10$.

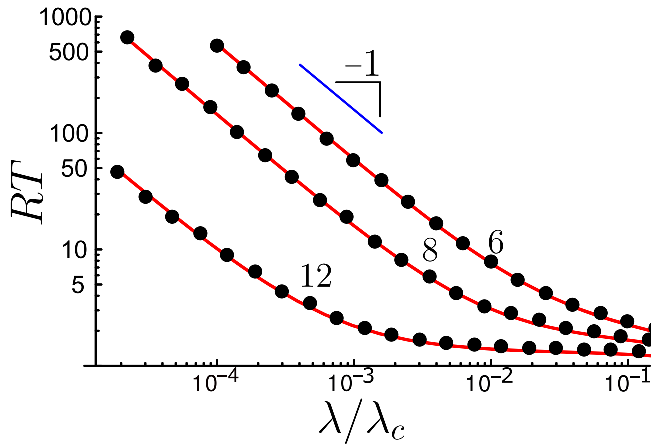

Figure 7. Dependence of the return time RT on λ/λ_c for various values of N.

In analogy with equation (9) for the fixation time, the time τ_n satisfies the recursion

$$\tau_n = (1-\lambda)\left[a_n\tau_{n+1} + b_n\tau_{n-1} + (1-a_n-b_n)\,\tau_n\right] \\ + \lambda\left[c_n\tau_{n+1} + d_n\tau_{n-1} + (1-c_n-d_n)\,\tau_n\right] + \delta t, \tag{25a}$$

with a_n, b_n given by (1a) and (1b), and c_n, d_n given by (19). We can rewrite this recursion in the canonical form of equation (9):

$$\tau_n = a'_n\tau_{n+1} + b'_n\tau_{n-1} + (1-a'_n-b'_n)\tau_n + \delta t, \tag{25b}$$

where $a'_n = (1-\lambda)a_n + \lambda c_n$ and $b'_n = (1-\lambda)b_n + \lambda d_n$. This recursion is valid for $\frac{1}{2}N < n < N$, while for $n = N$ the first term in (25b) is absent. This missing term acts as the effective reflecting boundary condition for $n = N$. At $n = \frac{1}{2}N$, the appropriate boundary condition is $\tau_{N/2} = 0$; namely, the balanced state corresponds to the end of the process. The result for τ_n for arbitrary n is given in (D.5), while the return time RT itself is given in (D.4) (see appendix D for details). Figure 7 shows that RT scales as λ^{-1} for $\lambda \to 0$. This behavior is the source of the long-lived temporal fluctuations in the composition of the population, as illustrated in figure 6.

6. Discussion

Much of the literature on populations of multiple cooperating species has focused on the continuous limit. Such a population is driven to an attractor state in which there are equal concentrations of each species. However, when finite-population stochasticity is incorporated, the true attractors of the dynamics of an isolated population are instead the fixation states, in which only one species exists. Because stochastic effects are relevant in real systems, a discrete approach that incorporates this stochasticity is necessary to describe the dynamics in a faithful way.

Within the discrete approach, we determined the probability for a finite population to reach a given fixation state as a function of the initial condition, as well as the time to reach fixation. The behaviors of these two quantities reflect the effective bias that drives the system to the quasi-steady state of equal concentrations of the two species. Namely, the fixation probability is nearly independent of the initial condition and the fixation time scales exponentially with population size N. As a consequence of this exponential dependence, fixation is not observable in a laboratory or in a simulational time scale for any reasonable population size.

It is worth mentioning that the statistical features of two-species cooperation share similarities with the vacillating voter model [40], despite their different microscopic update rules. In the vacillating voter model, agents (voters) can be in one of two voting states and their agreement or disagreement is influenced by the state of yet another neighbor. Here 'vacillation' refers to the possibility that a voter does not adopt the state of a randomly chosen neighbor, as in the standard voter model, but may adopt the state of this additional neighbor. The properties of this decision process drives the population toward a zero-magnetization state. This state is equivalent to the attractor in two-species cooperation, where both species are equally represented.

