Stochastic Gain in Population Dynamics

Arne Traulsen, Torsten Röhl, and Heinz Georg Schuster

Institut für Theoretische Physik und Astrophysik, Christian-Albrechts Universität, Olshausenstraße 40, 24098 Kiel, Germany
(Received 26 February 2004; published 9 July 2004)

We introduce an extension of the usual replicator dynamics to adaptive learning rates. We show that a population with a dynamic learning rate can gain an increased average payoff in transient phases and can also exploit external noise, leading the system away from the Nash equilibrium, in a resonancelike fashion. The payoff versus noise curve resembles the signal to noise ratio curve in stochastic resonance. This recalls the stochastic resonance effect [14–17], where the signal to noise ratio of a system is improved for intermediate noise intensities. In contrast to the usual stochastic resonance, a periodic force is not involved here, making the mechanism more similar to coherence resonance [18]. Seen in this broader context, we introduce another mechanism that exploits fluctuations in order to improve the performance of the system.

We consider two adaptive species $X$ and $Y$—each with different strategies—that are involved in a repeated game. Both populations have different objectives described by payoff matrices $P_x$ and $P_y$. The fraction of individuals $x_i, y_i$ that adopt a certain strategy $i$ grows proportional to the relative payoff of the strategy $i$; the same holds for $Y$. In the presence of noise, this coevolution can be described by the coupled replicator equations

$$
\dot{x}_i = x_i \eta_x (\Pi^x_i - \langle \Pi^x \rangle) + \xi^x_i,
\dot{y}_i = y_i \eta_y (\Pi^y_i - \langle \Pi^y \rangle) + \xi^y_i,
$$

where $\eta_x$ and $\eta_y$ are the learning rates of the populations. We assume for simplicity that the noise $\xi_i$ is Gaussian with autocorrelation $\langle \xi^x_i(t) \xi^x_j(s) \rangle = \sigma^2 \delta_{ij} \delta(t-s)$ as in Ref. [12]. We also follow Ref. [12] in choosing reflecting boundaries. The payoffs are defined as $\Pi^x_i = \langle P_x \cdot y \rangle_i$, $\Pi^y_i = \langle P_y \cdot x \rangle_i$, and similarly for $y$.

We extend the usual replicator dynamics by introducing adaptive learning rates as

$$
\eta_x = 1 - \tanh(\alpha_x \Delta \Pi),
$$

where $\Delta \Pi = \langle \Pi^x \rangle - \langle \Pi^y \rangle$ is the time dependent difference between the average payoffs of the populations and $\alpha_x \geq 0$ is a “perception ability” of the population. In order to maintain the basic features of the replicator dynamics, the learning rate must be a positive function with $\langle \eta \rangle = 1$, which is ensured by Eq. (2). For $\alpha_x > 0$ the population $X$ learns slower if it is currently in a good position; otherwise, it learns faster. The value of $\alpha_x$ determines how well a population can assess its current state. The adaptive learning rate leads to a faster escape
from unfavorable states, while on the other hand the population tends to remain in preferable states. Other choices for \(\eta\), which ensure these properties mentioned above will not alter our results. In the following, we focus on a setting where only one population has an adaptive learning rate \(\eta\) as in Eq. (2).

The noise introduced above drives the system away from the Nash equilibrium and leads for small amplitude to a positive gain of the population with an adaptive learning rate, whereas for large noise amplitudes the fluctuations smear out the trajectories in phase space so strongly that they can no longer be exploited. Hence, we expect an optimal noise effect for intermediate values of \(\sigma\). In order to be able to compare the payoffs of both populations, we assume that the dynamics starts from the Nash equilibrium.

As a first example, we consider the zero-sum game “matching pennies” [3,19]. Here both players can choose between two options \(\pm 1\). Player one wins if both players select the same option and player two wins otherwise. The game is described by the payoff matrices

\[
P_x = \begin{pmatrix} +1 & -1 \\ -1 & +1 \end{pmatrix} = -P_y.
\]

(3)

The replicator equations follow from Eqs. (1) and (3) as

\[
\dot{x} = -2\eta_x x(2y - 1)(x - 1) + \xi_x,
\]

\[
\dot{y} = +2\eta_y y(2x - 1)(y - 1) + \xi_y,
\]

