Stochastic differential equations for evolutionary dynamics with demographic noise and mutations

Arne Traulsen,[1] Jens Christian Claussen,[2] and Christoph Hauert[3]

[1]*Evolutionary Theory Group, Max-Planck-Institute for Evolutionary Biology, August-Thienemann-Strasse 2, 24306 Plön, Germany*
[2]*Institute for Neuro- and Bioinformatics, University of Lübeck, 23562 Lübeck, Germany*
[3]*Department of Mathematics, The University of British Columbia, 1984 Mathematics Road, Vancouver, B.C., Canada V6T 1Z2*

We present a general framework to describe the evolutionary dynamics of an arbitrary number of types in finite populations based on stochastic differential equations (SDEs). For large, but finite populations this allows us to include demographic noise without requiring explicit simulations. Instead, the population size only rescales the amplitude of the noise. Moreover, this framework admits the inclusion of mutations between different types, provided that mutation rates μ are not too small compared to the inverse population size $1/N$. This ensures that all types are almost always represented in the population and that the occasional extinction of one type does not result in an extended absence of that type. For $\mu N \ll 1$ this limits the use of SDEs, but in this case there are well established alternative approximations based on time scale separation. We illustrate our approach by a rock-scissors-paper game with mutations, where we demonstrate excellent agreement with simulation based results for sufficiently large populations. In the absence of mutations the excellent agreement extends to small population sizes.

I. INTRODUCTION

Populations evolve when different individuals have different traits or strategies that determine their reproductive success. In population genetic models, this success is typically constant, whereas in evolutionary game theory the fitness of an individual depends on interactions with other members of the population. Consequently, fitness depends on the relative proportions (or frequencies) of different strategic types and hence gives rise to so called frequency dependent selection. More successful strategies increase in abundance and may either take over the entire population or, as the strategy becomes increasingly common, may suffer from a decrease in fitness, which can result in the coexistence of two or more traits within the population. For example, in host-parasite coevolution, a rare parasite may be most successful, but once it reaches high abundance, selection pressure on the host increases and the host is likely to develop some defense mechanism. Consequently, the success of the parasite decreases. Such dynamics can be conveniently described by evolutionary game theory [1–6].

Traditionally, the mathematical description of evolutionary game dynamics is formulated in terms of the deterministic replicator equation [2]. This implies that population sizes are infinite and populations are unstructured. Only more recently the stochastic dynamics in finite populations attracted increasing attention [7,8]. Typically, the stochastic approach becomes deterministic in the limit of infinite populations. For large but finite populations the dynamics can be approximated by stochastic differential equations (SDEs) [8–10]. This analytic approach is a natural extension of the replicator dynamics, which is capable of bridging the gap between deterministic models and individual based simulations. Moreover, SDEs are typically computationally far less expensive than simulations because the execution time does not scale with population size.

II. MASTER EQUATION

In unstructured, finite populations of constant size N, consisting of d distinct strategic types and with a mutation rate μ, evolutionary changes can be described by the following class of birth-death processes: In each time step, one individual of type j produces a single offspring and displaces another randomly selected individual of type k. With probability $1 - \mu$, no mutation occurs and j produces an offspring of the same type. But with probability μ, the offspring of an individual of type i ($i \neq j$) mutates into a type j individual. This results in two distinct ways to increase the number of j types by one at the expense of decreasing the number of k types by one, hence keeping the population size constant. Biologically, keeping N constant implies that the population has reached a stable ecological equilibrium and assumes that this equilibrium remains unaffected by trait frequencies. The probability for the event of replacing a type k individual with a type j individual is denoted by T_{kj} and is a function of the state of the population $X = (X_1, X_2, \ldots, X_d)$, with X_n indicating the number of individuals of type n such that $\sum_{n=1}^{d} X_n = N$.

For this process it is straightforward to write down a Master equation [10] and, at this point, there is no need to further specify the transition probabilities T_{kj}:

$$P^{\tau+1}(X) = P^{\tau}(X) + \sum_{j,k=1}^{d} P^{\tau}(X_j^k) T_{kj}(X_j^k) - P^{\tau}(X) T_{jk}(X), \quad (1)$$

where $P^{\tau}(X)$ denotes the probability of being in state X at time τ and $X_j^k = (X_1, \ldots, X_j - 1, \ldots, X_k + 1, \ldots, X_d)$ represents a state adjacent to X.

III. FOKKER-PLANCK AND LANGEVIN EQUATIONS

While the master equation (1) is rather unwieldy, a Kramers-Moyal expansion yields a convenient approximation

for large but finite N in the form of a Fokker-Planck equation [11]

$$\dot{\rho}(\boldsymbol{x}) = -\sum_{k=1}^{d-1} \frac{\partial}{\partial x_k} \rho(\boldsymbol{x}) \mathcal{A}_k(\boldsymbol{x}) + \frac{1}{2} \sum_{j,k=1}^{d-1} \frac{\partial^2}{\partial x_k \partial x_j} \rho(\boldsymbol{x}) \mathcal{B}_{jk}(\boldsymbol{x}),$$
(2)

where $\boldsymbol{x} = \boldsymbol{X}/N$ represents the state of the population in terms of frequencies of the different strategic types and $\rho(\boldsymbol{x})$ is the probability density in state $\boldsymbol{x}$. Due to the normalization $\sum_{k=1}^{d} x_k = 1$, it suffices to consider $d-1$ elements of the deterministic drift vector $\mathcal{A}_k(\boldsymbol{x})$, $k = 1, \ldots, d-1$. Similarly, we only need to consider a diffusion matrix $\mathcal{B}_{jk}(\boldsymbol{x})$ with dimension $(d-1) \times (d-1)$, that is, $j,k = 1, \ldots, d-1$. The drift vector $\mathcal{A}_k(\boldsymbol{x})$ is given by

$$\mathcal{A}_k(\boldsymbol{x}) = \sum_{j=1}^{d} [T_{jk}(\boldsymbol{x}) - T_{kj}(\boldsymbol{x})] = -1 + \sum_{j=1}^{d} T_{jk}(\boldsymbol{x}). \quad (3)$$

For the second equality we have used $\sum_{j=1}^{d} T_{kj}(\boldsymbol{x}) = 1$, which simply states that a k-type individual transitions to some other type (including staying type k) with probability one. $\mathcal{A}_k(\boldsymbol{x})$ is bounded in $[-1, d-1]$ because the T_{jk} are probabilities.

