Dynamical mechanisms of adaptation in multiagent systems

K. Y. Michael Wong, S. W. Lim, and Zhuo Gao

Department of Physics, Hong Kong University of Science and Technology, Clear Water Bay, Hong Kong, China
(Received 22 March 2004; published 30 August 2004)

We consider multiagent systems whose agents compete for resources using strategies with adaptable preferences. Diversity of initial preferences of strategies is introduced by randomly assigning virtual points to the strategies of each agent. When diversity increases, the successive appearance of scaling, kinetic sampling and waiting mechanisms shows that agent cooperation becomes increasingly important. Analyses yield excellent agreement with simulations over nine decades of data.

Many natural and artificial systems involve interacting agents, each making independent decisions to compete for limited resources, but globally exhibit coordinated behavior through their mutual adaptation [1,2]. Examples include the competition of predators in ecology, buyers or sellers in economic markets, and routers in computer networks. While a standard approach is to analyze the steady state behavior of the system described by the Nash equilibria [3], it is interesting to consider the dynamics of how the steady state is approached. Dynamical effects are especially relevant in a changing environment, such as that in economics or distributed control.

Adaptation is a collective dynamical process. When it is achieved by small iterative steps taken by many agents, Hamiltonian functions can be used to analyze the steady-state behavior [4]. However, when it proceeds in large steps, a dynamical approach is more applicable than a game-theoretic approach. For example, when the complexity of the agents in a multiagent system is low, *a maladaptive* behavior takes place, in which there are bursts of the population's decisions due to their premature rush to certain states they perceive to be advantageous [5,6]. The resultant large fluctuations indicate that the agents fail to cooperate with each other.

Maladaptation originates from the uniformly zero preference of strategies in the initial condition. The dependence of initial conditions was noted in a statistical mechanical approach [4]. System efficiency can be improved by random initial conditions in systems driven by internal information [7], or external information [8], accompanied by hysteresis [9]. The same is valid in models with batch update [10] and their noisy extension [11].

In this Rapid Communication, we numerically and analytically study the dynamical mechanisms by which agents cooperate with each other to reach steady states. This is done by tuning the diversity of the agents in their initial preferences of strategies, so that the population of agents adapting to the environment at each step becomes increasingly sparse with diversity. When maladaptation gradually vanishes, we look for evidence that cooperative mechanisms among agents become favored or even indispensable. Concretely, we consider a prototype of multiagent systems, in which a population of N agents compete in an environment of limited resources, N being odd [2]. Each agent makes a decision $+$ or $-$ at each time step, and the minority group wins. The decisions of each agent are prescribed by *strategies*, which are Boolean functions mapping the *history* of the winning bits in the most recent m steps to decisions $+$ or $-$. Before the game starts, each agent randomly picks s strategies. Out of her s strategies, each agent makes decisions according to the most successful one at each step; the success of a strategy is measured by its *virtual point*, which increases (decreases) by 1 if it indicates a winning (losing) decision at a time step.

Diversity of initial preferences of strategies is introduced by randomly assigning R virtual points to the s strategies of each agent before the game starts. Hence the initial virtual point of each strategy obeys a multinomial distribution with mean R/s and variance $R(s-1)/s^2$. The ratio $\rho \equiv R/N$ is referred to as the *diversity*. In particular, for $s=2$ and odd R, no two strategies have the same virtual points throughout the process, and the dynamics is deterministic, resulting in highly precise simulation results useful for refined comparison with theories.

To describe the macroscopic dynamics of the system, we define the D-dimensional vector $A^\mu(t) \equiv \Sigma_i \, \xi_i^\mu(t)/N$, where ξ_i^μ is the decision of agent i at state μ, and $D \equiv 2^m$ is the number of histories. While only one of the D components corresponds to the historical state $\mu^*(t)$ of the system, the augmentation to D components is necessary to describe the attractor structure and the transient behavior of the system dynamics. The inset of Fig. 1 illustrates the convergence to the attractor for the visualizable case of $m=1$. The dynamics proceeds in the direction which tends to reduce the magnitude of the components of $A^\mu(t)$ [4]. However, a certain amount of maladaptation always exists in the system, so that the components of $A^\mu(t)$ overshoot, resulting in periodic attractors with period of $2D$. Every state μ appears as historical states two times in a steady-state period, yielding the winning bits $-$ and $+$ each exactly once. One occurence brings A^μ from positive to negative, and another bringing it back from negative to positive, thus completing a cycle. For $m=1$, the steady state is described by the sequence $\mu(t) = -, +, +, -$, where both states $-$ and $+$ are followed by $-$ and $+$ once each. For general values of m, the states in an attractor are given by a binary de Bruijn sequence of order $m+1$ [12].