(4)

where \(x = x_0\) and \(y = y_0\). Let us first consider the zero noise limit in the case \(\eta = \eta_0 = 1\). As for all zero-sum games, i.e., \(P_x = -P_y\), the system (1) without noise becomes Hamiltonian and has a constant of motion [20]. Here the constant is given by \(H(x, y) = -2\ln[x(1 - x) - 2\ln[y(1 - y)]]\). The trajectories oscillate around the Nash equilibrium at \(x = y = 1/2\). \(H(x, y)\) is connected to the temporal integral of the average payoff \(\langle \Pi_x \rangle = (x^t)' \cdot P_x \cdot y^t\) during a period with \(\langle \Pi^t \rangle > 0,

\[
\int_{t_0}^{t_1} \langle \Pi_x \rangle dt = -\frac{H(x_0, y_0) - H(1/2, 1/2)}{4},
\]

(5)

where \((x, y) = (x_0, y_0)\) at \(t_0\) and \((x, y) = (1/2, 1/2)\) at \(t_1\).

If we include adaptive learning rates (2) into the system, we find \(H(x, y) = -2\tanh(\alpha_x \Delta \Pi)\Delta \Pi \leq 0\), vanishing for \(\alpha_x = 0\). Hence, adaptive learning rates dampen the oscillations around the Nash equilibrium, and the trajectories in the \(x-y\) plane spiral towards the Nash equilibrium where \(\langle \Pi_x \rangle = \langle \Pi_y \rangle = 0\) (see Fig. 1). In addition, this leads to an increased payoff of one population. As the matrices (3) describe a zero-sum game, it is sufficient for a population if it knows its own current average payoff \(\langle \Delta \Pi \rangle = 2\langle \Pi_x \rangle\).

Numerical simulations for \(\alpha_x > 0\) show that the temporal integral of the payoff becomes

\[
\left(\int_{t_0}^{t_1} \langle \Pi_x \rangle dt\right)_{(x_0, y_0)} = -\frac{1}{8}[H(x_1, y_1) - H(x_0, y_0)].
\]

(6)

FIG. 1. Matching pennies: Comparison between the behavior of a population with a constant learning rate [i.e., \(\alpha_x = 0\) (thin lines)] and a population with an adaptive learning rate [perception ability \(\alpha_x = 10\) (thick lines)]. The opponent has in both cases a constant learning rate \(\eta = 1\). Left: Trajectories in strategy space. Arrows show the vector field of the replicator dynamics. Population \(X\) has positive (negative) average payoff in gray (white) areas. Right: Time development of the average payoff of the population \(X\). The adaptive learning rate increases the time intervals in which the corresponding population has a positive payoff, dampening the oscillations around the Nash equilibrium [21].

The averaged initial value \(H(x_0, y_0)\) can be calculated as \(\int_{t_0}^{t_1} dx_0 dy_0 H(x_0, y_0) = 8\). For \(t \rightarrow \infty\) the system relaxes to the Nash equilibrium where \(H = 8\ln 2\). Hence, we find for the average cumulated payoff with \(\langle \int_{t_0}^{t_1} \langle \Pi_x \rangle dt \rangle_{(x_0, y_0)} = -\frac{1}{8}(8\ln 2 - 8) = 0.307\). Numerical simulations yield \(0.308 \pm 0.005\) independent of \(\alpha\). We conclude that a population can increase its average payoff if it has an adaptive learning rate \(\alpha_x > 0\) and if the game does not start in the Nash equilibrium. The adaptation parameter \(\alpha\) influences only the time scale on which the Nash equilibrium is approached.

Small noise intensities drive the system away from the fixed point and the population with the adaptive learning rate gains an increased payoff. If the noise amplitude \(\sigma\) becomes too large, the trajectories will be smeared out homogeneously over the positive (gray) and negative (white) payoff regions in phase space (Fig. 1). This implies that the average gain of population one decreases to zero; cf. Fig. 2. Although the average payoff is very small even for the optimal noise intensity, the cumulated payoff increases linearly in time. This means that for long times the gained payoff accumulates to a profitable value.