The diffusion matrix $\mathcal{B}_{jk}(\boldsymbol{x})$ is defined as

$$\mathcal{B}_{jk}(\boldsymbol{x}) = -\frac{1}{N} [T_{jk}(\boldsymbol{x}) + T_{kj}(\boldsymbol{x})] \quad \text{for} \quad j \neq k, \quad (4a)$$

$$\mathcal{B}_{jj}(\boldsymbol{x}) = \frac{1}{N} \left\{ \sum_{l=1, l \neq j}^{d} [T_{jl}(\boldsymbol{x}) + T_{lj}(\boldsymbol{x})] \right\}. \quad (4b)$$

For a detailed derivation see, for example, [10] or [12]. Note that the diffusion matrix is symmetric, $\mathcal{B}_{jk}(\boldsymbol{x}) = \mathcal{B}_{kj}(\boldsymbol{x})$ and vanishes as $\sim 1/N$ in the limit $N \to \infty$. Moreover, we have for all diagonal elements $\mathcal{B}_{jj}(\boldsymbol{x}) \geqslant 0$ and for the nondiagonal elements $\mathcal{B}_{jk}(\boldsymbol{x}) \leqslant 0$. Note that $T_{jj}(\boldsymbol{x})$ cancels in Eq. (3) and does not appear in Eq. (4)—it is thus of no further concern.

The noise of our underlying process is uncorrelated in time and hence the Itô calculus [11] can be applied to derive a Langevin equation, which represents in our case a stochastic replicator-mutator equation,

$$\dot{x}_k = \mathcal{A}_k(\boldsymbol{x}) + \sum_{j=1}^{d-1} \mathcal{C}_{kj}(\boldsymbol{x}) \xi_j(t), \quad (5)$$

where the $\xi_j(t)$ represent uncorrelated Gaussian white noise with unit variance $\langle \xi_k(t) \xi_j(t') \rangle = \delta_{kj} \delta(t - t')$. The matrix $\mathcal{C}(\boldsymbol{x})$ is defined by $\mathcal{C}^T(\boldsymbol{x}) \mathcal{C}(\boldsymbol{x}) = \mathcal{B}(\boldsymbol{x})$ and its off-diagonal elements are responsible for correlations in the noise of different strategic types.

Fluctuations arising in finite populations are approximated by the stochastic term in the Langevin equation (5). For given transition probabilities the matrix $\mathcal{C}(\boldsymbol{x})$ provides a quantitative description of the fluctuations introduced by microscopic processes in finite populations. In the limit $N \to \infty$ the matrix $\mathcal{C}(\boldsymbol{x})$ vanishes with $\sim 1/\sqrt{N}$ and we recover a deterministic replicator-mutator equation. Note that the replicator equation does not impose an upper or lower bound on $\mathcal{A}$ [cf. Eq. (3)]. However, this difference merely amounts to a (constant) rescaling of time.

The multiplicative character of the noise and its strategy-strategy correlations are determined by the form of the matrix $\mathcal{C}(\boldsymbol{x})$. In order to determine $\mathcal{C}(\boldsymbol{x})$ we first diagonalize $\mathcal{B}(\boldsymbol{x})$. Because $\mathcal{B}(\boldsymbol{x})$ is real and symmetric it is diagonalizable by an orthogonal matrix $\mathcal{U}(\boldsymbol{x})$ with $\mathcal{U}(\boldsymbol{x})\mathcal{U}^T(\boldsymbol{x}) = \boldsymbol{1}$, where $\boldsymbol{1}$ denotes the identity matrix. Moreover, the normalized eigenvectors $\boldsymbol{f}_i(\boldsymbol{x})$ of $\mathcal{B}(\boldsymbol{x})$ form an orthonormal basis $\boldsymbol{f}_i(\boldsymbol{x}) \cdot \boldsymbol{f}_j(\boldsymbol{x}) = \delta_{ij}$. Thus we can construct the transformation matrix $\mathcal{U}(\boldsymbol{x}) = [\boldsymbol{f}_1(\boldsymbol{x}), \ldots, \boldsymbol{f}_{d-1}(\boldsymbol{x})]$ such that $\mathcal{B}(\boldsymbol{x}) = \mathcal{U}(\boldsymbol{x})\Lambda(\boldsymbol{x})\mathcal{U}^T(\boldsymbol{x})$, where $\Lambda(\boldsymbol{x})$ is a diagonal matrix with the eigenvalues $\lambda_i(\boldsymbol{x})$ of $\mathcal{B}(\boldsymbol{x})$ along its diagonal. From Eq. (4) follows that $\mathcal{B}(\boldsymbol{x})$ is positive definite and hence all eigenvalues $\lambda_i(\boldsymbol{x})$ are positive. Here we tacitly assume that all T_{jk} are nonzero; if certain transitions are excluded, $\mathcal{B}(\boldsymbol{x})$ is positive semidefinite and eigenvalues can be zero.

Finally, our matrix $\mathcal{C}(\boldsymbol{x})$ is given by

$$\mathcal{C}(\boldsymbol{x}) = \mathcal{U}(\boldsymbol{x}) \begin{pmatrix} \sqrt{\lambda_1} & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \sqrt{\lambda_{d-1}} \end{pmatrix} \mathcal{U}^T(\boldsymbol{x}). \quad (6)$$

This standard procedure to diagonalize matrices can be easily implemented numerically. However, the diffusion matrix $\mathcal{B}(\boldsymbol{x})$ obviously depends on the transition probabilities and thus on the abundances of all strategies. Consequently, the procedure to calculate $\mathcal{C}(\boldsymbol{x})$ must be continuously repeated as time progresses and the state $\boldsymbol{x}$ changes. This is computationally inconvenient and therefore it is desirable to calculate $\mathcal{C}(\boldsymbol{x})$ analytically.

The simplest case with $d = 2$ strategic types results in a one-dimensional Fokker-Planck equation [8]. Moreover, in special cases, for example in cyclic games such as the symmetric rock-scissors-paper game, the Fokker-Planck equation (2) can be approximated in polar coordinates [13]. However, such cases are nongeneric and here we focus on general, higher dimensional situations with $d \geqslant 3$. As a particular example to illustrate the framework numerically, we provide a detailed analysis of a generic rock-scissors-paper game.