The effects of introducing diversity is shown in Fig. 1, which plots the steady-state variance of the population for decision $+$, $\sigma^2/N \equiv N\langle [A^{\mu^*(t)}(t) - \langle A^{\mu^*(t)}(t)\rangle]^2\rangle/4$, as a func-
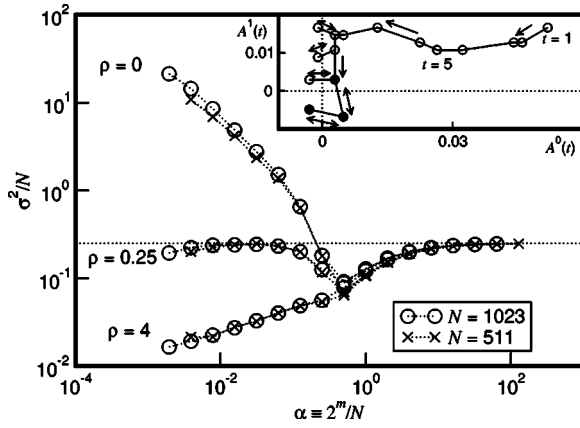
FIG. 1. The dependence of the variance of the population for decision + on the complexity for different diversities at $s=2$ averaged over 128 samples. The horizontal dotted line is the limit of random decisions. Inset: the state motion of a sample in the phase space for $m=1$. Solid dots: attractor states.

tion of the *complexity* $\alpha \equiv D/N$ [5]. We found that the variance decreases significantly with diversity in the *symmetric* phase where $\alpha < \alpha_c (\approx 0.3)$, and remains unaffected in the *asymmetric* phase with $\alpha > \alpha_c$ [13,14]. The dependence of the variance on the diversity is further shown in Fig. 2. Focusing on the physical picture of the dynamics [15], four regimes can be identified.

(a) *Multinomial regime.* When $\rho \sim N^{-1}$, $\sigma^2/N \sim N$ with proportionality constants dependent on m. To analyze this and other regimes, we let $S_{ab}(\omega)$ be the number of agents holding strategies a and b (with $a<b$), where the virtual point of strategy a is initially displaced by ω with respect to b. The average of $S_{ab}(\omega)$ over initial condition is proportional to the binomial distribution of virtual points, i.e., $\langle S_{ab}(\omega)\rangle = NC_{(R-\omega)/2}^R/2^{2D-1+R}$. The key to analyzing the system dynamics is the observation that the virtual points of a strategy displace by exactly the same amount for all agents. Hence for a given strategy pair, the profile of the virtual
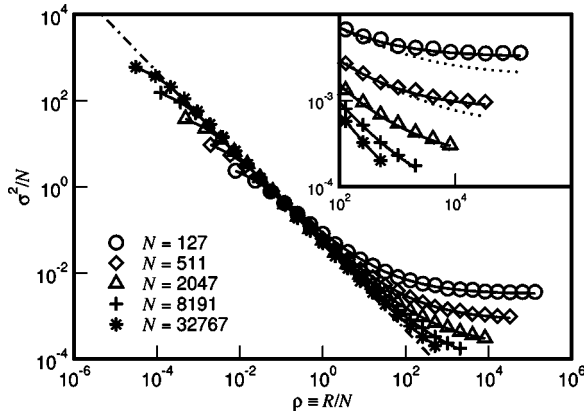

FIG. 2. The dependence of the variance of the population for decision + on the diversity at $m=1$ and $s=2$. Symbols: simulation results averaged over 1024 samples. Lines: theory. Dashed-dotted line: scaling prediction. Inset: Comparison between simulation results (symbols), theory with waiting effects included (lines) and excluded (dashed lines).

