Quasi-stationary states of game-driven systems: A dynamical approach

Sergey Denisov,[1,a] Olga Vershinina,[2] Juzar Thingna,[3] Peter Hänggi,[4] and Mikhail Ivanchenko[2]

AFFILIATIONS

[1] Department of Computer Science, Oslo Metropolitan University, N-0130 Oslo, Norway
[2] Department of Applied Mathematics, Lobachevsky University, 603950 Nizhny Novgorod, Russia
[3] Center for Theoretical Physics of Complex Systems (IBS), Daejeon 34126, South Korea
[4] Institut für Physik, Universität Augsburg, D-86135 Augsburg, Germany

[a] Author to whom correspondence should be addressed: sergiyde@oslomet.no

ABSTRACT

Evolutionary game theory is a framework to formalize the evolution of collectives ("populations") of competing agents that are playing a game and, after every round, update their strategies to maximize individual payoffs. There are two complementary approaches to modeling evolution of player populations. The first addresses essentially finite populations by implementing the apparatus of Markov chains. The second assumes that the populations are infinite and operates with a system of mean-field deterministic differential equations. By using a model of two antagonistic populations, which are playing a game with stationary or periodically varying payoffs, we demonstrate that it exhibits metastable dynamics that is reducible neither to an immediate transition to a fixation (extinction of all but one strategy in a finite-size population) nor to the mean-field picture. In the case of stationary payoffs, this dynamics can be captured with a system of stochastic differential equations and interpreted as a stochastic Hopf bifurcation. In the case of varying payoffs, the metastable dynamics is much more complex than the dynamics of the means.

Published under license by AIP Publishing. https://doi.org/10.1063/5.0019736

Although first formalized by Darroch and Seneta in 1964,[1] the best (in our opinion) outline of quasi-stationary states (QSs) was given by Yaglom when he asked in 1947:[2] "We are dealing with a stochastic process with an absorbing state (i.e., a state which cannot be left once the system got into it) so that the absorption happens with probability one. What is the distribution in the limit $t \to \infty$ provided that the absorption did not happen until time t?" In the Markov chain framework, the QSs are related to the probability distributions over the set of states from which all the absorbing[3] ones are excluded. Some quasi-stationary distributions could be very sustainable so that the time needed for their eventual "evaporation" into the absorbing states is much longer than all the relevant observation time scales.[4] In this case, the QS is also a *metastable* state.[5] The QS concept naturally applies to the game-driven evolution of finite populations.[6,7] Quasi-stationary states were recently interpreted as a mean[9] to resolve a conflict[6] between the stochastic approach based on the Markov chain ideology[7] and deterministic mean-field description.[8] Specifically, while the former tells that the asymptotic regime of any finite population is a fixation, the latter yields an asymptotic solution either in the form of a periodic cycle or a chaotic attractor or a fixed point (an "evolutionary stable strategy" in which different types of players are represented[8]). The gap between the two mutually exclusive pictures could be bridged with QSs by observing that, in the case of large but finite populations, the corresponding distributions are localized around the mean-field solutions and declaring that the mean-field solution is thus transformed into a transient (metastable) dynamics.[9–11] Here, we demonstrate that this is not always the case, and the metastable dynamics, underlying the QSs, can be very different from the mean-field solution.

I. INTRODUCTION

The agenda of Evolutionary Game Theory (EGT) is to explain the evolution of biological populations by formalizing two main driving factors, conflict and cooperation.[12] Current applications of the theory range far beyond the original biological context, and

it is used to model dynamics of financial markets[13] and interpret condensation phenomena in dissipative quantum systems.[14]

The central element of EGT is the interaction of players during a round of a game. Although formally any number of players can be involved into the interaction,[15] most of the existing results were obtained for two-player games. Players can choose strategies from fixed sets, and their interaction is mediated by payoffs corresponding to the choices they made. For a two-player game with two strategies per player, the payoffs can be arranged in a bi-matrix,[8]

$$\begin{bmatrix} a_{11}, b_{11} & a_{12}, b_{21} \\ a_{21}, b_{21} & a_{22}, b_{22} \end{bmatrix}. \tag{1}$$

For example, if player A chooses strategy 1 and player B chooses strategy 2, the former receives payoff a_{12} and the latter receives payoff b_{21}. After every round, players monitor the payoffs obtained by their peers (other members of the population) and try to adapt the strategy of the most successful ones. There are several formalizations of the strategy adaptation process; the Moran process[16,17] is currently the most popular one; see, e.g., Refs. 6, 9, 10, and 18.

Finite sizes of animal populations favor stochastic Markovian approaches[3,19] to modeling game-driven evolution. By assuming that players belonging to the same population are indistinguishable, the iterative process of matching and consequent strategy adaptation can be recast in the form of a Markov chain. The state of a population is then fully specified by the probability vector, which assigns probability to every possible arrangement of the strategies. In the absence of mutations[20] (the case we address here), a state corresponding to the situation when the whole population uses the same strategy is an *absorbing state*.[21] Once the population enters this state, a fixation has happened.[19]

An alternative approach, which was prevailing in the EGT field until recently, addresses the case of infinite populations. Its main tool is the celebrated replicator model, which is a system of nonlinear deterministic differential equations.[8] The variables governed by these equations are relative frequencies of the strategies, which are assumed to be continuous probabilities. Because of the non-linearity of its equations, the replicator model is able to exhibit a spectrum of dynamical regimes, ranging from fixed point solutions to periodic oscillations and chaos (when the number of players and/or number strategies per player are larger than two; see, e.g., Refs. 22–24).

In between these two limits, there is a land of large but finite populations. Many animal populations fall into this category; therefore, analysis of the corresponding models can provide some additional insight into complex real-life phenomena.[25] It is intuitive that finite size fluctuations play an important role in the evolutionary dynamics of such populations, but these populations are too big to be modeled with Markov chains.[26] The diffusion approximation of Markov processes[27] is an appealing tool to bridge the two approaches and explore the land between them. When implemented to Moran processes,[16,17] the approximation yields a system of stochastic differential equations (SDEs) with a multiplicative cross-correlated noise.[6,28] The noise strength scales down as the population size increases so that, in the thermodynamic limit, the equations reduce to the deterministic replicator model.

Because the size of the population enters the SDEs as a parameter, the diffusion approach provides a tremendous speed-up as

compared to Markov chain simulations and allows for simulating models of an arbitrary large size.[28] The SDE approach can be used to resolve QSs and capture the transient (before the fixation) dynamics, e.g., by replacing the absorbing boundary conditions with the reflecting ones[10] and then performing a routine ensemble averaging.

Here, we apply the concept of quasi-stationary distributions (QDs)[1,29] to analyze evolutionary dynamics of two antagonistic populations driven by a two-player game. We consider two variants of the game, with stationary payoffs and with payoffs changing from round to round, in a periodic manner (this choice is motivated by biological phenomena; see Sec. II).

We demonstrate that, in both cases, the metastable dynamics underlying the QDs is very different from the solutions of the corresponding replicator models. In the stationary case, the latter is a fixed point, while the metastable dynamics is manifested by a stochastic limit cycle, which, within the SDE framework, can be interpreted as a result of a stochastic Hopf bifurcation.[30–32] By employing the Floquet theory,[33,34] we generalize the notion of QD to evolutionary games with periodically varying payoffs. We demonstrate that the corresponding non-equilibrium states are the result of a complex dynamics, which cannot be reduced to the evolution of means.

