Santa Fe Institute Working Paper 04-08-025
arxiv.org:/abs/nlin.AO/0408039

Stability and Diversity in Collective Adaptation

Yuzuru Sato,[1,*] Eizo Akiyama,[2,†] and James P. Crutchfield[1,‡]

[1]*Santa Fe Institute, 1399 Hyde Park Road, Santa Fe, NM 87501, USA*
[2]*Institute of Policy and Planning Sciences, University of Tsukuba, Tennodai 1-1-1, Tsukuba, Ibaraki 305-8573, Japan*

We derive a class of macroscopic differential equations that describe collective adaptation, starting from a discrete-time stochastic microscopic model. The behavior of each agent is a dynamic balance between adaptation that locally achieves the best action and memory loss that leads to randomized behavior. We show that, although individual agents interact with their environment and other agents in a purely self-interested way, macroscopic behavior can be interpreted as game dynamics. Application to several familiar, explicit game interactions shows that the adaptation dynamics exhibits a diversity of collective behaviors, including stable limit cycles, quasiperiodicity, intermittency, and deterministic chaos. The simplicity of the assumptions underlying the macroscopic equations suggests that these behaviors should be expected broadly in collective adaptation. We also analyze the adaptation dynamics from an information-theoretic viewpoint and discuss self-organization induced by information flux between agents, giving a novel view of collective adaptation.

PACS numbers: 05.45.-a, 89.75.Fb 02.50.Le,
Keywords: collective adaptation, evolutionary dynamics, game theory, information theory, dynamical systems

I. INTRODUCTION

Collective behavior in groups of adaptive systems is an important and cross-cutting topic that appears under various guises in many fields, including biology, cognitive neurosciences, computer science, and social science. In all these adaptive systems, individual agents interact with one another and modify their behaviors according to the information they receive through those interactions. Often, though, collective behaviors emerge that are beyond the individual agent's perceptual capabilities and that sometimes frustrate satisfying the local goals. With competitive interactions dynamic adaptation can produce rich and unexpected behaviors. This kind of mutual adaptation has been discussed, for example, in studies of biological group interaction [1–3], interactive learning [4–6], large-scale adaptive systems [7–9], and learning in games [10, 11].

Here we develop a class of coupled differential equations for mutual adaptation in agent collectives—systems in which agents learn how to act in their environment and with other agents through reinforcement of their actions. We show that the adaptive behavior in agent collectives, in special cases, reduces to a generalized form of multipopulation replicator equations and, generally, can be viewed as a kind information-theoretic self-organization in a collective adaptive system.

Suppose that many agents interact with an environment and each independently attempts to adjust its behavior to the environment based on its sensory stimuli. The environment consists of other agents and other exogenous influences. The agents could be humans, animals, or machines, but we make no assumptions about their detailed internal structures. That is, the central hypothesis in the following is that collective adaptation is a dynamical behavior driven by agents' environment-mediated interactions. By separating the time scales of change in the environment, of agents' adaptation, and of agent-agent interactions, our models describe, not the action-taking itself, but the temporal change in the probability distribution of choices.

A. Related Work

This approach should be compared and contrasted with game theoretic view [12]. First, classical game theory often assumes that players have knowledge of the entire environmental structure and of other players' decision-making processes. Our adaptive agents, however, have no knowledge of a game in which they might be playing. Thus, unlike classical game theory, in our setting there is no bird's eye view for the entire collective that is available to the agents. Agents have only a myopic model of the environment, since any information external to them is given implicitly via the reinforcements for their action choices. Second, although we employ game-theoretic concepts such as Nash equilibria, we focus almost exclusively on *dynamics*—transients, attractors, and so on—of collective adaptation, while, naturally, making contact with the *statics* familiar from game theory. Finally, despite the differences, game structures can be introduced as a set of parameters corresponding to approximated static environments.

While replicator dynamics were introduced originally for evolutionary game theory [13–15], the relationship between learning with reinforcement and replicator equations has been discussed only recently [10, 11]. Briefly

*Electronic address: ysato@santafe.edu
†Electronic address: eizo@sk.tsukuba.ac.jp
‡Electronic address: chaos@santafe.edu

stated, in our model the state space represents an individual agent's probability distribution to choose actions and the adaptation equations describe the temporal evolution of choice probabilities as the agents interact. Here, we extend these considerations to collective adaptation, introducing the theory behind a previously reported model [16, 17]. The overall approach, though, establishes a general framework for dynamical-systems modeling and analysis of adaptive behavior in collectives. It is important to emphasize that our framework goes beyond the multipopulation replicator equations and asymmetric game dynamics since it does not require a static environment (cf. Ref. [18, 19] for dynamic environments) and it includes the key element of the temporal loss of memory.

We model adaptation in terms of the distribution of agents' choices, developing a set of differential equations that are a continuous-time limit of a discrete-time stochastic process; cf. Ref. [20]. We spend some time discussing the origin of action probabilities, since this is necessary to understand the model variables and also to clarify the limits that we invoke to arrive at our model. One is tempted to give a game-theoretic interpretation of the model and its development. For example, the mixed strategies in game play are often interpreted as weights over all (complete plans of) actions. However, the game-theoretic view is inappropriate for analyzing local, myopic adaptation and the time evolution of collective behavior.

Another interpretation of our use of action probabilities comes from regarding them as frequencies of action choices. In this view, one needs long-time trials so that the frequencies take on statistical validity for an agent. Short of this, they would be dominated by fluctuations, due to undersampling. In particular, one requires that stable limit distributions exist. Moreover, the underlying deterministic dynamics of adaptation should be ergodic and have strong mixing properties. Finally, considering agent-agent interactions, one needs to assume that their adaptation is very slow compared to interaction dynamics. For rapid, say, real-time, adaptation, these assumptions would be invalid. Nonetheless, they are appropriate for long-term reinforcement, as found in learning motion through iterated exercise and learning customs through social interaction.

B. Synopsis

The approach we take is ultimately phenomenological. We are reminded of the reaction-diffusion models of biological morphogenesis introduced originally in Ref. [21]. There, the detailed processes of biological development and pattern formation were abstracted, since their biochemical basis was (and still is) largely unknown, and a behavioral phenomenology was developed on this basis. Similarly, we abstract the detailed and unknown perceptual processes that underlie agent adaptation and construct a phenomenology that captures adaptive behavior at a larger scale, in agent collectives.

The phenomenology that we develop for this is one based on communications systems. Agents in a collective are confronted with the same three communication problems posed by Weaver in the founding work of information theory—*The Mathematical Theory of Communication* [22]: (i) communicating accurately, (ii) interpreting the transmitted information, and (iii) modifying future behavior based on the information. Shannon solved the first problem developing his theory of error-free transmission [22]. In his vocabulary adaptive agents are *information sources*. Each (i) receives information from the external environment, which includes other agents, (ii) interprets the received information and modifies its internal model accordingly, and (iii) then, making decisions based on those modifications, generates future behavior.

We will show that this information-theoretic view provides useful tools for analyzing collective adaptation and also an appropriate description for our assumed frequency dynamics. Using these we derive a new state space based on the self-informations of agent's actions and this allows one to investigate the dynamics of uncertainty in collective adaptation. It will become clear, though, that the assumption of global information maximization has limited relevance here, even for simple adaptation in a static environment. Instead, self-organization that derive from the information flux between agents gives us a new view of collective adaptation.

To illustrate collective adaptation, we present several simulations of example environments; in particular, those having frustrated agent-agent interactions [23]. Interestingly, for two agents with perfect memory interacting via zero-sum rock-scissors-paper interactions the dynamics exhibits Hamiltonian chaos [16]. With memory loss, though, the dynamics becomes dissipative and displays the full range of nonlinear dynamical behaviors, including limit cycles, intermittency, and deterministic chaos [17].

The examples illustrate that Nash equilibria often plays little or no role in collective adaptation. Sometimes the dynamics is explicitly excluded from reaching Nash equilibria, even asymptotically. Rather, it turns out that the network describing the switching between deterministic actions is a dominant factor in structuring the state-space flows. From it, much of the dynamics, including the origins of chaos becomes intuitively clear.

In the next section (Sec. II), we develop a dynamical system that models adaptive behavior in collectives. In Sec. III we introduce an information-theoretic view and coordinate-transformation for adaptive dynamics. To illustrate the rich range of behaviors, in the Sec. IV we give several examples of adaptive dynamics based on non-transitive interactions. Finally, in Sec. V we interpret our results and suggest future directions.

II. DYNAMICS FOR COLLECTIVE ADAPTATION

Before developing the full equations for a collective of adaptive agents, it is helpful to first describe the dynamics of how an individual agent adapts to the constraints imposed by its environment using the memory of its past behaviors. We then build up a description of how multiple agents interact, focusing only on the additional features that come from interaction. The result is a set of coupled differential equations that determine the behavior of adaptive agent collectives and are amenable to various kinds geometric, statistical, and information-theoretic analyses.

A. Individual Agent Adaptation

Here we develop a continuous-time model for adaptation in an environment with a single adaptive agent. Although the behavior in this case is relatively simple, the single-agent case allows us to explain several basic points about dynamic adaptation, without the complications of a collective and agent-agent interactions. In particular, we discuss how and why we go from a discrete-time stochastic process to a *continuous-time* limit. We also describe an agent's effective internal model of the environment and how we model its adaptation process via a *probability distribution* of action choices.

An agent takes one of N possible *actions*: $i = 1, 2, \ldots, N$ at each time step τ. Let the probability for the agent to chose action i be $x_i(\tau)$, where τ is the number of steps from the initial state $x_i(0)$. The agent's state vector—its *choice distribution*—at time τ is $\mathbf{x}(\tau) = (x_1(\tau), x_2(\tau), \ldots, x_N(\tau))$, where $\Sigma_n^N x_n(\tau) = 1$. In the following we call the temporal behavior of $\mathbf{x}(\tau)$ as the *dynamics of adaptation*.

Let $r_i(\tau)$ denote the reinforcement the agent receives for its taking action i at step τ. Denote the collection of these by the vector $\mathbf{r}(\tau) = (r_1(\tau), \ldots, r_N(\tau))$. The agent's memories—denoted $\mathbf{Q}(\tau) = (Q_1(\tau), \ldots, Q_N(\tau))$ —of past rewards from its actions are updated according to

$$
\begin{aligned}
Q_i(\tau+1) - Q_i(\tau) &= \frac{1}{T}\left[\delta_i(\tau) r_i(\tau) - \alpha Q_i(\tau)\right] \ , \qquad (1)\\
\delta_i(\tau) &= \begin{cases} 1, & \text{action } i \text{ chosen at step } \tau \\ 0, & \text{otherwise} \end{cases}
\end{aligned}
$$

with $i = 1, \ldots, N$ and $Q_i(0) = 0$. T is a constant that sets the agent-environment interaction time scale. $\alpha \in [0, 1)$ controls the agent's memory loss rate. For $\alpha = 0$, the agent has a perfect memory as the sum of the past reinforcements; for $\alpha > 0$ the memory is attenuated in that older reinforcements have less effect on the current Q_is and more recent reinforcements are given larger weight. One imagines that the agent constructs a histogram of past reinforcements and this serves as a simple internal memory of its environment.

An agent chooses its next action according to its choice distribution which is updated from the reinforcement memory according to:

$$
x_i(\tau) = \frac{e^{\beta Q_i(\tau)}}{\Sigma_{n=1}^N e^{\beta Q_n(\tau)}} \ , \qquad (2)
$$

where $i = 1, 2, \ldots, N$. $\beta \in [0, \infty]$ controls the adaptation rate: how much the choice distribution is changed by the memory of past reinforcements. For example, if $\beta = 0$, the choice distribution is unaffected by past reinforcements. Specifically, it becomes independent of $\mathbf{Q}$ and one has $x_i(\tau) = 1/N$. In this case, the agent chooses actions with uniform probability and so behaves completely randomly. In a complementary fashion, in the limit $\beta \to \infty$, an agent chooses that action i with the maximum $Q_i(\tau)$ and $x_i(\tau) \to 1$.

Given Eq. (2) the time evolution of agent's choice distribution) is:

$$
x_i(\tau+1) = \frac{x_i(\tau) e^{\beta(Q_i(\tau+1) - Q_i(\tau))}}{\Sigma_{n=1}^N x_n(\tau) e^{\beta(Q_n(\tau+1) - Q_n(\tau))}} \ . \qquad (3)
$$

where $i = 1, 2, \ldots, N$. This determines how the agent adapts its choice distribution using reinforcements it has received from the environment for its past actions.

This simple kind of adaptation was introduced as a principle of behavioral learning [24, 25] and as a model of stochastic learning [26]. It is sometimes referred to as reinforcement learning [27, 28]. Arguably, it is the simplest form of adaptation in which an agent develops relationships or behavior patterns through reinforcements from external stimuli.

Starting with the discrete-time model above, one can develop a continuous-time model that corresponds to the agent performing a large number of actions, iterates of Eq. (1), for each choice distribution update, iterate of Eq. (2). Thus, we recognize two different time scales: one for agent-environment interactions and one for adaptation of the agent's internal model based on its internal memory. We assume that the adaptation dynamics is very slow compared to interactions and so $\mathbf{x}$ is essentially constant during interactions. See Fig. 1.

Starting from Eq. (1), one can show that the continuous-time dynamics of memory updates is given by the differential equations

$$
\dot{Q}_i(t) = R_i(t) - \alpha Q_i(t) \ , \qquad (4)
$$

with $i = 1, 2, \ldots, N$ and $Q_i(0) = 0$. (see App. A.) Here R_i is the reward the environment gives to the agent choosing action i: the average of $r_i(\tau)$ during the time interval between updates of $\mathbf{x}$ at t and $t + dt$.

From Eq. (2) one sees that the map from $\mathbf{Q}(t)$ to $\mathbf{x}(t)$ at time t is given by

$$
x_i(t) = \frac{e^{\beta Q_i(t)}}{\Sigma_{n=1}^N e^{\beta Q_n(t)}} \ , \qquad (5)
$$

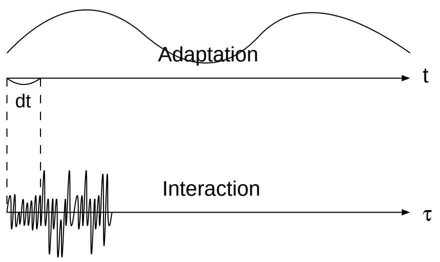

FIG. 1: The time scale (t) of a single agent interacting with its environment and the time scale (τ) of the agent's adaptation: $\tau \ll t$.

where $i = 1, 2, \ldots, N$. Differentiating Eq. (5) gives the continuous-time dynamics

$$\dot{x}_i(t) = \beta x_i(t)(\dot{Q}_i(t) - \Sigma_{n=1}^{N} \dot{Q}_n(t) x_n(t)) \ , \qquad (6)$$

with $i = 1, 2, \ldots, N$.

Assembling Eqs. (4), (5), and (6), one finds the basic dynamic that governs agent behavior on the adaptation time-scale:

$$\frac{\dot{x}_i}{x_i} = \beta(R_i - R) + \alpha(H_i - H) \ , \qquad (7)$$

where $i = 1, 2, \ldots, N$. Here

$$R = \Sigma_{n=1}^{N} x_n R_n \qquad (8)$$

is the net reinforcement averaged over the agent's possible actions. And, $H_i = -\log x_i$ is the *self-information* or degree of surprise when the agent takes action i [22]. The average self-information, or *Shannon entropy* of the choice distribution, also appears as

$$H = \Sigma_{n=1}^{N} x_n H_n = -\Sigma_{n=1}^{N} x_n \log x_n \ . \qquad (9)$$

These are the entropies of the agent's choice distribution measured, not in *bits* (binary digits), but in *nats* (natural digits), since the natural logarithm is used. The entropy measures the choice distribution's flatness, being maximized when the choices all have equal probability.

