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

(Dated: May 7th, 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 reasonance-like fashion. The payoff versus noise curve resembles the signal to noise ratio curve in stochastic resonance. Seen in this broad context, we introduce another mechanism that exploits fluctuations in order to improve properties of the system. Such a mechanism could be of particular interest in economic systems.

PACS numbers: 87.23.-n, 02.50.Le, 05.40.-a, 05.45.-a

Game theory [1] describes situations in which the success or payoff of an individual depends on its own action as well as on the actions of others. This paradigm can be applied to biological systems, as evolution through natural selection can be viewed as an optimization process in which the fitness-landscape changes with the state of the adaptive populations [2]. Evolutionary game theory focuses mainly on systems with a single fitness function for all individuals, which is identified with the payoff function of a game [3, 4, 5]. In nature often different populations with different ambitions interact with each other, as shoppers and sellers [6], attackers and defenders [6], or males and females [5]. Here, the payoff functions are different for the interacting populations. A mean-field description of such asymmetric conflicts is given by the coupled replicator equations [4, 5, 7]. These equations have a very rich dynamical behavior and can even display Hamiltonian chaos [8, 9]. In previous work [3, 4, 5] it has been tacitly assumed that both populations have the same adaptation mechanisms. But it seems to be natural that different mechanisms are applied by the interacting populations, e.g. different adaptation rates. Here, we analyze such systems for the case that both populations have slightly different adaptation mechanisms. We assume that one population can control its own adaptation rate. This alters the velocity when the system is approaching the stable Nash equilibria [10] in strategy space, leading to an increased average payoff.

In real systems fluctuations disturbing the system are to be expected. Such disturbances can arise from a variety of effects, e.g. errors of the players [11], deviations from a perfectly mixed population, or immigration of individuals with different strategy distributions. So far, stochastic extensions to the replicator dynamics have mainly been analyzed in the context of equilibrium selection [12, 13]. Here, we show that a population with adaptive learning rate can obtain an increased payoff if these fluctuations are present. For small noise intensities the average payoff increases, while very large fluctuations cannot longer be exploited, leading to a decrease of the average payoff. This recalls the stochastic resonance effect [14, 15, 16, 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 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,

$$\begin{aligned} \dot{x}_i &= x_i \eta_x \left[\Pi_i^x - \langle \Pi^x \rangle\right] + \xi_i^x \\ \dot{y}_i &= y_i \eta_y \left[\Pi_i^y - \langle \Pi^y \rangle\right] + \xi_i^y, \end{aligned} \tag{1}$$

where η_x and η_y are the learning rates of the populations. We assume for simplicity that the noise ξ_i is Gaussian with autocorrelation $\langle \xi_i^k(t) \xi_j^l(s) \rangle = \sigma^2 \delta_{ij} \delta_{kl} \delta(t-s)$ as in [12]. We also follow [12] choosing reflecting boundaries. The payoffs are defined as $\Pi_i^x = (P_x \cdot \mathbf{y})_i$, $\langle \Pi^x \rangle = \mathbf{x}^T \cdot P_x \cdot \mathbf{y}$, and similarly for y.

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

$$\eta_x = 1 - \tanh\left(\alpha_x \Delta \Pi\right), \tag{2}$$

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 α_x determines how well a population can assess its current state. The adaptive learning rate leads to a faster escape from unfavourable states, while on the other hand the population tends to remain in preferable states. Other choices for η_x which ensure these properties mentioned above will not alter our results. In the following we will focus on a setting

*traulsen@theo-physik.uni-kiel.de

where only one population has an adaptive learning rate η_x 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 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 σ. 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 ± 1. Player one wins if both players select the same option and player two wins otherwise. The game is described by the payoff matrices

$$P_x = \begin{pmatrix} +1 & -1 \\ -1 & +1 \end{pmatrix} = -P_y. \quad (3)$$

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

$$\begin{aligned} \dot{x} &= -2\eta_x x(2y-1)(x-1) + \xi_x \\ \dot{y} &= +2\eta_y y(2x-1)(y-1) + \xi_y, \end{aligned} \quad (4)$$

where $x = x_0$ and $y = y_0$. Let us first consider the zero noise limit in the case $\eta_x = \eta_y = 1$. As for all zero-sum games, i.e. $P_x = -P_y^T$, 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 $\langle \Pi_x \rangle = (\mathbf{x}^t)^T \cdot P_x \cdot \mathbf{y}^t$ during a period with $\langle \Pi^x \rangle > 0$,

$$\int_{t_0}^{t_1} \langle \Pi^x \rangle dt = -\frac{H\left(x_0, \frac{1}{2}\right) - H\left(\frac{1}{2}, \frac{1}{2}\right)}{4}, \quad (5)$$

where $(x,y) = (x_0, \frac{1}{2})$ at t_0 and $(x,y) = (\frac{1}{2}, x_0)$ at t_1.

