Stochastic differential equations for evolutionary dynamics with demographic noise and mutations

Arne Traulsen,1 Jens Christian Claussen,2 and Christoph Hauert3

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

(Received 3 November 2011; revised manuscript received 9 March 2012; published 3 April 2012)

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 $\mu$ 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.

DOI: 10.1103/PhysRevE.85.041901

PACS number(s): 87.23.—n, 89.65.—s, 02.50.Ey

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 $\mu$, 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 replaces 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 $\mu$, the offspring of an individual of type $j$ mutates into a type $k$ 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 $T_{kj}$:

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

where $P^\tau(X)$ denotes the probability of being in state $X$ at time $\tau$ and $X^j = (X_1, \ldots, X_j - 1, \ldots, X_k + 1, \ldots, X_d)$ 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]
\begin{equation}
\dot{\rho}(x) = -\sum_{k=1}^{d-1} \frac{\partial}{\partial x_k} \rho(x) A_k(x) + \frac{1}{2} \sum_{j,k=1}^{d-1} \frac{\partial^2}{\partial x_k \partial x_j} \rho(x) B_{jk}(x),
\end{equation}
\hspace{1cm} (2)

where $x = X/N$ represents the state of the population in terms of frequencies of the different strategic types and $\rho(x)$ is the probability density in state $x$. Due to the normalization $\sum_{x=1}^{d-1} x k = 1$, it suffices to consider $d - 1$ elements of the deterministic drift vector $A_k(x)$, $k = 1, \ldots, d - 1$. Similarly, we only need to consider a diffusion matrix $B_{jk}(x)$ with dimension $(d - 1) \times (d - 1)$, that is, $j,k = 1, \ldots, d - 1$. The drift vector $A_k(x)$ is given by
\begin{equation}
A_k(x) = \sum_{j=1}^{d} [T_{jk}(x) - T_{kj}(x)] = -1 + \sum_{j=1}^{d} T_{jk}(x). \quad (3)
\end{equation}

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

The diffusion matrix $B_{jk}(x)$ is defined as
\begin{align}
B_{jk}(x) &= -\frac{1}{N} [T_{jk}(x) + T_{kj}(x)] \quad \text{for} \quad j \neq k, \quad (4a) \\
B_{jj}(x) &= \frac{1}{N} \left\{ \sum_{l=1,l \neq j}^{d} [T_{jl}(x) + T_{lj}(x)] \right\}. \quad (4b)
\end{align}

For a detailed derivation see, for example, [10] or [12]. Note that the diffusion matrix is symmetric, $B_{jk}(x) = B_{kj}(x)$ and vanishes as $-1/N$ in the limit $N \rightarrow \infty$. Moreover, we have for all diagonal elements $B_{jj}(x) \geq 0$ and for the nondiagonal elements $B_{jk}(x) \leq 0$. Note that $T_{jk}(x)$ 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 Ito calculus [11] can be applied to derive a Langevin equation, which represents in our case a stochastic replicator-mutator equation,
\begin{equation}
\dot{x}_k = A_k(x) + \sum_{j=1}^{d-1} C_{kj}(x) \xi_j(t), \quad (5)
\end{equation}

where the $\xi_j(t)$ represent uncorrelated Gaussian white noise with unit variance $\langle \xi_j(t) \xi_j(t') \rangle = \delta_{j,j'} \delta(t - t')$. The matrix $C(x)$ is defined by $C^T(x) C(x) = B(x)$ 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 $C(x)$ provides a quantitative description of the fluctuations introduced by microscopic processes in finite populations. In the limit $N \rightarrow \infty$ the matrix $C(x)$ 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 $A$ [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 $C(x)$. In order to determine $C(x)$ we first diagonalize $B(x)$.

Because $B(x)$ is real and symmetric it is diagonalizable by an orthogonal matrix $U(x)$ with $U(x) U^T(x) = \mathbf{I}$, where $\mathbf{I}$ denotes the identity matrix. Moreover, the normalized eigenvectors $f_i(x)$ of $B(x)$ form an orthonormal basis $f_i(x) \cdot f_j(x) = \delta_{ij}$. Thus we can construct the transformation matrix $U(x) = [f_1(x), \ldots, f_{d-1}(x)]$ such that $B(x) = U(x) \Lambda(x) U^T(x)$, where $\Lambda(x)$ is a diagonal matrix with the eigenvalues $\lambda_i(x)$ of $B(x)$ along its diagonal. From Eq. (4) follows that $B(x)$ is positive definite and hence all eigenvalues $\lambda_i(x)$ are positive. Here we tacitly assume that all $T_{jk}$ are nonzero; if certain transitions are excluded, $B(x)$ is positive semidefinite and eigenvalues can be zero.