Fortunately, the basic dynamic captured by Eq. (7) is quite intuitive, being the balance of two terms on the right-hand side. The first term describes an adaptation dynamic, whose time scale is controlled by β. The second describes the loss of memory with a time scale controlled by α. That is, the adaptation in choice probabilities is driven by a balance between two forces: the tendency to concentrate the choice probability based on the reinforcement $\{R_i\}$ and the tendency to make choices equally likely. Finally, on the lefthand side, one has the logarithmic derivative of the choice probabilities: $\dot{x}_i/x_i = d/dt \log x_i$.

Note that each of the terms on the righthand side is a difference between a function of a particular choice and that function's average. Specifically, the first term $\Delta R_i \equiv R_i - R$ is the relative benefit in choosing action i compared to the mean reinforcement across all choices. Other things being held constant, if this term is positive, then action i is the better choice compared to the mean and x_i will increase. The second term $\Delta H_i \equiv H_i - H$ is the relative informativeness of taking action i compared to the average H. Thus, x_i decreases in proportion to the entropy at time t and so this term works to increase the uncertainty of agent's actions, flattening the choice distribution by increasing the probability of unlikely actions. When $x_i = N^{-1}$, the distribution is flat (purely random choices), $\Delta H = 0$, and memory loss effects disappear.

Mathematically, the adaptation equations have quite a bit of structure and this has important consequences, as we will see. Summarizing, the adaptation equations describe a dynamic that balances the tendency to concentrate on choices associated with the best action against the tendency to make the choices equally likely. The net result is to increase the choice uncertainty, subject to the constraints imposed by the environment via the reinforcements. Thus, the choice distribution is the least biased distribution consistent with environmental constraints and individual memory loss. We will return to discuss this mechanism in detail using information theory in the Sec. III.

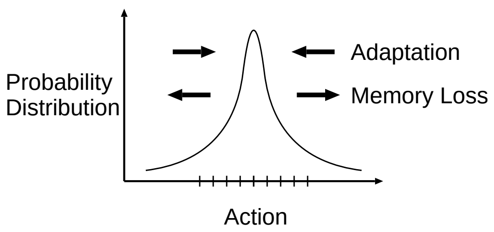

FIG. 2: A dynamic balance of adaptation and memory loss: Adaptation concentrates the probability distribution on the best actions. Memory loss of past history leads to a distribution that is flatter and has higher entropy.

Since the reinforcement determines the agent's interactions with the environment, there are, in fact, three different time scales operating: that for agent-environment interactions, that for each agent's adaptation, and that for changes to the environment. However, if the environment changes very slow compared to the agent's internal adaptation, the environment $r_i(t)$ can be regarded as effectively constant, as shown in Fig. 3.

In this case $r_i(t)$ can be approximated as a static relationship between an agent's actions and the reinforcements given by the environment. Let $r_i(t) = a_i$, where $\mathbf{a} = (a_1, \ldots, a_N)$ are constants that are normalized: $\Sigma_{n=1}^{N} a_n = 0$. Given this, the agent's time-average reinforcements are a_i ($R_i = a_i$) and the continuous-time

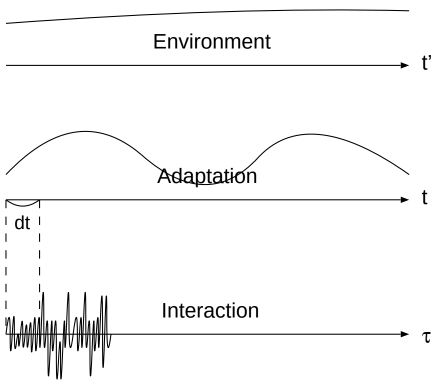

FIG. 3: The time scales of dynamic adaptation: Agent adaptation is slow compared to agent-environment interaction and environmental change is slower still compared to adaptation.

dynamic simplifies to:

$$\frac{\dot{x}_i}{x_i} = \beta(a_i - \Sigma_{n=1}^N a_n x_n) + \alpha(H_i - H) \ , \qquad (10)$$

where $i = 1, 2, \ldots, N$.

The behavior of single-agent adaptation given by Eq. (10) is very simple. When α is small, so that adaptation is dominant $x_i \to 1$, where i is the action with the highest reward a_i, and $x_j \to 0$ for $j \neq i$. The agent receives this information from the fixed environment and its behavior is simply to choose the action with the maximum reward and the choice distribution moves to the associated simplex vertex $(0, \ldots, x_i = 1, \ldots, 0)$. In the special case when $\alpha = 0$, it is known that for arbitrary $\mathbf{a}$ Eq. (10) moves $\mathbf{x}$ to the vertex corresponding to the maximum a_i [2]. In a complementary way, when α is large enough to overcome the relative differences in reinforcements—that is, when $\frac{\beta}{\alpha} \to 0$ memory loss dominates, the agent states goes to a uniform choice distribution ($x_i = N^{-1}$) and the system converges to the simplex center. Note that in machine learning this balance between local optimization and randomized behavior, which selects nonoptimal actions, is referred to as the *exploitation-exploration* trade-off [28].

For instance, consider an agent that takes $N = 3$ actions, $\{1, 2, 3\}$, in an environment described by $\mathbf{a} = (\frac{2}{3}\epsilon, -1 - \frac{1}{3}\epsilon, 1 - \frac{1}{3}\epsilon)$, with $\epsilon \in [-1, 1]$. In the perfect memory case ($\alpha = 0$), the choice distribution converges to a stable fixed point $(0, 0, 1)$. $\mathbf{x}^* = (\frac{1}{3}, \frac{1}{3}, \frac{1}{3})$ is an unstable hyperbolic fixed point. In the memory loss case ($\alpha > 0$), dynamics converges a stable fixed point inside the simplex. (These cases are illustrated in Fig. 4.)

Even when the environment is time-dependent, the agent's behavior can track the highest-reward action as long as the time scale of environment change is slow compared to the agent's adaptation. However, the situation is more interesting when environment change occurs at a rate near the time-scale set by adaptation. Mutual

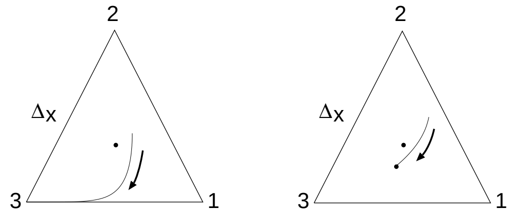

FIG. 4: Dynamics of single-agent adaptation: Here there are three actions, labeled 1, 2, and 3, and the environment gives reinforcements according to $\mathbf{a} = (\frac{2}{3}\epsilon, -1 - \frac{1}{3}\epsilon, 1 - \frac{1}{3}\epsilon)$. The figure shows two trajectories from simulations with $\epsilon = 0.5$ and $\beta = 0.1$ and with $\alpha = 0.0$ (right) and $\alpha = 0.3$ (left).

adaptation in agent collectives, the subject of the following sections, corresponds to just this situation. Other agents provide, thought their own adaptation, a dynamic environment to any given agent and if their times scales of adaptation are close the dynamics can be quite rich and difficult to predict and analyze.

B. Two Agent Adaptation

To develop equations of motion for adaptation in an agent collective we initially assume, for simplicity, that there are only two agents. The agents, denoted X and Y, at each moment take one of N or M actions, respectively. The agents states at time t are $\mathbf{x} = (x_1, \ldots, x_N)$ and $\mathbf{y} = (y_1, \ldots, y_M)$, with $\Sigma_{n=1}^N x_n = \Sigma_{m=1}^M y_m = 1$. $\mathbf{x}(0)$ and $\mathbf{y}(0)$ are the initial conditions. We view the time evolution of each agent's state vector in the simplices $\mathbf{x} \in \Delta_X$ and $\mathbf{y} \in \Delta_Y$ and the group dynamics in the *collective* state space Δ which is the product of the agent simplices:

$$\mathbf{X} = (\mathbf{x}, \mathbf{y}) \in \Delta = \Delta_X \times \Delta_Y \ . \qquad (11)$$

There are again three different time scales to consider: one for agent-agent interaction, one for each agent's internal adaptation, and one for the environment which now mediates agent interactions via the reinforcements given to the agents. Here we distinguish between the *global environment* experienced by the agents and the *external environment*, which is the global environment with the agent states removed. The external environment controls, for example, the degree of coupling between the agents. In contrast with the single-agent case, in the many agent setting each agent's behavior produces a dynamic environment for the other. This environment dynamics is particularly important when the adaptation time scales of each agent are close.

Following the single-agent case, though, we assume that the adaptation dynamic is very slow compared to that of agent-agent interactions and that the dynamics of the external environment changes very slowly compared to that of agents' mutual adaptation. Under these assumptions the agent state vectors $\mathbf{x}$ and $\mathbf{y}$ are effectively constant during the agent-agent interactions that occur between adaptation updates. The immediate consequence is that can describes the collective state space in terms of the frequencies of actions (the choice distributions). Additionally, the environment is essentially constant relative to changes in the states $\mathbf{x}$ and $\mathbf{y}$.

Denote the agents' memories by $\mathbf{Q}^X = (Q_1^X, \ldots, Q_N^X)$ for X and $\mathbf{Q}^Y = (Q_1^Y, \ldots, Q_M^Y)$ for Y and set $Q_i^X(0) = 0$ and $Q_j^Y(0) = 0$, for for $i = 1, \ldots, N$ and $j = 1, \ldots, M$. For the dynamic governing memory updates we have

$$\begin{aligned} \dot{Q}_i^X &= R_i^X - \alpha_X Q_i^X \ , \\ \dot{Q}_j^Y &= R_j^Y - \alpha_Y Q_j^Y \ , \end{aligned} \tag{12}$$

for $i = 1, 2, \ldots, N$ and $j = 1, 2, \ldots, M$. R_i^X is the reward for agent X choosing action i, averaged over agent Y's actions between adaptive updates; and R_j^Y is Y's. The parameters $\alpha_X, \alpha_Y \in [0, 1)$ control each agent's memory loss rate, respectively.

The map from $\mathbf{Q}^X(t)$ to $\mathbf{x}(t)$ and from $\mathbf{Q}^Y(t)$ to $\mathbf{y}(t)$ at time t is

$$\begin{aligned} x_i(t) &= \frac{e^{\beta_X Q_i^X(t)}}{\Sigma_{k=1}^N e^{\beta_X Q_k^X(t)}} \ , \\ y_j(t) &= \frac{e^{\beta_Y Q_j^Y(t)}}{\Sigma_{k=1}^M e^{\beta_Y Q_k^Y(t)}} \ , \end{aligned} \tag{13}$$

for $i = 1, \ldots, N$ and $j = 1, \ldots, M$. Here $\beta_X, \beta_Y \in [0, \infty]$ control the agents' adaptation rates, respectively. Differentiating Eq. (13) with respect to t, the continuous-time adaptation for two agents is governed by

$$\begin{aligned} \dot{x}_i &= \beta_X x_i (\dot{Q}_i^X - \Sigma_k^N \dot{Q}_k^X x_k) \ , \\ \dot{y}_j &= \beta_Y y_j (\dot{Q}_j^Y - \Sigma_k^M \dot{Q}_k^Y y_k) \ . \end{aligned} \tag{14}$$

for $i = 1, \ldots, N$ and $j = 1, \ldots, M$.

Putting together Eqs. (12), (13), and (14), one finds the coupled adaptation equations for two agents:

$$\begin{aligned} \frac{\dot{x}_i}{x_i} &= \beta_X (R_i^X - R^X) + \alpha_X (H_i^X - H^X) \ , \\ \frac{\dot{y}_j}{y_j} &= \beta_Y (R_j^Y - R^Y) + \alpha_Y (H_j^Y - H^Y) \ , \end{aligned} \tag{15}$$

for $i = 1, \ldots, N$ and $j = 1, \ldots, M$ and where

$$\begin{aligned} R^X = \Sigma_{k=1}^N x_k R_k^X, \quad R^Y = \Sigma_{k=1}^M y_k R_k^Y, \\ H^X = \Sigma_{k=1}^N x_k H_k^X, \quad H^Y = \Sigma_{k=1}^M y_k H_k^Y \ . \end{aligned} \tag{16}$$

The interpretations of the $\Delta R = R_i - R$ and $\Delta H = H_i - H$ terms are not essentially different from those introduced to describe the single-agent case. That is, the behavior of each agent is a dynamic balance between (i) adaptation: concentrating the choice probability on the best action at t and (ii) memory loss: increasing the choice uncertainty. What is new here is that there are two (and eventually more) agents attempting to achieve this balance together using information that comes from their interactions.

As given, the adaptation equations include the possibility of a time-dependent environment, which would be implemented (say) using a time-dependent reinforcement scheme. However, as with the single-agent case, it is helpful to simplify the model by assuming a static external environment and, in particular, static relationships between the agents.

Assume that the external environment changes slowly compared to the dynamics of mutual adaptation, as illustrated in Fig. 3. This implies a (near) static relationship between pairs of action choices (i, j) and reinforcements r_{ij} for both agents. Since the environmental dynamics is very slow compared to each agents' adaptation, $r_{ij}^X(t)$ and $r_{ji}^Y(t)$ are essentially constant during adaptation. The rs can be approximated then as constant:

$$\begin{aligned} r_{ij}^X(t) &= a_{ij} \ , \\ r_{ij}^Y(t) &= b_{ji} \ , \end{aligned} \tag{17}$$

for $i = 1, \ldots, N$ and $j = 1, \ldots, M$. a_{ij} and b_{ji} are normalized over j and i so that when summing over all actions the reinforcements vanish.

$$\begin{aligned} \Sigma_{n=1}^N a_{nj} = 0 \ , \\ \Sigma_{m=1}^M b_{mi} = 0 \ , \end{aligned} \tag{18}$$

Given the form of ΔR in the adaptation equations, this normalization does not affect the dynamics.

Assume further that $\mathbf{x}$ and $\mathbf{y}$ are independently distributed. This is equivalent to agents never having a global view of the collective or their interactions with the environment (other agents). Each agent's knowledge of the environment is uncorrelated, at each moment, with the state of the other agents. The time-average rewards for X and Y now become

$$\begin{aligned} R_i^X &= \Sigma_{m=1}^M a_{im} y_m = (A\mathbf{y})_i \\ R_j^Y &= \Sigma_{n=1}^N b_{jn} x_n = (B\mathbf{x})_j \ , \end{aligned} \tag{19}$$

for $i = 1, \ldots, N$ and $j = 1, \ldots, M$. In this restricted case, the continuous-time dynamic is given by the coupled adaptation equations

$$\begin{aligned} \frac{\dot{x}_i}{x_i} &= \beta_X [(A\mathbf{y})_i - \mathbf{x} \cdot A\mathbf{y}] \\ &+ \alpha_X [-\log x_i + \Sigma_k^N x_k \log x_k], \\ \frac{\dot{y}_j}{y_j} &= \beta_Y [(B\mathbf{x})_j - \mathbf{y} \cdot B\mathbf{x}] \\ &+ \alpha_Y [-\log y_j + \Sigma_k^M y_k \log y_k]. \end{aligned} \tag{20}$$

for $i = 1, \ldots, N$ and $j = 1, \ldots, M$. A is an $N \times M$ matrix and B is a $M \times N$ matrix with $(A)_{ij} = a_{ij}$ and $(B)_{ji} = b_{ji}$, respectively. $\mathbf{x} \cdot A\mathbf{y}$ is the inner product between $\mathbf{x}$ and $A\mathbf{y}$: $\mathbf{x} \cdot A\mathbf{y} = \Sigma_n^N \Sigma_m^M a_{nm} x_n y_m$.