As a second application we analyze the effect of adaptive learning rates and noise on the prisoner’s dilemma. We use the standard payoff matrix [22]

\[
P_x = \begin{pmatrix} 3 & 0 \\ 5 & 1 \end{pmatrix} = P_y.
\]

(7)

where rows and columns are placed in the order “cooperate” and “defect.” As this game is not a zero-sum game, the population with the adaptive learning rate must be able to compare its own average payoff with the opponent’s average payoff. The replicator dynamics of this system is determined by Eqs. (1) and (7),

\[
\dot{x} = x \eta_x (x - 1)(1 + y) + \xi_x,
\]

\[
\dot{y} = y \eta_y (y - 1)(1 + x) + \xi_y.
\]

(8)
There is a stable fixed point in the Nash equilibrium $x = y = 0$ where both players defect and an unstable fixed point for mutual cooperation, i.e., $x = y = 1$.

The average payoff difference under the influence of noise is similar as in matching pennies. Small fluctuations lead the system slowly away from the Nash equilibrium and tend to increase the payoff. If the fluctuations are too large, they disturb the population with adaptive learning rates and the payoff decreases again (see Fig. 3). Interestingly enough, here too much noise even leads to a decreasing payoff difference.

In order to describe the “stochastic gain” effect analytically, we introduce a simplified model. A linearization of Eq. (8) around the stable Nash equilibrium leads for constant learning rates to $\dot{x} = -\eta_x x + \xi_x$ and $\dot{y} = -\eta_y y + \xi_y$. We now analyze a game in which the replicator dynamics is given by these linear equations and include adaptive learning rates based on the payoffs for the prisoner’s dilemma. With $\Delta \Pi = -5(x - y)$ the adaptive learning rate $\eta_x$ becomes $\eta_x = 1 + \tanh[5\alpha(x - y)] = 1 + 5\alpha(x - y)$ for $\alpha, x, y \ll 1$. The simplified system can be viewed as a small noise expansion of the prisoner’s dilemma, where the trajectory stays close to the Nash equilibrium. For $\eta_x = 1$ the simplified noisy replicator equations read

$$\dot{x} = -x - \alpha' x(x - y) + \xi_x,$$  \hspace{1cm} (9a)  
$$\dot{y} = -y + \xi_y,$$  \hspace{1cm} (9b)

where $\alpha' = 5\alpha$. The effect of different constant learning rates is discussed in Ref. [23]. The mechanism we introduce here is more intricate, as the adaptive learning rate leads to a dynamical adjustment of the learning rate, and the average of $\eta_x = 1 + \alpha'(x - y)$ over all possible strategies is $\eta_x = 1$.

Equation (9b) describes an Ornstein-Uhlenbeck process [24]; here the dynamics is restricted to $0 \leq y \leq 1$. The Fokker-Planck equation [25] for $p_y = p_y(y, t|y_0, t_0)$,

$$p_y = \frac{d}{dy} \left( y p_y + \frac{\sigma^2}{2} \frac{d}{dy} p_y \right).$$  \hspace{1cm} (10)

has the stationary solution $p_y^* = \mathcal{N}_y e^{-y^2/\sigma^2}$, where $\mathcal{N}_y^{-1} = \int_0^1 e^{-y^2/\sigma^2} dy$. We find the mean value $\langle y(\sigma) \rangle$ as

$$\langle y \rangle = \int_0^1 dy p_y y = \frac{\sigma(1 - e^{-\sigma^2})}{\sqrt{\pi} \text{erf}(1/\sigma)}.$$  \hspace{1cm} (11)

$y$ is a correlated stochastic process which appears in Eq. (9a) as a multiplicative noise. Numerical simulations indicate that we may neglect the stochastic nature of $y$ and replace it by $\langle y \rangle$ for small $\alpha$. This leads to an approximated Fokker-Planck equation for $p_x = p_x(x, t|x_0, 0)$,

$$\dot{p}_x = \frac{d}{dx} \left[ -a(x) p_x + \frac{\sigma^2}{2} \frac{d}{dx} p_x \right].$$  \hspace{1cm} (12)

where $a(x) = -x - x\alpha'(x - y)$. Since $x$ is (similarly to $y$) also restricted to $0 \leq x \leq 1$, we find the stationary solution

$$p_x^* = \mathcal{N}_x \exp \left[ -\frac{x^2}{\sigma^2} - \frac{2\alpha' x^3}{3\sigma^2} + \frac{\alpha'(y)x^2}{\sigma^2} \right]$$  \hspace{1cm} (13)