point distribution remains unchanged, but the peak position shifts with the dynamics. If the virtual point displacement of strategy a at time t is $\Omega_a(t)$, then the agents holding strategies a and b make decisions according to strategy a if $\omega + \Omega_a(t) - \Omega_b(t) > 0$, and strategy b otherwise. At time t, we can write $\Omega_a(t) = \Sigma_\mu k_\mu(t)\xi_a^\mu$, where $k_\mu(t)$ is the number of wins minus losses of decision + up to time t at state μ, and ξ_a^μ is the corresponding decision of strategy a. Consider the difference $A^\mu(t) - A^\mu(0) = \Sigma_{a<b} S_{ab}(\omega)(\xi_a^\mu - \xi_b^\mu)[\Theta(\omega + \Omega_a(t) - \Omega_b(t)) - \Theta(\omega)]/N$. Its average can be found by introducing the average $\langle S_{ab}(\omega)\rangle$, writing the step function Θ as a sum over Kronecka deltas and introducing their integral representation, using the identity $e^{ik\theta(\xi_a^\mu - \xi_b^\mu)} = \cos^2 k\theta + i \sin k\theta \cos k\theta(\xi_a^\mu - \xi_b^\mu) + \sin^2 k\theta\xi_a^\mu\xi_b^\mu$, and noting that $\Sigma_a \xi_a^\mu = 0$. The final result is

$$\langle A^\mu(t) - A^\mu(0)\rangle = \int_0^{2\pi} \frac{d\theta}{2\pi}\cos^R\theta\frac{\sin k_\mu(t)\theta}{\sin \theta}$$

$$\times \cos k_\mu(t)\theta\prod_{\nu\neq\mu} \cos^2 k_\nu(t)\theta. \quad (1)$$

When $\rho \sim N^{-1}$, $\langle A^\mu(t+1) - A^\mu(t)\rangle \sim O(1)$ and is self-averaging. Since $A^\mu(0)$ is Gaussian with variance N^{-1}, the values of $A^\mu(t)$ at the attractors can be computed, and the variance found. For example, for $m=1$, $\sigma^2/N = N[7(c_{R+1})^2 - 2c_{R+1}c_{R+3} + 7(c_{R+3})^2]$, where $c_n = 2^{-n}C_{n/2}^n$ for even integer n.

(b) *Scaling regime.* When $\rho \sim 1$, $\sigma^2/N \sim \rho^{-1}$ with proportionality constants effectively independent of m for m not too large [16]. In this case, Eq. (1) can be simplified to $\langle A^\mu(t) - A^\mu(0)\rangle = k_\mu(t)\sqrt{2/\pi R}$. The average step size becomes $\langle A^\mu(t+1) - A^\mu(t)\rangle = \sqrt{2/\pi R}\delta_{\mu\mu^*(t)} \sim O(N^{-1/2})$ and is self-averaging. To interpret this result, we note that changes in $A^\mu(t)$ are only contributed by *fickle* agents with marginal preferences of their strategies. That is, those with $\omega + \Omega_a(t) - \Omega_b(t) = \pm 1$ and $\xi_a^\mu - \xi_b^\mu = \mp 2$ sgn $A^\mu(t)$ for $\mu = \mu^*(t)$. For large R, the binomial virtual point distribution among agents of a given strategy pair is effectively a Gaussian with variance R. Hence the number of agents switching strategies at time t scales as the height of the Gaussian, which is $\sqrt{2/\pi R}$. Thus, by spreading the virtual point distribution, diversity reduces the step size and hence maladaptation.