C. Collective Adaptation

Generalizing to an arbitrary number of agents at this point should appear straightforward. It simply requires extending Eqs. (15) to a collection of adaptive agents. Suppose there are S agents labeled $s = 1, 2, \ldots, S$ and each agent can take one of N^s actions, then we have

$$\frac{\dot{x_i^s}}{x_i^s} = \beta_s(R_i^s - R^s) + \alpha_s(H_i^s - H^s) \ . \tag{21}$$

for $i = 1, \ldots, N_s$ and $s = 1, \ldots, S$. Equations (21) constitute our general model for adaptation in agents collective.

One describes the time evolution of the agents' state vectors in the simplices $\mathbf{x}^1 \in \Delta_1$, $\mathbf{x}^2 \in \Delta_2$, ..., and $\mathbf{x}^S \in \Delta_S$. The adaptation dynamics in the higher-dimensional *collective* state space occurs within

$$\mathbf{X} = (\mathbf{x}^1, \mathbf{x}^2, \ldots, \mathbf{x}^S) \in \Delta = \Delta_1 \times \Delta_2 \times \ldots \Delta_S \ . \tag{22}$$

Note that the general model includes heterogeneous network settings with local interactions; see App. E.

With three agents X, Y, and Z, one obtains, for example:

$$\begin{aligned} \frac{\dot{x}_i}{x_i} &= \beta_X(R_i^X - R^X) + \alpha_X[H_i^X - H^X] \\ \frac{\dot{y}_j}{y_j} &= \beta_Y(R_j^Y - R^Y) + \alpha_Y[H_j^Y - H^Y] \\ \frac{\dot{z}_k}{z_k} &= \beta_Z(R_k^Z - R^Z) + \alpha_Z[H_k^Z - H^Z] \ , \end{aligned} \tag{23}$$

for $i = 1, \ldots, N$, $j = 1, \ldots, M$, and $k = 1, \ldots, L$. The static environment version reduces to

$$\begin{aligned} \frac{\dot{x}_i}{x_i} &= \beta_X[(A\mathbf{yz})_i - \mathbf{x} \cdot A\mathbf{yz}] \\ &+ \alpha_X[-\log x_i + \Sigma_n^N x_n \log x_n], \\ \frac{\dot{y}_j}{y_j} &= \beta_Y[(B\mathbf{zx})_j - \mathbf{y} \cdot B\mathbf{zx}] \\ &+ \alpha_Y[-\log y_j + \Sigma_m^M y_m \log y_m], \\ \frac{\dot{z}_k}{z_k} &= \beta_Z[(C\mathbf{xy})_k - \mathbf{z} \cdot C\mathbf{xy}] \\ &+ \alpha_Z[-\log z_k + \Sigma_l^L z_l \log z_l] \ , \end{aligned} \tag{24}$$

for $i = 1, \ldots, N$, $j = 1, \ldots, M$, and $k = 1, \ldots, L$, and with tensors $(A)_{ijk} = a_{ijk}$, $(B)_{jki} = b_{jki}$, $(C)_{kij} = c_{kij}$. Here $(A\mathbf{yz})_i = \Sigma_m^M \Sigma_l^L a_{iml} y_m z_l$ and $\mathbf{x} \cdot A\mathbf{yz} = \Sigma_n^N \Sigma_m^M \Sigma_l^L a_{nml} x_n y_m z_l$ and similarly for Y and Z.

D. Evolutionary Dynamics and Game Theory

We now interrupt the development to discuss the connections between the model developed this far and models from population dynamics and game theory. There are interesting connections and also some important distinctions that need to be kept in mind, before we can move forward.

The special case that allows us to make contact with evolutionary dynamics and game theory is the restriction to agents with perfect memory interacting in a static environment. (For further details see App. B.) In the two-agent case we set $\alpha_X = \alpha_Y = 0$ and equal adaptation rates, $\beta_X = \beta_Y$. Under these assumptions our model, Eqs. (20), reduces to what is either called *multi-population replicator equations* [14] or *asymmetric game dynamics* [10, 14]. The equations are:

$$\begin{aligned} \frac{\dot{x}_i}{x_i} &= (A\mathbf{y})_i - \mathbf{x} \cdot A\mathbf{y} \ , \\ \frac{\dot{y}_j}{y_j} &= (B\mathbf{x})_j - \mathbf{y} \cdot B\mathbf{x} \ . \end{aligned} \tag{25}$$

From the perspective of game theory, one regards the interactions determined by A and B, respectively, as X's and Y's *payoff matrices* for a linear game in which X plays action i against Y's action j. Additionally, $\mathbf{x}$ and $\mathbf{y}$, the agent state vectors, are interpreted as the *mixed strategies*. In fact, $\mathbf{x} \cdot A\mathbf{y}$ and $\mathbf{y} \cdot B\mathbf{x}$ in Eqs. (25) formally satisfy von Neumann-Morgenstern utilities [12]. If they exist in the interior of the collective simplices Δ_X and Δ_Y, interior Nash equilibria of the game (A, B) are the fixed points determined by the intersections of the x- and y-nullclines of Eqs. (25).

One must be careful, though, in drawing parallels between our general dynamic setting and classical game theory. In the idealized economic agents, it is often assumed that agents have knowledge of the entire game structure and of other agents' decision-making processes. Its central methodology derives how these *rational players* should act. Our adaptive agents, in contrast, have no knowledge of a game in which they might be playing, only a myopic model of the environment and, even then, this is given only implicitly via the reinforcements the agents receive from the environment. In particular, the agents do not know whether they are playing a game or not, how many agents there are beyond themselves, or even whether other agents exist or not. Our model of dynamic adaptation under such constraints is appropriate nonetheless for many real world learning systems, whether animal, human, or economic agent collectives [29]. The bi-matrix game (A, B) appears above as a description of the collective's global dynamic only under the assumptions that the external environment changes very slowly.

The connection with evolutionary dynamics is formal and comes from the fact that Eqs. (25) are the well known replicator equations of population dynamics [2]. However, the interpretation of the variables is rather different. Population dynamics views $\mathbf{x}$ and $\mathbf{y}$ as two separate, but interacting (infinite size) groups. These two populations are described as distributions of various organismal phenotypes. The equations of motion determine the evolution of these populations over time and through interaction. In our model, in contrast, $\mathbf{x}$ and $\mathbf{y}$ represent the probability of agents choosing actions. The equations of motion describe their dynamic adaptation to each other through interaction.

Despite the similarities that one can draw in this special case, it is important to emphasize that our framework goes beyond the multipopulation replicator equations and asymmetric game dynamics. First, the reinforcement scheme R need not lead to linear interactions. Second, the model does not require a static environment described by a constant bi-matrix (A, B). Finally, the occurrence of the memory loss term is entirely new and not found in game theory or evolutionary dynamics.

III. INFORMATION, UNCERTAINTY, AND DYNAMIC ADAPTATION

We now shift away from a dynamical systems view and, as promised earlier, begin to think of the agent collective as a communication system. Although, this initially will appear unrelated, we will show that there is a close connection between the dynamical and information theoretic perspectives—connections that have both mathematical and pragmatic consequences.

We consider the adaptive agents in the collective to be information sources. Each agent receives information from its environment, which includes other agents. Each agent interprets the received information and modifies its behavior accordingly, changing from $\mathbf{x}(t)$ to $\mathbf{x}(t+dt)$. Each agent generates a series of messages (actions) based on its updated internal model and introduces this new behavior back into the environment. This is a different interpretation of the interaction process in the collective which we motivated up to now only as a dynamical process. Now we discuss the adaptive dynamics from information theoretic viewpoint.

A. Dynamics in Information Space

In this section we introduce a new state space that directly represents the uncertainties of agent actions. First, as before, for clarity we focus on the two-agent static-environment case, Eqs. (20). Since the components of the agents' states are probabilities, the quantities

$$\begin{aligned} \xi_i &= -\log x_i \ , \\ \eta_j &= -\log y_j \ , \end{aligned} \tag{26}$$

are the *self-informations* of agents X and Y choosing actions i and j, respectively. When x_i is small, for example, the self-information ξ_i is large since action i is rarely chosen by agent X. Consider the resulting change in coordinates in $\mathbf{R}^N_+ \times \mathbf{R}^M_+$:

$$\boldsymbol{\Xi} = (\boldsymbol{\xi}, \boldsymbol{\eta}) = (\xi_1, \ldots, \xi_N) \times (\eta_1, \ldots, \eta_M) \ . \tag{27}$$

The normalization conditions—$\Sigma_{n=1}^N x_n = \Sigma_{m=1}^M y_m = 1$—that restrict the agent states to lie in simplices become $\Sigma_{n=1}^N e^{-\xi_n} = \Sigma_{m=1}^M e^{-\eta_m} = 1$ in $\boldsymbol{\Xi}$.

In this space the equations of motion become:

$$\begin{aligned} \dot{\xi}_i &= -\beta_X [(Ae^{-\boldsymbol{\eta}})_i - e^{-\boldsymbol{\xi}} \cdot Ae^{-\boldsymbol{\eta}}] - \alpha_X [\xi_i - e^{-\boldsymbol{\xi}} \cdot \boldsymbol{\xi}] \\ \dot{\eta}_j &= -\beta_Y [(Be^{-\boldsymbol{\xi}})_j - e^{-\boldsymbol{\eta}} \cdot Be^{-\boldsymbol{\xi}}] - \alpha_Y [\eta_j - e^{-\boldsymbol{\eta}} \cdot \boldsymbol{\eta}] \ , \end{aligned} \tag{28}$$

for $i = 1, \ldots, N$ and $j = 1, \ldots, M$ and where $e^{-\boldsymbol{\xi}} = (e^{-\xi_1}, \ldots, e^{-\xi_N})$ and $e^{-\boldsymbol{\eta}} = (e^{-\eta_1}, \ldots, e^{-\eta_N})$.

Recall that both the ΔR interaction term and the ΔH memory loss term are differences from means. This suggests yet another transformation to remove these comparisons to the mean:

$$\begin{aligned} u_i &= \xi_i - N^{-1} \sum_{k=1}^{N} \xi_k \ , \\ v_j &= \eta_j - M^{-1} \sum_{k=1}^{M} \eta_k \ , \end{aligned} \tag{29}$$

with $i = 1, \ldots, N$ and $j = 1, \ldots, M$. This leads to the normalized space in $\mathbf{R}^N \times \mathbf{R}^M$:

$$\mathbf{U} = (\mathbf{u}, \mathbf{v}) = (u_1, \ldots, u_N) \times (v_1, \ldots, v_M) \ , \tag{30}$$

with the constraints $\sum_{n=1}^N u_n = \sum_{m=1}^M v_m = 0$. $\mathbf{u}$ and $\mathbf{v}$ are the normalized self-informations relative to their means. We refer to this space as *information space*.

The combined coordinate transformation, Eq. (29) composed with Eq. (26), gives the well known *centered log-ratio* coordinates [30]. The inverse transformation is:

$$\begin{aligned} x_i &= \frac{e^{-u_i}}{\Sigma_k^N e^{-u_k}} \ , \\ y_i &= \frac{e^{-v_i}}{\Sigma_k^M e^{-v_k}} \ . \end{aligned} \tag{31}$$

The resulting transformed adaptation equations directly model the dynamics of uncertainties of agents' behavior:

$$\begin{aligned} \dot{\mathbf{u}} &= -\beta_X \left[A\mathbf{y} - \sum_k^N (A\mathbf{y})_k \right] - \alpha_X \mathbf{u} \ , \\ \dot{\mathbf{v}} &= -\beta_Y \left[B\mathbf{y} - \sum_k^N (B\mathbf{y})_k \right] - \alpha_Y \mathbf{v} \ . \end{aligned} \tag{32}$$

When the interaction matrices are normalized to zero mean, the equations simplify even further to

$$\begin{aligned} \dot{\mathbf{u}} &= -\beta_X A\mathbf{y} - \alpha_X \mathbf{u} \ , \\ \dot{\mathbf{v}} &= -\beta_Y B\mathbf{x} - \alpha_Y \mathbf{v} \ . \end{aligned} \tag{33}$$

The origin $\mathbf{O} = (0, 0, \ldots, 0)$ of the normalized information space $\mathbf{U}$ corresponds to random behavior: $(\mathbf{x}, \mathbf{y}) = (1/N, \ldots, 1/N, 1/M, \ldots, 1/M)$. The Shannon entropy of the choice distribution is maximized at this point. In contrast, when agents choose an action with probability 1 the entropy vanishes and the agent state is located in Δ at the simplex vertices and in $\mathbf{U}$ at infinity.

In Eqs. (33) the first term is related to information influx to an agent from outside; i.e., from other agents and the environment. The second term is related to the information dissipation due to internal memory loss. Eqs. (33) are useful for theory, for analysis in certain limits, as we will shortly demonstrate, and for numerical stability during simulation, which we will illustrate when considering example collectives below. Note that Eqs. (20), Eqs. (28), and Eqs. (32) are topologically orbit equivalent.

B. Self-organization Induced by Dynamics of Uncertainty

Equations (32) describe a dynamics of uncertainty between deterministic and random behavior. Information influx occurs when the agents adapt to environmental constraints and accordingly change their choice distribution. Information dissipation occurs when memory loss dominates and the agents increase their uncertainty to behave more randomly with less regard to the environmental constraints. The dissipation rate γ of the dynamics in $\mathbf{U}$ is controlled entirely by the memory loss rate α:

$$\begin{aligned} \gamma &= \sum_{k-1}^{N} \frac{\partial \dot{u}_k}{\partial u_k} + \sum_{k-1}^{M} \frac{\partial \dot{v}_k}{\partial v_k} \\ &= -N\alpha_X - M\alpha_Y \ . \end{aligned} \tag{34}$$

Therefore, Eqs. (33) are volume preserving in $\mathbf{U}$ when $\alpha_X = \alpha_Y = 0$.

In the case that agents behave without memory loss ($\alpha_X = \alpha_Y = 0$), if the interaction specified by (A, B) is zero-sum, $B = -A^T$, and if, in addition, it determines an interior Nash equilibrium $(\mathbf{x}^*, \mathbf{y}^*)$ (see App. B), then the collective has a constant of motion:

$$E = \beta_X^{-1} D(\mathbf{x}^* \parallel \mathbf{x}) + \beta_Y^{-1} D(\mathbf{y}^* \parallel \mathbf{y}) \ , \tag{35}$$

where $D(\mathbf{p} \parallel \mathbf{q}) = \Sigma_k p_k \log(p_k/q_k)$ is the *relative entropy* or the *information gain* which measures the similarity between probability distributions $\mathbf{p}$ and $\mathbf{q}$ [31]. (Appendix C gives the derivation of Eq. (35).) Since the constant of motion E is a linear sum of relative entropies, the collective maintains the information-theoretic distance between the interior Nash equilibrium and each agent's state. Thus, in the perfect memory case ($\alpha = 0$), by the inequality $D(\mathbf{p} \parallel \mathbf{q}) \geq 0$, the interior Nash equilibrium cannot be reached unless the initial condition itself starts on it (Fig. 5). This is an information-theoretic interpretation of the constant of motion noted in Ref. [32].