When migration into the system can also occur, the population now ostensibly reaches a steady state. An unanticipated feature of this steady state for small migration rate λ is that this state is not genuinely steady, because there are long-term stochastic fluctuations that drive the population from one near-fixation state to the opposite near-fixation state, with the population spending long time periods in these near-fixation states. Such macroscopic wanderings are reflected in the steady-state abundance distribution, which is strongly peaked at these near-fixation states for sufficiently small λ. The time scale associated with these fluctuations increases rapidly as the migration rate decreases. Thus observations of a cooperative system have to be sufficiently long to incorporate many such wanderings so as to ensure that true average behavior is actually probed. We found that the return time—the time between successive instants where the number of A's and B's are equal—allows us to quantify these temporal fluctuations in a precise way.

This observation of large fluctuations also has important ramifications for populations that consist of more than two cooperating species. Depending the immigration rate, the population size N, and the number of distinct species S, the number of species that are actually present at any given time could be much less than S. Thus a *typical* steady state could have a very different character than the *average* steady state that is predicted by the time-independent density distribution. Moreover, a multispecies population will also exhibit large fluctuations in the actual composition of the species that are present. This intriguing issue has recently been investigated in the context of multispecies Lotka–Volterra models [41, 42]. The backward Kolmogorov equation offers the possibility of obtaining new insights into these large fluctuations because of the relative simplicity of neutral models of cooperating species.

The predictions presented in our study can also be tested in experimental settings that are based on microfluidic chambers, where small populations of cells can be maintained so that long-term monitoring can be performed. These small-scale devices offer a unique opportunity to explore the impact of population size and validate the approximations presented in this work. Both well-mixed populations, as well as two- and there-dimensional spatial populations, could be maintained in constant numbers over time. Future extensions of our model, such as including cell death or environmental noise, would be helpful to design experimental protocols and explore the conditions required to maintain stable cooperative cell assemblies [43, 44].

Acknowledgments

The authors thank Deepak Bhat and Jacopo Grilli for fruitful discussions, and Eric A Blair for inspiring ideas. JP and RS were supported by the Botín Foundation and the Spanish Ministry of Economy and Competitiveness through Grant FIS2015-67616-P MINEICO/AEI/FEDER. JP is also supported by 'María de Maeztú' fellowship MDM-2014-0370-17-2. SRs research was supported in part by NSF Grant DMR-1910736. JP and RS thank the hospitality of the Santa Fe Institute, where this project began.

Appendix A. The fixation probability

We want to solve equation (7) for the fixation probability:

$$E_n = a_n E_{n+1} + b_n E_{n-1} + (1 - a_n - b_n)\, E_n.$$

This calculation is standard (see, e.g., [45]) and we provide it here so that our presentation is self contained. It is convenient to rewrite the above equation as $a_n(E_{n+1} - E_n) = b_n(E_n - E_{n-1})$, and then define $u_n \equiv E_n - E_{n-1}$

and $r_n \equiv b_n/a_n$ to recast it as the following first-order recursion for the u_n:

$$u_n = r_{n-1}u_{n-1} = r_{n-1}r_{n-2}r_{n-2}\dots r_1 u_1,$$

with $u_1 = E_1 - E_0 = E_1$. We now define $R_n = \prod_{m=1}^{n} r_m$ so that the equation for u_n becomes

$$u_n = R_{n-1}u_1 = R_{n-1}E_1. \tag{A.1}$$

Since the u_n are successive differences of the E_n, we determine E_n by summing the u_n. Thus

$$E_n = \sum_{m=1}^{n} u_m = \sum_{m=1}^{n} R_{m-1}E_1 = \sum_{m=0}^{n-1} R_m E_1, \tag{A.2}$$

where we need to define $R_0 = 1$ for consistency. We now determine the unknown E_1 by using the boundary condition $E_N = 1$ in (A.2) to give $E_N = \sum_{m=0}^{N-1} R_m E_1 = 1$. Having found E_1, the general solution for E_n is