with the normalization constant $\mathcal{N}_x$. Since $x$ is typically of the order of $\sigma$ for $\sigma \ll 1$, the term $x^2/\sigma^2$ is finite. Therefore, we can expand Eq. (13) for $\alpha' \ll 1$ and obtain by expanding $\langle x \rangle$ again an analytical expression for $\langle \Delta \Pi \rangle = -5\langle x - y \rangle$,

$$\langle \Delta \Pi \rangle = -5\alpha' \frac{d}{d\alpha'}(x) = 5\alpha \left[ \frac{\sigma^2}{2} - \delta^3 \sigma \gamma (1 - \gamma)^2 + \delta^2 (1 - \gamma) \right] - \gamma \left[ \frac{5}{3} \gamma - \frac{7}{6} \sigma^2 (1 - \gamma) \right] - \gamma \left[ \frac{2}{3\sigma} + \gamma \right],$$  \hspace{1cm} (14)

where $\delta = \frac{1}{\sqrt{\text{erf}(1/\sigma)}}$ and $\gamma = e^{-1/\sigma^2}$. The asymptotics of Eq. (14) can be computed as $\langle \Delta \Pi \rangle = \alpha'/(24\sigma^2)$ for $\sigma \gg 1$ and

\[028701-3\]
FIG. 4. Simplified model: Comparison of the average payoff difference $\langle \Delta \Pi \rangle$ from a simulation of Eqs. (9a) and (9b) and the analytical function Eq. (14) ($\Delta t = 0.01$, $\alpha' = 5 \alpha = 0.1$, averages over $4 \times 10^4$ time steps and $4 \times 10^4$ realizations).

$\langle \Delta \Pi \rangle = \alpha' (\frac{35}{3} - \frac{35}{\sigma^2}) \sigma^2$ for $\sigma \ll 1$. We stress that this simplified system which consists of a stable fixed point with a linear adaptive learning rate in the presence of noise is the simplest possible model that describes the stochastic gain effect. Figure 4 shows a comparison between the analytical payoff difference Eq. (14) and a simulation of Eqs. (9a),(9b).

To summarize, we have introduced an extension to the usual replicator dynamics that modifies the learning rates using a simple “win stay–lose shift” rule. In this way, a population optimizes the payoff difference to a competing population optimizes the payoff difference to a competing population with an adaptive learning rate under the influence of external noise. This simple rule leads to a convergence towards the mixed Nash equilibrium for the game of matching pennies [26]. Even in games with stable Nash equilibria as the prisoner’s dilemma, transient phases can be exploited, although the basins of attraction are not altered, as, e.g., in Ref. [23]. Weak external noise drives the system into the transient regime and leads to an increased gain for one adaptive population.

In conclusion, we have found a learning process which improves the gain of the population with an adaptive learning rate under the influence of external noise. Fluctuations lead to an increased payoff for intermediate noise intensities in a resonancelike fashion. This phenomenon could be of particular interest in economics, where interactions are always subject to external disturbances [6,13,27].

We thank J. C. Claussen for stimulating discussions and comments on this manuscript. A.T. acknowledges support by the Studienstiftung des deutschen Volkes.

*Electronic address: traulsen@theo-physik.uni-kiel.de

[21] To ensure the conservation of $H(x,y)$ after a discretization of Eq. (4), symplectic algorithms have to be applied. The algorithm proposed by Hofbauer [20] can be written as

$$x_{i+1} = x_i + x_i [\Pi^+_i - \langle \Pi^+_i \rangle] + \frac{1}{C/\eta_1 + \langle \Pi^+_i \rangle} \cdot \eta_1 \cdot (\Pi^+_i)^+$$

$$y_{i+1} = y_i + y_i [\Pi^-_i - \langle \Pi^-_i \rangle] + \frac{1}{C/\eta_2 + \langle \Pi^-_i \rangle} \cdot \eta_2 \cdot (\Pi^-_i)^+,$$

with $C \gg 1$ and where $\langle \Pi^+_i \rangle = (P_y \cdot x^{i+1})_i$, $\langle \Pi^-_i \rangle = (y^T \cdot P_y \cdot x^{i+1})_i$. Here we choose $C = 100$.
[26] A comparison with “rock-paper-scissors” reveals that this is not necessarily the case, as there are periodic attractors with $\Delta \Pi = 0$ leading to constant $\eta_1$. However, for the chaotic extension of this game [8], these attractors vanish and the trajectory converges to the Nash equilibrium, as expected by Sato et al. [8].