Moreover, when $N = M$ the dynamics has a symplectic structure in $\mathbf{U}$ with the Hamiltonian E given in Eq. (35) [32]. In this case, Eqs. (32) are described quite simply,

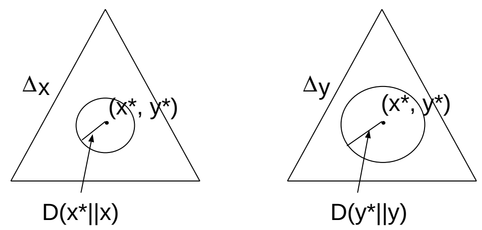

FIG. 5: Dynamics of zero-sum interaction without memory loss: Constant of motion $E = \beta_X^{-1} D(\mathbf{x}^* \parallel \mathbf{x}) + \beta_Y^{-1} D(\mathbf{y}^* \parallel \mathbf{y}$ keeps the linear sum of distance between the interior Nash equilibrium and each agent's state.

$$\dot{\mathbf{U}} = J \nabla_U E, \text{with} \ \ J = \begin{pmatrix} O & P \\ -P^T & O \end{pmatrix} \ , \tag{36}$$

where $P = -\beta_X \beta_Y A$. (Again, see App. C.)

When the bi-matrix interaction (A, B) satisfies $B = A^T$, E is a Lyapunov function of dynamics and decreases to 0 over time [2]. In this case, $(\mathbf{x}^*, \mathbf{y}^*)$ may not be in the interior of the collective simplices Δ.

In some cases when neither $B = -A^T$ nor $B = A^T$, E increases non-monotonically, the dynamics in $\mathbf{U}$ diverges, and the Shannon entropies of agents' choice distribution asymptotically decreases. (See Figs. 17 and 19 below.) Note that in single-agent adaptation with state $\mathbf{x}$ and normalizing the environment's reinforcements to a probability distribution $\mathbf{p}_e$, $D(\mathbf{p}_e \parallel \mathbf{x})$ is a Lyapunov function of the dynamics. It always decreases monotonically. In mutual adaptation, though, agents adapt to a dynamic environment that includes the other agents. As a result, in some cases, E, a linear sum of agent relative entropies, will itself exhibit nontrivial dynamics and, in addition, the uncertainties of agents' choices will asymptotically decrease.

When agents adapt with memory loss ($\alpha > 0$), the dynamics is dissipative. Since the memory loss terms induce information dissipation, the dynamics remains well inside the simplices.

Notably, when the agents attempt to achieve this balance together by interacting and, in particular, when the interaction has *nontransitive* structure, the dynamics can persistently wander in a bounded area in information space. Since, in some cases, mutual adaptation and memory loss produce successive stretching and folding, deterministic chaos can occur with decreasing α, even with only two agents. A schematic view of the flow in mutual adaptation is given in Fig. 6. Roughly, information space locally splits into subspaces governed by effects of mutual adaptation (information influx) and memory loss (information dissipation). Information influx and information dissipation correspond to unstable and stable flow directions. In this case, "weak" uncertainty of behavior plays an important role in organizing the collective's behavior. Small fluctuations in behavior are amplified through repeated mutual adaptation with competitive interactions.

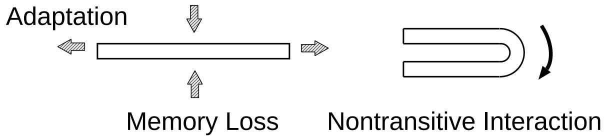

FIG. 6: Horseshoe in mutual adaptation: Effect of mutual adaptation and memory loss produce unstable and stable directions. The nontransitive structure of interactions leads to state-space folding.

Summarizing, in single-agent adaptation, information flows unidirectionally from the environment to the agent and the agent adapts its behavior to the environmental constraints. Adaptation leads to $D(\mathbf{p}_e \parallel \mathbf{x}) \to 0$. For mutual adaptation in an agent collective, however, information flow is multidirectional since each agent obtains information from its environment and organizes its behavior based on that information. In this situation, E need not be a Lyapunov function for the dynamics. As we will see, when the dynamics is chaotic, global information maximization is of doubtful utility. A dynamic view of collective adaptation is more appropriate in these cases.

Now consider many agents interacting. In the perfect memory case, when the game is zero-sum and has an interior Nash equilibrium $(\mathbf{x}^{1*}, \mathbf{x}^{2*}, \ldots, \mathbf{x}^{S*})$, following Eq. (35), the following constant of motion exists:

$$\begin{aligned} E &= \Sigma_{s=1}^{S} \frac{1}{\beta}_s D(\mathbf{x}^{s*} \parallel \mathbf{x}^s) \\ &= \Sigma_{s=1}^{S} \frac{1}{\beta}_s \left(\Sigma_k^{N^s} x_k^{s*} \log \frac{x_k^{s*}}{x_k} \right) . \end{aligned} \tag{37}$$

Although, strictly speaking, Hamiltonian dynamics and the associated symplectic structure of information space occurs only for two agents, one can describe multiple agent dynamics as a generalized Hamiltonian system [33].

In the general case with $\alpha > 0$, dissipative dynamics and high-dimensional chaotic flows can give rise to several unstable directions, since information influx has a network structure relative the other agents. At least S stable directions are expected since memory loss comes from each individual's internal dynamics. A detailed information-theoretic analysis along these lines will be reported elsewhere.

IV. EXAMPLES

To illustrate collective adaptation, we now give several examples of the dynamics in a static environment with two and three agents interacting via versions of Matching Pennies and Rock-Scissors-Paper, games with nontransitive structures. Appendix D gives the details of the reinforcement schemes for these cases. The agents will have equal adaptation rates ($\beta_X = \beta_Y = \cdots$) and the same number of actions ($N = M = L = \cdots$). In these simplified cases, the equations of motion for two agents are given by

$$\begin{aligned} \frac{\dot{x}_i}{x_i} &= [(A\mathbf{y})_i - \mathbf{x} \cdot A\mathbf{y}] + \alpha_X [H_i - H^X] \\ \frac{\dot{y}_j}{y_j} &= [(B\mathbf{x})_j - \mathbf{y} \cdot B\mathbf{x}] + \alpha_Y [H_j - H^Y] , \end{aligned} \tag{38}$$

for $i = 1, \ldots, N$ and $j = 1, \ldots, N$. A detailed analysis of this case with zero memory loss ($\alpha = 0$) is given in Ref. [2] in terms of asymmetric game dynamics. We will present results for zero and positive memory loss rates.

We then consider three agents, for which the adaptation equations are

$$\begin{aligned} \frac{\dot{x}_i}{x_i} &= [(A\mathbf{yz})_i - \mathbf{x} \cdot A\mathbf{yz}] + \alpha_X [H_i - H^X] \\ \frac{\dot{y}_j}{y_j} &= [(B\mathbf{zx})_j - \mathbf{y} \cdot B\mathbf{zx}] + \alpha_Y [H_j - H^Y] \\ \frac{\dot{z}_k}{z_k} &= [(C\mathbf{xy})_k - \mathbf{z} \cdot C\mathbf{xy}] + \alpha_Z [H_k - H^Z] . \end{aligned} \tag{39}$$

for $i, j, k = 1, \ldots, N$. We again will describe cases with and without memory loss.

Computer simulations are executed in the information space $\mathbf{U}$ and the results are shown in the state space X. We ignore the dynamics on the boundary of the simplex and concentrate the case that all variables are greater than 0 and less than 1.

A. Two Agents Adapting under Matching Pennies Interaction

In the matching pennies game, agents play one of two actions: heads (H) or tail (T). Agent X wins when the plays do not agree; agent Y wins when they do. Agent X's state space is $\Delta_X = (x_1, x_2)$ with $x_i \in (0, 1)$ and $x_1 + x_2 = 1$. That is, x_1 is the probability that agent X plays heads; x_2, tails. Agent Y is described similarly. Thus, each agent's state space is effectively one dimensional and the collective state space $\Delta = \Delta_X \times \Delta_Y$, two dimensional.

The environment for two agents interacting via the matching pennies game leads to the following matrices for Eqs. (38):

$$A = \begin{bmatrix} -\epsilon_X & \epsilon_X \\ \epsilon_X & -\epsilon_X \end{bmatrix} \text{ and } B = \begin{bmatrix} -\epsilon_Y & \epsilon_Y \\ \epsilon_Y & -\epsilon_Y \end{bmatrix} , \tag{40}$$

where $\epsilon_X \in (0.0, 1.0]$ and $-\epsilon_Y \in (0.0, 1.0]$.

Figure 7 shows a heteroclinic cycle of adaptation dynamics on the boundary of Δ when the αs vanish. Flows on the border occur only when agents completely ignore an action at the initial state; that is, when $x_i(0) = 0$ or $y_j(0) = 0$ for at least one i or j. Each vertex of the simplex is a saddle since the interaction is non-transitive.

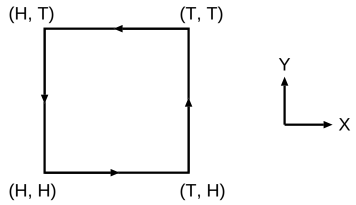

FIG. 7: Flows on the boundary in Matching Pennies interaction: Actions H and T correspond to "heads" and "tails", respectively. Arrows indicate the direction of adaptation dynamics on the boundary of the state space Δ.

The Nash equilibrium $(\mathbf{x}^*, \mathbf{y}^*)$ of the Matching Pennies game is in the center of Δ: $(\mathbf{x}^*, \mathbf{y}^*) = (\frac{1}{2}, \frac{1}{2}, \frac{1}{2}, \frac{1}{2})$ and this is also a fixed point of the adaptation dynamics. The Jacobian at $(\mathbf{x}^*, \mathbf{y}^*)$ is

$$J = \begin{pmatrix} -\frac{\alpha_X}{2}(1+\log 2) & -\frac{\epsilon_X}{2} \\ -\frac{\epsilon_Y}{2} & -\frac{\alpha_Y}{2}(1+\log 2) \end{pmatrix} \qquad (41)$$

and its eigenvalues are

$$\begin{aligned} \frac{4\lambda_i}{1+\log 2} &= -(\alpha_X + \alpha_Y) \\ &\pm \sqrt{(\alpha_X - \alpha_Y)^2 + 4\epsilon_X\epsilon_Y/(1+\log 2)^2} \end{aligned} \qquad (42)$$

In the perfect memory case ($\alpha_X = \alpha_Y = 0$), trajectories near $(\mathbf{x}^*, \mathbf{y}^*)$ are neutrally stable periodic orbits, since $\lambda_i = \pm\frac{1}{2}\sqrt{\epsilon_X\epsilon_Y}$ are pure imaginary. In the memory loss case ($\alpha_X > 0$ and $\alpha_Y > 0$), $(\mathbf{x}^*, \mathbf{y}^*)$ is globally asymptotically stable, since $\mathrm{Re}(\lambda_1)$ and $\mathrm{Re}(\lambda_2)$ are strictly negative. Examples of the trajectories in these two cases are given in Figure 8.

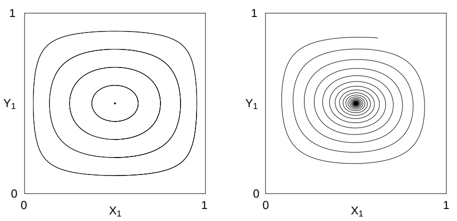

FIG. 8: Adaptation dynamics in Matching Pennies interaction: Here $\epsilon_X = 0.5$ and $\epsilon_Y = -0.3$ with (left) $\alpha_X = \alpha_Y = 0$ and (right) $\alpha_X = 0.02$ and $\alpha_Y = 0.01$.

B. Three Agents Adapting with Even-Odd Interaction

Now consider extending Matching Pennies for two agents so that it determines the interactions between three. Here we introduce the *Even-Odd* interaction in which there are again two actions, H and T, but agents win according to whether or not the number of heads in the group of three plays by the agents is even or odd. The environment now is given by, for agent X,

$$a_{ijk} = \begin{cases} \epsilon_X, & \text{number of } H\text{s is even} \\ -\epsilon_X, & \text{otherwise} \end{cases} \qquad (43)$$

with actions for agents X, Y, and Z given by $i, j, k = \{H, T\}$ and $\epsilon_X \in (0.0, 1.0]$. The interaction matrices b_{jki} and c_{kij} for agents Y and Z, respectively, are given similarly, but with $\epsilon_Y \in (0.0, 1.0]$ and $\epsilon_Z \in [-1.0, 0.0)$. Appendix D gives the details of the reinforcement scheme.

Following the reasoning used in Matching Pennies, the collective state space $\Delta = \Delta_X \times \Delta_Y \times \Delta_Z$ is now a solid three-dimensional cube. Figure 9 shows a heteroclinic network of adaptation dynamics on the boundary of Δ when αs vanish. Flows on Δ's boundary is shown in Fig. 9.

Δ is partitioned into four prism-shaped subspaces. Each prism subspace has a heteroclinic cycle on the face that is also a face of Δ.

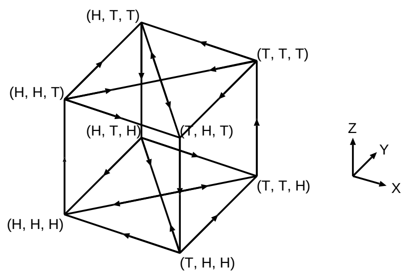

FIG. 9: Flows on the state space boundary under the Even-Odd interactions: H and T correspond to "heads" and "tails", respectively. Arrows indicate the direction of adaptation dynamics on Δ's boundary when the αs vanish.

The Nash equilibrium of the Even-Odd interaction is $(\mathbf{x}^*, \mathbf{y}^*, \mathbf{z}^*) = (\frac{1}{2}, \frac{1}{2}, \frac{1}{2}, \frac{1}{2}, \frac{1}{2}, \frac{1}{2})$ at the center of Δ and this is also a fixed point of the adaptation dynamics. The Jacobian there is

$$J = \begin{pmatrix} -\alpha_X & 0 & 0 \\ 0 & -\alpha_Y & 0 \\ 0 & 0 & -\alpha_Z \end{pmatrix} . \qquad (44)$$

Its eigenvalues are $\lambda = -\alpha_X, -\alpha_Y, -\alpha_Z$. Thus, in complete memory case ($\alpha_X = \alpha_Y = \alpha_Z = 0$), trajectories near $(\mathbf{x}^*, \mathbf{y}^*, \mathbf{z}^*)$ are neutrally stable periodic orbits. With memory decay ($\alpha_X, \alpha_Y, \alpha_Z > 0$), the $(\mathbf{x}^*, \mathbf{y}^*, \mathbf{z}^*)$ is globally asymptotically stable. The hyperbolic fixed points in the top and bottom faces are unstable in all cases. Examples of the trajectories are given in Figure 10.