IV. MULTIPLICATIVE NOISE FOR SPECIFIC PROCESSES

In general, the diffusion matrix $\mathcal{B}(\boldsymbol{x})$ depends not only on the frequencies $\boldsymbol{x}$ but also on the payoffs (fitness) of the different strategic types. In this case, an analytic representation of $\mathcal{C}(\boldsymbol{x})$ is only of limited use because it would be valid just for one particular game. However, $\mathcal{B}(\boldsymbol{x})$ becomes payoff independent if $T_{jk}(\boldsymbol{x}) + T_{kj}(\boldsymbol{x})$ is payoff independent. Fortunately this holds for the broad and relevant class of pairwise comparison processes. In these processes, a focal individual f and a model m with payoffs π_f and π_m are picked at random and a payoff comparison determines whether the focal individual switches its strategy.

Let $\gamma(\pi_f, \pi_m)$ be the probability that the focal individual adopts the strategy of the model (see, e.g., [14,15]) and assume that every mutation leads to a *different* strategy. Then the transition probabilities from type k to type j read

$$T_{kj}(\boldsymbol{x}) = (1 - \mu)x_k x_j \gamma(\pi_j, \pi_k) + \mu x_k \frac{1}{d-1} \quad (7)$$

for $j \neq k$. Consequently, any pairwise comparison process with

$$\gamma(\pi_j, \pi_k) + \gamma(\pi_k, \pi_j) = \text{const.} \qquad (8)$$

leads to a payoff independent diffusion matrix $\mathcal{B}(\boldsymbol{x})$. If Eq. (8) is fulfilled, the noise term in Eq. (5) is independent of the evolutionary game. In particular, it allows the consideration of multiplayer games in which the payoff functions are nonlinear [16]. Examples for evolutionary processes that fulfill Eq. (8) include the local update process [8] with $\gamma(\pi_j, \pi_k) = \frac{1}{2} + w(\pi_j - \pi_k)$, where w indicates the strength of selection acting on payoff differences between different strategic types. For $w \ll 1$ selection is weak, payoff differences amount to little changes in fitness and the process is dominated by random updating. For larger w selection strength increases but an upper limit is imposed on w by the requirement $0 \leqslant \gamma(\pi_j, \pi_k) \leqslant 1$. Another example is the Fermi process [17–19] with $\gamma(\pi_j, \pi_k) = \{1 + \exp[-w(\pi_j - \pi_k)]\}^{-1}$. Again w indicates the selection strength but without an upper bound. In the limit $w \to \infty$, that is, $\gamma(\pi_j, \pi_k) = \Theta[\pi_j - \pi_k]$, where $\Theta[x]$ denotes the Heaviside step function, the imitation dynamics [20,21] is recovered. A further example, where $\gamma(\pi_j, \pi_k)$ does not simply depend on the difference between its arguments, is $\gamma(\pi_j, \pi_k) = \frac{\pi_j}{\pi_j + \pi_k}$ [7,22]. Incidentally, all examples above satisfy $\gamma(\pi_j, \pi_k) + \gamma(\pi_k, \pi_j) = 1$ but, for example, there might be resilience to change in the local update process such that $\gamma(\pi_j, \pi_k) = \alpha[\frac{1}{2} + w(\pi_j - \pi_k)]$ with $0 < \alpha < 1$ and hence $\gamma(\pi_j, \pi_k) + \gamma(\pi_k, \pi_j) = \alpha < 1$.

Last but not least, an example of an important process that does not lead to a payoff independent $\mathcal{B}(\boldsymbol{x})$ is given by the standard frequency dependent Moran process [7,23,24] or its linearized equivalent [12,25].

In the following we concentrate on cases with $\gamma(\pi_j, \pi_k) + \gamma(\pi_k, \pi_j) = 1$. This leads to

$$\mathcal{B}_{jk}(\boldsymbol{x}) = \frac{1}{N}\left[-x_j x_k (1-\mu) - \frac{\mu}{d-1}(x_j + x_k) \right] \quad \text{for } j \neq k \qquad (9)$$

and

$$\mathcal{B}_{jj}(\boldsymbol{x}) = \frac{1}{N}\left\{ x_j(1-x_j)(1-\mu) + \frac{\mu}{d-1}[1 + x_j(d-2)] \right\}. \qquad (10)$$

Unfortunately, even in this case, a full derivation of $\mathcal{C}(\boldsymbol{x})$ is difficult for general d. Thus we focus on the more manageable but highly illustrative case of $d = 3$. Henceforth, we set $x_1 = x$ and $x_2 = y$ (and $x_3 = 1 - x - y$) for convenience.

A. No mutations, $\mu = 0$

For $d = 3$ and in the absence of mutations, we can give a relatively compact analytic expression for $\mathcal{C}(\boldsymbol{x})$,

$$\mathcal{B}(\boldsymbol{x}) = \frac{1}{N}\begin{pmatrix} x(1-x) & -xy \\ -xy & y(1-y) \end{pmatrix}. \qquad (11)$$

The eigenvalues of this matrix are

$$\lambda_{\pm}(\boldsymbol{x}) = \frac{K_{\pm} \pm L}{2N}, \qquad (12)$$

where $K_{\pm} = x(1-x) \pm y(1-y)$ and $L = \sqrt{K_-^2 + 4x^2y^2}$. The eigenvectors are given by

$$\boldsymbol{f}_{\pm}(\boldsymbol{x}) = \frac{1}{\mathcal{N}}(-K_- \pm L, 2xy), \qquad (13)$$

where $\mathcal{N} = \sqrt{4x^2y^2 + (-K_- \pm L)^2}$ is a normalization factor. From this, it is straightforward to construct the transformation matrix $\mathcal{U}(\boldsymbol{x}) = [\boldsymbol{f}_+(\boldsymbol{x}), \boldsymbol{f}_-(\boldsymbol{x})]$. The product $\mathcal{U}(\boldsymbol{x})\sqrt{\Lambda(\boldsymbol{x})}\,\mathcal{U}^T(\boldsymbol{x})$ then yields our matrix $\mathcal{C}(\boldsymbol{x})$, which can be written as