II. MODEL

Motivated by the original biological context of EGT, we decided to use the celebrated "Battle of Sexes" model by Dawkins[35] as an example. This model formalizes a sex conflict over the parental investment,[36] and two antagonistic populations are represented by females and males of a species (see sketch in Fig. 1). Players of each sex have two strategies, and payoffs are given by a 2×2 bi-matrix. We generalize this model by introducing periodic time variations in the payoffs. We are not only inspired by the possibility to encounter a more complex evolutionary dynamics but also by several new findings in ecology.

In recent years, it has been found that in many species, mating strategies and preferences are not constant in time but are season-dependent.[37–40] When courting (selecting) a mate, a female or male of the species faces a complex choice problem when benefits of a choice depend on the season and have to be traded off against each other in the context of current environmental conditions. For example, the rate of the hormonal activity, courtship strategies, and mate selection of Carolina anole lizards (*Anolis carolinensis*), both of females and males, are regulated by the temperature and photoperiod and thus are strongly season dependent.[41] Even the amount of different types of muscle fibers that control the vibrations of a red throat fan (dewlap)—which males employ during courtship—is season dependent.[42] Within the "Battle of Sexes" framework, this can be modeled by introducing periodically varying payoffs, as illustrated in Fig. 1.

Specifically, players A (males) and B (females) form two populations of a fixed size N, with each of them having two strategies $s = \{1, 2\}$. Payoffs, $\{a_{ss'}\}$ and $\{b_{s's}\}$, $s, s' = \{1, 2\}$, may change from round to round. During a single round, members of the antagonistic populations are randomly matched and then simultaneously play N games. After that, a strategy adaptation phase takes place in every populations. Then, the process is iterated.
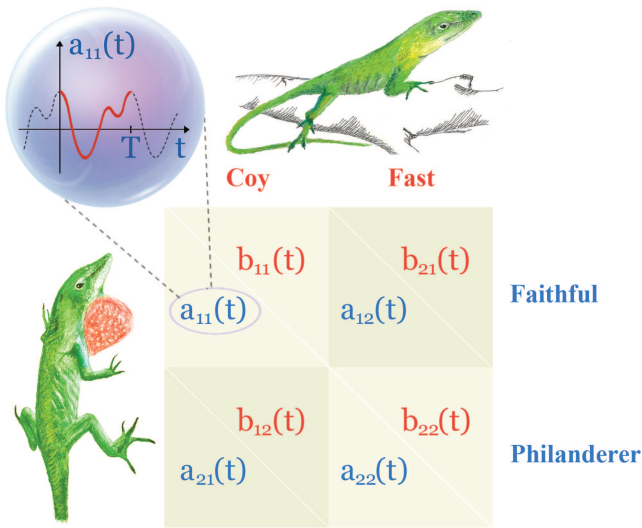
FIG. 1. "Battle of Sexes" with seasonal variations of mate preferences. It can model, e.g., the process of mate selection in a population of Carolina anole lizards. A female of the species could be either *coy* (and prefer a long and arduous courtship in order to be sure that a mate will contribute to parental care) or *fast* (and not worry much about parental care). A male could be either *faithful* (and ready to assure a potential partner that he will be a "faithful husband" by performing a long courtship) or *philanderer* (and thus would prefer to shorten the courtship stage). Depending on the strategies, s (played by the female) and s' (played by her mate), the female gets payoff $b_{ss'}$ (red) and the male gets $a_{s's}$ (blue). Both females and males are season-constrained in their preferences of opposite-sex mates. These seasonal constraints are modeled via time-periodic variations of the payoffs.

This evolution is an essentially discrete-time process and its rounds are labeled with index m. In order to be able to compare the discrete-time evolution to the mean-field dynamics, we introduce time variable t, which is counted from zero and incremented by Δt after every round. Now, we define time-periodic payoffs, $c_{ss'}(t) = c_{ss'}(t + T)$, $c = \{a, b\}$, where $\Delta t = T/M$. The payoffs can be represented as sums of constant and zero-mean time-periodic components, $c_{ss'}(t) = \bar{c}_{ss'} + \tilde{c}_{ss'}(t)$, $\langle \tilde{c}_{ss'}(t)\rangle_T = 0$. After M rounds, the payoffs return to their initial values.

To introduce the strategy adaptation stage, we use the Moran process,[16,17] which works in the following way. The state of the populations after the mth round is specified by the number of players playing first strategy from their strategy list, i (males with $s = 1$) and j (females with $s' = 1$), $0 \leq i, j \leq N$. The payoffs obtained by the players using strategy s are

$$\pi_s^A(j, t) = a_{s1}(t)\frac{j}{N} + a_{s2}(t)\frac{(N - j)}{N}, \qquad (2)$$

$$\pi_s^B(i, t) = b_{s1}(t)\frac{i}{N} + b_{s2}(t)\frac{(N - i)}{N}. \qquad (3)$$

Payoffs determine the probabilities for a player to be chosen for reproduction, e.g., for population A,

$$P_s^A(i, j, t) = \frac{1}{N} \cdot \frac{1 - w + w\pi_s^A(j, t)}{1 - w + w\bar{\pi}^A(i, j, t)}, \qquad (4)$$

where $\bar{\pi}^A(i, j, t) = [i\pi_1^A(j, t) + (N - i)\pi_2^A(j, t)]/N$ is the average payoff for the population A.

The baseline fitness $w \in [0, 1]$ is a tunable parameter of the game-driven evolution.[6,17] For example, when $w = 0$, the probability to be chosen for reproduction does not depend on player's payoff and is uniform across the population. After the choice for reproduction has been made, another member of the population is chosen completely randomly and replaced with an offspring of the chosen player, i.e., with a player that uses the same strategy as its parent.[43] This mechanism is acting simultaneously in both populations, A and B, so that one male and one female are chosen for reproduction. The mating pair then produces two offspring, a male and female, which then introduced into the corresponding populations. The size N of both population is therefore preserved.

A single round can be considered a two-dimensional Markov chain, a generalization of the one-dimensional chain introduced in Refs. 6 and 28. Before formalizing the Markov chain, we first define the transition rates for two populations of players. For example, the probability for population A to get one more player with strategy 1 after one round (and, correspondingly, one player with strategy 2 less) is given by

$$T_A^+(i, j, t) = \frac{1 - w + w\pi_1^A(t)}{1 - w + w\bar{\pi}^A}\frac{i}{N}\frac{N - i}{N}. \qquad (5)$$

The probability to get one more player with strategy 2 (and, correspondingly, one player with strategy 1 less) is

$$T_A^-(i, j, t) = \frac{1 - w + w\pi_2^A(t)}{1 - w + w\bar{\pi}^A}\frac{N - i}{N}\frac{i}{N}. \qquad (6)$$

The probability for population B is defined in a similar manner. These four probabilities are building blocks to construct the transition matrix for the two-dimensional Markov process (see the Appendix).