Notably, when a single agent (say, Z) has memory loss and others have perfect memory, the crossed lines given by $\{z = x = 0.5,\ z = y = 0.5\}$ become an invariant subspace and trajectories are attracted to points in this subspace. Thus, there are infinitely many neutrally stable points. With $\alpha_X = \alpha_Y = 0$ and $\alpha_Z = 0.01$, for example, the adaptive dynamics alternates between a Matching Pennies interaction between agents X and Z by one between agents Y and Z during the transient relaxation to a point on the invariant subspace.

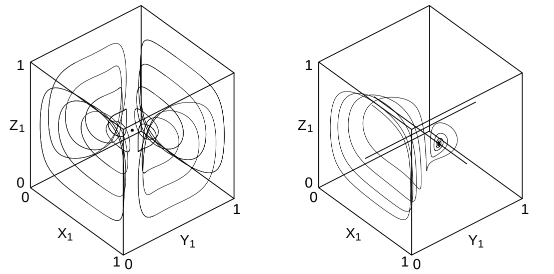

FIG. 10: Dynamics of adaptation in the Even-odd interaction: $\epsilon_X = 0.5$, $\epsilon_Y = 0.2$, and $\epsilon_Z = -0.3$ with $\alpha_X = \alpha_Y = \alpha_Z = 0$ in (left) and with $\alpha_X = \alpha_Y = 0$ and $\alpha_Z = 0.01$ in (right). Note that the neutral subspace is shown as the horizontal cross. The trajectory chosen illustrates the attraction to a point in this subspace.

C. Two Agents Adapting under Rock-Scissors-Paper Interactions

In this subsection, we give an example of an environment in which agents have three actions. One of the most commonly studied games with three actions is the Rock-Scissors-Paper (RSP) game, in which an agent playing Rock beats one playing Scissors, which in turn beats an agent playing Paper, which finally beats Rock.

First we examine two agents, which is a straightforward implementation of the RSP game and then extend the RSP interaction to three agents and analyze the higher-dimensional behavior. The interaction matrices for these cases are given in App. D.

Under the RSP interaction each agent has the option of playing one of three actions: "rock" (R), 'scissors" (S), and "paper" (P). Agent X's probability of playing these are denoted x_1, x_2, and x_3 and $x_1 + x_2 + x_3 = 1$. Agent Y probabilities are given similarly. Thus, the agent state spaces, Δ_X and Δ_Y, are each two dimensional simplices, and the collective state space $\Delta = \Delta_X \times \Delta_Y$ is four dimensional.

For two agents the environment is given by the interaction matrices

$$A = \begin{bmatrix} \epsilon_X & 1 & -1 \\ -1 & \epsilon_X & 1 \\ 1 & -1 & \epsilon_X \end{bmatrix} \text{ and } B = \begin{bmatrix} \epsilon_Y & 1 & -1 \\ -1 & \epsilon_Y & 1 \\ 1 & -1 & \epsilon_Y \end{bmatrix} , \quad (45)$$

where $\epsilon_X, \epsilon_Y \in [-1.0, 1.0]$ are the rewards for ties and normalized to

$$A' = \begin{bmatrix} \frac{2}{3}\epsilon_X & 1 - \frac{1}{3}\epsilon_X & -1 - \frac{1}{3}\epsilon_X \\ -1 - \frac{1}{3}\epsilon_X & \frac{2}{3}\epsilon_X & 1 - \frac{1}{3}\epsilon_X \\ 1 - \frac{1}{3}\epsilon_X & -1 - \frac{1}{3}\epsilon_X & \frac{2}{3}\epsilon_X \end{bmatrix} \quad (46)$$

and

$$B' = \begin{bmatrix} \frac{2}{3}\epsilon_Y & 1 - \frac{1}{3}\epsilon_Y & -1 - \frac{1}{3}\epsilon_Y \\ -1 - \frac{1}{3}\epsilon_Y & \frac{2}{3}\epsilon_Y & 1 - \frac{1}{3}\epsilon_Y \\ 1 - \frac{1}{3}\epsilon_Y & -1 - \frac{1}{3}\epsilon_Y & \frac{2}{3}\epsilon_Y \end{bmatrix} . \quad (47)$$

Note that the reinforcements are normalized to zero mean and that this does not affect the dynamics.

The flow on Δ's boundary is shown in Fig. 11. This represents the heteroclinic network of adaptation dynamics on Δ's edges when the αs vanish. Each vertex is a saddle since the interaction has non-transitive structure.

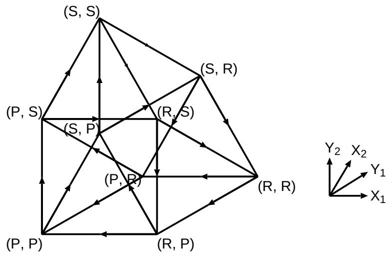

FIG. 11: Flows on the boundary of the simplex in the Rock-Scissors-Paper interaction for two agents: R, S, and P denote "rock", "scissors", and "paper", respectively. The arrows indicate the direction of the adaptation dynamics on the boundary of the collective state space Δ when the αs vanish.

The Nash equilibrium $(\mathbf{x}^*, \mathbf{y}^*)$ is given by the centers of the simplex:

$$(\mathbf{x}^*, \mathbf{y}^*) = (\frac{1}{3}, \frac{1}{3}, \frac{1}{3}, \frac{1}{3}, \frac{1}{3}, \frac{1}{3}) . \quad (48)$$

This is also a fixed point of the adaptation dynamics. The Jacobian there is

$$J = \begin{pmatrix} -\alpha_X & 0 & \frac{1+\epsilon_X}{3} & \frac{2}{3} \\ 0 & -\alpha_X & -\frac{2}{3} & \frac{-1+\epsilon_X}{3} \\ \frac{1+\epsilon_Y}{3} & \frac{2}{3} & -\alpha_Y & 0 \\ -\frac{2}{3} & \frac{-1+\epsilon_Y}{3} & 0 & -\alpha_Y \end{pmatrix} . \quad (49)$$

Its eigenvalues are

$$2\lambda_i = -(\alpha_X + \alpha_Y)$$

$$\pm\sqrt{(\alpha_X - \alpha_Y)^2 + \frac{4\left(\epsilon_X\epsilon_Y - 3 \pm \sqrt{-3(\epsilon_X + \epsilon_Y)^2}\right)}{9}} . \quad (50)$$

Thus, when (A, B) is zero-sum ($\epsilon_X + \epsilon_Y = 0$) and agents have complete memory ($\alpha_X = \alpha_Y = 0$), trajectories near $(\mathbf{x}^*, \mathbf{y}^*)$ are neutrally stable periodic orbits since all λ's are pure imaginary. The dynamics is Hamiltonian in this case. With memory decay ($\alpha_X, \alpha_Y > 0$), and $|\alpha_X - \alpha_Y| < \frac{2}{3}(\epsilon_X^2 + 3)$, $(\mathbf{x}^*, \mathbf{y}^*)$ is globally asymptotically stable.

For the nonzero-sum case, we will give examples of dynamics with $\epsilon_X = 0.5$, $\epsilon_Y = -0.3$, $\alpha_Y = 0.01$. In this case, when $\alpha_X > \alpha_c$, $(\mathbf{x}^*, \mathbf{y}^*)$ is globally asymptotically stable. At the point $\alpha_c \sim 0.055008938$, period-doubling bifurcation occurs. The example of two agents adapting in the Rock-Scissors-Paper interaction adaptation dynamics illustrates various types of low-dimensional chaos. We now explore several cases.

1. *Hamiltonian Limit*

When the agent memories are perfect ($\alpha_X = \alpha_Y = 0$) and the game is zero-sum ($\epsilon_X = -\epsilon_Y$), the dynamics in the information space $\mathbf{U}$ is Hamiltonian with a function consists of relative entropy $E = D(\mathbf{x}^* \parallel \mathbf{x}) + D(\mathbf{y}^* \parallel \mathbf{y})$. The left columns of Figs. 12 and 13 give trajectories in the collective state space Δ, while the plots given in the middle and right columns are these trajectories projected onto the individual agent simplices, Δ_X and Δ_Y. The trajectories were generated using a 4th-order symplectic integrator [34] in $\mathbf{U}$.

When $\epsilon_X = -\epsilon_Y = 0.0$ it appears that the dynamics is integrable, since only quasiperiodic tori exist for all initial conditions. With some initial conditions, the tori is knotted to form trefoil. Otherwise, when $\epsilon_X = -\epsilon_Y > 0.0$, Hamiltonian chaos occurs with positive-negative pairs of Lyapunov exponents. (See Table I.) The game-theoretic behavior of this example was investigated briefly in Ref. [16]. The dynamics is very rich. For example, there are infinitely many distinct behaviors near the fixed point at the center—the interior Nash equilibrium—and a periodic orbit arbitrarily close to any chaotic one.

A more detailed view of the complex dynamics is given in Figure 14 which shows Poincaré sections of Eqs. (38)'s trajectories. The Poincaré section is given by $\dot{u}_3 > 0$ and $\dot{v}_3 = 0$. In $(\mathbf{x}, \mathbf{y})$ space the section is determined by the constraints:

$$\begin{aligned} (1-\epsilon_X)y_1 \;&-\; (1+\epsilon_X)y_2 + \frac{2}{3}\epsilon_X < 0 \ , \\ (1-\epsilon_Y)x_1 \;&-\; (1+\epsilon_Y)x_2 + \frac{2}{3}\epsilon_Y = 0 \ . \end{aligned} \tag{51}$$

These sections are indicated as the straight lines drawn in the Δ_X simplices of Figs. 12 and 13. In Figure 14, when $\epsilon_X = -\epsilon_Y = 0.0$, closed loops depending on the initial conditions exhibits tori in the Poincaré section. When $\epsilon_X = -\epsilon_Y = 0.5$, some tori collapse and become chaotic. The scatter of dots among the remaining closed loop shows characteristic Hamiltonian chaos.

Table I shows Lyapunov spectra in $\mathbf{U}$ for dynamics with $\epsilon_X = -\epsilon_Y = 0.0$ and $\epsilon_X = -\epsilon_Y = 0.5$ with initial condition $(\mathbf{x}(0), \mathbf{y}(0)) = (x_1, 0.35, 0.65 - x_1, 0.1, y_2, 0.9 - y_2)$ with $E = E_0 = 0.74446808$ fixed. (x_1, y_2) satisfies

$$\frac{e^{-3(E_0 + 2\log 3)}}{0.035} = x_1(0.65 - x_1)y_2(0.9 - y_2). \tag{52}$$

When $x_1(0) = 0.05$, the initial condition is (B): $(\mathbf{x}, \mathbf{y}) = (0.05, 0.35, 0.6, 0.1, 0.2, 0.7)$, which we gave in the preceding examples.

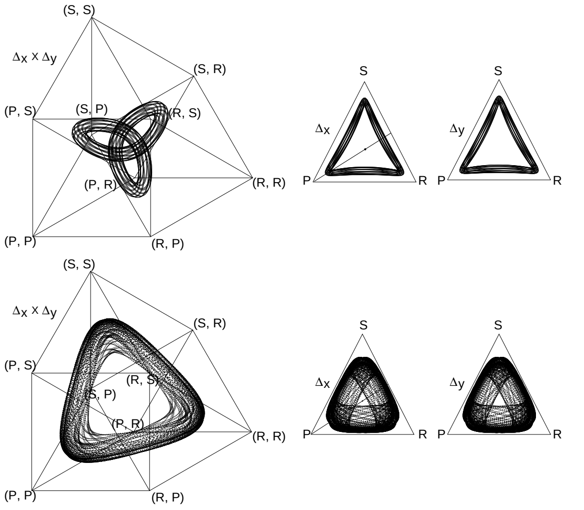

FIG. 12: Quasiperiodic tori: Collective dynamics in Δ (left column) and individual dynamics projected onto Δ_X and Δ_Y respectively (right two columns). Here $\epsilon_X = -\epsilon_Y = 0.0$ and $\alpha_X = \alpha_Y = 0$. The initial condition is (A): $(\mathbf{x}, \mathbf{y}) = (0.26, 0.113333, 0.626667, 0.165, 0.772549, 0.062451)$ for the top and (B): $(\mathbf{x}, \mathbf{y}) = (0.05, 0.35, 0.6, 0.1, 0.2, 0.7)$ for the bottom. The constant of motion (Hamiltonian) is $E = 0.74446808 \equiv E_0$. The Poincaré section used for Fig. 14 is given by $x_1 = x_2$ and $y_1 < y_2$ and is indicated here as the straight diagonal line in agent X's simplex Δ_X.

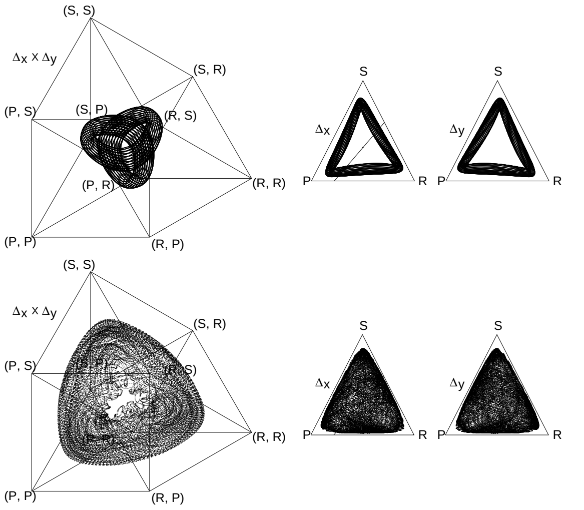

FIG. 13: Quasiperiodic tori and chaos: Collective dynamics in Δ (left column) and individual dynamics projected onto Δ_X and Δ_Y, respectively (right two columns). Here $\epsilon_X = -\epsilon_Y = 0.5$ and $\alpha_X = \alpha_Y = 0$. The initial conditions are the same as in Fig. 12, (A) for top row and (B) for bottom rows, respectively. Also, the constant of motion is the same: $E = E_0$. The Poincaré section is given by $3x_1 - x_2 - 2/3 = 0$ and $y_1 - 3y_2 + 2/3 < 0$ and this is indicated as a straight line in Δ_X.

When $\epsilon_X = 0.5$, the Lyapunov exponents indicate positive-negative pairs for $x_1(0) = 0.05, 0.06$ and 0.08, which clearly show Hamiltonian chaos. Note that $\lambda_2 \simeq 0.0$, $\lambda_3 \simeq 0.0$, and $\lambda_4 \simeq -\lambda_1$, as expected.

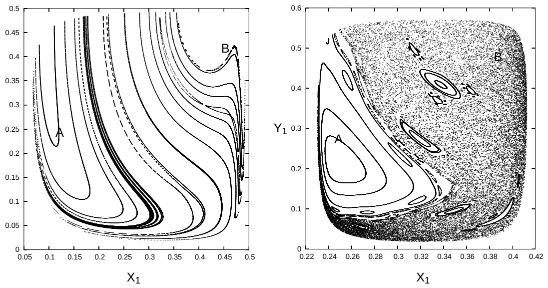

FIG. 14: Poincaré sections of the behavior in the preceding two figures. That is, $\epsilon_X = -\epsilon_Y = 0.0$ (left) and $\epsilon_X = -\epsilon_Y = 0.5$ (right). The Poincaré section is given by $x_1 = x_2$ and $y_1 < y_2$ (left) and $3x_1 - x_2 - 2/3 = 0$ and $y_1 - 3y_2 + 2/3 < 0$ (right). There are 25 randomly selected initial conditions, including the two, (A) and (B), used in Figs. 12 and 13. The constant of motion ($E = E_0$) forms the outer border of the Poincaré sections.