$$\mathcal{C}_{xx}(\boldsymbol{x}) = \frac{1}{\sqrt{2N}}\left(\frac{\sqrt{Q_{+-}}}{1 + \left(\frac{2xy}{Q_{--}}\right)^2} + \frac{\sqrt{Q_{++}}}{1 + \left(\frac{2xy}{Q_{-+}}\right)^2} \right), \qquad (14a)$$

$$\mathcal{C}_{xy}(\boldsymbol{x}) = \mathcal{C}_{yx} = \frac{xy}{\sqrt{2N}L}(\sqrt{Q_{+-}} - \sqrt{Q_{++}}), \qquad (14b)$$

$$\mathcal{C}_{yy}(\boldsymbol{x}) = \frac{1}{\sqrt{2N}}\left(\frac{\sqrt{Q_{+-}}}{1 + \left(\frac{2xy}{Q_{-+}}\right)^2} + \frac{\sqrt{Q_{++}}}{1 + \left(\frac{2xy}{Q_{--}}\right)^2} \right), \qquad (14c)$$

where $Q_{\pm\pm} = K_{\pm} \pm L$. It is obvious that $\mathcal{C}_{xx}(\boldsymbol{x}) \geqslant 0$ and $\mathcal{C}_{yy}(\boldsymbol{x}) \geqslant 0$. Since $L \geqslant 0$, we have $Q_{++} \geqslant Q_{+-}$ and therefore $\mathcal{C}_{xy}(\boldsymbol{x}) \leqslant 0$. It is remarkable that even for this simple $\mathcal{B}(\boldsymbol{x})$, the matrix $\mathcal{C}(\boldsymbol{x})$ already takes a rather complicated form that is not easy to interpret. Therefore, the elements $\mathcal{C}_{xx}(\boldsymbol{x})$ and $\mathcal{C}_{xy}(\boldsymbol{x})$ are plotted in Fig. 1. The remaining element $\mathcal{C}_{yy}(\boldsymbol{x})$ follows from the symmetry of the matrix, that is, in Figs. 1(a) and 1(b), the simplex needs to be mirrored along the vertical axis $x = y$. Recall that the matrix $\mathcal{C}(\boldsymbol{x})$ is independent of the evolutionary game that is played and does not depend on the microscopic evolutionary process, as long as Eq. (8), $\gamma(\pi_j, \pi_k) + \gamma(\pi_k, \pi_j) = \text{const.}$, holds. While analytical calculations of $\mathcal{C}$ are also, in principle, feasible for $d = 4$ and $d = 5$, they lead to very lengthy expressions.

B. With mutations, $\mu > 0$

The procedure to derive $\mathcal{C}(\boldsymbol{x})$ remains the same when including mutations just starting with

$$\mathcal{B}(\boldsymbol{x}) = \frac{1-\mu}{N}\begin{pmatrix} x(1-x) & -xy \\ -xy & y(1-y) \end{pmatrix} + \frac{\mu}{2N}\begin{pmatrix} 1+x & -x-y \\ -x-y & 1+y \end{pmatrix}. \qquad (15)$$

Unfortunately, the analytic expressions for the elements of $\mathcal{C}(\boldsymbol{x})$ grow to unwieldy proportions. A graphical illustration of the elements of $\mathcal{C}(\boldsymbol{x})$ for $\mu > 0$ is shown in Fig. 1. Compared to $\mu = 0$, the entries of $\mathcal{C}_{xx}(\boldsymbol{x})$ become larger and the noise terms no longer vanish at the boundaries. For example, at the corners of the simplex we have for $\mathcal{C}_{xx}(\boldsymbol{x})$,

$$\mathcal{C}_{xx}(1,0,0) = \frac{3}{\sqrt{10}}\sqrt{\frac{\mu}{N}},$$

$$\mathcal{C}_{xx}(0,1,0) = \sqrt{\frac{2}{5}}\sqrt{\frac{\mu}{N}},$$

$$\mathcal{C}_{xx}(0,0,1) = \frac{1}{\sqrt{2}}\sqrt{\frac{\mu}{N}}.$$
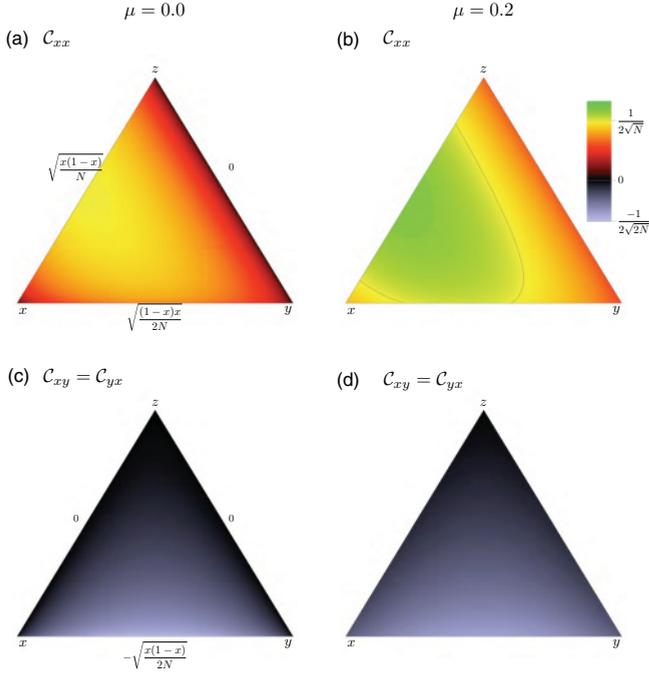
FIG. 1. (Color online) Elements of the noise matrix $\mathcal{C}(x, y, z)$ for processes fulfilling Eq. (8) with $d = 3$ strategies for $\mu = 0$ (left) and $\mu > 0$ (right). (a) The element $\mathcal{C}_{xx}$ determines how the noise in the x direction affects the x coordinate. In the case of $\mu = 0$, this noise vanishes for $x \to 0$. For $y \to 0$ and $z \to 0$ we recover the usual multiplicative noise from one-dimensional evolutionary processes [8]. (b) For nonvanishing mutations $\mathcal{C}_{xx}$ increases at $x \to 0$ and $x \to 1$ and exceeds the value $\frac{1}{2\sqrt{N}}$ (contour line) in the interior of the simplex, which is the maximum value without mutations. The element $\mathcal{C}_{yy}$ follows from the transformation $x \leftrightarrow y$. (c) The element $\mathcal{C}_{xy}$ determines how the noise in the x direction affects the y coordinate (or vice versa). This term is always negative, as explained in the text. For $x \to 0$ or $y \to 0$, $\mathcal{C}_{xy}$ vanishes in the case of $\mu = 0$. For $z \to 0$ we find $\mathcal{C}_{xy} = -\mathcal{C}_{xx}$, which ensures that the sum of the noise in the x coordinate $\mathcal{C}_{xx} + \mathcal{C}_{xy}$ vanishes on this edge of the simplex: If the noise increases x for $z = 0$ it has to decrease y by the same amount. (d) In the case of $\mu > 0$, the noise no longer vanishes at the boundaries of the simplex but the expressions are too cumbersome to display explicitly.