If we include adaptive learning rates (2) into the system, we find $\dot{H}(x,y) = -2\tanh(\alpha_x \Delta\Pi)\Delta\Pi \leq 0$, 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

$$\langle \int_{t_0}^{t_1} \langle \Pi_x \rangle dt \rangle_{(x_0,y_0)} = -\frac{1}{8}\left(H(x_1,y_1) - H(x_0,y_0)\right). \quad (6)$$

The averaged initial value $H(x_0,y_0)$ can be calculated as $\iint_0^1 dx_0 dy_0 H(x_0,y_0) = 8$. For $t \to \infty$ the system relaxes to the Nash equilibrium where $H = 8\ln 2$. Hence, we find for the average cumulated payoff $\langle \int_{t_0}^{\infty} \langle \Pi_x \rangle dt \rangle_{(x_0,y_0)} \leq -\frac{1}{8}\left(8\ln 2 - 8\right) \approx 0.307$. Numerical simulations yield 0.308 ± 0.005 independent of α. 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 α influences only the time scale on which the Nash equilibrium is approached.

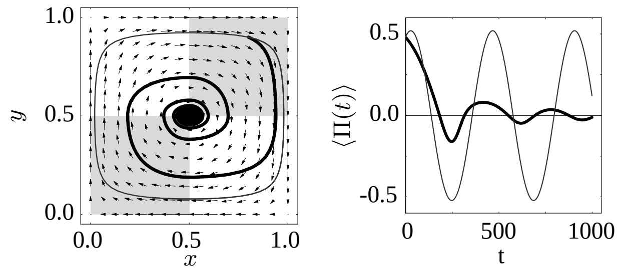

FIG. 1: Matching pennies: Comparison between the behavior of a population with constant learning rate, i.e. $\alpha_x = 0$, (thin lines) and a population with adaptive learning rate (perception ability $\alpha_x = 10$, thick lines). The opponent has in both cases a constant learning rate $\eta_y = 1$. 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 [22].

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 σ 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. 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 [21],

$$P_x = \begin{pmatrix} 3 & 0 \\ 5 & 1 \end{pmatrix} = P_y, \quad (7)$$

where rows and columns are placed in the order "cooperate", "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),

$$\begin{aligned} \dot{x} &= x\eta_x(x-1)(1+y) + \xi_x \\ \dot{y} &= y\eta_y(y-1)(1+x) + \xi_y. \end{aligned} \quad (8)$$

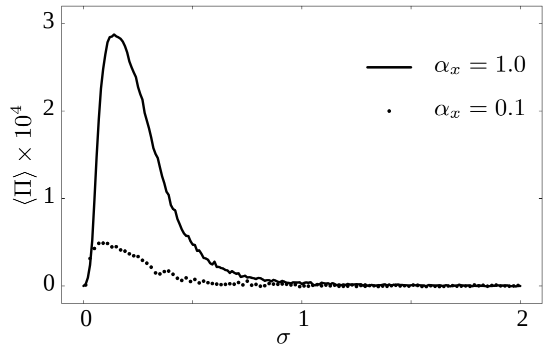

FIG. 2: Matching pennies: Average payoff of a population with adaptive learning rate against a population with constant learning rate under the influence of noise for different noise intensities ($\alpha_y = 0$, averages over 2×10^4 initial conditions and 2×10^4 time steps, see [22] for further details).

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 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.