ϵ_X	λ	$x_1(0)$=0.05	0.06	0.07	0.08	0.09	0.10
	λ_1	+0.881	+0.551	+0.563	+0.573	+0.575	+0.589
0.0	λ_2	+0.436	+0.447	+0.464	+0.467	+0.460	+0.461
	λ_3	−0.436	−0.447	−0.464	−0.467	−0.460	−0.461
	λ_4	−0.881	−0.551	−0.563	−0.573	−0.575	−0.589
	λ_1	**+36.4**	**+41.5**	+0.487	**+26.3**	+0.575	+0.487
0.5	λ_2	+0.543	+0.666	+0.204	+0.350	+0.460	+0.460
	λ_3	−0.637	−0.666	−0.197	−0.338	−0.460	−0.467
	λ_4	**−36.3**	**−41.5**	−0.494	**−26.3**	−0.575	−0.480

TABLE I: Lyapunov spectra for different initial conditions (columns) and different values of the tie breaking parameter ϵ_X. The initial conditions are $(x_1, x_2, x_3, y_1, y_2, y_3) = (x_1, 0.35, 0.65 - x_1, 0.1, y_2, 0.9 - y_2)$ with $E = E_0 = 0.74446808$ fixed. We choose the initial conditions $(x_1, y_2) = (0.05, 0.2)$, $(0.06, 0.160421)$, $(0.07, 0.135275)$, $(0.08, 0.117743)$, $(0.09, 0.104795)$, $(0.10, 0.0948432)$. The Lyapunov exponents are multiplied by 10^3. Note that $\lambda_2 \simeq 0.0$, $\lambda_3 \simeq 0.0$ and $\lambda_4 \simeq -\lambda_1$ as expected. The Lyapunov exponents indicating chaos are shown in boldface.

2. *Conservative Dynamics*

With perfect memory ($\alpha_X = \alpha_Y = 0$) and a game that is not zero-sum ($\epsilon_X \neq -\epsilon_Y$) the dynamics is conservative in $\mathbf{U}$ and one observes transients that are attracted to heteroclinic networks in the state space X. (See Fig. 15.)

When $\epsilon_X + \epsilon_Y < 0$, the behavior is intermittent and orbits are guided by the flow on Δ's edges, which describes a network of possible heteroclinic cycles. Since action ties are not rewarded there is only one such cycle. It is shown in the top row of Fig. (15): $(R, P) \rightarrow (S, P) \rightarrow (S, R) \rightarrow (P, R) \rightarrow (P, S) \rightarrow (R, S) \rightarrow (R, P)$. Note that during the cycle each agent switches between almost deterministic actions in the order $R \rightarrow S \rightarrow P$. The agents are out of phase with respect to each other and they alternate winning each turn.

With $\epsilon_X + \epsilon_Y > 0$, however, the orbit is an infinitely persistent chaotic transient [35]. Since, in this case, agent X can choose a tie, the cycles are not closed. For example, with $\epsilon_X > 0$, at (R, P), X has the option of moving to (P, P) instead of (S, P) with a positive probability. This embeds an instability along the heteroclinic cycle and so orbits are chaotic. (See Fig. 15, bottom row.)

Figure 16 shows the time series for these behaviors. Usually, in transient relaxation to heteroclinic cycle, the duration over which orbits stay near saddle vertices increases exponentially. However, for our case, it appears to increase subexponentially. This is because of the very small exponent; $(1 + \delta)^n \sim 1 + n\delta + \dots$ $(\delta << 1)$. In the second chaotic transient case, it still increases subexponentially, but the visited vertices change irregularly.

Figure 17 shows the behavior of H^X, H^Y, and E. For both cases E eventually increases monotonically and H^X and H^Y asymptotically decrease. The agents show a tendency to decrease choice uncertainty and to switch between almost deterministic actions. H^X and H^Y oscillate over the range $[0, \log 2]$ for $\epsilon_X = -0.1$ and $\epsilon_Y = 0.05$ and over $[0, \log 3]$ for $\epsilon_X = 0.1$ and $\epsilon_Y = -0.05$.

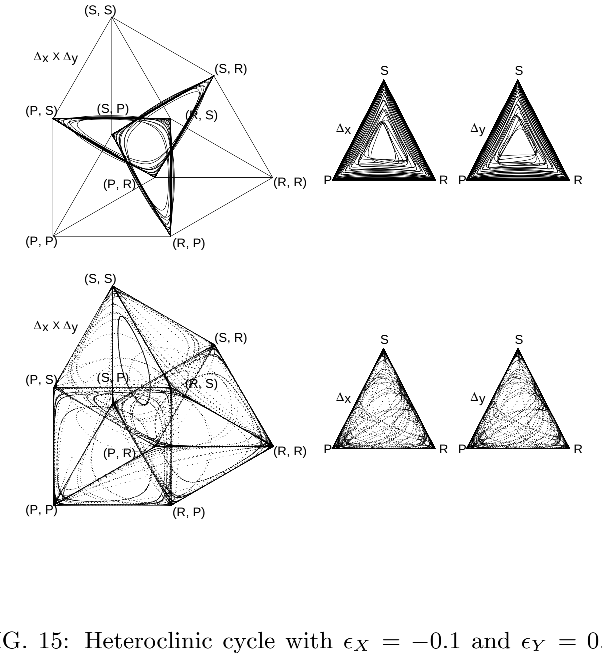

FIG. 15: Heteroclinic cycle with $\epsilon_X = -0.1$ and $\epsilon_Y = 0.05$ (top row). Chaotic transient to a heteroclinic network (bottom row) with $\epsilon_X = 0.1$ and $\epsilon_Y = -0.05$). For both $\alpha_X = \alpha_Y = 0$.

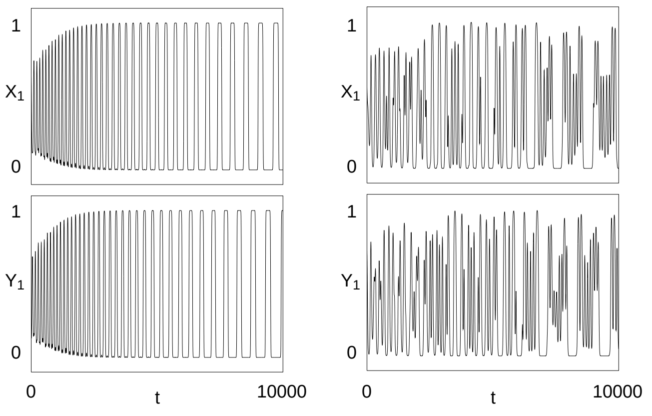

FIG. 16: Time series of action probabilities during the heteroclinic cycles of Fig. 15. $\epsilon_X = -0.1$ and $\epsilon_Y = 0.05$ for the left column. The right column shows the chaotic transient to a possible heteroclinic cycles when $\epsilon_X = 0.1$ and $\epsilon_Y = -0.05$. For both $\alpha_X = \alpha_Y = 0$.

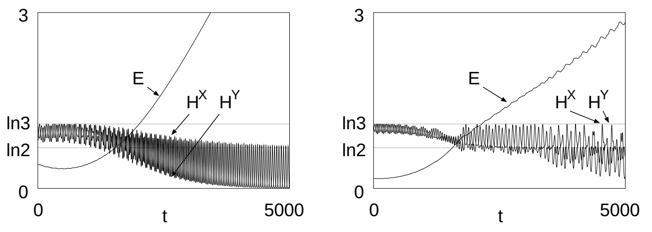

FIG. 17: Dynamics of H^X, H^Y and E in conservative adaptive dynamics: $\epsilon_X = -0.1$ and $\epsilon_Y = 0.05$ for the left plot and $\epsilon_X = 0.1$ and $\epsilon_Y = -0.05$ for the right. For both $\alpha_X = \alpha_Y = 0$. Note that E increases asymptotically and H^X and H^Y tend to decrease.

3. Dissipative Dynamics

If the memory loss rates (α_X and α_Y) are positive, the dynamics becomes dissipative in information space $\mathbf{U}$ and exhibits limit cycles and chaotic attractors. (See Fig. 18.) Figure 20 (top) shows a diverse range of bifurcations as a function of α_X. It shows the dynamics on the surface specified by $\dot{u}_3 < 0$ and $\dot{v}_3 = 0$ projected onto v_3. The fixed point $(\mathbf{x}^*, \mathbf{y}^*)$ becomes unstable when α_X is larger than $\alpha_c \approx 0.055008938$. Typically, period-doubling bifurcation to chaos occurs with decreasing α_X. Chaos can occur only when $\epsilon_X + \epsilon_Y > 0$ [17].

Figure 19 shows dynamics of H^X, H^Y, and E in dissipative adaptive dynamics. For both cases shown E does not diverge due to memory loss. When $\alpha_X = 0.025$, H^X and H^Y converge to oscillations over the range $[\log 2, \log 3]$. When $\alpha_X = 0.01$, H^X and H^Y exhibit chaotic behavior over the range $[0, \log 3]$.

Figure 20 (bottom) shows that the largest Lyapunov exponent in $\mathbf{U}$ is positive across a significant fraction of the parameter space; indicating that chaos is common. The dual aspects of chaos, coherence and irregularity, imply that agents may behave cooperatively or competitively (or switch between both). This ultimately derives from agents' successive mutual adaptation and memory loss in non-transitive interactions, such as in the RSP game; as was explained in Sec. III. Note that such global behavior organization is induced by each agents' self-interested and myopic adaptation and "weak" uncertainty of their environment.

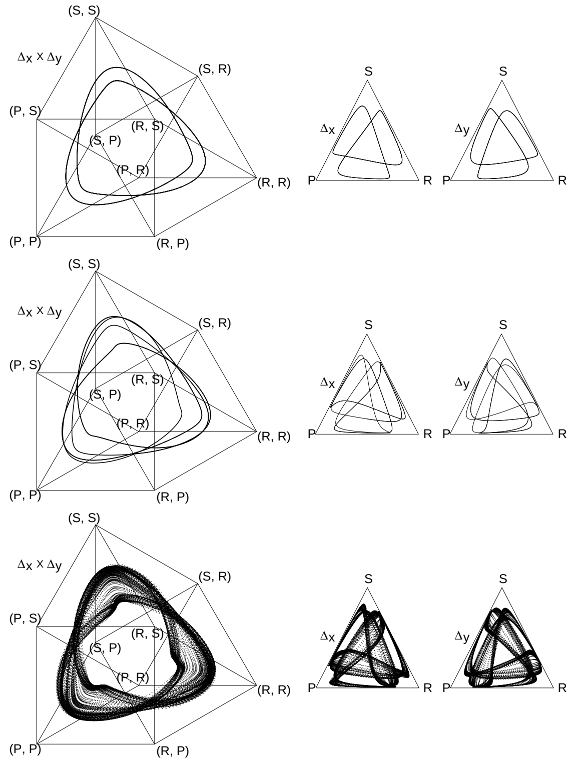

FIG. 18: Dissipative adaptive dynamics: Stable limit cycle for $\alpha_X = 0.025$ (top), $\alpha_X = 0.021$ (middle) and chaotic attractors with $\alpha_X = 0.0198$ (bottom). All cases have $\epsilon_X = 0.5$, $\epsilon_Y = -0.3$ and $\alpha_Y = 0.01$. Period-doubling bifurcation to chaos occurs with decreasing α_X.

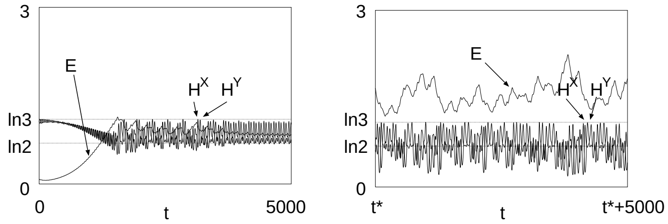

FIG. 19: Dynamics of H^X, H^Y, and E in dissipative adaptive dynamics: $\epsilon_X = 0.5$, $\epsilon_Y = -0.3$, and $\alpha_Y = 0.01$ for both. $\alpha_X = 0.025$ for the left plot and $\alpha_X = 0.01$ for the right. $t^* \approx 10^8$ in the right figure is the (rather long) transient time. In both cases E does not diverge due to memory loss.

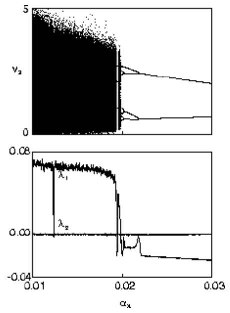

FIG. 20: Bifurcation diagram (top) of dissipative dynamics (adapting with memory loss) projected onto coordinate v_3 from the Poincaré section ($\dot{u}_3 > 0$, $\dot{v}_3 = 0$) and the largest two Lyapunov exponents λ_1 and λ_2 (bottom) as a function of $\alpha_Y \in [0.01, 0.03]$. Here with $\epsilon_X = 0.5$, $\epsilon_Y = -0.3$ and $\alpha_Y = 0.01$. Simulations show that λ_3 and λ_4 are always negative.

D. Three Agents Adapting under Rock-Scissors-Paper Interactions

Consider three agents adapting via (an extension of) the RSP interaction. Here the environment is given by the following interaction

$$a_{ijk} = \begin{cases} 2 & \text{Win over the others.} \\ -2 & \text{Lose to the other two.} \\ 1 & \text{Win over one other.} \\ -1 & \text{Lose to one other.} \\ \epsilon_X & \text{Tie.} \end{cases} \quad (53)$$

and similarly for b_{jki} and c_{kij}, with $i, j, k = \{R, S, P\}$. Here $\epsilon_X, \epsilon_Y, \epsilon_Z \in (-1.0, 1.0)$. (See App. D for the detailed listing of the reinforcement scheme.) As before we use normalized a'_{ijk}, b'_{jki}, and c'_{kij}:

$$a'_{ijk} = \begin{cases} 2 - \frac{\epsilon_X}{5} & \text{Win over the others.} \\ -2 - \frac{\epsilon_X}{5} & \text{Lose to the other two.} \\ 1 - \frac{\epsilon_X}{5} & \text{Win over one other.} \\ -1 - \frac{\epsilon_X}{5} & \text{Lose to one other.} \\ \frac{4}{5}\epsilon_X & \text{Tie.} \end{cases} \quad (54)$$

The normalization does not affect the dynamics.