Finally, our matrix $C(x)$ is given by
\begin{equation}
C(x) = U(x) \left( \begin{array}{cccc}
\sqrt{\lambda_1} & \cdots & 0 \\
\vdots & \ddots & \vdots \\
0 & \cdots & \sqrt{\lambda_{d-1}}
\end{array} \right) U^T(x). \quad (6)
\end{equation}

This standard procedure to diagonalize matrices can be easily implemented numerically. However, the diffusion matrix $B(x)$ obviously depends on the transition probabilities and thus on the abundances of all strategies. Consequently, the procedure to calculate $C(x)$ must be continuously repeated as time progresses and the state $x$ changes. This is computationally inconvenient and therefore it is desirable to calculate $C(x)$ 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 $d \geq 3$. 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 $B(x)$ depends not only on the frequencies $x$ but also on the payoffs (fitness) of the different strategic types. In this case, an analytic representation of $C(x)$ is only of limited use because it would be valid just for one particular game. However, $B(x)$ becomes payoff independent if $T_{jk}(x) + T_{kj}(x)$ 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 $\pi_f$ and $\pi_m$ 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

\begin{equation}
T_{kj}(x) = (1 - \mu) x_k x_j \gamma(\pi_j, \pi_k) + \mu x_k \frac{1}{d - 1}. \quad (7)
\end{equation}

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 $d \geq 3$. As a particular example to illustrate the framework numerically, we provide a detailed analysis of a generic rock-scissors-paper game.
for \( j \neq k \). Consequently, any pairwise comparison process with
\[
\gamma(\pi_j, \pi_k) + \gamma(\pi_k, \pi_j) = \text{const.}
\tag{8}
\]
leads to a payoff independent diffusion matrix \( B(x) \). 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 \( 0 \leq \gamma(\pi_j, \pi_k) \leq 1 \). 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 \( \gamma(\pi_j, \pi_k) = \frac{1}{\pi_j - \pi_k} \) [7,22]. Incidentally, all examples above satisfy \( \gamma(\pi_j, \pi_k) + \gamma(\pi_k, \pi_j) = 1 \) but, for example, there might be resistance to change in the local update process such that \( \gamma(\pi_j, \pi_k) = a[\frac{1}{2} + w(\pi_j - \pi_k)] \) with \( 0 < a < 1 \) and hence \( \gamma(\pi_j, \pi_k) + \gamma(\pi_k, \pi_j) = a < 1 \).