By following the procedure from Ref. 6, it can be shown that, in the limit $N \to \infty$ and $dt = \frac{1}{N} \to 0$, the dynamics of the variables $x = i/N$ and $y = j/N$ is governed by the adjusted replicator equations,[12,44]

$$\dot{x} = x[1 - x][\triangle^A(t) + \Sigma^A(t)y]\frac{1}{\Gamma + \bar{\pi}^A(x, y, t)}, \qquad (7)$$

$$\dot{y} = y[1 - y][\triangle^B(t) + \Sigma^B(t)x]\frac{1}{\Gamma + \bar{\pi}^B(x, y, t)}, \qquad (8)$$

where $\triangle^C = c_{12} - c_{22}$, $\Sigma^C = c_{11} + c_{22} - c_{12} - c_{21}$, $\Gamma = \frac{1-w}{w}$, and $C(c) = \{A(a), B(b)\}$.

Within the stochastic framework, a single round can be represented as the multiplication[45] of the state **p**, expressed as a $N \times N$ matrix with elements $p(i, j)$, with the transition fourth-order tensor **S**, with elements $S(i, j, i', j')$ (see the Appendix). By using the bijection $k = (N - 1)j + i$, we can unfold the probability matrix $p(i, j)$ into the vector $\tilde{p}(k)$, $k = 0, \ldots, N^2$, and the tensor $S(i, j, i', j')$ into the

matrix $\tilde{S}(k,l)$ with $k,l = 0,1,\ldots,N^2$. This reduces the problem to a Markov chain,[46] $\tilde{\mathbf{p}}^{m+1} = \tilde{\mathbf{S}}^m \tilde{\mathbf{p}}^m$, where m is the round to be played and every state is fully specified now with a single integer k.

The four states ($i = \{0,N\}, j = \{0,N\}$) are absorbing states because the transition probabilities leading out of them, Eqs. (5) and (6), are identically zero. The absorbing states are, therefore, attractors (sinks) of the evolutionary dynamics, and it is evident that finite-size fluctuations will eventually drive a population to one of these states.

We are interested in the dynamics before the absorption; therefore, we merge the four states into a *single* absorbing state by summing the corresponding incoming rates. The boundary states, ($i = \{0,N\}, j \in \{1,\ldots,N-1\}$) and ($i \in \{1,\ldots,N-1\}, j = \{0,N\}$), can also be merged into this absorbing "super-state." The reason for including these additional states into the "absorbing" state is that once the population gets to the boundary, it will only move toward one of the two nearest absorbing states. By labeling the absorbing super-state with index $k = 0$, we end up with a stochastic $(L+1) \times (L+1)$ matrix

$$\tilde{\mathbf{S}}^m = \begin{bmatrix} 1 & \varrho_0{}^m \\ \mathbf{0} & \tilde{\mathbf{Q}}^m \end{bmatrix}, \tag{9}$$

where $L = (N-2)^2$, $\varrho_0{}^m$ is the row vector given by the set of incoming transition probabilities to the absorbing super-state, $\mathbf{0}$ is a $L \times 1$ column vector with all entries zero, and $\tilde{\mathbf{Q}}^m$ is a $L \times L$ reduced transition matrix.

With Eq. (9), we arrive at the setup used by Darroch and Seneta to formulate the concept of QDs.[1] Specifically, there exists a vector d, which represent a quasi-stationary distribution (QD). This distribution can be very sustainable and remains near invariant under the action of the matrix $\tilde{\mathbf{S}}^m$. In this case, it is a metastable state.[5] The QD can be obtained from the right eigenvector d of the reduced transition matrix $\tilde{\mathbf{Q}}^m$ corresponding to the maximum (by modulus) eigenvalue λ.[1] By virtue of the Perron–Frobenius theorem, λ is real and vector d is real and non-negative.[46] By applying the inverse bijection and then performing the normalization, we obtain two-dimensional probability distribution $\mathbf{d}(x,y)$.

The most straightforward way to evaluate the dynamics underlying the QDs is to perform Monte-Carlo sampling. To address the limit $N \gg 1$, we follow the recipe given in Refs. 6 and 28 and derive stochastic differential equations (SDEs), which can be then used to approximate the evolution. As the first step, we introduce continuous variables $x = i/N$, $y = j/N$, and $t = m/N$. Next, we define probability density $\varrho(x,y,t) \simeq N \cdot p_m(i,j)$. Finally, by Taylor expanding the Markov chain in orders of $1/N$ and truncating the expansion after the second order, we obtain a Fokker–Planck equation,[48]

$$\dot{\rho}(x,y,t) = -\sum_{k=A,B} \frac{\partial}{\partial x_k} \cdot \left[v_k(x,y,t)\rho(x,y,t) \right]$$
$$+ \frac{1}{2} \sum_{k,l=A,B} \frac{\partial^2}{\partial x_k x_l} \cdot \left[d_{kl}(x,y,t)\rho(x,y,t) \right], \tag{10}$$

where

$$v_k(x,y,t) = T_k^+(x,y,t) - T_k^-(x,y,t), \tag{11}$$

$$d_{kk}(x,y,t) = \frac{T_k^+(x,y,t) + T_k^-(x,y,t)}{N}, \tag{12}$$

$$d_{kl}(x,y,t) = \frac{v_k(x,y,t) \cdot v_l(x,y,t)}{N}, \quad k \neq l, \tag{13}$$

where $k,l \in \{A,B\}$. The corresponding stochastic differential equations can be represented in the Langevin form,[48]

$$\dot{x}(t) = v_A(x,y,t) + \sum_{l=A,B} g_{Al} \cdot \xi_l(t), \tag{14}$$

$$\dot{y}(t) = v_B(x,y,t) + \sum_{l=A,B} g_{Bl} \cdot \xi_l(t), \tag{15}$$

where $\xi(t)$ is an uncorrelated Gaussian white noise of variance one. The 2×2 matrix $G = \begin{pmatrix} g_{AA} & g_{AB} \\ g_{BA} & g_{BB} \end{pmatrix}$ can be expressed in terms of the diffusion matrix $D = \begin{pmatrix} d_{AA} & d_{AB} \\ d_{BA} & d_{BB} \end{pmatrix}$ via $G^T G = D$. We integrate Eqs. (14) and (15) by using the standard Euler–Maruyama method[49] with time step $dt = 10^{-4}$. In order to obtain matrix G, on every step, we diagonalize 2×2 diffusion matrix $D = U\Lambda U^T$ and use its eigenvectors to construct matrix $G = U\sqrt{\Lambda}U^T$. A conditional sampling of QD is performed by integrating the Langevin equations up to time t_{fin} and updating the histogram with the final points of the trajectories.

III. STATIONARY PAYOFFS: A STOCHASTIC BIFURCATION

We first consider the case when all payoffs are constant. As an example, we use a game with payoffs a_{11}, a_{22}, b_{12}, and b_{21} equal to 1, and -1 for the rest of strategies (this choice corresponds to the "Matching Pennies" game[50]). Figure 2 presents the metastable states of the corresponding evolution. In order to find them, we numerically obtain the maximal-eigenvalue eigenvector of the matrix $\tilde{\mathbf{Q}}^1$ ($m = 1$ since the payoffs are constant) and perform sampling by using the SDEs, Eqs. (14) and (15).