For $\mathcal{C}_{xy}(\boldsymbol{x})$ we obtain at these points

$$\mathcal{C}_{xy}(1,0,0) = \mathcal{C}_{xy}(0,1,0) = -\frac{1}{\sqrt{10}}\sqrt{\frac{\mu}{N}},$$

$$\mathcal{C}_{xy}(0,0,1) = 0.$$

In addition, the functional form of the noise term is altered with increasing μ and becomes significantly more complex than in the case of no mutations, $\mu = 0$.

For $\mu > 0$ it is important to mention that for small mutation rates serious mathematical intricacies arise in the vicinity of absorbing boundaries and saddle-node fixed points [26,27]. Intuitively, the reason for these complications arises from the diffusion approximation that underlies the derivation of the Fokker-Planck equation (2). In systems with such absorbing boundaries (or subspaces where one or more strategic types are absent) finite populations spend nonnegligible time on these boundaries in the limit $\mu \ll 1/N$. In contrast, the diffusion

approximation is based on a continuum approximation, which reflects the limit $N \to \infty$, and adds finite size corrections for large, finite N. For example, the state of the population $\boldsymbol{x}$ is a continuous variable and hence can be located arbitrarily close to a boundary in both the Fokker-Planck or Langevin formalisms, Eqs. (2) and (5), but this is impossible in finite populations. More specifically, this implies that the case of $\mu \ll 1/N$, for which the population could get trapped on an absorbing boundary for an extended period of time, cannot be captured by the diffusion approximation. Consequently, the quality of the approximation is expected to decrease for small $\mu > 0$ and to get worse if N is small too. Interestingly, this failure of the continuum approximation has not been raised in population genetics, where similar considerations have been made, but typically the description is only made on the level of the Fokker-Planck equation and not based on stochastic differential equations [28]. However, population geneticists are typically interested in simpler scenarios, where each type has a fixed fitness. In this case, the corresponding deterministic system has no generic trajectories that are close to absorbing boundaries. Moreover, the usual diffusion approximation in population genetics, where the selection intensity scales to zero while the population size diverges, leads to substantial noise, such that these effects have a minor impact.

For evolutionary games with small mutation rates one is typically not forced to apply this relatively complex approximation. Instead, a time scale separation between mutation and selection allows to approximate the dynamics based on a pairwise consideration of strategies by considering an embedded Markov chain over the (quasiabsorbing) homogenous states of the population [29–33].

V. APPLICATION TO THE ROCK-PAPER-SCISSORS GAME

To compare our approach based on stochastic differential equations in $d - 1$ dimensions to individual based simulations with d strategies, we focus on the case of $d = 3$ where we have obtained closed analytical results above for $\mu = 0$. As an example, we consider the cyclic dynamics in the rock-scissors-paper game, which is not only a popular children's game, but also relevant in biological [34–37] and social systems [38,39]. Moreover, the evolutionary dynamics of this game is theoretically very well understood, both in infinite as well as in finite populations [2,13,14,40–42]. Here we focus on a generic rock-scissors-paper game with payoff matrix

$$
\begin{array}{c}
\begin{array}{ccc} R & S & P \end{array} \\
\begin{array}{c} R \\ S \\ P \end{array}
\begin{pmatrix}
0 & \frac{s}{2} & -1 \\
-1 & 0 & 2+s \\
\frac{1+s}{3} & -1 & 0
\end{pmatrix}
\end{array}
\tag{16}
$$

which avoids artificial symmetries. According to the replicator equation, the game exhibits saddle node fixed points at $x = 1$, $y = 1$, and $z = 1 - x - y = 1$ as well as an interior fixed point at $\hat{\boldsymbol{x}} = (\frac{1}{2}, \frac{1}{3}, \frac{1}{6})$, independent of the parameter s but s controls the stability of $\hat{\boldsymbol{x}}$. For $s > 1$, $\hat{\boldsymbol{x}}$ is a stable focus and an unstable focus for $s < 1$. We do not consider the nongeneric case of $s = 1$, which exhibits closed orbits [2].
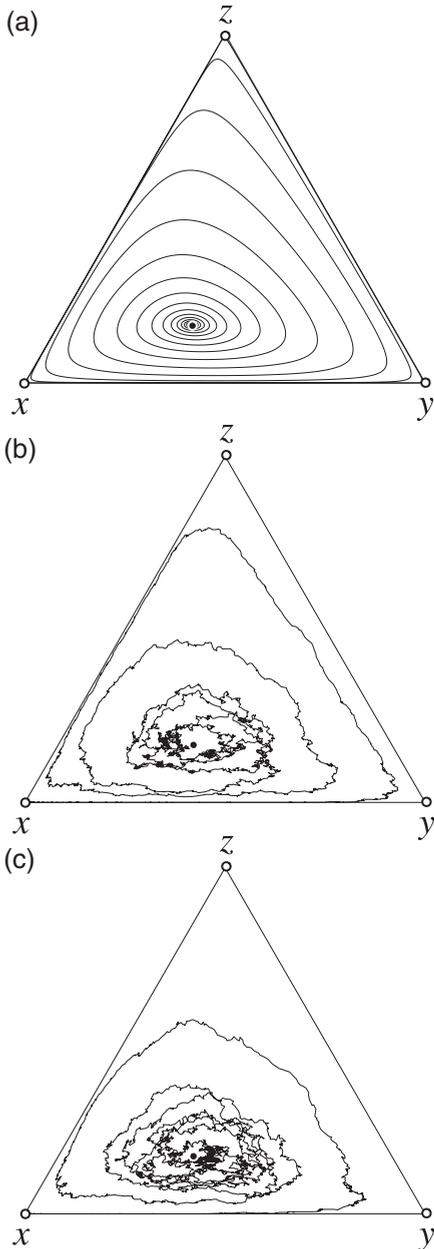
FIG. 2. Comparison of trajectories for the (a) deterministic case (replicator equation), (b) stochastic differential equation [Eq. (5)], and (c) individual based simulations for an RSP interaction with weakly attracting interior fixed point $\hat{x}$ with $s = 1.4$ [cf. Eq. (16)] and $N = 1000$ [in (b) and (c)]. Stochastic trajectories start close to $\hat{x}$ and end at time $T = 349.9$ (b) and $T = 395.8$ (c). Simulations can also be performed online at [43].