$$E_n = \sum_{m=0}^{n-1} R_m \Bigg/ \sum_{m=0}^{N-1} R_m. \tag{A.3}$$

To simplify the above expression, we start with a_n and b_n defined in equation (1):

$$a_n = \frac{n}{N}\left(1 - \frac{n}{N}\right)^2 \qquad b_n = \left(\frac{n}{N}\right)^2\left(1 - \frac{n}{N}\right),$$

so that $r_n = b_n/a_n = n/(N-n)$. Then

$$R_n = r_1 r_2 \dots r_n = \frac{n!(N-n-1)!}{(N-1)!} = \left[\binom{N-1}{n}\right]^{-1}.$$

Substituting this representation for R_n in (A.3) gives equation (8).

For completeness, we also give the continuum solution for the fixation probability. We take the continuum limit of equation (7) by letting $(n \pm 1)/N \to x \pm \mathrm{d}x$, with $\mathrm{d}x = \frac{1}{N}$, and then expanding this equation to second order in $\mathrm{d}x$. This gives

$$E'' + 2N(1-2x)E' = 0, \tag{A.4}$$

where the prime denotes differentiation with respect to x. This equation is subject to the boundary conditions $E(0) = 0$ and $E(1) = 1$. As in the discrete formulation, the first condition corresponds to the impossibility of reaching a population of all A's if the initial state contains no A's, while the second condition corresponds to the initial state coinciding with the desired final state. The solution to (A.4), subject to the given boundary conditions is

$$E(x) = \frac{\int_0^x \mathrm{d}u\, \mathrm{e}^{2N(u^2-u)}}{\int_0^1 \mathrm{d}u\, \mathrm{e}^{2N(u^2-u)}} = \frac{1}{2}\left[1 + \frac{\mathrm{erfi}\left(\sqrt{2N}\left(x - \frac{1}{2}\right)\right)}{\mathrm{erfi}(\sqrt{N/2})}\right], \tag{A.5}$$

where erfi is the imaginary error function. This expression from the continuum approximation agrees well with the exact discrete result (8) (see figure 3). However, the continuum approach is no longer accurate for the fixation time (see the next section).

Appendix B. The fixation time

We now solve the recursions (9) for the fixation time:

$$T_n = a_n T_{n+1} + b_n T_{n-1} + (1 - a_n - b_n)\, T_n + \delta t.$$

Following the same steps that led to (A.1), we obtain, for the difference $v_n \equiv T_n - T_{n-1}$,

$$v_n = r_{n-1}v_{n-1} - \alpha_{n-1}, \tag{B.1}$$

where $\alpha_n \equiv \delta t/a_n$. Notice that $v_1 = T_1 - T_0 = T_1 \equiv R_0 T_1$.

We develop the recursion (B.1) to give

$$\begin{aligned} v_n &= r_{n-1}r_{n-2}v_{n-2} - \alpha_{n-1} - r_{n-1}\alpha_{n-2} \\ &= r_{n-1}r_{n-2}r_{n-3}v_{n-3} - \alpha_{n-1} - r_{n-1}\alpha_{n-2} - r_{n-1}r_{n-2}\alpha_{n-3} \\ &\quad\vdots \end{aligned}$$

continuing this development to the end leads to

$$v_n = R_{n-1}v_1 - Q_{n-1} = R_{n-1}T_1 - Q_{n-1}, \tag{B.2}$$

where

$$Q_n \equiv \alpha_n + r_n\alpha_{n-1} + r_n r_{n-1}\alpha_{n-2} + \cdots + r_n r_{n-1}\ldots r_2\alpha_1,$$

and $Q_0 = 0$ by virtue of (B.2).

Finally, we sum the v_n to obtain the fixation time:

$$T_n = \sum_{m=1}^{n} v_m = \sum_{m=0}^{n-1} R_m T_1 - \sum_{m=1}^{n-1} Q_m. \tag{B.3}$$

To eliminate the unknown T_1, we use the boundary condition $T_N = 0$ to give

$$T_1 = \sum_{m=1}^{N-1} Q_m \Big/ \sum_{m=0}^{N-1} R_m.$$

Substituting this in (B.3) and noting that $\sum_{m=0}^{n-1} R_m / \sum_{m=0}^{N-1} R_m$ is just the fixation probability E_n, we obtain the result quoted in equation (10).