The means (first moments of the QD),

$$\bar{x} = \sum_{i,j=1}^{N-1} \frac{i \cdot d(i,j)}{(N-2)^2}; \quad \bar{y} = \sum_{i,j=1}^{N-1} \frac{j \cdot d(i,j)}{(N-2)^2}, \tag{16}$$

coincide with the Nash equilibrium $\left(\frac{1}{2}, \frac{1}{2}\right)$ for any N. Yet, the distributions obtained for $N = 100$ and $N = 200$ are not simply peaked but have "craters" on their tops. Stochastic Markovian simulations reveal the existence of a long transient trajectory going along the ridge of the craters and encircling the Nash equilibrium; see Fig. 3. The SDE integration yields a near identical dynamics. However, the SDE framework offers an interpretation of the dynamics as a limit cycle resulted from a stochastic Hopf bifurcation.[30–32,51] The latter is a result of a marginal stability of the fixed point $\left(\frac{1}{2}, \frac{1}{2}\right)$ of the replicator system (7) and (8).[52]
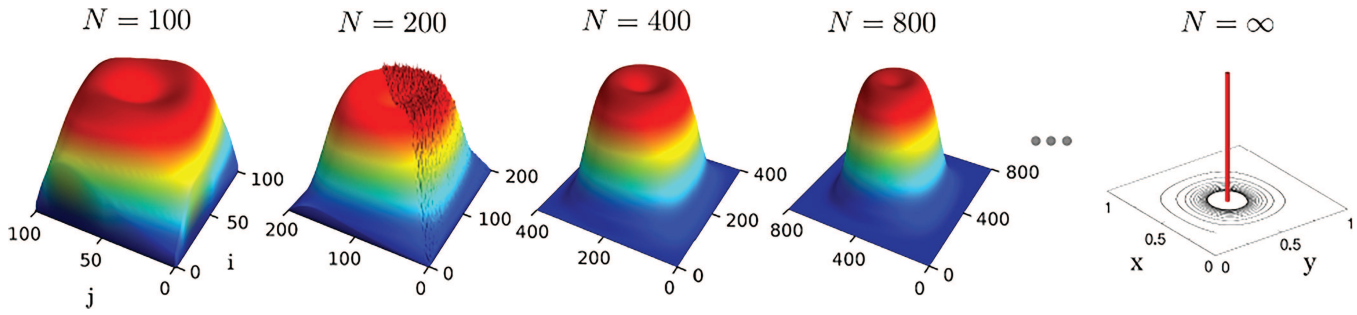
FIG. 2. Quasi-stationary probability distributions of the "Battle of Sexes" with constant payoffs. For all values of N, the distribution was obtained by finding the maximal-eigenvalue eigenvector of the reduced transient matrix $\tilde{\mathbf{Q}}^1$, Eq. (9).[47] For $N = 200$, the plot combines the distribution obtained by reshuffling the eigenvector (left half) and the histogram sampled by using Langevin equations (14) and (15), which were propagated to time $t_{fin} = 10^6$ (right half). The total number of realization is 10^7. In the thermodynamic limit $N = \infty$ (right panel), the quasi-stationary distribution shrinks to the Nash equilibrium $(\frac{1}{2}, \frac{1}{2})$, which thus becomes a fixed point of the mean-field dynamics. The baseline fitness $w = 0.3$.

The limit-cycle character of the dynamics can be validated by tracking the trajectory polar angle as a function of time and noting a well pronounced linear dependence; see the inset in Fig. 4. We also calculate radius R_{MC} of the metastable cycle and its period T_{MC} by performing both Monte-Carlo and SDE samplings. The radius scales proportionally to the fluctuation strength, $\sim \frac{1}{\sqrt{N}}$, and the period is near independent of N, as expected for a stochastic Hopf bifurcation.[51]

The average lifetime of the metastable dynamics can be defined by using the largest eigenvalue λ of the reduced transition matrix $\tilde{\mathbf{Q}}^1$, as $t_{life} = 1/(1 - \lambda)$.[1] It can then be compared with the average

extinction time,[53] obtained by performing Monte-Carlo sampling of the Markovian dynamics. The two times are in a perfect agreement; see Fig. 5.

IV. PERIODICALLY VARYING PAYOFFS: A METASTABLE FLOQUET DYNAMICS

By introducing periodic variations into the payoffs of the "Battle of Sexes," we find that the mean-field dynamics, Eqs. (7) and (8),
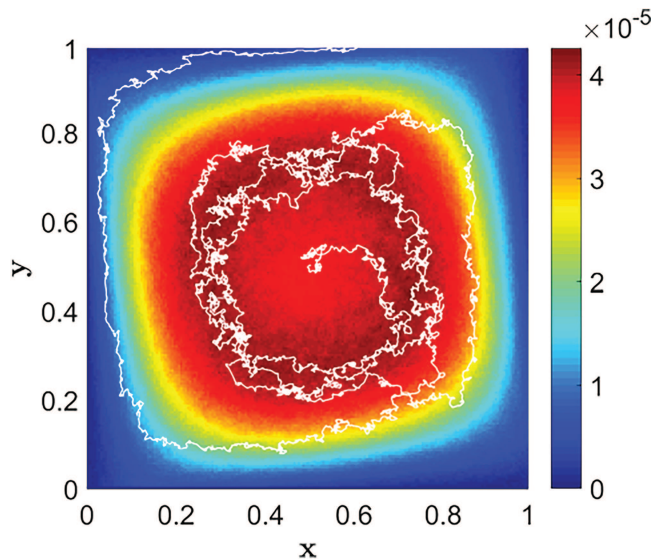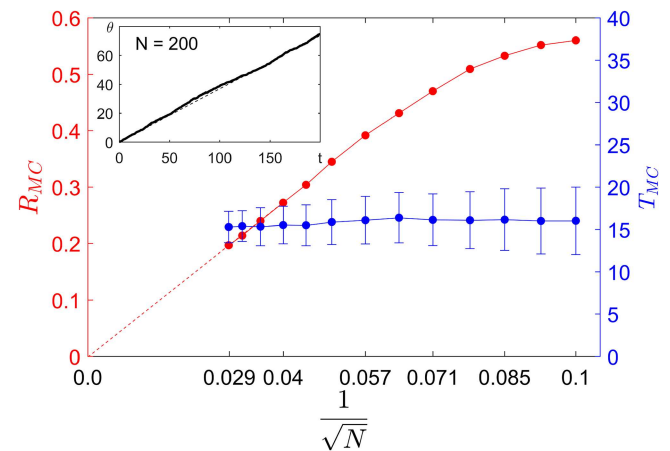

FIG. 3. Transient evolution of two populations driven by the "Battle of Sexes" game with constant payoffs. The color indicates the quasi-stationary distribution $\mathbf{d}(x, y)$. The white line corresponds to a single realization of the Markov process initiated at the point $(\frac{N}{2}, \frac{N}{2})$. $N = 200$. Other parameters are as in Fig. 2.