Let us first analyze the fixation probabilities and the fixation times for the case without mutations $\mu = 0$ for a game with $s = 1.4$, such that $\hat{x}$ is an attractor. The replicator equation, obtained in the limit $N \to \infty$, predicts that fixation never occurs and that the population evolves toward $\hat{x}$, see Fig. 2(a). In contrast, in the stochastic system ultimately two strategies are lost, see Figs. 2(b) and 2(c). Most importantly, the sample trajectories generated by the stochastic differential equation (5) in Fig. 2(b) do not exhibit any qualitative differences when compared to individual based simulations [Fig. 2(c)]. As
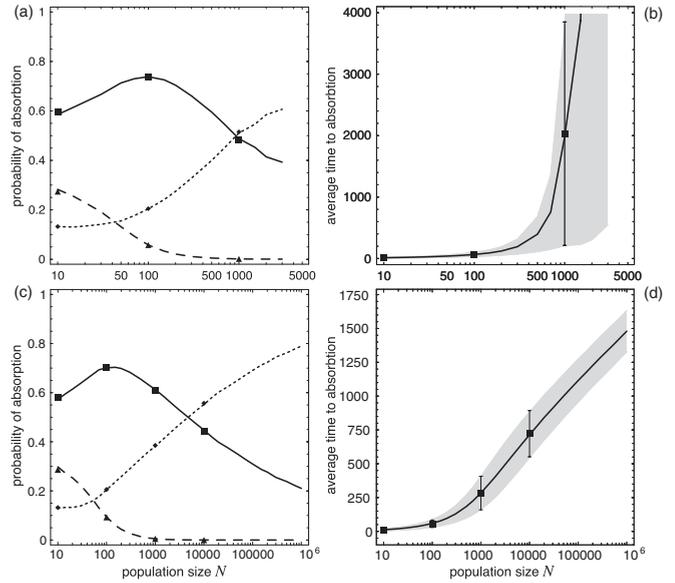
FIG. 3. Probability of absorption and time to absorption (in the absence of mutations, $\mu = 0$). (a) and (c) Probability to reach one of the three absorbing homogenous states of *rock* (solid line, ■), *scissors* (dashed line, ▲), or *paper* (dotted line, ♦) as well as (b) and (d) the associated time to absorption for *rock* as a function of the population size N with a stable (top row, $s = 1.4$) and unstable (bottom row, $s = 0.8$) interior fixed point $\hat{x}$. If $\hat{x}$ is an attractor [(a) and (b)] it may take exceedingly long times for large N until the population reaches an absorbing state—even though this will inevitably occur. Because of this results are only shown up to $N = 3000$. No such limitations occur if $\hat{x}$ is a repellor. The symbols indicate results from individual based simulations. Error bars and gray shaded areas indicate the standard deviation of the mean for simulations and stochastic calculations, respectively. Simulation results and results based on the stochastic differential equation [Eq. (5)] were both averaged over 10^5 independent runs.

a complementary scenario, we consider a game with $s = 0.8$, such that $\hat{x}$ is a repellor. According to the replicator equation trajectories now spiral away from $\hat{x}$ toward the boundaries of the simplex and approach a heteroclinic cycle. The probabilities that the population reaches any one of the three homogenous absorbing states in either scenario is shown in Fig. 3 together with the average time to fixation.

In all cases excellent agreement between the Langevin framework (5) and individual based simulations is obtained. Note that larger N not only improve the approximation, Eq. (5), but also result in a performance gain as compared to individual based simulations. More specifically, the computational effort scales quadratically with population size for simulations but remains constant when integrating Eq. (5) numerically (see Fig. 5(b)).