The Nash equilibrium $(\mathbf{x}^*, \mathbf{y}^*, \mathbf{z}^*)$ is at the simplex center:

$$(\mathbf{x}^*, \mathbf{y}^*, \mathbf{z}^*) = \frac{1}{3}(1,1,1,1,1,1,1,1,1) \ . \quad (55)$$

It is also a fixed point of the adaptation dynamics. The Jacobian there is

$$J = \begin{pmatrix} -\alpha_X & 0 & \frac{1}{3} & \frac{2}{3} & \frac{1}{3} & \frac{2}{3} \\ 0 & -\alpha_X & -\frac{2}{3} & -\frac{1}{3} & -\frac{2}{3} & -\frac{1}{3} \\ \frac{1}{3} & \frac{2}{3} & -\alpha_Y & 0 & \frac{1}{3} & \frac{2}{3} \\ -\frac{2}{3} & -\frac{1}{3} & 0 & -\alpha_Y & -\frac{2}{3} & -\frac{1}{3} \\ \frac{1}{3} & \frac{2}{3} & \frac{1}{3} & \frac{2}{3} & -\alpha_Z & 0 \\ -\frac{2}{3} & -\frac{1}{3} & -\frac{2}{3} & -\frac{1}{3} & 0 & -\alpha_Z \end{pmatrix} . \quad (56)$$

When $\alpha_X = \alpha_Y = \alpha_Z = \alpha$, its eigenvalues are

$$\lambda_i + \alpha = \frac{i}{\sqrt{3}}(-1, -1, -2, 1, 1, 2) \ . \quad (57)$$

In the perfect memory case ($\alpha_X = \alpha_Y = \alpha_Z = 0$), trajectories near $(\mathbf{x}^*, \mathbf{y}^*, \mathbf{z}^*)$ are neutrally stable periodic orbits, since the λs are pure imaginary. In the memory loss case ($\alpha_X, \alpha_Y, \alpha_Z > 0$), $(\mathbf{x}^*, \mathbf{y}^*, \mathbf{z}^*)$ is asymptotically stable, since all $\mathrm{Re}(\lambda_i)$ are strictly negative. One expects multiple attractors in this case.

The collective state space Δ is now 6 dimensional, being the product of three two-dimensional agent simplices $\Delta = \Delta_X \times \Delta_Y \times \Delta_Z$. The flow on Δ's boundary is shown in Fig. 21, giving the adaptation dynamics on the edges of Δ when the αs vanish.

We give two examples with $\alpha_X = \alpha_Y = \alpha_Z = 0.01$, $\epsilon_X = 0.5$, $\epsilon_Y = -0.365$, $\epsilon_Z = 0.8$ (top: limit cycle) and $\epsilon_X = 0.5$, $\epsilon_Y = -0.3$, $\epsilon_Z = 0.6$ (bottom: chaos) in Fig. 22. Chaos is typically observed when $\epsilon_X + \epsilon_Y + \epsilon_Z > 0$. Limit cycles are highly complex manifolds depending on the 6-dimensional heteroclinic network on the simplex boundary.

The Lyapunov spectrum for the chaotic dynamics is $(\lambda_1, \ldots, \lambda_6) = (+45.2, +6.48, -0.336, -19.2, -38.5, -53.6) \times 10^{-3}$. The dynamics has two positive Lyapunov exponents. Note that this dynamics could have many neutrally stable subspaces in three or more dimensions. These subspaces act as quasistable attractors and may even have symplectic structure. These properties of high-dimensional dynamics will be reported elsewhere.

V. CONCLUDING REMARKS

We developed a class of dynamical systems for collective adaptation. We started with very simple agents, whose adaptation was a dynamic balance between adaptation to environmental constraints and memory loss. A macroscopic description of a network of adaptive agents was produced. In one special case we showed that the dynamical system reduces to replicator equations, familiar in evolutionary game theory and population biology. In a more general setting, we investigated several of the resulting periodic, intermittent, and chaotic behaviors in which agent-agent interactions were explicitly given as game interactions. Self-organization induced by information flux was discussed using an information-theoretic viewpoint.

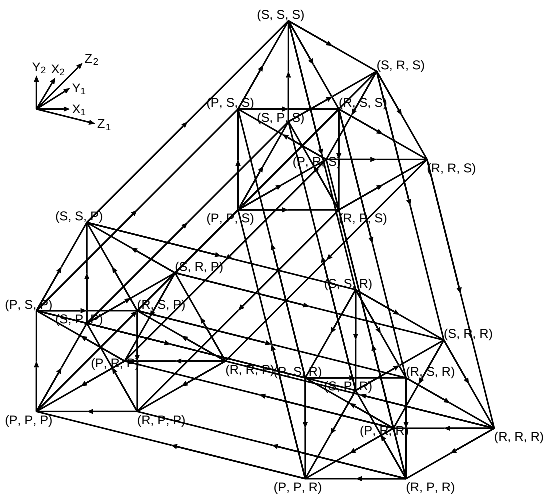

FIG. 21: Flows on the simplex edges in three-agent RSP: Arrows indicate the direction of adaptation dynamics on Δ's boundary when the αs vanish.

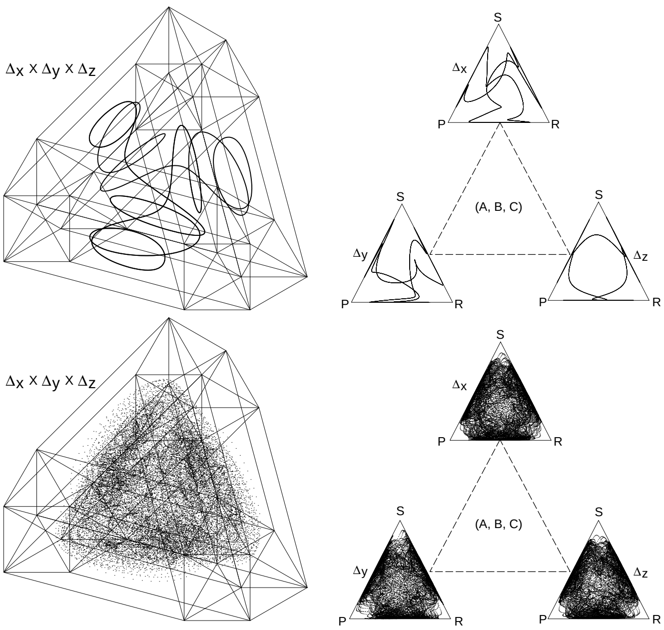

FIG. 22: Periodic orbit (top: $\epsilon_X = 0.5$, $\epsilon_Y = -0.365$, $\epsilon_Z = 0.8$) and chaotic orbit (bottom: $\epsilon_X = 0.5$, $\epsilon_Y = -0.3$, $\epsilon_Z = 0.6$) with the other parameters are $\alpha_X = \alpha_Y = \alpha_Z = 0.01$. The Lyapunov spectrum for chaotic dynamics is $(\lambda_1, \ldots, \lambda_6) = (+45.2, +6.48, -0.336, -19.2, -38.5, -53.6) \times 10^{-3}$.

We pointed out that unlike single-agent adaptation, information flow is multidirectional in collective adaptation and a dynamical analysis is required. We also noted that for nontransitive interactions deterministic chaos occurs due to the agents' local adaptation which amplifies fluctuations in behavior and to memory loss stabilizing behavior. Since deterministic chaos occurs even in this simple setting, one expects that in higher-dimensional and heterogeneous adaptive systems intrinsic unpredictability would become a dominant collective behavior.

We close by indicating some future directions in which to extent the model. First, as we alluded to during the development, there are difficulties of scaling the model to large numbers of agents. We focused on collectives with global coupling between all agents. However, in this case, the complexity of interaction terms grows exponentially with number of agents, which is both impractical from the viewpoints of analysis and simulation and unrealistic for nature systems that are large collectives. The solution to this, given in App. E, is to develop either spatially distributed agents collectives or to extend the equations to include explicit communication networks between agents. Both of these extensions will be helpful in modeling the many adaptive collectives noted in the introduction.

Second, important for applications, is to develop the stochastic generalization of the deterministic equations of motion which accounts for the effects of finite and fluctuating numbers of agents and also finite histories for adaptation. Each of these introduces its own kind of sampling stochasticity and will require a statistical dynamics analysis reminiscent of that found in population genetics [36].

One necessary and possibly difficult extension will be to agents that adapt continuous-valued actions—say, learning the spatial location of objects—to their environments. Mathematically, this requires a continuous-space extension of the adaptation equations (Eq. (15)) and this results in models that are described by PDEs [37].

Finally, another direction, especially useful if one attempts to quantify global function in large collectives, will be structural and information-theoretic analyses of local and global adaptive behaviors [38, 39]. Analyzing the stored information and the causal architecture [40, 41] in each agent versus that in the collective, communication in large networks, and emerging hierarchical structures in collective adaptation are projects now made possible using this framework.

Acknowledgments

The authors thank D. Albers, K. Engo, D. Krakauer, I. Tsuda, and J. D. Farmer for helpful discussions. YS was supported by the Special Postdoctoral Researchers Program at RIKEN. EA was supported by the Grant-in-Aid for Young Scientists (B) No.14780342, The Ministry of Education, Culture, Sports, Science and Technology, Japan. This work was also supported at the Santa Fe Institute under the Network Dynamics Program by Intel Corporation and under DARPA contract F30602-00-2-0583.

APPENDIX A: CONTINUOUS TIME

Here we give the derivation of the continuous-time limits that lead to the differential equations from the original stochastic discrete-time adaptation model.

Denote the agent-agent interaction time scale, number of interactions per adaptation interval, and adaptation time scale as $d\tau$, T, and t, respectively. We assume that adaptation is very slow compared to agent-agent interactions and take the limits $d\tau \to 0$ and $T \to \infty$, keeping $dt = T d\tau$ finite. Then we take the limit $dt \to 0$ to get the derivative of the vector $\mathbf{Q}(t)$.

With Eq. (2) and $Q_i^X(0) = 0$, we have

$$Q_i^X(T) = \frac{1}{T}\left[\Sigma_{k=1}^T \delta_{ij}(k) r_{ij}^X(k) - \alpha_X \Sigma_{k=1}^T Q_i^X(k)\right] . \tag{A1}$$

Thus, for continuous-time, when action i is chosen at step n,

$$\begin{aligned} &\frac{Q_i^X(t+dt) - Q_i^X(t)}{dt} \\ &= \frac{1}{Tdt}\left[\Sigma_{k=Tt}^{T(t+dt)} \delta_{ij}(\frac{k}{T}) r_{ij}^X(\frac{k}{T}) - \alpha_X \Sigma_{k=Tt}^{T(t+dt)} Q_i^X(\frac{k}{T})\right] . \end{aligned} \tag{A2}$$

Taking $T \to \infty$ and $d\tau \to 0$, we have

$$\begin{aligned} &\frac{Q_i^X(t+dt) - Q_i^X(t)}{dt} \\ &= \frac{1}{dt}\int_t^{t+dt} \delta_{ij}(s) r_{ij}^X(s) ds - \frac{\alpha_X}{dt}\int_t^{t+dt} Q_i^X(s) ds . \end{aligned} \tag{A3}$$

Assuming $r_{ij}(t)$ changes as slowly as the adaptive dynamics, $r_{ij}(t)$ is constant during the adaptation interval $t \sim t + dt$. If we assume in addition that the behaviors of two agents X and Y are statistically independent at time t, then the law of the large numbers gives

$$\frac{1}{dt}\int_t^{t+dt} \delta_{ij}(s) r_{ij}^X(s) ds \to \Sigma_{j=1}^M r_{ij}(t) y_j(t) \equiv R_i^X(t) . \tag{A4}$$

Now take $dt \to 0$. Eqs. (A4) and (A4) together give

$$\dot{Q}_i^X(t) = R_i^X(t) - \alpha_X Q_i^X(t) \tag{A5}$$

for the continuous-time updating of the reinforcement memory. When environmental change is slow compared to adaptation, then

$$R_i^X(t) = \Sigma_k^N a_{ik} y_i(t) . \tag{A6}$$

The single-agent case is given by letting $\mathbf{y} = (1, 0, 0, \ldots, 0)$ and $a_{i1} = a_i$, $i = 1, \ldots, N$.

APPENDIX B: NASH EQUILIBRIA

The *Nash equilibria* $(\mathbf{x}^*, \mathbf{y}^*)$ of the interaction (game) (A, B) are those states in which all players can do no better by changing state; that is,

$$\mathbf{x}^* A \mathbf{y}^* \geq \mathbf{x} A \mathbf{y}^* \text{ and } \mathbf{y}^* B \mathbf{x}^* \geq \mathbf{y} B \mathbf{x}^* , \tag{B1}$$

for all $(\mathbf{x}, \mathbf{y}) \in \Delta_X \times \Delta_Y$. If they exist in the interior, the solutions of the following simultaneous equations are Nash equilibria:

$$\begin{aligned} (A\mathbf{y})_i &= A(\mathbf{y})_1 \text{ and } (B\mathbf{x})_j = (B\mathbf{x})_1 \\ &\iff (A\mathbf{y})_i - \mathbf{x}A\mathbf{y} = 0 \text{ and } (B\mathbf{x})_j - \mathbf{y}B\mathbf{x} = 0 , \end{aligned} \tag{B2}$$

where $\Sigma_n^N x_n = \Sigma_m^M y_m = 1$.

It is known that $N = M$ is a necessary condition for the existence of a unique Nash equilibrium in the interior of Δ. With $N = M$ in the perfect memory case ($\alpha_X = \alpha_Y = 0$), the unique Nash equilibrium, if it exists, is the fixed point given by the intersection of the x- and y-nullclines of Eqs. (20).

This Nash equilibrium is not asymptotically stable, but the time average of trajectories converges to it. To see this, suppose that $x_i(t) > \delta$ for all t sufficiently large, we have

$$\begin{aligned} \frac{d}{dt}(\log x_i) &= \frac{\dot{x}_i}{x_i} = (A\mathbf{y})_i - \mathbf{x}A\mathbf{y} , \\ \frac{d}{dt}(\log y_i) &= \frac{\dot{y}_i}{y_i} = (B\mathbf{x})_i - \mathbf{y}B\mathbf{x} . \end{aligned} \tag{B3}$$

Integrating the both sides from 0 to T and dividing by T, we get

$$\begin{aligned} \frac{\log x_i(T) - \log x_i(0)}{T} &= \Sigma_j a_{ij} \overline{y}_i - S_A , \\ \frac{\log y_i(T) - \log y_i(0)}{T} &= \Sigma_j b_{ij} \overline{x}_i - S_B , \end{aligned} \tag{B4}$$

where $\overline{x}_i = T^{-1}\int_0^T x_i dt$, $\overline{y}_i = T^{-1}\int_0^T y_i dt$, $S_A = T^{-1}\int_0^T \mathbf{x}A\mathbf{y} dt$, and $S_B = T^{-1}\int_0^T \mathbf{y}B\mathbf{x} dt$. Letting $T \to \infty$, the left-hand sides converge to 0. Thus, $\overline{\mathbf{x}}$ and $\overline{\mathbf{y}}$ are a solution of Eqs. (B3). (This proof follows Ref. [42].)

APPENDIX C: HAMILTONIAN DYNAMICS

Consider a game (A, B) that admits an interior Nash equilibrium $(\mathbf{x}^*, \mathbf{y}^*)$ and is zero-sum ($B = -A^T$), then,

$$E = \beta_X^{-1} D(\mathbf{x}^* \parallel \mathbf{x}) + \beta_Y^{-1} D(\mathbf{y}^* \parallel \mathbf{y}) \tag{C1}$$

is a constant of the motion. This follows by direct calculation:

$$\begin{aligned} \frac{dE}{dt} &= -\frac{1}{\beta_X}\Sigma_k^N \frac{\dot{x}_k}{x_k} - \frac{1}{\beta_Y}\Sigma_k^M \frac{\dot{y}_k}{y_k} \\ &= -(\mathbf{x}^* A\mathbf{y} - \mathbf{x}A\mathbf{y}) - (\mathbf{y}^* B\mathbf{x} - \mathbf{y}B\mathbf{x}) \\ &= (\mathbf{x}^* - \mathbf{x})A(\mathbf{y}^* - \mathbf{y}) + (\mathbf{y}^* - \mathbf{y})B(\mathbf{x}^* - \mathbf{x}) \\ &= 0 . \end{aligned} \tag{C2}$$

(C3)

This holds for any number of agents. Give the agents equal numbers of actions ($N = M$) and set α to zero (perfect memory) and make all βs finite and positive.