FIG. 4. Radius of the metastable limit cycle (red) and its period (blue) for the "Battle of Sexes" with constant payoffs. For $N \leq 300$ ($1/\sqrt{N} \geq 0.057$), both quantities were obtained from quasi-stationary distributions and Monte-Carlo sampling. For $N > 300$, a sampling with the SDEs, Eqs. (14) and (15), was used. The inset shows the evolution of the polar angle (black line) of the stochastic trajectory of the stochastic differential equations Eqs. (14) and (15). The black dashed line corresponds to linear dependence $\langle\Omega\rangle t$, where $\langle\Omega\rangle = 2\pi/T_{MC}$ for $N = 200$. The red dashed line indicates an inverse dependence of the cycle radius with respect to the fluctuation strength $\sqrt{N}$. The radius scaling and the period independence of N are signatures of a stochastic Hopf bifurcation.[51] All parameters are the same as in Fig. 3.
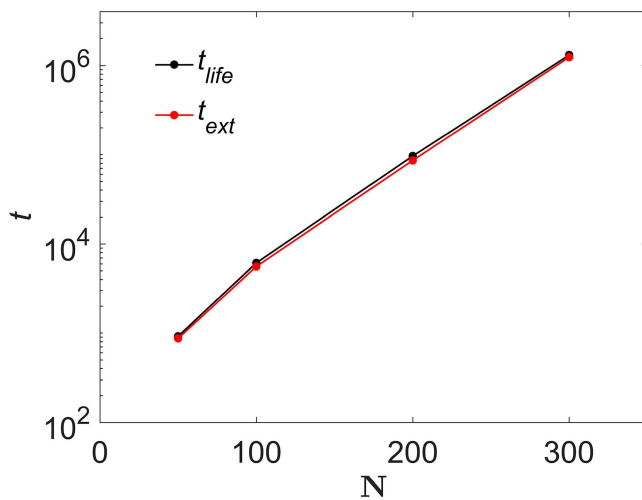
FIG. 5. The average lifetime t_{life} vs average extinction time t_{ext} for the "Battle of Sexes" with constant payoffs. The average extinction time was sampled with 10^5 realizations per non-absorbing state. All parameters are the same as in Fig. 3.

does not exhibit substantial changes even for a relatively large variation amplitude. For example, for a choice when $\varepsilon(t) = \tilde{a}_{11}(t) = \tilde{b}_{22}(t) = f\cos(\omega t)$, with $f = 0.5$ and $\omega = 2\pi/T$ and all other payoffs kept constant, we observe a period-one limit cycle localized near the Nash equilibrium of the constant payoff case; see Fig. 6(a). In the limit $\omega \ll 1$, the cycle collapses to a set of adiabatic

Nash equilibria [dashed black lines in Fig. 6(a)]. Here, an adiabatic Nash equilibrium is defined as the Nash equilibrium of a stationary game with "frozen" instantaneous values of payoffs, $\{x_{NE}(\varepsilon) = \frac{2-\varepsilon}{4-\varepsilon}, y_{NE}(\varepsilon) = \frac{2}{4+\varepsilon}\}$.

Figure 7(a) shows the average extinction time t_{ext} as a function of variation strength f for $\omega = 0.1$. In order to compare the evolution for different population size N, we keep T constant and scale the time step as $\Delta t = \frac{2\pi}{N}$. Another words, in terms of index m, the number of the rounds needed for one variation cycle is $M = \omega N$.

As it can be seen from Fig. 7(a), changes of t_{ext} with f are not substantial for all three considered values of N. However, the extinction times are more than the order of magnitude shorter than in the case of constant payoffs (compare to Fig. 5). The Monte-Carlo sampling reveals the cycling transient dynamics that is much less localized around the Nash equilibrium than the limit cycle corresponding to the mean-field description; see Fig. 6(b). Evidently, the dynamics of the means is not able to unreveal the whole complexity of the transient dynamics. We are going now to generalize the idea of QDs to the case of periodically varying payoffs and demonstrate that the corresponding quasi-stationary state provides a deeper insight into the transient dynamics.

The transition matrices, Eq. (9), are round-specific now and form a set $\{\tilde{\mathbf{S}}^m\}$, $m = 1, \ldots, M$ (recall that after $M = T/\Delta t$ rounds, the payoffs return to their initial values). The propagator over the interval $[0, t]$, $0 < t < T$, is the product $\tilde{\mathbf{U}}(0, t) = \prod_{k=1}^{m} \tilde{\mathbf{S}}^k$ with $m = t/\Delta t$. All the propagators, including the period-one propagator $\tilde{\mathbf{U}}(0, T) = \tilde{\mathbf{U}}_T$, have the same structure as the matrix in Eq. (9).

We introduce the quasi-stationary distribution of $\tilde{\mathbf{U}}_T$, $\mathbf{d}[0]$. It is a *Floquet state*[33,34] of the reduced period-one propagator $\tilde{\mathbf{U}}_T^r$, obtained by replacing, in the definition of the propagator,
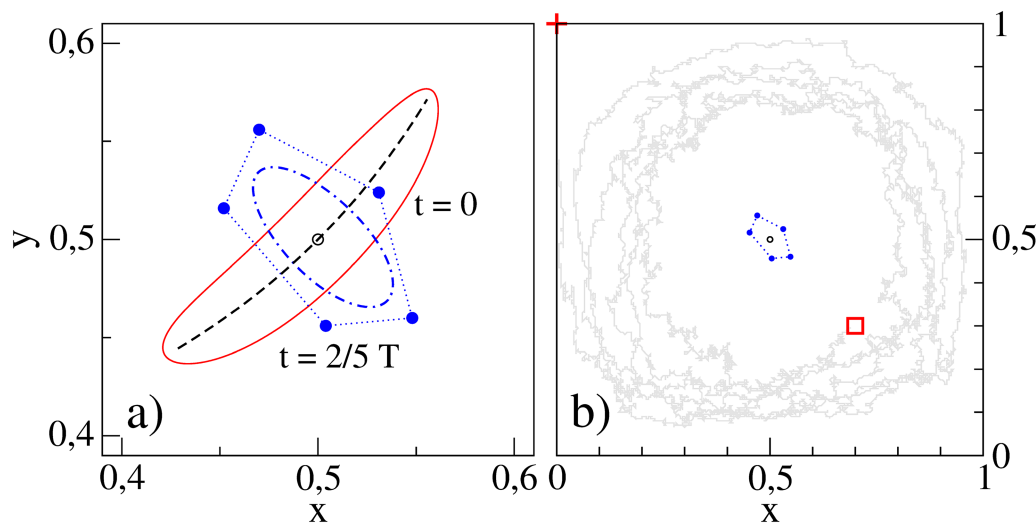