As a result, each state of the attractor is confined in a D-dimensional hypercube of size $\sqrt{2/\pi R}$, irrespective of the initial position of the A^μ components. Starting from the initial state $A^\mu(0)$, the state changes in steps of size $\sqrt{2/\pi R}$ until it reaches the attractor, whose $2D$ historical states are given by $\sqrt{2/\pi R}\lceil\sqrt{\pi R/2}A^\mu(0)\rceil$ and $\sqrt{2/\pi R}\{\lceil\sqrt{\pi R/2}A^\mu(0)\rceil - 1\}$, where $[x]$ represents the decimal part of x [17]. Averaging over $A^\mu(0)$, which are Gaussian numbers with mean 0 and variance $1/N$, the variance of decisions become $\sigma^2/N = f(\rho)/2\pi\rho$, where $f(\rho)$ approaches $(1-1/4D)/3$ for $\rho \gg 1$. Note that $f(\rho)$ is a smooth function of ρ, since σ^2/N depends on ρ mainly through the step size factor $1/2\pi\rho$, whereas $f(\rho)$ merely provides a higher order correction to the functional dependence. This accounts for the scaling regime in Fig. 2. Furthermore, we note that $f(\rho)$ rapidly approaches $1/3$ when

m increases. Hence for general values of m, $\sigma^2/N \rightarrow 1/6\pi\rho$, provided that m is not too large.

(c) *Kinetic sampling regime*. When $\rho \sim N$, σ^2/N deviates above the scaling with ρ^{-1}, and is given by $\sigma^2/N = f_m(\Delta)/N$, where $\Delta \equiv \sqrt{2N/\pi\rho}$ is the *kinetic step size*, and f_m is a function dependent on the memory size m. Here $A^\mu(t+1) - A^\mu(t)$ scales as N^{-1} and is no longer self-averaging. Rather, it is equal to $2/N$ times the number of agents who switch strategies at time t, which is Poisson distributed with a mean of $N/\sqrt{2\pi R}$. However, since the attractor is formed by steps which *reverse the sign* of A^μ, the average step size in the attractor is *larger* than that in the transient state. To see this, we consider the probability of $P_{att}(\Delta \mathbf{A})$ of step sizes $\Delta \mathbf{A}$ in the attractor. Assuming that all states of the phase space are equally likely to be accessed, we have $P_{att}(\Delta \mathbf{A}) = \Sigma_{\mathbf{A}} P_{att}(\Delta \mathbf{A}, \mathbf{A})$, where $P_{att}(\Delta \mathbf{A}, \mathbf{A})$ is the probability of finding the position $\mathbf{A}$ with displacement $\Delta \mathbf{A}$ in the attractor. Consider the example of $m=1$ in the inset of Fig. 1. The sign reversal condition implies that $P_{att}(\Delta \mathbf{A}, \mathbf{A}) = P_{Poi}(\Delta \mathbf{A})\Pi_\mu \Theta[-A^\mu(A^\mu + \Delta A^\mu)]$, where $P_{Poi}(\Delta \mathbf{A})$ is the Poisson distribution of step sizes, yielding $P_{att}(\Delta \mathbf{A}) = P_{Poi}(\Delta \mathbf{A})\Pi_\mu \Delta A^\mu$. Thus the attractor averages $\langle(\Delta A^\pm)^2\rangle_{att}$, which are required for computing the variance of decisions, are given by $\langle(\Delta A^\pm)^2 \Delta A^+ \Delta A^-\rangle_{Poi}/\langle\Delta A^+ \Delta A^-\rangle_{Poi}$. In other words, the sampling of the step sizes is weighted by the attractor sizes due to the kinetics. The result for $m=1$ is $\sigma^2/N = (14\Delta^3 + 105\Delta^2 + 132\Delta + 24)/96N(2\Delta+1)$.

(d) *Waiting regime*. When $\rho \gg N$, σ^2/N deviates above the predictions of kinetic sampling. Here the agents are so diverse that the average step size is approaching 0. At each state in the phase space, the system remains stationary for many time steps, waiting for some agent to reduce the magnitude of her virtual point until strategy switching can take place. This waiting effect modifies the composition of the group of fickle agents who contribute to the state transitions, and consequently increase the step sizes and variance above those predicted by kinetic sampling. Consider the example of $m=1$. As shown in the inset of Fig. 1, the attractor consists of both vertical and horizontal hops, and detailed analysis shows that only one type of agent can complete both hops. Since fewer and fewer agents contribute to the switching of states in the limit $\rho \gg N$, a single agent of this type will dominate the game dynamics, and one would expect that σ^2/N approaches $0.25/N$. However, when waiting is possible, agents not of this correct type can wait for other agents to complete the hops in the attractor, even though one would expect that the probability of finding more than one fickle agent is drastically less than that for one. In fact, simulation shows that only 8% of the attractors consist of a single fickle agent, and σ^2/N approaches $0.43/N$ rather than $0.25/N$. As shown in the inset of Fig. 2, lengthy analytic results including waiting effects significantly improve the agreement with simulations over the kinetic sampling prediction.