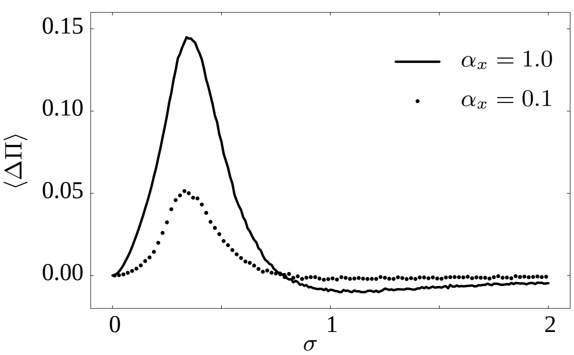

FIG. 3: Prisoner's dilemma: Average payoff difference of a population with adaptive learning rate against a population with constant learning rate for different noise intensities. The negative payoffs arise from the fact that we have $\eta_x < \eta_y$ for $x < y$ ($\Delta t = 0.01$, $\alpha_y = 0$, averages over 2×10^4 initial conditions and 2×10^4 time steps).

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 η_x becomes $\eta_x = 1 + \tanh(5\alpha(x - y)) \approx 1 + 5\alpha(x - y)$ 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 $\eta_y = 1$ the simplified noisy replicator equations read

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

where $\alpha' = 5\alpha$. The effect of different constant learning rates is discussed in [23]. The mechanism we introduce here is more intricated, 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 $\eta_y = 1$.

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 $p_y = p_y(y, t|y_0, t_0)$,

$$\dot{p}_y = \frac{d}{dy}\left(y p_y + \frac{\sigma^2}{2}\frac{d}{dy}p_y\right), \tag{10}$$

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

$$\langle y\rangle = \int_0^1 dy p_y y = \frac{\sigma\left(1 - e^{-\sigma^{-2}}\right)}{\sqrt{\pi}\mathrm{Erf}(\frac{1}{\sigma})}. \tag{11}$$

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 α. This leads to an approximated Fokker-Planck equation for $p_x = p_x(x, t|x_0, 0)$

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

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

$$p_x^s = \mathcal{N}_x \exp\left[-\frac{x^2}{\sigma^2} - \frac{2\alpha' x^3}{3\sigma^2} + \frac{\alpha'\langle y\rangle x^2}{\sigma^2}\right] \tag{13}$$

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

$$\begin{aligned}\langle\Delta\Pi\rangle &= -5\alpha'\frac{d}{d\alpha'}\langle x\rangle = 5\alpha'\left[\frac{\sigma^2}{2} - \delta^3\sigma\gamma(1-\gamma)^2\right. \\ &\quad \left. + \delta^2(1-\gamma)\left(\frac{5}{3}\gamma - \frac{7}{6}\sigma^2(1-\gamma)\right) - \delta\gamma\left(\frac{2}{3\sigma} + \sigma\right)\right],\end{aligned} \tag{14}$$

where $\delta = \frac{1}{\sqrt{\pi}\mathrm{Erf}(1/\sigma)}$ and $\gamma = e^{-1/\sigma^2}$. The asymptotics of Eq. (14) can be computed as $\langle\Delta\Pi\rangle = \alpha'/\left(24\sigma^2\right)$ for $\sigma \gg 1$ and $\langle\Delta\Pi\rangle = \alpha'\left(\frac{5}{2} - \frac{35}{6\pi}\right)\sigma^2$ for $\sigma \ll 1$. We stress that this simplified system which consists of a stable fixed point with linear adaptive learning rate in the presence of noise is the

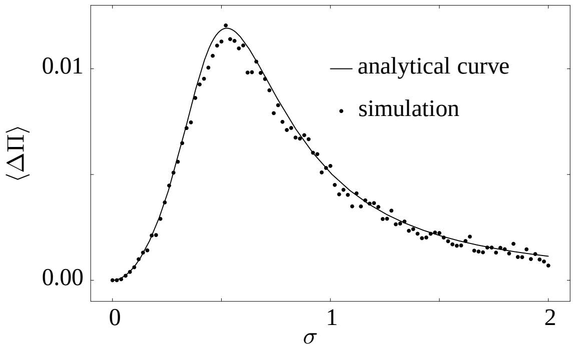

FIG. 4: Simplified model: Comparison of the average payoff difference $\langle\Delta\Pi\rangle$ from a simulation of Eqs. (9a,9b) and the analytical function Eq. (14) ($\Delta t = 0.01$, $\alpha' = 5\alpha = 0.1$, averages over 4×10^4 time steps and 4×10^4 realizations).

simplest possible model that describes the stochastic gain effect. Fig. 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 the learning rates using a simple "win stay—lose shift" rule. In this way, a population optimizes the payoff difference to a competing population. 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 [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 adaptive learning rate under the influence of external noise. Fluctuations lead to an increased payoff for intermediate noise intensities in a resonance-like 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.