FIG. 6. Evolution driven by a "Battle of Sexes" with periodically varying payoffs. (a) Period-one limit cycles of the mean-field dynamics, $N \to \infty$ with $\omega = 0.1$ (solid blue line) and $\omega = 0.01$ (solid red line) are localized near the point $(\frac{1}{2}, \frac{1}{2})$ (arrows indicate the direction of the motion). In the adiabatic limit, $\omega \to 0$, the mean-field cycle collapses to the set of the instantaneous Nash equilibria (dashed black line). Averages of the Floquet quasi-stationary distribution $\mathbf{d}[t]$ for $N = 200$, $\bar{x}(t)$ and $\bar{y}(t)$, evolve along a limit cycle (blues dots) localized near the Nash equilibrium $(\frac{1}{2}, \frac{1}{2})$. (b) A stochastic trajectory (gray line) reveals the existence of a very de-localized transient dynamics. The trajectory was initiated at an interior point (square) and ended up at an absorbing state (cross). The parameters are $f = 0.5$, $\omega = 0.1$, and other parameters as in Fig. 2.
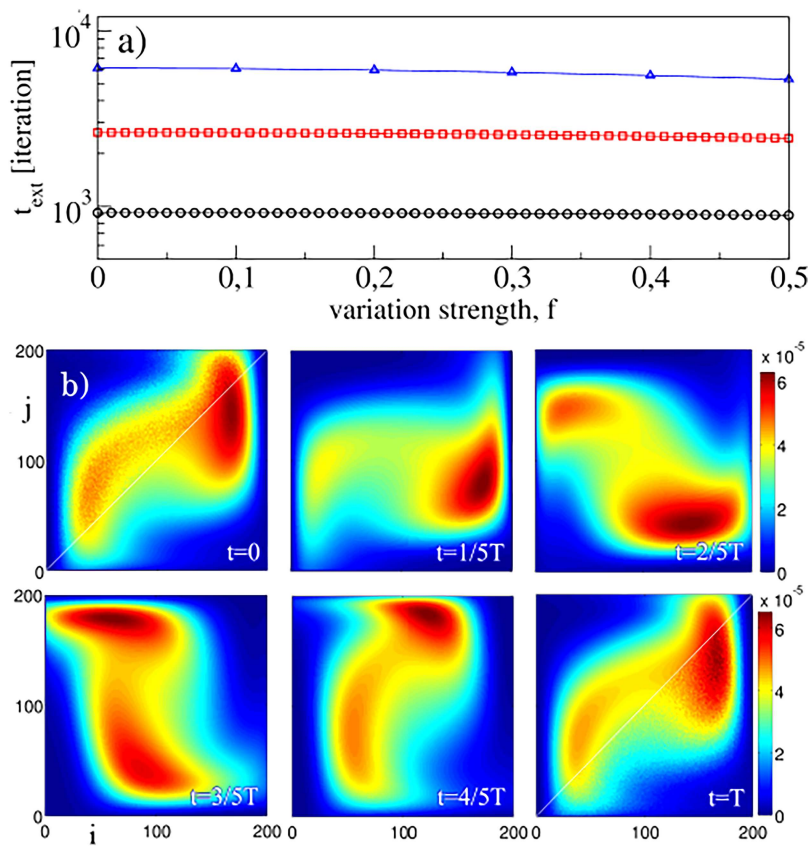
FIG. 7. Metastable Floquet dynamics. (a) The average extinction time t_{ext} as a function of the modulation strength f for the population size $N = 50$ (dots), 100 (squares), and 200 (triangles). For every values of the variation strength f, t_{ext} was sampled with 10^5 realizations per non-absorbing state. (b) Stroboscopic snapshots of quasi-stationary Floquet distributions $\mathbf{d}[t]$ for different time instances, $t = m \cdot \Delta t \in [0, T]$. Distributions are obtained by diagonalizing the reduced propagators $\tilde{\mathbf{Q}}^m$ for $N = 200$. Other parameters are as in Fig. 2.

transition matrix $\tilde{\mathbf{S}}^{m'}$ with reduced transition matrix $\tilde{\mathbf{Q}}^{m'}$. The maximal-eigenvalue eigenvector of the matrix, after the reshuffling, leads to probability distribution $\mathbf{d}[0] = \mathbf{d}[T]$. This is a stroboscopic snapshot of Floquet QD, which periodically evolves, $\mathbf{d}[t] = \mathbf{d}[t + T]$, being locked by payoff variations; see Fig. 7(a). For any value of t, $0 < t < T$, QD $\mathbf{d}[t]$ can be obtained by acting with the reduced propagator $\tilde{\mathbf{U}}^r(0, t)$ on the distribution $\mathbf{d}[0]$.

The structure and dynamics of the Floquet QD explain the dramatic shortening of the average extinction times. Specifically, while in the constant limit, QD $\mathbf{d}(x, y)$ was localized near the Nash equilibrium; i.e., at the maximal distance from the absorbing boundaries, both maxima of the Floquet QD skim very close to the boundary [see Fig. 7(b)]. In the ecological context, this can be considered a periodic sequence of population bottlenecks.[35] A better articulated biological interpretation of this effect is intriguing but goes outside the scope of our work.

Finally, we compare an average lifetime extracted from the Floquet QD with the result of the sampling of the average extinction time. The average lifetime of $\mathbf{d}[t]$ is defined by using the largest eigenvalue λ_T, $0 < \lambda_T < 1$, of matrix $\tilde{\mathbf{U}}^r_T$. In order to compare it with the lifetime introduced for the case of constant payoffs, we define the mean single-round exponent as $\bar{\lambda}_T = \sqrt[T]{\lambda_T}$. The average lifetime is defined then as $t_{life} = 1/(1 - \bar{\lambda}_T)$. The obtained values are in a perfect agreement (within the sampling error) with the values presented in Fig. 7(a).

Different from the constant payoff case, the SDE approach, when realized numerically by using the first order Euler–Maruyama scheme, works poorly in this case. In particular, the trajectory often crosses the boundary of the unit square, thus making results meaningless.[54] Even though the noise strength is strictly zero along the adsorbing boundary, it is not enough to prevent the trajectory from crossing the latter. A numerical integration of stochastic equations with multiplicative noise and time-dependent coefficients is a non-trivial task, and its realization demands more sophisticated higher-order schemes.[55]

V. CONCLUSIONS

By implementing the idea of quasi-stationary distributions,[1,29] we demonstrated that the transient (before the fixation) game-driven evolution of a large but finite population can be very different from the mean-field replicator picture. In the case of a game with constant payoffs, the transient evolution corresponds to a noisy cycling dynamics, which can be reproduced with a system of stochastic differential equations. Within this framework, the dynamics can be interpreted as a stochastic Hopf bifurcation.[30–32,51]

There are ongoing discussions on the role of "stochasticity in population cycling,"[61–63] and the phenomenon of "noise-sustained oscillations around otherwise stable equilibria" was addressed recently in the ecological context.[62]

It is important, therefore, to demarcate this type of cycling from the well-known phenomena observed in many EGT models,[8] especially in a system driven by games with so-called "cycling dominance"[56,57] (Rock–paper–scissors is perhaps the most celebrated example). This phenomenon goes under different names, depending on the interpretation, e.g., "evolutionary cycling,"[58] "cycles of cooperation and defection,"[59] and "oscillating tragedy of the commons."[60] At the formal level, the corresponding models demonstrate cycling behavior in the thermodynamics limit, and the origin of the cycling is a limit-cycle dynamics exhibited by the corresponding mean-field system.

In contrast, in our models, oscillations are absent in the mean-field picture due to the marginal stability of the fixed point. Our findings can be considered an extension of the results by McKane and Newman.[61] Namely, McKane and Newman highlighted the existence of a cycling behavior in a finite predator–prey model, while the corresponding mean-field system is characterized, similar to adjusted replicator equations we considered, Eqs. (7) and (8), by a single marginally stable fixed point (an attracting point with zero Lyapunov exponents). Their main finding is that the mechanism responsible for the appearance of the cycling is internal, and it is a "demographic stochasticity inherent in discrete birth, death, and predation events."[61] We consider a complimentary aspect, i.e., the transient character of the cycling dynamics and the finiteness of its lifetime.