For nonvanishing mutation rates, $\mu > 0$, absorption is no longer possible and the boundary of the simplex becomes repelling. Figure 4 depicts the average frequencies of each strategic type as a function of μ. For small μ the population can still get trapped for considerable time along the boundary, which results in systematic deviations between the stochastic differential equation (5) and individual based simulations [Fig. 4]. More specifically, for $\mu < 1/N$ the deviations in-
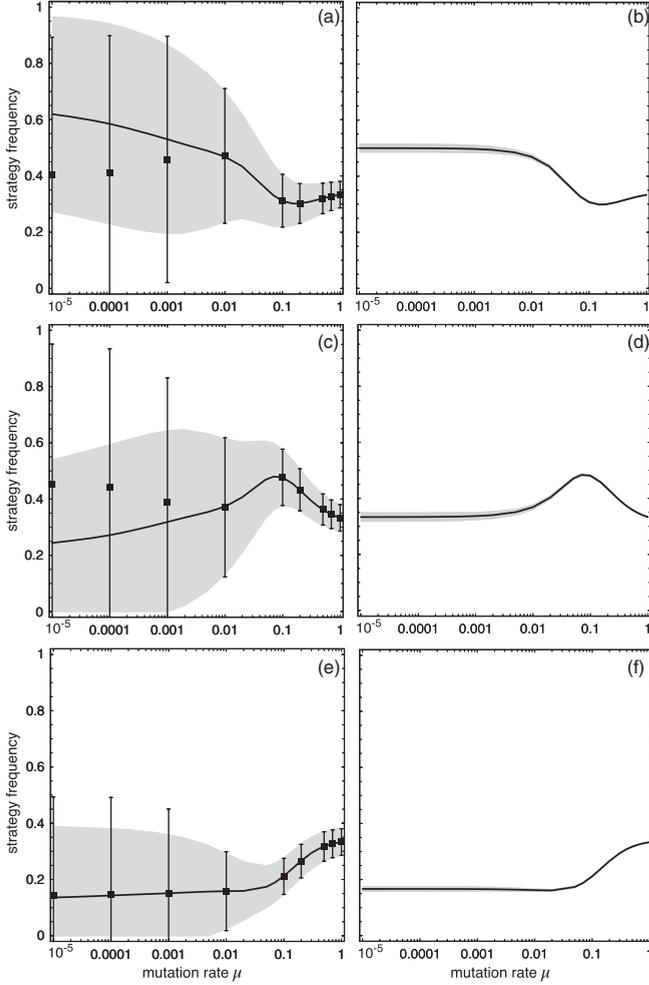
FIG. 4. Average frequency (solid line) and standard deviation (gray shaded area) of the three strategies *rock* [(a) and (b)], *scissors* [(c) and (d)], and *paper* [(e) and (f)] as a function of the mutation rate μ for small (left column, $N = 100$) and large populations (right column, $N = 10^5$) with a stable interior fixed point $\hat{x}$, $s = 1.4$, based on Eq. (5). For comparison, symbols and error bars depict results from individual based simulations. No simulation results are shown for $N = 10^5$ because of prohibitive computational efforts. Simulations and Eq. (5) are in excellent agreement for $\mu \gtrsim 1/N$ but for smaller μ substantial deviations occur (see text for details). Stochastic fluctuations decrease with increasing N and for $N = 10^5$ the population spends most of the time in the close vicinity of $\hat{x}$. Simulation results and results based on the stochastic differential equation [Eq. (5)] were both averaged over 10^{10} time steps after a relaxation time of 10^6 steps ($dt = 0.01$ for the SDE).

crease with decreasing μ but the agreement remains excellent for $\mu > 1/N$. This threshold simply means that for larger μ the time spent along the boundary can be neglected. In the limit $N \to \infty$ the replicator-mutator Eq. (5) exhibits a stable limit cycle [44].

VI. DISCUSSION

Evolutionary dynamics can be implemented in multiple ways. In particular, if only two types are present, the dynamics reduces to a single dimension and is typically solvable
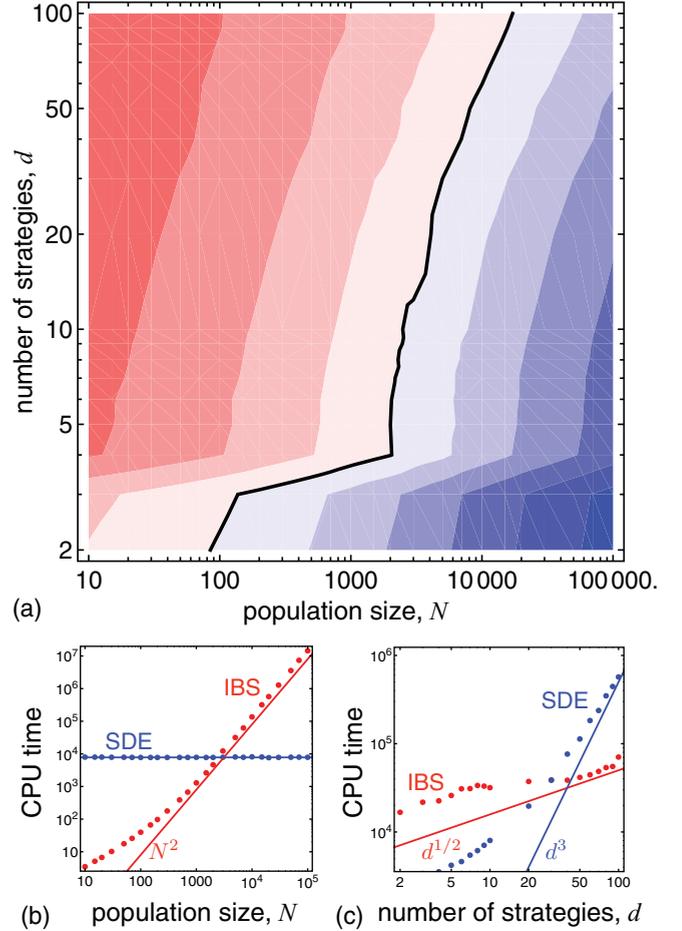
FIG. 5. (Color online) Performance comparison of individual based simulations (IBSs) vs stochastic differential equations (SDEs). (a) Ratio of the CPU times $CPU_{\text{SDE}}/CPU_{\text{IBS}}$ as a function of the population size N and the number of strategic types d. The bold contour indicates equal performance. For small N and large d IBSs are faster (left of solid line), but for larger N and smaller d SDEs are faster (right of solid line). Each contour indicates a performance difference of one order of magnitude. Note that IBSs for $d = 2, 3$ are based on analytical calculations of the eigenvalues and eigenvectors of $\mathcal{B}(\boldsymbol{x})$, which requires substantially less time and hence explains the discontinuity in the contours. (b) Computational time with $d = 10$ as a function of N for IBSs and SDEs. As a reference for the scaling N^2 and a constant are shown. (c) Computational time with $N = 5000$ as a function of d for IBSs and SDEs. As a reference for the scaling $d^{1/2}$ and d^3 are shown. For a proper scaling argument much larger d are required but already $d = 100$ far exceeds typical evolutionary models and hence is only of limited relevance in the current context. All comparisons use a constant payoff matrix and the local update process [8] [such that $\mathcal{A}_k(\boldsymbol{x}) = 0$ and $\gamma(\pi_j, \pi_k) = 1/2$], a mutation rate of $1/N$ and are based on at least 1000 time steps as well as at least 1 min running time. CPU time is measured in milliseconds required to calculate 1000 time steps. The time increment for the SDEs is $dt = 0.01$.