[1] J. v. Neumann and O. Morgenstern, *Theory of Games and Economic Behavior* (Princeton University Press, Princeton, 1953).
[2] M. A. Nowak and K. Sigmund, Science **303**, 793 (2004).
[3] J. W. Weibull, *Evolutionary Game Theory* (MIT Press, Cambridge, 1995).
[4] J. Hofbauer and K. Sigmund, *Evolutionary Games and Population Dynamics* (Cambridge Univ. Press, Cambridge, 1998).
[5] J. M. Smith, *Evolution and the Theory of Games* (Cambridge Univ. Press, Cambridge, 1982).
[6] H. Gintis, *Game Theory Evolving* (Princeton Univ. Press, Princeton, 2000).
[7] P. Schuster and K. Sigmund, Animal behaviour **29**, 186 (1981).
[8] Y. Sato, E. Akiyama, and J. D. Farmer, Proc. Nat. Acad. Sc. **99**, 4748 (2002).
[9] Y. Sato and J. P. Crutchfield, Phys. Rev. E **67**, 015206 (2002).
[10] C. A. Holt and A. E. Roth, Proc. Nat. Acad. Sc. **101**, 3999 (2004).
[11] M. A. Nowak and K. Sigmund, Games and Econ. Behaviour **11**, 364 (1995).
[12] D. Foster and P. Young, Theo. Pop. Biol. **38**, 219 (1990).
[13] A. Cabrales, Int. Econ. Rev. **41**, 451 (2000).
[14] R. Benzi, A. Sutera, and A. Vulpiani, J. Phys. A **14**, L453 (1981).
[15] L. Gammaitoni, P. Hänggi, P. Jung, and F. Marchesoni, Rev. Mod. Phys. **70**, 223 (1998).
[16] *Noise in Nonlinear Dynamical Systems*, edited by F. Moss and P. V. E. McClintock (Cambridge Univ. Press, Cambridge, 1989).
[17] *Stochastic Resonance in Physics and Biology*, edited by F. Moss, A. Bulsara, and M. F. Shlesinger (Special issue, Journal of Statistical Physics, 70, 3/4, Plenum, NY, 1993).
[18] A. S. Pikovsky and J. Kurths, Phys. Rev. Lett. **78**, 775 (1997).
[19] H. G. Schuster, *Complex Adaptive Systems* (Scator Pub., Saarbrücken, 2002).
[20] J. Hofbauer, J. Math. Biol. **34**, 675 (1996).
[21] R. Axelrod, *The Evolution of Cooperation* (Basic Books, New York, 1984).
[22] To ensure the conservation of $H(x, y)$ after a discretization of Eq. (4) symplectic algorithms have to be applied. The algorithm proposed by J. Hofbauer [20] can be written as

$$\begin{aligned} x_i^{t+1} &= x_i^t + x_i^t \left[\Pi_i^x - \langle\Pi^x\rangle\right] \frac{1}{C/\eta_x + \langle\Pi^x\rangle} \\ y_i^{t+1} &= y_i^t + y_i^t \left[(\Pi_i^y)^+ - \langle\Pi^y\rangle^+\right] \frac{1}{C/\eta_y + \langle\Pi^y\rangle^+} \end{aligned} \qquad (15)$$

with $C \gg 1$ and where $(\Pi_i^y)^+ = \left(P_y \cdot \mathbf{x}^{t+1}\right)_i$, $\langle\Pi^y\rangle^+ = \left(\mathbf{y}^t\right)^T \cdot P_y \cdot \mathbf{x}^{t+1}$. Here we choose $C = 100$.
[23] C. T. Bergstrom and M. Lachmann, Proc. Nat. Acad. Sc. **100**, 593 (2003).
[24] G. E. Uhlenbeck and L. S. Ornstein, Phys. Rev. **36**, 823 (1930).
[25] C. W. Gardiner, *Handbook of Stochastic Methods*, 2nd ed. (Springer, Berlin, 1985).
[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 η_x. 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].
[27] D. Fudenberg and E. Maskin, American Econ. Rev. **80**, 274 (1990).