Appendix C. WKB approximation

A comprehensive review of the application of the WKB method for large deviations in stochastic populations can be found in [26]. In this section we summarize the basic steps to reach equation (13). We make the ansatz $\tilde{P}(x) \sim \exp\left\{NS_0(x) + S_1(x) + O\left(N^{-1}\right)\right\}$, and expand to linear order in δx to give

$$\tilde{P}(x \pm \delta x) \simeq \tilde{P}(x)\exp\left\{\pm S_0' + \delta x\left(\frac{S_0''}{2} \pm S_1'\right)\right\} \simeq \tilde{P}(x)\mathrm{e}^{\pm S_0'}\left[1 + \delta x\left(\frac{S_0''}{2} \pm S_1'\right)\right],$$

with $\delta x = 1/N$. We now define $\Lambda \equiv \mathrm{e}^{S_0'}$ and substitute this into (12) to obtain

$$\left(a - a'\delta x\right)\Lambda^{-1}\left[1 + \delta x\left(\frac{S_0''}{2} - S_1'\right)\right] + \left(b + b'\delta x\right)\Lambda\left[1 + \delta x\left(\frac{S_0''}{2} + S_1'\right)\right] = a + b,$$

which can be separated into terms of $O(1)$ and terms of $O(\delta x)$. For the former, we have

$$a\Lambda^{-1} + b\Lambda = a + b, \tag{C.1}$$

which has the two solutions $\Lambda_0 = 1$ and $\Lambda = \frac{a}{b}$. The Λ_0 solution corresponds to $S_0' = 0$, and the resulting constant can be absorbed by the normalization condition on $\tilde{P}(x)$. The second solution is

$$S_0' = \log\left(\frac{a}{b}\right), \tag{C.2}$$

which, after integration, results in the first expression in (13). For the $O(\delta x)$ terms we must solve

$$-a'\Lambda^{-1} + a\Lambda^{-1}\left(\frac{S_0''}{2} - S_1'\right) + b'\Lambda + b\Lambda\left(\frac{S_0''}{2} + S_1'\right) = 0, \tag{C.3}$$

which, after substitution of $\Lambda = \frac{a}{b}$, yields

$$b\left(\frac{S_0''}{2} - S_1'\right) + a\left(\frac{S_0''}{2} + S_1'\right) = b\frac{a'}{a} - a\frac{b'}{b}. \tag{C.4}$$

On the other hand, $S_0'' = \frac{a'}{a} - \frac{b'}{b}$, which reduces the previous equation to

$$(a-b)S_1' = \frac{1}{2}\left[b\frac{a'}{a} - a\frac{b'}{b} - (a' - b')\right] = -\frac{1}{2}\left(\frac{a'}{a} + \frac{b'}{b}\right)(a-b). \tag{C.5}$$

The first term in the brackets on the right-hand side corresponds to the derivative of $\log(ab)$. Hence, after integration, we obtain the second expression in (13).

Appendix D. The return time

We want to solve the recursion for the return time τ_n, defined as the time for the population to first reach the state with equal numbers of A's and B's when the initial state contains $n > \frac{1}{2}N$ A's. The state with $n = N$ plays the role of an effective reflecting boundary condition. The system of equations that we wish to solve is (25b):

$$\tau_n = a'_n \tau_{n+1} + b'_n \tau_{n-1} + (1 - a'_n - b'_n)\tau_n + \delta t, \tag{D.1a}$$

for $1 + N/2 \leqslant n < N$, while for $n = N$ the appropriate equation is

$$\tau_N = b'_B \tau_{N-1} + (1 - a'_N - b'_N)\tau_N + \delta t. \tag{D.1b}$$

Using the fact that $a'_N = 0$, this last equation can be rewritten as

$$\tau_N - \tau_{N-1} \equiv v_N = \frac{\delta t}{b'_N}. \tag{D.2}$$

The remaining $\frac{N}{2} - 1$ equations (D.1a) are of the same type as (9) for the fixation time. Thus the solution for τ_n has the same form as (B.3):