Many properties of the system dependent on the transient dynamics also depend on its diversity. For example, since diversity reduces the fraction of agents switching strategies at each time step, it also slows down the convergence to the steady state. Hence in the scaling regime, the convergence time scales as $\rho^{1/2}$. Specifically, when $\rho \gg 1$, the average con-

vergence time becomes $(2+\sqrt{2})\sqrt{\rho}$ for $m=1$. Similarly, the distribution of payoffs among the *frozen* agents (that is, agents who do not switch their strategies at the steady state) also depends on the transient. Since the system dynamics reaches a periodic attractor, they have constant average payoffs at the steady state. Hence any spread in their payoff distribution is a consequence of the transient dynamics. Thus, in the scaling regime, the mean square payoff scales as ρ. Specifically, when $\rho \gg 1$, the mean square payoff becomes $\pi\rho$ for $m=1$. Simulation results of both the convergence time and the mean square payoff have an excellent agreement with the theory [15].

The results presented here can be generalized to other cases. Consider the *exogenous* dynamics, in which the information $\mu(t)$ was randomly and independently drawn at each time step t [4]. The picture that the states of the game are hopping between hypercubes in the phase space remains valid. At the steady state, the attractor consists of hoppings among all vertices of a hyperpolygon enclosing the origin in the phase space, analogous to the present *endogenous* case, in which a fraction of hyperpolygon vertices belong to the attractor. In the scaling regime, the behavior depends on the scaling of the step sizes with diversity, rather than the actual sequence of the steps. Consequently, the behavior is similar to that of the endogenous dynamics.

The present results can be extended to higher values of m [15]. For $m=2$, analysis using the de Bruijn sequence explicitly yields excellent results. For higher m, we approximate the attractor of the exogenous dynamics by a hyperpolygon enclosing the origin of the phase space. Using a generating function approach, and taking into account the scaling of step sizes and kinetic sampling, the computed variance of decisions agrees qualitatively with simulations, except for values of α close to α_c.

We can also make qualitative predictions about the transition from the symmetric to asymmetric phase when the complexity α increases [13]. From Eq. (1), the average displacement in the phase space is given by

$$\langle A^\mu(t) - A^\mu(0)\rangle \approx k_\mu(t)\sqrt{\frac{2}{\pi(R + 2D\langle k^2\rangle)}}, \qquad (2)$$

where $\langle k^2\rangle$ represents the mean of $k_\nu(t)^2$ for all $\nu \lesssim D$. For $\rho \sim \alpha \sim 1$, it can be verified that $A^\mu(t) - A^\mu(0)$ is self-averaging. Suppose the dynamics leads to an attractor near the origin, with $\langle A^\mu(t)\rangle \rightarrow 0$. Noting that $\langle A^\mu(0)^2\rangle \sim 1/N$, we obtain the self-consistent relation $\langle k^2\rangle = \rho/2(\alpha_c - \alpha)$, where $\alpha_c = 1/\pi \approx 0.318$. This means that when α approaches α_c, the average step size approaches 0 in the asymptotic limit. There is a critical slow down since the convergence time diverges. When α exceeds α_c, the average step size vanishes before the system reaches the attractor near the origin, so that the state of the system is trapped at locations with at least some components being nonzero. The interpretation is that when α is large, the distributions of strategies become so sparse that motions in the phase space cannot be achieved by the switching of strategies. Note that the value of α_c is close to the value of 0.337 obtained by the continuum approximation [4] or batch update [10].