By introducing periodically varying payoffs, we demonstrated the existence of a metastable, periodically changing probability distribution, which cannot be deduced from the evolution of the means. To be more specific, the mean-field system yields the limit cycle strongly localized near the Nash equilibrium of the average (in terms of the payoffs) game; see Fig. 6(b). In the case of finite-size dynamics with $N \gg 1$, it is usually expected that the corresponding probability distribution is localized on the cycle and the localization becomes stronger upon the increase of N. This is not so in the model with periodically varying payoffs: The corresponding probability distribution is multi-modal, and its first moments, $\bar{x}(t)$ and $\bar{y}(t)$, Eq. (16), do not reflect the complex shape and dynamics of the distribution. For example, during a single round of modulations, maxima of the distribution pass very close to the absorbing border [see Fig. 7(b)] so that the probability of extinction for the corresponding player

species is high. This, however, cannot be guessed by looking at the evolution of $\bar{x}(t)$ and $\bar{y}(t)$.

The idea that the Floquet theory can be used to model the evolution of ecological systems subjected to time-periodic environmental variations has recently been emphasized in Refs. 64 and 65. However, it was used to analyze asymptotic regimes of models described with a set of deterministic linear differential equations in the spirit of the conventional Floquet theory.[33]

Other potential applications of quasi-stationary Floquet distributions include modeling of periodically modulated finite systems with complex kinetics. For example, transient Floquet states can be related to a transient single-cell gene-expression dynamics when the chemical kinetics of molecule ensembles is modulated by the inner-cell circadian rhythm.[66,67]

Finally, we would like to mention a possibility of extension of the QS concept to quantum systems. Quantum channels,[68] also known as "completely positive trace-preserving maps" and "quantum operations,"[69] which can transform a quantum state (density operator) into another quantum state while preserving some important properties, are conventionally accepted as a quantum generalization of Markov chains. Formalization of quasi-stationary quantum states and figuring out their relations to metastable[70] and Floquet[71] states of open quantum systems is a thought-provoking perspective.

APPENDIX: THE TRANSITION TENSOR

Here, we describe the transition fourth-order tensor $S^m(i,j,i',j')$ in terms of the rates $[T_A^{+,-}(i,j,t)$ and $T_B^{+,-}(i,j,t)]$ for populations A and B, given by Eqs. (5) and (6) in the main text. The stochastic Moran process can be expressed as a Markov chain

$$
\begin{aligned}
p^{m+1}(i,j) =& \left[1 - T_A^+(i,j) - T_A^-(i,j)\right]\left[1 - T_B^+(i,j) - T_B^-(i,j)\right]p^m(i,j) + T_B^-(i,j+1)\left[1 - T_A^-(i,j+1) - T_A^+(i,j+1)\right]p^m(i,j+1) \\
&+ T_B^+(i,j-1)\left[1 - T_A^-(i,j-1) - T_A^+(i,j-1)\right]p^m(i,j-1) + T_A^-(i+1,j)\left[1 - T_B^-(i+1,j) - T_B^+(i+1,j)\right]p^m(i+1,j) \\
&+ T_A^+(i-1,j)\left[1 - T_B^-(i-1,j) - T_B^+(i-1,j)\right]p^m(i-1,j) + T_A^+(i+1,j+1)T_B^-(i+1,j+1)p^m(i+1,j+1) \\
&+ T_A^+(i-1,j+1)T_B^-(i-1,j+1)p^m(i-1,j+1) + T_A^-(i+1,j-1)T_B^+(i+1,j-1)p^m(i+1,j-1) \\
&+ T_A^+(i-1,j-1)T_B^+(i-1,j-1)p^m(i-1,j-1).
\end{aligned}
\tag{A1}
$$

Above, we have suppressed the time index $t = m\Delta t$ for all the rates.
Equation (A1) can be recast into

$$
p^{m+1}(i,j) = \sum_{i'j'} S^m(i,j,i',j')p^m(i',j'),
\tag{A2}
$$

where the fourth-order tensor $S^m(i,j,i',j')$ is given by

$$S^m(i,j,i',j') = \left[1 - T_A^+(i',j') - T_A^-(i',j')\right]\left[1 - T_B^+(i',j') - T_B^-(i',j')\right]\delta_{i',i}\,\delta_{j',j} + T_B^-(i',j')\left[1 - T_A^-(i',j') - T_A^+(i',j')\right]\delta_{i',i}\,\delta_{j',j+1}$$

$$+ T_B^+(i',j')\left[1 - T_A^-(i',j') - T_A^+(i',j')\right]\delta_{i',i}\,\delta_{j',j-1} + T_A^-(i',j')\left[1 - T_B^-(i',j') - T_B^+(i',j')\right]\delta_{i',i+1}\,\delta_{j',j}$$

$$+ T_A^+(i',j')\left[1 - T_B^-(i',j') - T_B^+(i',j')\right]\delta_{i',i-1}\,\delta_{j',j} + T_A^-(i',j')T_B^-(i',j')\delta_{i',i+1}\,\delta_{j',j+1} + T_A^+(i',j')T_B^-(i',j')\delta_{i',i-1}\,\delta_{j',j+1}$$

$$+ T_A^-(i',j')T_B^+(i',j')\delta_{i',i+1}\,\delta_{j',j-1} + T_A^+(i',j')T_B^+(i',j')\delta_{i',i-1}\,\delta_{j',j-1}, \tag{A3}$$

with the indices i, j, i', and $j' \in \{0, \dots, N\}$. Using the bijection $k = (N-1)j + i$ and $l = (N-1)j' + i'$, we obtain the required matrix form, Eq. (9), in the main text.

DATA AVAILABILITY

The data that support the findings of this study are available from the corresponding author upon reasonable request.

REFERENCES

[1] J. N. Darroch and E. Seneta, J. Appl. Prob. **2**, 88 (1965).

[2] A. Yaglom, Dokl. Akad. Nauk. SSSR **56**, 795 (1947).

[3] D. L. Isaacson and R. W. Madsen, *Markov Chains: Theory and Applications* (Wiley, 1976).

[4] E. A. van Doorn and P. K. Pollett, Eur. J. Oper. Res. **230**, 1 (2013).

[5] A. Bianchi and A. Gaudillière, Stoch. Proc. Appl. **126**, 1622 (2015).

[6] A. Traulsen, J. C. Claussen, and C. Hauert, Phys. Rev. Lett. **95**, 238701 (2005).

[7] A. Traulsen and C. Hauert, "Stochastic evolutionary game dynamics," in *Reviews of Nonlinear Dynamics and Complexity*, edited by H. G. Schuster (Wiley-VCH, Weinheim, 2009), Vol. 2, pp. 25–61.

[8] J. Hofbauer and K. Sigmund, *Evolutionary Games and Population Dynamics* (Cambridge University Press, Cambridge, 1998).

[9] D. Zhou, B. Wu, and H. Ge, J. Theor. Biol. **264**, 874 (2010).

[10] D. Zhou and H. Qian, Phys. Rev. E **84**, 031907 (2011).

[11] T. Herpich, J. Thingna, and M. Esposito, Phys. Rev. X **8**, 031056 (2018)

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

[13] D. Friedman, Quant. Finance **1**, 177 (2001).