Then the adaptive dynamics is Hamiltonian in the information space $\mathbf{U} = (\mathbf{u}, \mathbf{v})$ with the above constant of motion E and has Poisson structure J:

$$J = \begin{pmatrix} O & P \\ -P^T & O \end{pmatrix} , \tag{C4}$$

with $P = -\beta_X \beta_Y A$.
Proof:

$$\begin{aligned} \frac{\partial E}{\partial u_i} &= \frac{\partial}{\partial u_i} \left[\beta_X^{-1} \Sigma_n^N x_n^* \log x_n^* + \beta_Y^{-1} \Sigma_n^N y_n^* \log y_n^* \right. \\ & -\beta_X^{-1} \left(\Sigma_n^N x_n^* u_n - \log(\Sigma_n^N e^{-u_n}) \right) \\ &- \left. \beta_Y^{-1} \left(\Sigma_n^N y_n^* v_n - \log(\Sigma_n^N e^{-v_n}) \right) \right] \\ &= \beta_X^{-1} (x_i^* - \frac{e^{-u_i}}{\Sigma_n^N e^{-u_n}}) \\ &= \beta_X^{-1} (x_i^* - x_i) \ . \end{aligned} \tag{C5}$$

Similarly, $\partial E / \partial v_j = \beta_Y^{-1} (y_j^* - y_j)$. Since $(\mathbf{x}^*, \mathbf{y}^*)$ is an interior Nash equilibrium, with Eq. (17), $(A\mathbf{y}^*)_i = (B\mathbf{x}^*)_j = 0$. Thus,

$$\begin{aligned} A \frac{\partial E}{\partial \mathbf{v}} &= -\frac{1}{\beta_Y} A\mathbf{y} \ , \\ B \frac{\partial E}{\partial \mathbf{u}} &= -\frac{1}{\beta_X} B\mathbf{x} \ . \end{aligned} \tag{C6}$$

and

$$\begin{aligned} J \nabla_{\mathbf{U}} E &= \begin{bmatrix} -\beta_X \beta_Y A \frac{\partial E}{\partial \mathbf{v}} \\ -(-\beta_X \beta_Y A)^T \frac{\partial E}{\partial \mathbf{u}} \end{bmatrix} \\ &= \begin{bmatrix} -\beta_X A\mathbf{y} \\ -\beta_Y B\mathbf{x} \end{bmatrix} \\ &= \begin{bmatrix} \dot{\mathbf{u}} \\ \dot{\mathbf{v}} \end{bmatrix} \\ &= \dot{\mathbf{U}} \ . \end{aligned} \tag{C7}$$

We can transform $\mathbf{U} = (\mathbf{u}, \mathbf{v})$ to canonical coordinates $\mathbf{U}' = (\mathbf{p}, \mathbf{q})$:

$$\dot{\mathbf{U}}' = S \nabla_{\mathbf{U}'} E \ , \tag{C8}$$

with

$$S = \begin{pmatrix} O & -I \\ I & O \end{pmatrix} \tag{C9}$$

where I is an $N \times N$ identity matrix and with a linear transformation $\mathbf{U}' = M\mathbf{U}$ to the Hamiltonian form. □

APPENDIX D: REINFORCEMENT SCHEMES AND INTERACTION MATRICES

Here we give the reinforcement scheme interaction matrices for the constant-environment collectives investigated in the main text.

1. Matching Pennies

This game describes a non-transitive competition. Each agent chooses a coin, which turns up either heads (H) or tails (T). Agent X wins when the coins differ, otherwise agent Y wins. Table II gives the reinforcement scheme for the various possible plays. Note that the ϵs determine the size of the winner's rewards. When $\epsilon_X + \epsilon_Y = 0$, the game is zero-sum. The Nash equilibrium is $\mathbf{x}^* = \mathbf{y}^* = (1/2, 1/2)$.

X Y	r^X	r^Y
H H	$-\epsilon_X$	$-\epsilon_Y$
H T	ϵ_X	ϵ_Y
T H	ϵ_X	ϵ_Y
T T	$-\epsilon_X$	$-\epsilon_Y$

TABLE II: The two-person Matching Pennies game: $\epsilon_X \in (0.0, 1.0]$ and $\epsilon_Y \in [-1.0, 0.0)$.

Various extensions of Matching Pennies to more than two players are known. We give the *Even-Odd* game as an example for three agents X, Y, and Z in a collective. All flip a coin. Agents X and Y win when the number of heads is even, otherwise Z wins. Table III gives the reinforcement scheme. When the ϵs add to zero, the game is zero-sum. The unique mixed Nash equilibrium is $\mathbf{x}^* = \mathbf{y}^* = \mathbf{z}^* = (\frac{1}{2}, \frac{1}{2}, \frac{1}{2})$—the simplex center.

X Y Z	r^X	r^Y	r^Z
H H H	$-\epsilon_X$	$-\epsilon_Y$	$-\epsilon_Z$
H H T	ϵ_X	ϵ_Y	ϵ_Z
H T H	ϵ_X	ϵ_Y	ϵ_Z
H T T	$-\epsilon_X$	$-\epsilon_Y$	$-\epsilon_Z$
T H H	ϵ_X	ϵ_Y	ϵ_Z
T H T	$-\epsilon_X$	$-\epsilon_Y$	$-\epsilon_Z$
T T H	$-\epsilon_X$	$-\epsilon_Y$	$-\epsilon_Z$
T T T	ϵ_X	ϵ_Y	ϵ_Z

TABLE III: The three-player Even-Odd game: $\epsilon_X \in (0.0, 1.0]$ and $\epsilon_Y, \epsilon_Z \in [-1.0, 0.0)$.

2. Rock-Scissors-Paper

This game describes a non-transitive three-sided competition between two agents: rock (R) beats scissors (S), scissors beats paper (P), but paper beats rock. Table IV gives the reinforcement scheme. The ϵs here control the rewards for ties. When they add to zero, the game is zero-sum. The unique mixed Nash equilibrium is $\mathbf{x}^* = \mathbf{y}^* = (\frac{1}{3}, \frac{1}{3}, \frac{1}{3})$—again, the center of the simplex.

The extension of RSP interaction to three agents is straightforward. The reinforcement scheme is given in

X	Y	r^X	r^Y
R	R	ϵ_X	ϵ_Y
R	S	1	-1
R	P	-1	1
S	R	-1	1
S	S	ϵ_X	ϵ_Y
S	P	1	-1
P	R	1	-1
P	S	-1	1
P	P	ϵ_X	ϵ_Y

TABLE IV: The two-person Rock-Scissors-Paper game: $\epsilon_X, \epsilon_Y \in (-1.0, 1.0)$.

Table V. When $\epsilon_X + \epsilon_Y + \epsilon_Z = 0$, the game is zero-sum. The Nash equilibrium is $\mathbf{x}^* = \mathbf{y}^* = \mathbf{z}^* = (1/3, 1/3, 1/3)$.

X	Y	Z	r^X	r^Y	r^Z	X	Y	Z	r^X	r^Y	r^Z	X	Y	Z	r^X	r^Y	r^Z
R	R	R	ϵ_X	ϵ_Y	ϵ_Z	S	R	R	-2	1	1	P	R	R	2	-1	-1
R	R	S	1	1	-2	S	R	S	-1	2	-1	P	R	S	ϵ_X	ϵ_Y	ϵ_Z
R	R	P	-1	-1	2	S	R	P	ϵ_X	ϵ_Y	ϵ_Z	P	R	P	1	-2	1
R	S	R	1	-2	1	S	S	R	-1	-1	2	P	S	R	ϵ_X	ϵ_Y	ϵ_Z
R	S	S	2	-1	-1	S	S	S	ϵ_X	ϵ_Y	ϵ_Z	P	S	S	-2	1	1
R	S	P	ϵ_X	ϵ_Y	ϵ_Z	S	S	P	1	1	-2	P	S	P	-1	2	-1
R	P	R	-1	2	-1	S	P	R	ϵ_X	ϵ_Y	ϵ_Z	P	P	R	1	1	-2
R	P	S	ϵ_X	ϵ_Y	ϵ_Z	S	P	S	1	-2	1	P	P	S	-1	-1	2
R	P	P	-2	1	1	S	P	P	2	-1	-1	P	P	P	ϵ_X	ϵ_Y	ϵ_Z

TABLE V: The 3-person Rock-Scissors-Paper game: $\epsilon_X, \epsilon_Y, \epsilon_Z \in (-1.0, 1.0)$.

APPENDIX E: NETWORK INTERACTIONS

We can describe heterogeneous network interactions within our model. We give an example of a model for lattice interactions here. Agents $s = 1, 2, \ldots, S$ are on a spatial lattice: agent s interacts with agent $s-1$ through bi-matrices (A^s, B^{s-1}) and agent $s+1$ through (B^s, A^{s+1}). Each bi-matrix is 2×2. See Fig. 23.

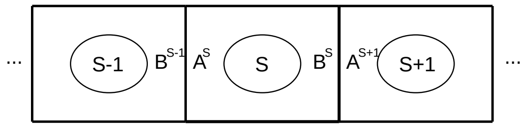

FIG. 23: Agent s interacts with agent $s-1$ through bi-matrices (A^s, B^{s-1}) and agent $s+1$ through (B^s, A^{s+1}).

Agents choose actions among the 2×2 action pairs for both the right and left neighboring agents. The action pairs are $(1,1), (1,2), (2,1), (2,2)$ and are weighted with probabilities $x_1, \ldots, x_4$. Inserting the interaction bi-matrices into the S-agent adaptive dynamics of Eq. (21) gives

$$\begin{aligned} \frac{\dot{x}_i^s}{x_i^s} &= \beta_s \left[(A^s \mathbf{x}^{s-1})_i - \mathbf{p}^s \cdot A^s \mathbf{x}^{s-1} \right. \\ &+ \left. (B^s \mathbf{x}^{s+1})_i - \mathbf{q}^s \cdot B^s \mathbf{x}^{s+1} \right] \\ &+ \alpha_s (-\log x_i^s - \Sigma_n^4 x_n^s \log x_n^s). \end{aligned} \tag{E1}$$

where $\Sigma x_i^s = 1$ and $\mathbf{p}^s = (x_1^s + x_2^s, x_3^s + x_4^s)$, $\mathbf{q}^s = (x_1^s + x_3^s, x_2^s + x_4^s)$.

In a similar way, arbitrary network interactions can be described by our adaptive dynamics.

[1] A. T. Winfree, *The Geometry of Biological Time* (Springer Verlag, 1980).
[2] J. Hofbauer and K. Sigmund, *Evolutionary Games and Population Dynamics* (Cambridge University Press: Cambridge, 1988).
[3] S. Camazine, J.-L. Deneubourg, N. R. Franks, J. Sneyd, G. Theraulaz, and E. Bonabeau, eds., *Self-Organization in Biological Systems* (Princeton University Press, Princeton, 2001).
[4] G. Bateson, *Steps to an ecology of mind : collected essays in anthropology, psychiatry, evolution, and epistemology* (Northvale, N.J. : Aronson, 1987).
[5] O. E. Rossler, Ann. NY Acad, Sci. **504**, 229 (1987).
[6] M. Taiji and T. Ikegami, Physica **D134**, 253 (1999).
[7] H. A. Simon, *The Sciences of the Artificial*, Karl Taylor Compton Lectures (MIT Press, Cambridge, 1996), first edition 1969.
[8] R. Brooks and L. Steele, *The Artificial Life Route to Artificial Intelligence: Building Embodied, Situated Agents* (Lawrence Erlbaum Associatges, New York, 1995).
[9] J. H. Holland, *Adaptation in Natural and Artificial Systems* (MIT Press, Cambridge, MA, 1992), second edition (First edition, 1975).
[10] T. Borgers and R. Sarin, J. Econ. Th. **77**, 1 (1997).
[11] D. Fudenberg and D. K. Levine, *Theory of Learning in Games* (MIT Press, 1998).
[12] J. von Neumann and O. Morgenstern, *Theory of Games and Economic Behavior* (Princeton University Press, 1944).
[13] P. D. Taylor and L. B. Jonker, Mathematical Biosciences **40**, 145 (1978).
[14] P. D. Taylor, J. Appl. Probability **16**, 76 (1979).
[15] J. W. Weibull, *Evolutionary Game Theory* (MIT Press, 1995).
[16] Y. Sato, E. Akiyama, and J. D. Farmer, Proc. Natl. Acad. Sci. USA **99**, 4748 (2002).
[17] Y. Sato and J. P. Crutchfield, Phys. Rev. **E67**, 015206(R) (2003).
[18] E. Akiyama and K. Kaneko, Phyisca D **147**, 221 (2000).
[19] E. Akiyama and K. Kaneko, Phyisca D **167**, 36 (2002).

[20] Y. Ito, J. Appl. Prob. **16**, 36 (1979).
[21] A. M. Turing, Phil. Trans. R. Soc. London **B 237**, 37 (1952).
[22] C. E. Shannon, *The Mathematical Theory of Communication* (The University of Illinois Press, 1949).
[23] W. S. McCulloch, Bulletin of Mathematical Biophysics **5**, 115 (1945).
[24] B. F. Skinner, *The behavior of organisms: an experimental analysis* (Appleton-Century / New York, 1938).
[25] D. O. Hebb, *The Organization of Behaviour* (John Wiley and Sons, New York, 1949).
[26] M. F. Norman, *Markov Processes and Learning Models* (Academic Press / New York, 1972).
[27] A. L. Samuel, IBM Journal on Research and Development **11**, 601 (1967).
[28] R. S. Sutton and A. G. Barto, *Reinforcement learning: An introduction* (MIT Press, 1998).
[29] D. Kahneman and A. Tversky, eds., *Choices, Values, and Frames* (Cambridge University Press, 2000).
[30] J. Aitchison, *The Statistical Analysis of Compositional Data* (Blackburn Press, Caldwell, New Jersey, 1986).
[31] T. M. Cover and J. A. Thomas, *Elements of Information Theory* (John Wiley and Sons, Inc., 1991).
[32] J. Hofbauer, J. Math. Biol. **34**, 675 (1996).
[33] A. M. Perelomov, *Integrable Systems of Classical Mechanics and Lie Algebras* (Birkhauser, 1990).
[34] H. Yoshida, Phys. Lett. **A150**, 262 (1990).
[35] T. Chawanya, Prog. Theo. Phys. **94**, 163 (1995).
[36] E. van Nimwegen, J. P. Crutchfield, and M. Mitchell, Theoret. Comp. Sci. **229**, 41 (1999).
[37] J. Hofbauer, V. Hutson, and G. T. Vickers, Nonlinear Analysis, Theory, Methods, and Applications **30**, 1235 (1997).
[38] R. Shaw, *The Dripping Faucet as a Model Chaotic System* (Aerial Press, Santa Cruz, California, 1984).
[39] J. P. Crutchfield and K. Young, Phys. Rev. Lett. **63**, 105 (1989).
[40] J. P. Crutchfield and C. R. Shalizi, Physical Review E **59**, 275 (1999).
[41] J. P. Crutchfield and D. P. Feldman, Chaos **13**, 25 (2003).
[42] P. Schuster, K. Sigmund, J. Hofbauer, and R. Wolff, Biol. Cybern. **40**, 1 (1981).