$$\tau_n = \sum_{m=N/2}^{n-1} R'_m \tau_{1+N/2} - \sum_{m=1+N/2}^{n-1} Q'_m, \tag{D.3}$$

where

$$R'_n = \prod_{m=1+N/2}^{n} \frac{b'_n}{a'_n} \equiv \prod_{m=1+N/2}^{n} r'_n$$

and

$$Q'_n = \alpha'_n + r'_n \alpha'_{n-1} + r'_n r'_{n-1} \alpha'_{n-2} + \cdots + r'_n r'_{n-1} \ldots r'_{2+N/2} \alpha'_{1+N/2},$$

with $\alpha'_n = \delta t / a'_n$ and $Q'_{N/2} = 0$.

To eliminate the unknown $\tau_{1+N/2}$, we now write (D.3) for the special cases of $n = N$ and $n = N - 1$:

$$\tau_N = \sum_{m=N/2}^{N-1} R'_m \tau_{1+N/2} - \sum_{m=1+N/2}^{N-1} Q'_m \qquad \tau_{N-1} = \sum_{m=N/2}^{N-2} R'_m \tau_{1+N/2} - \sum_{m=1+N/2}^{N-2} Q'_m.$$

Their difference is

$$\tau_N - \tau_{N-1} = R'_{N-1} \tau_{1+N/2} - Q'_{N-1} = \frac{\delta t}{b'_N},$$

so that $\tau_{1+N/2}$ is given by

$$\tau_{1+N/2} = \frac{Q'_{N-1} + \delta t / b'_N}{R'_{N-1}}. \tag{D.4}$$

Substituting this expression for $\tau_{1+N/2}$ in (D.3) gives the average time to reach the balanced state of equal numbers of A's and B's when starting from a population that contains n A's with a reflecting boundary condition at $n = N$:

$$\tau_n = \sum_{m=N/2}^{n-1} R'_m \frac{Q'_{N-1} + \delta t / b'_N}{R'_{N-1}} - \sum_{m=1+N/2}^{n-1} Q'_m. \tag{D.5}$$

What we want, however, is, the return time, defined as the average time to start at the balanced state and first return to this state. This is $RT = \frac{1}{N} + \tau_{1+N/2}$.

ORCID iDs

Jordi Piñero https://orcid.org/0000-0002-4183-3733
S Redner https://orcid.org/0000-0001-5906-4904
Ricard Solé https://orcid.org/0000-0001-6974-1008

References

[1] Goel N S, Maitra S C and Montroll E W 1971 *Rev. Mod. Phys.* **43** 231
[2] Murray J D 2001 *Mathematical Biology: I. An Introduction* (New York: Springer)
[3] May R M 2001 *Stability and Complexity in Model Ecosystems* (Princeton, NJ: Princeton University Press)