[14] J. Knebel, M. F. Weber, T. Krüger, and E. Frey, Nat. Commun. **6**, 6977 (2015).

[15] C. Gokhale and A. Traulsen, Proc. Natl. Acad. Sci. U.S.A. **107**, 5500 (2010).

[16] P. A. P. Moran, *The Statistical Processes of Evolutionary Theory* (Clarendon Press, Oxford, 1962).

[17] M. A. Nowak, A. Sasaki, C. Taylor, and D. Fudenberg, Nature **428**, 646 (2004).

[18] L. Hindersin, B. Wu, A. Traulsen, and J. García, Sci. Rep. **9**, 6946 (2019).

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

[20] A mutation is a change of the strategy outside the adaptation processes.

[21] D. A. Levin and Y. Peres, *Markov Chains and Mixing Times* (AMS, 2017).

[22] B. Skyrms, J. Log. Lang. Inf. **1**, 11 (1992).

[23] Y. Sato, E. Akiyama, and J. D. Farmer, Proc. Natl. Acad. Sci. U.S.A. **99**, 4748 (2002).

[24] J. B. T. Sanders, J. D. Farmer, and T. Galla, Sci. Rep. **8**, 4902 (2018).

[25] A. Traulsen, J. C. Claussen, and C. Hauert, Phys. Rev. E **74**, 011901 (2006).

[26] For a population with N players and n possible strategies, the number of states is equal to the number of different ways to distribute N indistinguishable balls into n distinguishable boxes, that is, $C(N + n - 1, n - 1)$. In the case of two antagonistic populations playing against each other, this number squares. In the simplest case $n = 2$, with two populations of the size $N = 200$ each, we have to deal with a Markov chain with $200^2 = 40\,000$ states.

[27] C. W. Gardiner, *Handbook of Stochastic Methods* (Springer, New York, NY, 2004).

[28] A. Traulsen, J. C. Claussen, and C. Hauert, Phys. Rev. E **85**, 041901 (2012).

[29] P. Collet, S. Martínez, and J. S. Martín, *Quasi-stationary Distributions* (Springer, 2013).

[30] L. Arnold, *Random Dynamical Systems* (Springer, New York, NY, 2003).

[31] P. H. Baxendale, Probab. Theory Relat. Fields **99**, 581 (1994).

[32] A. Zakharova, T. Vadivasova, V. Anishchenko, A. Koseska, and J. Kurths, Phys. Rev. E **81**, 011106 (2010).

[33] G. Floquet, Ann. Ecole Norm. Sup. **12**, 47 (1883).

[34] M. Grifoni and P. Hänggi, Phys. Rep. **304**, 229 (1998).

[35] R. Dawkins, *The Selfish Gene* (Oxford University Press, Oxford, 1976).

[36] M. Andersson, *Sexual Selection* (Princeton University Press, Princeton, NJ, 1994).

[37] A. Qvarnström, T. Pärt, and B. C. Sheldon, Nature **405**, 344 (2000).

[38] R. N. C. Milner, T. Detto, M. D. Jennions, and P. R. Y. Backwell, Behav. Ecol. **21**, 311 (2010).

[39] A. A. Borg, E. Forsgren, and T. Amudsen, Anim. Behav. **72**, 763 (2006).

[40] K. U. Heubel and J. Schlupp, Behav. Ecol. **19**, 1080 (2008).

[41] D. Crews, Science **189**, 1059 (1975).

[42] M. M. Holmes, C. L. Bartrem, and J. Wade, Physiol. Behav. **91**, 601 (2007).

[43] These two consecutive steps, death and birth, can be reinterpreted as a single step of *adaptation*.[19]

[44] J. Hofbauer and K. H. Schlag, J. Evol. Econ. **10**, 523 (2000).

[45] We follow the convention that the stochastic matrix acts on the probability vector from the right.

[46] E. Seneta, *Non-negative Matrices and Markov Chains* (Springer, New York, NY, 2006).

[47] In order to find the maximal-eigenvalue eigenvector of the reduced transition matrix, we use *Python* routine scipy.sparse.linalg.eigs, https://docs.scipy.org/doc/scipy/reference/tutorial/arpack.html.

[48] H. Risken, *The Fokker-Planck Equation* (Springer, 1984).

[49] P. E. Kloeden and E. Platen, *Numerical Solution of Stochastic Differential Equations* (Springer, Berlin, 1992).

[50] J. von Neumann and O. Morgenstern, *Theory of Games and Economic Behaviour* (Princeton University Press, Princeton, NJ, 1944).

[51] L. Arnold, S. Namachchivaya, and K. R. Schenk-Hoppé, J. Bifur. Chaos Appl. Sci. Engrg. **6**, 1947 (1996).

[52] S. Denisov, O. Vershinina, and M. Ivanchenko, unpublished.

[53] The extinction time for a *specific* initial state $t_ext(i,j)$ is defined as the average number of rounds before the absorption to the boundary. The average extinction time t_ext is obtained by averaging $t_ext(i,j)$ over all initial states different from the boundary states.

[54] We were not able to avoid the crossing for $N = 400$ even with the integration step $dt = 10^{-5}$.

[55] P. E. Kloeden and E. Platen, *Numerical Solution of Stochastic Differential Equations* (Springer, 1992).

[56] A. Szolnoki *et al.*, J. R. Soc. Interface **11**, 20140735 (2014).

[57] G. Szabo and I. Borsos, Phys. Rep. **624**, 1 (2016).

[58] U. Dieckmann, P. Marrow, and R. Law, J. Theor. Biol. **176**, 91 (1995).

[59] L. A. Imhof, D. Fudenberg, and M. A. Nowak, Proc. Natl. Acad. Sci. U.S.A. **102**, 10797 (2005).

[60] J. Weitz *et al.*, Proc. Natl. Acad. Sci. U.S.A. **113**, E7518 (2016).

[61] A. J. McKane and T. J. Newman, Phys. Rev. Lett. **94**, 218102 (2005).

[62] F. Barraquand *et al.*, Ecol. Lett. **20**, 1074 (2017).

[63] J. H. Myers, Proc. R. Soc. B: Biol. Sci. **285**, 20172841 (2018).

[64] C. A. Klausmeier, Theor. Ecol. **1**, 153 (2008).

[65] G. J. Boender, A. A. de Koeijer, and E. A. J. Fischer, Acta Biotheor. **60**, 303 (2012).

[66] S. S. Golden, V. M. Cassone, and A. Li Wang, Nat. Struct. Mol. Biol. **14**, 362 (2007).

[67] A. Sancar, Nat. Struct. Mol. Biol. **15**, 23 (2008).

[68] M. M. Wolf, see http://www-m5.ma.tum.de/foswiki/pub/M5/Allgemeines/MichaelWolf/QChannelLecture for "Quantum Channels and Operations: Guided Tour."

[69] M. A. Nielsen and I. L. Chuang, *Quantum Computation and Quantum Information* (Cambridge University Press, 2000).

[70] K. Macieszczak, M. Guţă, I. Lesanovsky, and J. P. Garrahan, Phys. Rev. Lett. **116**, 240404 (2016).

[71] M. Hartmann, D. Poletti, M. Ivanchenko, S. Denisov, and P. Hänggi, New J. Phys. **19**, 083011 (2017).