Last but not least, an example of an important process that does not lead to a payoff independent \( B(x) \) 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
\[
B_{jk}(x) = \frac{1}{N} \left[ -x_j x_k (1 - \mu) - \frac{\mu}{d-1} (x_j + x_k) \right]
\tag{9}
\]
for \( j \neq k \) and
\[
B_{jj}(x) = \frac{1}{N} \left[ x_j (1 - x_j) (1 - \mu) + \frac{\mu}{d-1} [1 + x_j (d-2)] \right].
\tag{10}
\]
Unfortunately, even in this case, a full derivation of \( \mathcal{C}(x) \) 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 \( \mathcal{C}(x) \),
\[
\mathcal{B}(x) = \frac{1}{N} \begin{pmatrix} x(1-x) & -xy & -xy \\ -xy & y(1-y) & 0 \end{pmatrix}.
\tag{11}
\]
The eigenvalues of this matrix are
\[
\lambda_{\pm}(x) = \frac{K_{\pm} \pm L}{2N},
\tag{12}
\]
where \( K_{\pm} = x(1-x) \pm y(1-y) \) and \( L = \sqrt{K_{\pm}^2 + 4x^2y^2} \). The eigenvectors are given by
\[
f_{\pm}(x) = \frac{1}{N} (-K_{\pm} \pm 2xy),
\tag{13}
\]
where \( N = \sqrt{4x^2+y^2} + (K_{\pm} \pm L)^2 \) is a normalization factor. From this, it is straightforward to construct the transformation matrix \( \mathcal{U}(x) = [f_{+}(x), f_{-}(x)] \). The product \( \mathcal{U}(x)\sqrt{\Lambda(x)}\mathcal{U}^{T}(x) \) then yields our matrix \( \mathcal{C}(x) \), which can be written as
\[
C_{zx}(x) = \frac{1}{\sqrt{2N}} \left( \sqrt{Q_{+}^+} + \sqrt{Q_{++}} \right),
\tag{14a}
\]
\[
C_{zy}(x) = C_{yz}(x) = \frac{xy}{\sqrt{2NL}} \left( \sqrt{Q_{--}} - \sqrt{Q_{++}} \right),
\tag{14b}
\]
\[
C_{yy}(x) = \frac{1}{\sqrt{2N}} \left( \sqrt{Q_{--}} + \sqrt{Q_{++}} \right),
\tag{14c}
\]
where \( Q_{\pm\pm} = K_{\pm} \pm L \). It is obvious that \( C_{zz}(x) > 0 \) and \( C_{zy}(x) > 0 \). Since \( L \geq 0 \), we have \( Q_{+} \geq Q_{-} \) and therefore \( C_{yy}(x) < 0 \). It is remarkable that even for this simple \( B(x) \), the matrix \( \mathcal{C}(x) \) already takes a rather complicated form that is not easy to interpret. Therefore, the elements \( C_{zz}(x) \) and \( C_{zy}(x) \) are plotted in Fig. 1. The remaining element \( C_{yy}(x) \) 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 \( \mathcal{C}(x) \) 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 \( \mathcal{C}(x) \) remains the same when including mutations just starting with
\[
\mathcal{B}(x) = \frac{1}{N} \begin{pmatrix} x(1-x) & -xy & -xy \\ -xy & y(1-y) & 0 \end{pmatrix}
\tag{11}
\]
\[
+ \frac{\mu}{2N} \begin{pmatrix} 1 + x & -x - y & 1 + y \\ -x - y & x & 1 + y \end{pmatrix}.
\tag{15}
\]
Unfortunately, the analytic expressions for the elements of \( \mathcal{C}(x) \) grow to unwieldy proportions. A graphical illustration of the elements of \( \mathcal{C}(x) \) for \( \mu > 0 \) is shown in Fig. 1. Compared to \( \mu = 0 \), the entries of \( \mathcal{C}_{zz}(x) \) become larger and the noise terms no longer vanish at the boundaries. For example, at the corners of the simplex we have for \( \mathcal{C}_{zz}(x) \),
\[
C_{zz}(1,0,0) = \frac{3}{\sqrt{10}} \sqrt{\frac{\mu}{N}},
\]
\[
C_{zz}(0,1,0) = \sqrt{\frac{2}{5}} \sqrt{\frac{\mu}{N}},
\]
\[
C_{zz}(0,0,1) = \frac{1}{\sqrt{2}} \sqrt{\frac{\mu}{N}}.
\]
this noise vanishes for
in the
processes \[8\]. (b) For nonvanishing mutations
the usual multiplicative noise from one-dimensional evolutionary
coordinate
and
of the simplex but the expressions are too cumbersome to display
explicitly.

For \(C_{xy}(x)\) we obtain at these points

\[
C_{xy}(1,0,0) = C_{xy}(0,1,0) = -\frac{1}{\sqrt{10}} \sqrt{\frac{\mu}{N}}.
\]

\[
C_{xy}(0,0,1) = 0.
\]

In addition, the functional form of the noise term is altered with
increasing \(\mu\) 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 \(x\)
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 approxi-
mation. Instead, a time scale separation between mutation and
selection allows to approximate the dynamics based on a pair-
wise 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

\[
\begin{pmatrix}
R & S & P \\
S & -1 & 2 + s \\
P & 2/s & -1
\end{pmatrix}
\]

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 \(\hat{x} = \left(\frac{1}{2}, \frac{1}{2}, \frac{1}{2}\right)\), independent of the parameter \(s\) but \(s\) controls
the stability of \(\hat{x}\). For \(s > \frac{1}{2}\), \(\hat{x}\) 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 \[2\].
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 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 $\mu$. For small $\mu$ 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-
Figure 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 $\mu$ 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 $\mu \gg 1/N$ but for smaller $\mu$ 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).

Increase with decreasing $\mu$ but the agreement remains excellent for $\mu > 1/N$. This threshold simply means that for larger $\mu$ 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 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 \( \sim d^{3/2} \), 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 \( B(x) \). In theory, analytical solutions are available for \( d \leq 5 \) 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.

ACKNOWLEDGMENTS

We thank P. M. Altrock and B. Werner for helpful comments. A.T. acknowledges support by the Emmy-Noether program of the DFG and by the MPG. C.H. acknowledges support by the Natural Sciences and Engineering Research Council of Canada (NSERC).