[4] May R M, Levin S A and Sugihara G 2008 *Nature* **451** 893–4
[5] Bronstein J L 2015 *Mutualism* (Oxford: Oxford University Press)
[6] Perc M 2016 *Phys. Lett.* A **380** 2803–8
[7] Perc M, Jordan J J, Rand D G, Wang Z, Boccaletti S and Szolnoki A 2017 *Phys. Rep.* **687** 1–51
[8] Mitri S, Clarke E and Foster K R 2016 *ISME J.* **10** 1471–82
[9] Nadell C D, Drescher K and Foster K R 2016 *Nat. Rev. Microbiol.* **14** 589–600
[10] Müller M J I, Neugeboren B I, Nelson D R and Murray A W 2014 *Proc. Natl Acad. Sci.* **111** 1037–42
[11] Shou W, Ram S and Vilar J M G 2007 *Proc. Natl Acad. Sci.* **104** 1877–82
[12] Amor D R, Montañez R, Duran-Nebreda S and Solé R 2017 *PLoS Comput. Biol.* **13** e1005689
[13] Amor D R and Dal Bello M 2019 *Life* **9** 22
[14] Nowak M A 2006 *Evolutionary Dynamics: Exploring the Equations of Life* (Cambridge, MA: Harvard University Press)
[15] Antal T and Scheuring I 2006 *Bull. Math. Biol.* **68** 1923–44
[16] Altrock P M and Traulsen A 2009 *New J. Phys.* **11** 013012
[17] Black A J, Traulsen A and Galla T 2012 *Phys. Rev. Lett.* **109** 028101
[18] Nadell C D, Foster K R and Xavier J B 2010 *PLoS Comput. Biol.* **6** e1000716
[19] Rakoff-Nahoum S, Foster K R and Comstock L E 2016 *Nature* **533** 255–9
[20] Foster K R, Schluter J, Coyte K Z and Rakoff-Nahoum S 2017 *Nature* **548** 43–51
[21] Friedman J and Gore J 2017 *Curr. Opin. Syst. Biol.* **1** 114–21
[22] Vega N M and Gore J 2018 *Curr. Opin. Syst. Biol.* **45** 195–202
[23] Elgart V and Kamenev A 2004 *Phys. Rev.* E **70** 041106
[24] Kessler D A and Shnerb N M 2007 *J. Stat. Phys.* **127** 861–86
[25] Assaf M and Meerson B 2010 *Phys. Rev.* E **81** 021116
[26] Assaf M and Meerson B 2017 *J. Phys. A: Math. Theor.* **50** 263001
[27] Krapivsky P L, Redner S and Ben-Naim E 2010 *A Kinetic View of Statistical Physics* (New York: Cambridge University Press)
[28] Van Kampen N G 1992 *Stochastic Processes in Physics and Chemistry* vol 1 (Amsterdam: Elsevier)
[29] Redner S 2001 *A Guide to First-Passage Processes* (New York: Cambridge University Press)
[30] Lee S *et al* 2012 *Mathematics Stack Exchange* https://math.stackexchange.com/questions/151441/calculate-sums-of-inverses-of-binomial-coefficients
[31] Moran P A P *et al* 1962 *The Statistical Processes of Evolutionary Theory* (Oxford: Oxford University Press)
[32] Kimura M 1983 *The Neutral Theory of Molecular Evolution* (Cambridge: Cambridge University Press)
[33] Ewens W J 2012 *Mathematical Population Genetics 1: Theoretical Introduction* (New York: Springer)
[34] Bender C M and Orszag S A 1999 *Advanced Mathematical Methods for Scientists and Engineers I: Asymptotic Methods and Perturbation Theory* (New York: Springer)
[35] Gardiner C W *et al* 1985 *Handbook of Stochastic Methods* 3rd edn (Berlin: Springer)
[36] Fichthorn K, Gulari E and Ziff R 1989 *Phys. Rev. Lett.* **63** 1527
[37] Considine D, Redner S and Takayasu H 1989 *Phys. Rev. Lett.* **63** 2857
[38] Carro A, Toral R and Miguel M S 2016 *Sci. Rep.* **6** 24775
[39] Herrerías-Azcué F and Galla T 2019 *Phys. Rev.* E **100** 022304
[40] Lambiotte R and Redner S 2007 *J. Stat. Mech.* L10001
[41] Bunin G 2017 *Phys. Rev.* E **95** 042414
[42] Pearce M T, Agarwala A and Fisher D S 2020 *Proc. Natl Acad. Sci. USA* **117** 14572–83
[43] Ferry M S, Razinkov I A and Hasty J 2011 *Methods Enzymol.* **497** 295–372
[44] Luke C S, Selimkhanov J, Baumgart L, Cohen S E, Golden S S, Cookson N A and Hasty J 2016 *ACS Synth. Biol.* **5** 8–14
[45] Karlin S and Taylor H M 1975 *A First Course in Stochastic Processes* (San Diego, CA: Academic)