analytically, even if the population is finite and demographic noise is present. If more than two types are present, it is significantly more challenging to describe the dynamics analytically. When the mutation rates are sufficiently small [33], there are typically at most two types present at any

time, thus leading to situations which justify simpler tools based on the interaction of two types. In the limit of infinite populations, deterministic differential equations arise and can be used to describe the system even when the population size is large, but finite. However, it is more challenging to incorporate noise in this case. Here we have proposed a way to address this issue by deriving a Langevin equation for more than two types. For large populations, it is typically much more efficient to solve these equations numerically than to resort to individual based simulations. A detailed performance comparison is shown in Fig. 5. Computational costs of simulations increase slowly with the number of strategic types $\sim d^{1/2}$, but increase with the population size as N^2. In contrast, SDEs are essentially unaffected by N, but computational costs arise from repeatedly solving for eigenvalues and eigenvectors of $\mathcal{B}(\boldsymbol{x})$. In theory, analytical solutions are available for $d \leqslant 5$ but are probably meaningful in practice only for $d = 2,3$ and numerical methods are required for $d > 3$, which scale with d^3 [45]. To illustrate this, for the data point $N = 10\,000$ in Figs. 3(c) and 3(d) the simulations required approximately two months (1418 hours) to complete as compared to 49 min for the corresponding calculation based on stochastic differential equations, which corresponds to a 1700-fold performance gain. However, if only a small number of one type is present in a population, our approach does not work well unless there are no mutations or the product of the population size and the mutation rate is sufficiently high, $N\mu > 1$, such that the time spent along the boundary becomes negligible [cf. Fig. 4].

In summary, our approach establishes a transparent link between deterministic models and individual-based simulations of evolutionary processes. Moreover, the resulting stochastic differential equations provide substantial speed-up compared to simulations, and therefore may serve well in the investigation of multidimensional evolutionary dynamics.

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

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

[3] M. A. Nowak, *Evolutionary Dynamics* (Harvard University Press, Cambridge, MA, 2006).

[4] G. Szabó and G. Fáth, Phys. Rep. **446**, 97 (2007).

[5] C. P. Roca, J. A. Cuesta, and A. Sanchez, Phys. Life Rev. **6**, 208 (2009).

[6] M. Perc and A. Szolnoki, BioSystems **99**, 109 (2010).

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

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

[9] D. Helbing, TheoryDecision **40**, 149 (1996).

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

[11] C. W. Gardiner, *Handbook of Stochastic Methods*, 3rd ed. (Springer, New York, 2004).

[12] A. J. Bladon, T. Galla, and A. J. McKane, Phys. Rev. E **81**, 066122 (2010).

[13] T. Reichenbach, M. Mobilia, and E. Frey, Phys. Rev. E **74**, 051907 (2006).

[14] W. H. Sandholm, *Population Games and Evolutionary Dynamics* (MIT Press, Cambridge, MA, 2010).

[15] B. Wu, P. M. Altrock, L. Wang, and A. Traulsen, Phys. Rev. E **82**, 046106 (2010).

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

[17] L. E. Blume, Games Econ. Behavior **5**, 387 (1993).

[18] G. Szabó and C. Tőke, Phys. Rev. E **58**, 69 (1998).

[19] A. Traulsen, M. A. Nowak, and J. M. Pacheco, Phys. Rev. E **74**, 011909 (2006).

[20] K. H. Schlag, J. Econ. Theory **78**, 130 (1998).

[21] A. Traulsen, C. Hauert, H. De Silva, M. A. Nowak, and K. Sigmund, Proc. Natl. Acad. Sci. USA **106**, 709 (2009).

[22] C. Hauert, Int. J. Bifurcation Chaos Appl. Sci. Eng. **12**, 1531 (2002).

[23] A. Traulsen, N. Shoresh, and M. A. Nowak, Bull. Math. Biol. **70**, 1410 (2008).

[24] P. M. Altrock and A. Traulsen, Phys. Rev. E **80**, 011909 (2009).

[25] J. C. Claussen, Eur. Phys. J. B **60**, 391 (2007).

[26] F. A. Chalub and M. O. Souza, Theor. Popul. Biol. **76**, 268 (2009).

[27] F. A. C. C. Chalub and M. O. Souza, Commun. Math. Sci. **7**, 489 (2009).

[28] W. J. Ewens, *Mathematical Population Genetics* (Springer, New York, 2004).

[29] D. Fudenberg and L. A. Imhof, J. Econ. Theory **131**, 251 (2006).

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

[31] S. Van Segbroeck, F. C. Santos, T. Lenaerts, and J. M. Pacheco, Phys. Rev. Lett. **102**, 058105 (2009).

[32] K. Sigmund, H. De Silva, A. Traulsen, and C. Hauert, Nature (London) **466**, 861 (2010).

[33] B. Wu, C. S. Gokhale, L. Wang, and A. Traulsen, J. Math. Biol. **1**, (2012), doi: 10.1007/s00285-011-0430-8.

[34] B. Sinervo and C. M. Lively, Nature (London) **380**, 240 (1996).

[35] B. Sinervo *et al.*, Proc. Natl. Acad. Sci. USA **103**, 7372 (2006).

[36] T. L. Czaran, R. F. Hoekstra, and L. Pagie, Proc. Natl. Acad. Sci. USA **99**, 786 (2002).

[37] B. Kerr, M. A. Riley, M. W. Feldman, and B. J. M. Bohannan, Nature (London) **418**, 171 (2002).

[38] C. Hauert, S. De Monte, J. Hofbauer, and K. Sigmund, Science **296**, 1129 (2002).

[39] D. Semmann, H. J. Krambeck, and M. Milinski, Nature (London) **425**, 390 (2003).

[40] G. Szabó and C. Hauert, Phys. Rev. E **66**, 062903 (2002).

[41] T. Reichenbach, M. Mobilia, and E. Frey, Nature (London) **448**, 1046 (2007).

[42] J. C. Claussen and A. Traulsen, Phys. Rev. Lett. **100**, 058104 (2008).

[43] C. Hauert, [http://www.evoludo.org] (2012).

[44] A. Arenas, J. Camacho, J. A. Cuesta, and R. J. Requejo, J. Theor. Biol. **279**, 113 (2011).

[45] J. W. Demmel, *Applied Numerical Linear Algebra* (Society for Industrial and Applied Mathematics, Philadelphia, PA, 1997).