Game theorists may ask whether the introduction of diversity assists the system to reach a Nash equilibrium. It has been verified that Nash equilibria consist of pure strategies [4]. Hence all frozen agents have no incentives to switch their strategies. Thus, the Nash equilibrium is approached in the sense that the fraction of fickle agents decreases with increasing diversity. In the extremely diverse limit, there are cases that only one fickle agent switches strategy at each step in the attractor, so that even that fickle agent cannot increase her payoff, since on switching she always remains on the majority side and loses. Then a Nash equilibrium is reached exactly. For $m=1$, for example, simulations show that a Nash equilibrium is reached in this way with probability 0.57.

In summary, we have studied the dynamical mechanisms of adaptation in a multiagent system with adaptive agents competing for finite resources. In the scaling regime, we find that agents making an adaptive move at each step are sufficiently numerous that agent cooperation can be described at the level of statistical distributions of strategy preferences. At high diversity, we find that the scaling mechanism is supplemented by kinetic sampling, a mechanism self-imposed by the requirement to stay in the attractor. In other words, the attractor dynamics favors the cooperation of larger clusters of agents. In extremely diverse systems, we discover further a waiting mechanism, when agents who are unable to complete the attractor dynamics alone wait for other agents to collaborate with them. When waiting is present, cooperation between individual types of agents becomes indispensable in reaching the steady state behavior. Together, these mechanisms yield theoretical predictions with excellent agreement with simulations over 9 decades of data. By introducing diversity, the variance of decisions in the symmetric phase decreases, showing that the maladaptive behavior is reduced.

The combination of scaling, kinetic sampling and waiting in accounting for the steady state properties of the system illustrates the importance of dynamical considerations. We anticipate that these dynamical effects will play a crucial role in the entire symmetric phase, since when α increases, the state motion in a high dimensional phase space can easily shift the tail of the virtual point distributions to the verge of strategy switching, leading to the sparseness condition where kinetic sampling and waiting effects are relevant. In general, due to the generic nature of these effects, we expect that they are relevant to multiagent systems with different payoff functions and updating rules. Besides, these generic phenomena may have relevance to biological evolution, since the waiting mechanism is analogous to the coordination of limited genetic changes which result in bigger changes.

---

[1] P. W. Anderson, K. J. Arrow, and D. Pines, *The Economy as an Evolving Complex System* (Addison Wesley, Redwood City, CA, 1988).

[2] D. Challet and Y. C. Zhang, Physica A **246**, 407 (1997).

[3] E. Rasmusen, *Games and Information* (Basil Blackwell, Oxford, 1989).

[4] D. Challet, M. Marsili, and R. Zecchina, Phys. Rev. Lett. **84**, 1824 (2000).

[5] R. Savit, R. Manuca, and R. Riolo, Phys. Rev. Lett. **82**, 2203 (1999).

[6] N. F. Johnson, M. Hart, and P. M. Hui, Physica A **269**, 1 (1999).

[7] J. P. Garrahan, E. Moro, and D. Sherrington, Phys. Rev. E **62**, R9 (2000).

[8] E. Moro, e-print cond-mat/0402651.

[9] D. Sherrington, E. Moro, and J. P. Garrahan, Physica A **311**, 527 (2002).

[10] J. A. F. Heimel and A. C. C. Coolen, Phys. Rev. E **63**, 056121 (2001).

[11] A. C. C. Coolen, J. A. F. Heimel, and D. Sherrington, Phys. Rev. E **65**, 016126 (2001).

[12] S. W. Golomb, *Shift Register Sequences, Revised Edition* (Aegean Park Press, Laguna Hills, CA, 1982).

[13] D. Challet and M. Marsili, Phys. Rev. E **60**, R6271 (1999).

[14] This result agrees qualitatively with Fig. 11 of [8] and Fig. 3 of [10]; here we use the data collapse for different N to explicitly show that the variance is a function of ρ at $\rho \sim 1$.

[15] Details to be published elsewhere.

[16] This is the regime considered in [10] and [11], but results differ due to model differences.

[17] The hypercube picture implies that both the standard deviation and the average of A^μ are bounded by the step size, as evident in the inset of Fig. 1. Hence the variance is a sufficient measure of system efficiency.